Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América

Dirección General de Estudios de Posgrado
Facultad de Ingeniería de Sistemas e Informática
Unidad de Posgrado

De la Crisis al Algoritmo: Modelos de Machine Learning para Predecir el Riesgo de Crédito en Perú

TESIS

Para optar el Grado Académico de Doctor en Ingeniería de Sistemas e Informática

AUTOR

Jomark Pablo NORIEGA ZAPATA

ASESOR

Dr. José Alfredo HERRERA QUISPE

Lima, Perú

2025



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

https://creativecommons.org/licenses/by-nc-sa/4.0/

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Noriega, J. (2025). *De la Crisis al Algoritmo: Modelos de Machine Learning para Predecir el Riesgo de Crédito en Perú*. [Tesis doctorado, Universidad Nacional Mayor de San Marcos, Facultad de Ingeniería de Sistemas e Informática, Unidad de Posgrado]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Jomark Pablo Noriega Zapata
Tipo de documento de identidad	DNI
Número de documento de identidad	00252325
URL de ORCID	https://orcid.org/0000-0002-8602-2766
Datos de asesor	
Nombres y apellidos	José Alfredo Herrera Quispe
Tipo de documento de identidad	DNI
Número de documento de identidad	40362859
URL de ORCID	https://orcid.org/0000-0002-8207-9714
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Pedro Martin Lezama Gonzales
Tipo de documento	DNI
Número de documento de identidad	09656793
Miembro del jurado 1	
Nombres y apellidos	Hugo David Calderon Vilca
Tipo de documento	DNI
Número de documento de identidad	01317748
Miembro del jurado 2	
Nombres y apellidos	Hugo Froilán Vega Huerta
Tipo de documento	DNI
Número de documento de identidad	06147737
Datos de investigación	

Línea de investigación	No aplica.
Grupo de investigación	No aplica.
Agencia de financiamiento	Sin financiamiento.
Ubicación geográfica de la investigación	Lugar: Universidad Nacional Mayor de San Marcos Latitud: -12.05540 Longitud: -77.08452
Año o rango de años en que se realizó la investigación	Mayo 2023 - diciembre 2024
URL de disciplinas OCDE	Ciencias de la computación https://purl.org/pe-repo/ocde/ford#1.02.01 Estadísticas, Probabilidad https://purl.org/pe-repo/ocde/ford#1.01.03 Econometría https://purl.org/pe-repo/ocde/ford#5.02.02

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
Universidad del Perú. Decana de América
Facultad de Ingeniería de Sistemas e Informática
Vicedecanato de Investigación y Posgrado
Unidad de Posgrado

ACTA DE SUSTENTACIÓN DE TESIS PARA OPTAR EL GRADO ACADÉMICO DE DOCTOR EN INGENIERÍA DE SISTEMAS E INFORMÁTICA

A los diez (10) días del mes de noviembre de 2025, siendo las 11:00 a.m., se reunieron en el Auditorio Alfredo Celso Alva Bravo, el Jurado de Tesis conformado por los siguientes docentes:

Dr. Pedro Martin Lezama Gonzales (Presidente)
Dr. Hugo David Calderon Vilca (Miembro)
Dr. Hugo Froilán Vega Huerta (Miembro)
Dr. José Alfredo Herrera Quispe (Miembro Asesor)

Se inició la Sustentación invitando al candidato a Doctor ***JOMARK PABLO NORIEGA ZAPATA****, para que realice la exposición oral de la tesis para optar el Grado Académico de Doctor en Ingeniería de Sistemas e Informática, siendo la Tesis intitulada:*

"De la Crisis al Algoritmo: Modelos de Machine Learning para Predecir el Riesgo de Crédito en Perú".

Concluida la exposición, los miembros del Jurado de Tesis procedieron a formular sus preguntas que fueron absueltas por el graduando; acto seguido se procedió a la evaluación correspondiente, habiendo obtenido la siguiente calificación:

Diecinueve (19) Muy Bueno

Por tanto, el presidente del Jurado, de acuerdo con el Reglamento General de Estudios de Posgrado, otorga al Magíster ***JOMARK PABLO NORIEGA ZAPATA*** *el Grado Académico de Doctor en Ingeniería de Sistemas e Informática.*

Siendo las 12:30 *horas, el presidente del Jurado de Tesis, da por concluido el acto académico de Sustentación de Tesis.*

Dr. Pedro Martin Lezama Gonzales
(Presidente)

Dr. Hugo David Calderon Vilca
(Miembro)

Dr. Hugo Froilán Vega Huerta
(Miembro)

Dr. José Alfredo Herrera Quispe
(Miembro Asesor)

UNIVERSIDAD NACIONAL MAYOR DE
SAN MARCOS
Universidad del Perú, Decana de América

CERTIFICADO DE SIMILITUD

Yo, JOSÉ ALFREDO HERRERA QUISPE en mi condición de asesor acreditado con

DICTAMEN N° 000042-2023-UPG-VDIP-FISI/UNMSM de la tesis cuyo título es

"De la Crisis al Algoritmo: Modelos de Machine Learning para Predecir el Riesgo de Crédito en Perú", presentado por el Magíster Jomark Pablo Noriega Zapata

para optar el grado de Doctor en Ingeniería de Sistemas e Informática,

CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante la plataforma PIDS, el documento evaluado cuenta con el 1% de similitud, nivel **PERMITIDO** para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional.**

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del grado correspondiente.

Lima, 2 de Mayo de 2025



Firma del Asesor

Nombres y Apellidos del asesor: JOSÉ ALFREDO HERRERA QUISPE

DNI 40362859

DEDICATORIA

A mi familia por siempre estar allí, a mi esposa Rosa e hijo Sebastián por siempre apoyarme; a mis compañeros y profesores del programa doctoral, por compartir sus conocimientos y camaradería; a mis compañeros de trabajo, especialmente a Estevan, Jorge, Marco, José, y Edgar; a mi equipo de colaboradores; a los estudiantes de pregrado que me asistieron en las investigaciones; a todas las personas que nos han apoyado como familia, directa o indirectamente, a realizar este sueño. Papá Zenón y mamá Irene, este reconocimiento es para ustedes.

AGRADECIMIENTOS

A mi asesor el Dr. José Herrera por darme su confianza y creer en este proyecto, muchas gracias por su acompañamiento, consejo y comprensión. Al Profesor Jesús Aguilar por sus valiosos aportes a esta investigación. Finalmente, a la Facultad de Ingeniería de Sistemas e Informática de la UNMSM por el excelente trabajo en alentar la divulgación científica.

Índice General

RESUMEN XIV
ABSTRACT XV

I Introducción **1**
1.1 Situación problemática 1
1.2 Formulación del problema 3
1.2.1 Problema general 3
1.2.2 Problemas específicos 3
1.3 Justificación teórica de la investigación 4
1.4 Justificación práctica de la investigación 4
1.5 Objetivos de la investigación 5
1.5.1 Objetivo general 5
1.5.2 Objetivos específicos 6
1.6 Hipótesis de la investigación 6
1.6.1 Hipótesis general 6
1.6.2 Hipótesis específicas 6

II Marco Teórico **7**
2.1 Marco Teórico 7
2.1.1 Fundamentos sobre riesgo crediticio y aprendizaje automático 7
2.1.2 Uso de los términos predicción y clasificación en la literatura sobre riesgo crediticio 8
2.1.3 Antecedentes de investigación 9
2.1.4 Estado del arte y benchmarking de modelos ante factores exógenos 17
2.2 Bases Teóricas 19
2.2.1 Incertidumbre 19
2.2.2 Riesgo de crédito. 19
2.2.3 Aprendizaje automático. 20
2.2.4 Balanceo de clases. 21
2.2.5 Matriz de confusión. 22

2.2.6 Análisis de grandes volúmenes de datos estructurados y no estructurados . . . 24
2.2.7 Aprendizaje automático y análisis de grandes volúmenes de datos estructurados y no estructurados . . . 25
2.2.8 Inteligencia computacional y análisis de grandes volúmenes de datos estructurados y no estructurados . . . 26
2.2.9 Modelos de *ML* . . . 26
2.2.10 Métodos de Interpretabilidad . . . 32
2.2.11 *Cross-Industry Standard Process for Data Mining* . . . 33
2.2.12 Test de Estacionariedad de *Dickey-Fuller* . . . 34
2.2.13 Test de Causalidad de *Granger* . . . 36
2.2.14 Prueba t de Student para muestras emparejadas . . . 37
2.3 Marco Sistémico . . . 38
2.4 Marco Ético . . . 40
2.5 Marco Filosófico. . . . 43
2.6 Glosario. . . . 44

III Metodología **46**
3.1 Tipos de investigación. . . . 46
3.1.1 Nivel de investigación. . . . 46
3.1.2 Tipo de la investigación. . . . 46
3.1.3 Diseño de la investigación. . . . 47
3.1.4 Método de la investigación . . . 47
3.2 *Dataset* y partición. . . . 48
3.2.1 *Dataset*. . . . 48
3.2.2 Partición. . . . 54
3.2.3 Técnica de muestreo de partición. . . . 54
3.3 Operacionalización de variables . . . 54
3.3.1 Matriz de operacionalización de variables . . . 55
3.4 Técnicas e instrumentos de recolección de datos. . . . 56
3.4.1 Técnicas . . . 56
3.4.2 Procedimientos de análisis y modelado . . . 56
3.4.3 Validación y confiabilidad de los modelos de *ML* . . . 60
3.4.4 Procedimientos . . . 60

IV Fase Experimental **64**
4.1 Análisis de *datasets* . . . 64
4.1.1 Ventajas del *dataset* seleccionado . . . 66
4.1.2 Limitaciones del *dataset* seleccionado . . . 67
4.1.3 Aplicación de *Data Augmentation* con *FE* . . . 67

4.2 Modelo Propuesto 75
4.3 Implementación de Modelos 77

V Resultados y discusión **82**
5.1 Resultados. 82
5.1.1 *FE* COVID 86
5.1.2 *FE* COVID con desplazamiento de serie temporal correspondiente a fallecidos, un periodo a la derecha 97
5.1.3 Evaluación de los niveles de significancia obtenidos para los modelos de *ML* aplicando el *FE* COVID 106
5.1.4 *FE* Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos 107
5.2 Respuesta a preguntas de investigación. 121
5.2.1 ¿Cómo influyó la pandemia de *COVID-19* , a través de factores como los casos positivos y las tasas de mortalidad, en los impagos de créditos en diferentes actividades económicas? . . 122
5.2.2 ¿En qué medida los indicadores del cambio climático, como las anomalías de temperatura y los bloqueos de carreteras debido al clima, afectan los patrones de morosidad crediticia? 128
5.2.3 ¿Cuál es la relación entre la morosidad crediticia y la convulsión social, considerando las perturbaciones de las actividades económicas y la estabilidad social? 132
5.2.4 ¿Cómo contribuyen los efectos combinados de *FE* (*COVID-19* , cambio climático y convulsión social) a las variaciones en la morosidad crediticia, y cuáles son los factores más influyentes?. 137
5.3 Desafíos Presentados en la Investigación 140
5.4 Discusión 142

VI Conclusiones y Recomendaciones **147**
6.1 Conclusiones 147
6.2 Recomendaciones 148

Referencias bibliográficas **151**

Anexos **160**

A Pseudocódigo de las funciones de evaluación de modelos *ML* y variables usadas **161**

B Código fuente para la aplicación de Test de Estacionariedad de Dickey-Fuller y Test de Causalidad de Granger **171**

C Código fuente para el proceso de enriquecimiento de Dataset con datos de *FE* 190

D Dataset enriquecido con datos derivados de *FE COVID-19*, convulsión social y anomalías por factores climáticos 196

E Código fuente del proceso de entrenamiento y validación para *FE COVID-19* 208

F Código fuente del proceso de entrenamiento y validación para *FE* convulsión social y anomalías por factores climáticos 223

G Código fuente del proceso validación de significancia estadística 240

H Artículo: Machine Learning for Credit Risk Prediction: A Systematic Literature Review 243

I Artículo: From Crisis to Algorithm: Credit Delinquency Prediction in Peru Under Critical External Factors Using Machine Learning 261

Lista de Tablas

Tabla 1. Tabla Estadística 14
Tabla 2. Comparativa de modelos aplicados a la predicción del riesgo crediticio 17
Tabla 3. Confusion Matrix with Evaluation Metrics 22
Tabla 4. Términos y Definiciones 44
Tabla 5. Descripción de variables en el conjunto de datos de series temporales sobre morosidad crediticia por actividad económica . . 49
Tabla 6. Resumen de variables utilizadas en los experimentos de modelado 50
Tabla 7. Estadísticas descriptivas del conjunto de datos crediticio utilizado para el modelado 52
Tabla 8. Descripción de variables en el conjunto de datos *COVID-19* Positive Cases 52
Tabla 9. Descripción de variables en el conjunto de datos *COVID-19* Death Cases 53
Tabla 10. Descripción de variables en el conjunto de datos Road Blockages (Social Unrest) 53
Tabla 11. Descripción de variables en el conjunto de datos Temperature Anomaly 53
Tabla 12. Matriz de operacionalización de variables 55
Tabla 13. Datasets. 58
Tabla 14. Matriz de operacionalización de variables. 59
Tabla 15. Resumen de variables empleadas en el modelado predictivo . . 66
Tabla 16. Estacionariedad de la Mora de Actividades Económicas 70
Tabla 17. Algoritmos evaluados y referencias en la literatura sobre riesgo crediticio 78
Tabla 18. Hiperparámetros de modelos clásicos y ensambladores implementados 80
Tabla 19. Hiperparámetros de modelos de redes neuronales profundas implementados 81
Tabla 20. Acrónimos de Escenarios Evaluados y sus resultados 85

Tabla 21. Promedio de *AUC* & *ACC* en 10 folds de modelos de *ML* con el *FE COVID-19* , incluyendo los créditos con apoyo gubernamental, según el escenario evaluado y los grupos de relevancia. 89

Tabla 22. Promedio de *AUC* & *ACC* en 10 folds de modelos de *ML* con el *FE COVID-19* , excluyendo los créditos con apoyo gubernamental, según el escenario evaluado y los grupos de relevancia. 93

Tabla 23. Modelos con mejoras estadísticamente significativas en las métricas *AUC* y *ACC* con el *FE COVID-19* , considerando escenarios con créditos con beneficios gubernamentales y sin créditos con beneficios gubernamentales y grupos de relevancia - escenarios. 96

Tabla 24. Promedio de *AUC* & *ACC* en 10 folds de modelos de *ML* con el *FE COVID-19* [*COVID MOV*], incluyendo los créditos con apoyo gubernamental, según el escenario evaluado y los grupos de relevancia. 98

Tabla 25. Promedio de *AUC* & *ACC* en 10 folds de modelos de *ML* con el *FE COVID-19* [*COVID MOV*], excluyendo los créditos con apoyo gubernamental, según el escenario evaluado y los grupos de relevancia. 102

Tabla 26. Modelos con mejoras estadísticamente significativas en las métricas *AUC* y *ACC* con el *FE COVID-19* [*COVID MOV*], considerando escenarios con créditos con beneficios gubernamentales y sin créditos con beneficios gubernamentales y grupos de relevancia - escenarios. 105

Tabla 27. *FE* Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos: Promedio de *AUC* en 10 *folds* de los modelos de *ML* según grupos de relevancia, incluyendo los créditos con apoyo gubernamental. 108

Tabla 28. *FE* Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos: Promedio de *ACC* en 10 *folds* de los modelos de *ML* según grupos de relevancia, incluyendo los créditos con apoyo gubernamental. 111

Tabla 29. *FE* Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos: Promedio de *AUC* en 10 *folds* de los modelos de *ML* según grupos de relevancia, excluyendo los créditos con apoyo gubernamental. 114

Tabla 30. *FE* Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos: Promedio de *ACC* en 10 *folds* de los modelos de *ML* según grupos de relevancia, excluyendo los créditos con apoyo gubernamental. 116
Tabla 31. *FE* Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos: Modelos que mostraron una mejora significativa en los indicadores, agrupados por relevancia, incluyendo los créditos con apoyo gubernamental. 119
Tabla 32. *FE* Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos: Modelos que mostraron una mejora significativa en los indicadores, agrupados por relevancia, excluyendo los créditos con apoyo gubernamental. 121
Tabla 33. Modelos más relevantes agrupados por Inc. CG y Exc. CG, destacando los mejores valores en los escenarios analizados (P, F, F+P) y la mejora alcanzada en comparación con *SFE*. 122
Tabla 34. Modelos más relevantes agrupados por Inc. CG y Exc. CG en los escenarios W, T y T+W, destacando las métricas *ACC* y *AUC* y las mejoras alcanzadas en comparación con *SFE*. 132
Tabla 35. Modelos más relevantes agrupados por Inc. CG y Exc. CG en los escenarios U, U+T, U+W y U+T+W, destacando las métricas *ACC* y *AUC* y las mejoras alcanzadas en comparación con *SFE*. 136
Tabla 36. Resultados *ACC* para los modelos con mayores valores en los escenarios seleccionados 138
Tabla 37. Resultados *AUC* para los modelos con mayores valores en los escenarios seleccionados 138
Tabla 38. Comparación de desempeño entre modelos tradicionales y de aprendizaje automático avanzados 142
Tabla 39. Comparación de contribuciones de esta tesis con estudios del estado del arte 145
Tabla 40. Pseudocódigo de las funciones de evaluación de modelos incluyendo análisis LIME, SHAP para escenarios con y sin FE . . . 161
Tabla 41. Variables por Escenario 166
Tabla 42. Descripción de registros de Dataset enriquecido 196

Lista de Figuras

Figura 1. Factores del riesgo de crédito. 20
Figura 2. Balanceo de datos aplicando *oversampling* y *undersampling*. . 22
Figura 3. Flujo de enriquecimiento del conjunto de datos con variables exógenas 51
Figura 4. Flujo de reducción y depuración del dataset: integración de series temporales, filtros aplicados y tamaños resultantes. 65
Figura 5. Series temporales de *FE* y actividades económicas 71
Figura 6. Relevancia de Actividad Económica. 72
Figura 7. Simplificación de afectación, a partir del Test de Granger. . . . 73
Figura 8. Actividades Económicas impactadas por *FE* COVID. Esta figura presenta métricas de impacto en el riesgo crediticio derivadas del contexto pandémico, utilizando las siguientes notaciones: **F** = Fallecimientos por *COVID-19*, **P** = Casos positivos de *COVID-19*, **F+P** = Impacto combinado de fallecimientos y casos positivos por *COVID-19*, **F+Mov** = Impacto de fallecimientos movido un periodo a la derecha. 74
Figura 9. Actividades Económicas impactadas por *FE* Convulsión Social, Anomalía de Temperatura y Bloqueo de vías por factor Climático. Esta figura presenta métricas de impacto en el riesgo crediticio derivadas de FE, utilizando las siguientes notaciones: **U** = Convulsión social, **T** = Anomalía de temperatura, **W** = Bloqueos viales por condiciones climáticas, **U+T+W** = Efecto combinado de convulsión social, anomalía de temperatura y bloqueos viales. 75

Figura 10. Modelo *CRISP-DM* para el análisis de *FE*. Este diagrama ilustra cómo las fases del modelo CRISP-DM estructuraron el flujo experimental, desde la comprensión y preparación de los datos —incluyendo la alineación mensual e integración de variables exógenas en series temporales— hasta el modelado basado en escenarios (con y sin factores externos) y la aplicación de técnicas de inteligencia artificial explicable (SHAP y LIME) durante la fase de evaluación. Las variables `DíasMora_MMYYYY` y `ImpactoFactorExterno_MMYYYY` fueron construidas para permitir un análisis temporal del comportamiento crediticio y su relación con shocks externos validados en distintos sectores económicos. 76

Figura 11. Periodos de evaluación de series temporales 83

Figura 12. Comparación del promedio en 10 folds de métricas *AUC* y *ACC* en modelos de *ML* con el *FE COVID-19*, incluyendo créditos con apoyo gubernamental, al 100 % de las actividades económicas. 90

Figura 13. Comparación del promedio en 10 folds de métricas *AUC* y *ACC* en modelos de *ML* con el *FE COVID-19*, excluyendo créditos con apoyo gubernamental, al 100 % de las actividades económicas. 94

Figura 14. Comparación del promedio en 10 folds de métricas *AUC* y *ACC* en modelos de *ML* con el *FE COVID-19 MOV*, incluyendo créditos con apoyo gubernamental, al 100 % de las actividades económicas. 99

Figura 15. Comparación del promedio en 10 folds de métricas *AUC* y *ACC* en modelos de *ML* con el *FE COVID-19 MOV*, excluyendo créditos con apoyo gubernamental, al 100 % de las actividades económicas. 103

Figura 16. Comparación del promedio en 10 folds de métricas *AUC* en modelos de *ML* con el *FE* Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos, incluyendo créditos con apoyo gubernamental, al 100 %, 99 % y 80 % de las actividades económicas. 109

Figura 17. Comparación del promedio en 10 folds de métricas *ACC* en modelos de *ML* con el *FE* Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos, incluyendo créditos con apoyo gubernamental, al 100 %, 99 % y 80 % de las actividades económicas. 112

Figura 18. Comparación del promedio en 10 folds de métricas *AUC* en modelos de *ML* con el *FE* Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos, excluyendo créditos con apoyo gubernamental, al 100 %, 99 % y 80 % de las actividades económicas. 115

Figura 19. Comparación del promedio en 10 folds de métricas *ACC* en modelos de *ML* con el *FE* Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos, excluyendo créditos con apoyo gubernamental, al 100 %, 99 % y 80 % de las actividades económicas. 117

Figura 20. Top 10 factores entre los escenarios *SFE* y F+P en **XNN**: (**a**) *SHAP*. (**b**) *LIME*. Leyenda: F = Casos confirmados de COVID-19; P = Fallecidos por COVID-19; F+P = Escenarios combinados; F+Mov = Movimiento de F un periodo a la derecha; WGB_CR = Con crédito respaldado por el gobierno; WOGB_CR = Sin crédito respaldado por el gobierno. 125

Figura 21. Top 10 factores entre los escenarios *SFE* y F+P en **CNN**: (**a**) *SHAP*. (**b**) *LIME*. Leyenda: F = Casos confirmados de COVID-19; P = Fallecidos por COVID-19; F+P = Escenarios combinados; F+Mov = Movimiento de F un periodo a la derecha; WGB_CR = Con crédito respaldado por el gobierno; WOGB_CR = Sin crédito respaldado por el gobierno. 126

Figura 22. Top 10 factores entre los escenarios *SFE* y T+W: (**a**) *SHAP* en XGB. (**b**) *LIME* en *XGB*. Leyenda: SFE = Escenario sin FE; T = Anomalía de temperatura; W = Bloqueo por condiciones climáticas; T+W = Escenarios combinados; WGB_CR = Con crédito respaldado por el gobierno; WOGB_CR = Sin crédito respaldado por el gobierno. 130

Figura 23. *SHAP* XGB y *CNN*: Factores determinantes en escenarios de convulsión social. Leyenda: U = Convulsión social, representada por eventos de protesta civil e inestabilidad política registrados en 2023; WGB_CR = Con crédito respaldado por el gobierno; WOGB_CR = Sin crédito respaldado por el gobierno. . 134

Figura 24. SHAP XGB, *CNN* y XNN: Factores determinantes en escenarios *U+T+W*. Leyenda: U = Convulsión social (por ejemplo, protestas, bloqueos de vías); T = Anomalías de temperatura; W = Bloqueos por condiciones climáticas; U+T+W = Escenario combinado que incluye los tres factores externos; WOGB_CR = Cartera de riesgo crediticio excluyendo créditos respaldados por el gobierno. 139

RESUMEN

La predicción del riesgo crediticio en economías emergentes exige cada vez más la integración de factores externos (FE) fuera del control de los prestatarios. Esta tesis propone una metodología basada en escenarios que incorpora variables críticas —seleccionadas por su impacto demostrado en la estabilidad financiera— como la severidad del *COVID-19* (número de contagios y fallecidos), anomalías climáticas (temperaturas extremas y bloqueos viales por motivos climáticos) y convulsión social (bloqueos viales generados por factores humanos) en modelos de aprendizaje automático (ML) aplicados a la morosidad crediticia. Estos factores fueron incorporados como indicadores exógenos en modelos de *ML*, a partir de series temporales nacionales y sectoriales validadas mediante pruebas de estacionariedad (Dickey-Fuller) y causalidad (Granger). El análisis se desarrolló bajo el marco *CRISP-DM* y utilizó técnicas de explicabilidad como *SHAP* y *LIME*. Se entrenaron y evaluaron modelos sobre más de 8.2 millones de registros, incluyendo 367,000 créditos individuales otorgados por una institución financiera regulada en Perú (FMOD) entre enero de 2020 y septiembre de 2023. Los resultados muestran que la inclusión de *FE* mejoró el desempeño de los modelos *CNN* y *XGB* en hasta 7 puntos porcentuales en métricas *AUC* y *ACC*, especialmente en actividades económicas de alta sensibilidad. Destaca el escenario *COVID MOV*, en el cual la curva de mortalidad fue desplazada un mes hacia adelante, permitiendo capturar su impacto diferido y modelar el desfase temporal entre el *shock* sanitario y su efecto en la morosidad. Asimismo, los escenarios compuestos (por ejemplo, clima más protestas) intensificaron el riesgo de incumplimiento entre un 10 % y 18 % respecto a los escenarios sin factores externos, principalmente en carteras sin respaldo estatal. Esta investigación aporta un marco replicable y escalable para modelar riesgos crediticios ante *shocks* sistémicos, con aplicaciones directas para reguladores, gestores de riesgo e instituciones financieras que operan en contextos volátiles.

Palabras clave: loan; credit risk; prediction; machine learning; external factors.

ABSTRACT

In predicting credit risk in emerging economies, we increasingly recognize the need to integrate external factors (EF) beyond borrowers' control. In this thesis, we propose a scenario-based methodology that incorporates critical variables —selected for their proven impact on financial stability— such as the severity of *COVID-19* (number of confirmed cases and deaths), climate anomalies (extreme temperatures and weather-related road blockages), and social unrest (human-induced road blockages) into machine learning (ML) models for credit delinquency. We include these factors as exogenous indicators in *ML* models, derived from national and sectoral time series validated through stationarity tests (Dickey-Fuller) and causality analysis (Granger). We conduct the analysis within the *CRISP-DM* framework and employ explainability techniques such as *SHAP* and *LIME*. We train and evaluate models on more than 8.2 million records, including 367,000 individual loans granted by a regulated financial institution in Peru (FMOD) between January 2020 and September 2023. Our results show that incorporating *EF* improves the performance of *CNN* and *XGB* models by up to 7 percentage points in *AUC* and *ACC* metrics, particularly in highly sensitive economic activities. Notably, in the *COVID MOV* scenario —where we shift the mortality curve one month forward— we effectively capture its delayed impact and model the temporal lag between the health shock and its effect on delinquency. Furthermore, we find that compound scenarios (e.g., climate plus protests) increase default risk by between 10 % and 18 % compared to scenarios without external factors, especially in portfolios without government backing. Through this research, we provide a replicable and scalable framework for modeling credit risk under systemic shocks, offering direct applications for regulators, risk managers, and financial institutions operating in volatile environments.

Keywords: loan; credit risk; prediction; machine learning; external factors.

Capítulo I

Introducción

1.1. Situación problemática

En la actualidad, la inteligencia artificial (IA), el aprendizaje automático (ML) y la digitalización de procesos constituyen pilares fundamentales en la transformación digital del sector financiero. Esta transformación se aceleró drásticamente durante la pandemia de *COVID-19* (Lombardo et al., 2022), obligando a las organizaciones a replantear sus modelos operativos y adoptar tecnologías avanzadas para enfrentar la incertidumbre. En particular, el sector financiero experimentó un crecimiento exponencial en el uso de créditos digitales, lo que generó volúmenes masivos de datos y una necesidad urgente de herramientas tecnológicas para la gestión del riesgo crediticio (Wen et al., 2021; Shih et al., 2022).

La evolución de los modelos de predicción de riesgo crediticio ha estado impulsada por innovaciones orientadas al análisis intensivo de datos (*Data-Driven Innovation, DDI*) y por la adopción del paradigma de *Big Data* (Hani et al., 2022; Feng, 2022; Z. Chen et al., 2020). A pesar de los avances en la integración de modelos de ML en los sistemas financieros, persisten limitaciones significativas, especialmente en contextos caracterizados por alta incertidumbre y disrupciones externas. Eventos como la crisis financiera global de 2008 (Lombardo et al., 2022), la pandemia de *COVID-19* (Telg et al., 2023), y los efectos del cambio climático (Ahmed, 2023), han revelado debilidades críticas en la evaluación del riesgo crediticio, debido a la incapacidad de los modelos tradicionales para incorporar de forma efectiva factores externos no lineales.

Una de las principales brechas en la literatura actual es la limitada consideración de los factores externos críticos (FE) en los modelos de predicción de riesgo crediticio. Eventos disruptivos como pandemias, anomalías climáticas y convulsión social pueden alterar profundamente los patrones de pago y morosidad de los prestatarios (Hani et al., 2022; C. Chen et al., 2022). No obstante, la mayoría de los modelos de ML han sido diseñados bajo el supuesto de estabilidad de los datos históricos, lo cual los hace poco

robustos frente a eventos exógenos (Pavitha y Sugave, 2024). Además, el uso creciente de modelos opacos o de tipo "caja negra" ha generado preocupación entre reguladores y responsables de toma de decisiones, reforzando la necesidad de desarrollar modelos explicables y transparentes (Nallakaruppan et al., 2024; Bastos y Matos, 2022).

En este contexto, la presente investigación plantea la incorporación de variables asociadas a *FE* en modelos de ML aplicados a la evaluación del riesgo crediticio, priorizando la explicabilidad de los resultados. Para ello, se integró datos históricos sobre *FE* mediante pruebas de causalidad y estacionariedad (Ye et al., 2015), con el objetivo de identificar patrones relevantes vinculados a la morosidad. Se utilizaron series temporales provenientes de fuentes públicas, complementadas con información crediticia interna de una institución financiera peruana, en adelante referida como "Financiera Modelo" (FMOD).

El caso de estudio en Perú resulta particularmente pertinente debido a su elevada exposición a múltiples *FE*. Durante la pandemia de *COVID-19*, el país registró la tasa de mortalidad más alta por cada 100,000 habitantes a nivel mundial (BBC News Mundo, 2020), lo cual impactó significativamente la estabilidad económica y el comportamiento financiero de hogares y empresas Ontón y Reilly (2024). Asimismo, fenómenos climáticos asociados al evento de El Niño generan desastres naturales recurrentes, afectando la infraestructura, interrumpiendo las cadenas de suministro y restringiendo el acceso a servicios financieros (Ahmed, 2023; Flori et al., 2021). A ello se suma la reciente convulsión social tras la crisis política de diciembre de 2022 (CNN en Español, 2024), la cual produjo un clima de incertidumbre política y volatilidad en los mercados de crédito.

Frente a esta problemática, la investigación propone analizar la relación entre los días de mora y la presencia de *FE*, mediante el uso de modelos de ML que incorporen técnicas de explicabilidad como *SHapley Additive Explanations* (SHAP) y *Local Interpretable Model-Agnostic Explanations* (LIME) (Pavitha y Sugave, 2024; Ma et al., 2023). La integración de estos enfoques busca no solo mejorar la precisión predictiva, sino también dotar a los modelos de mayor transparencia, facilitando su comprensión por parte de reguladores, analistas y gestores financieros.

Adicionalmente, la investigación plantea una metodología replicable orientada a responder dos interrogantes clave: ¿cómo afectan los distintos *FE* la estabilidad del sistema financiero?, y ¿qué variables tienen mayor peso explicativo en la predicción del riesgo crediticio? Esta aproximación tiene el potencial de generar nuevas estrategias de mitigación del riesgo en escenarios de incertidumbre extrema, fortaleciendo la resiliencia institucional frente a eventos disruptivos.

Finalmente, el presente estudio contribuye al cuerpo de conocimiento mediante el desarrollo de un marco analítico que integra de forma sistemática los *FE* en modelos de predicción de riesgo crediticio, con énfasis en economías emergentes como la pe-

ruana. Si bien existen múltiples estudios sobre aplicaciones de ML en la evaluación crediticia, la mayoría de ellos se centran en variables financieras convencionales, sin considerar *shocks* sistémicos exógenos como pandemias, cambios climáticos o convulsiones sociales (Alzamora et al., 2022; Elkhayat y ElBannan, 2018; Madeira, 2022) . La literatura reciente destaca la complejidad de los *shocks* compuestos —caracterizados por su ocurrencia simultánea y efectos no lineales— y su impacto en la conducta de los prestatarios, la calidad de los portafolios y la resiliencia sistémica (Kemp et al., 2022).

En respuesta a esta realidad, el estudio incorpora indicadores exógenos derivados de series temporales —como muertes por COVID-19, anomalías de temperatura y bloqueos viales climáticos— en una variedad de modelos de ML, evaluando tanto la precisión como la interpretabilidad de los resultados. Se espera que este enfoque aporte herramientas prácticas para anticipar riesgos sistémicos y diseñar políticas de respuesta más efectivas en contextos complejos y dinámicos.

1.2. Formulación del problema

1.2.1. Problema general

Existe incertidumbre en la predicción del riesgo crediticio ante la interacción de *FE* como la pandemia, el cambio climático y la convulsión social. En este contexto, surge la pregunta de investigación: *¿De qué manera la interacción de factores externos como la pandemia, el cambio climático y la convulsión social afecta la precisión y la robustez de los modelos de predicción del riesgo crediticio en economías emergentes como la peruana?*.

1.2.2. Problemas específicos

- ¿Cómo influyó la pandemia de *COVID-19*, a través de factores como los casos positivos y las tasas de mortalidad, en los impagos de créditos en diferentes actividades económicas?
- ¿En qué medida los indicadores del cambio climático, como las anomalías de temperatura y los bloqueos de carreteras debido al clima, afectan los patrones de morosidad crediticia?
- ¿Qué relación existe entre la convulsión social —caracterizada por perturbaciones en las actividades productivas y alteraciones en la estabilidad social— y la morosidad crediticia?

- ¿Cómo contribuyen los efectos combinados de factores externos (*COVID-19*, cambio climático y convulsión social) a las variaciones en la morosidad crediticia, y cuáles son los factores más influyentes?.

1.3. Justificación teórica de la investigación

Según Bueff et al. (2022), el uso de aprendizaje automático en entidades financieras presenta numerosos casos de uso orientados a la resolución de problemas de clasificación. Sin embargo, como destacan Hariri et al. (2019), persisten importantes desafíos, especialmente en lo que respecta a la incertidumbre en la evaluación de resultados cuando se procesan grandes volúmenes de datos. A estos desafíos se suma la necesidad de explicabilidad en los modelos de predicción del riesgo crediticio. En entornos financieros, donde las decisiones de crédito tienen impactos económicos y sociales significativos, es crucial comprender no solo la precisión de los modelos, sino también las razones detrás de sus predicciones. Modelos como *SHAP* y *LIME* han demostrado ser herramientas útiles para desglosar la influencia de cada variable en la morosidad crediticia (Pavitha y Sugave, 2024; Ma et al., 2023).

El Banco Mundial (Banco Mundial, 2025) resalta que el sector bancario desempeña un papel esencial en la mejora del bienestar social y en el crecimiento económico. En contextos de crisis, resulta imperativo que las instituciones financieras se reinventen, desarrollando productos y servicios innovadores que les permitan mantenerse competitivas, generar valor para los accionistas y mejorar la experiencia del cliente (Deloitte, 2020). Sin embargo, la implementación de modelos predictivos sin mecanismos de explicabilidad puede generar riesgos regulatorios y de adopción, dado que las decisiones financieras requieren transparencia y auditabilidad.

En este marco, nuestra investigación propone una metodología que integra *FE* —tales como pandemia, cambio climático y convulsión social— en modelos de aprendizaje automático, incorporando mecanismos de interpretabilidad que permitan a las instituciones financieras y a los reguladores comprender, auditar y confiar en las predicciones generadas por los algoritmos. De esta manera, se contribuye a cerrar la brecha entre la capacidad técnica de los modelos y las exigencias de transparencia en entornos financieros regulados.

1.4. Justificación práctica de la investigación

Las instituciones financieras enfrentan desafíos crecientes derivados de diversos *FE*, como lo son la convulsión social, la pandemia por *COVID-19* y los efectos del cambio climático. Estos eventos no solo impactan la estabilidad del mercado, sino que también

generan efectos colaterales que incrementan los riesgos financieros, como el aumento de la morosidad y la vulnerabilidad en el cumplimiento de las obligaciones crediticias (Pastor, 2024).

Convulsión Social: La agitación civil puede afectar negativamente los rendimientos del mercado bursátil (Espinosa-Méndez, 2022). Entre sus efectos se encuentran el aumento del endeudamiento, especialmente en los hogares de bajos ingresos, lo que incrementa el riesgo de morosidad (Madeira, 2022).

Pandemia de *COVID-19*: Este evento sin precedentes representó un desafío crítico para la economía global y la estabilidad financiera de los hogares, aumentando el riesgo de crédito y el incumplimiento de compromisos financieros (Madeira, 2022). Los efectos económicos de la pandemia se extendieron a través de las cadenas de suministro, afectando también la cadena de pagos (Teles et al., 2021; Ağca et al., 2023).

Factores Climáticos: Las instituciones financieras enfrentan un creciente riesgo de activos debido al cambio climático. Eventos como sequías, inundaciones y tormentas incrementan la probabilidad de incumplimientos crediticios, ya que los clientes tienden a priorizar necesidades básicas como la alimentación por sobre sus obligaciones financieras (Ahmed, 2023).

Además del impacto de estos *FE* en la morosidad crediticia, la falta de explicabilidad en los modelos de predicción representa un obstáculo clave para su adopción en el sector financiero. La toma de decisiones en instituciones bancarias requiere modelos que no solo sean precisos, sino también comprensibles y auditables por reguladores y stakeholders. En este sentido, la aplicación de enfoques explicativos, como *SHAP* y *LIME*, permitió identificar los *FE* que tienen mayor influencia en los impagos crediticios, facilitando la toma de decisiones informadas y la adaptación de políticas de mitigación de riesgos (Pavitha y Sugave, 2024; Ma et al., 2023).

La presente investigación podría contribuir a la implementación de modelos más robustos y explicables en la predicción del riesgo crediticio, permitiendo a las instituciones financieras desarrollar estrategias más efectivas para gestionar la morosidad en escenarios de incertidumbre extrema.

1.5. Objetivos de la investigación

1.5.1. Objetivo general

Reducir la incertidumbre en la predicción del riesgo de crédito ante la ocurrencia de *FE*, proporcionando explicabilidad en los resultados para mejorar la transparencia y la toma de decisiones en el sector financiero.

1.5.2. Objetivos específicos

- Analizar la relación entre los días de mora y la pandemia por *COVID-19*, considerando factores como los casos positivos y las tasas de mortalidad en diferentes actividades económicas.
- Evaluar el impacto del cambio climático en la morosidad crediticia, analizando indicadores como las anomalías de temperatura y los bloqueos de carreteras debido a eventos climáticos extremos.
- Examinar la relación entre la convulsión social —caracterizada por perturbaciones en las actividades productivas y alteraciones en la estabilidad social— y la morosidad crediticia.
- Determinar la influencia combinada de la pandemia por *COVID-19*, el cambio climático y la convulsión social en las variaciones de la morosidad crediticia e identificar los factores más influyentes.

1.6. Hipótesis de la investigación

1.6.1. Hipótesis general

El uso de analítica y *ML*, junto con técnicas de explicabilidad como *SHAP* y *LIME*, reduce la incertidumbre en la predicción del riesgo de crédito ante la ocurrencia de *FE*, mejorando la precisión, confiabilidad y transparencia de los modelos predictivos.

1.6.2. Hipótesis específicas

- El uso de analítica y *ML* contribuye a identificar la relación entre los días de mora y la pandemia por *COVID-19*, permitiendo interpretar su impacto a través de técnicas explicativas.
- La integración de analítica y *ML* permite evaluar la influencia del cambio climático en la morosidad crediticia, proporcionando modelos explicables y auditables.
- El uso de *ML* facilita la identificación de la relación entre la convulsión social y los días de mora, permitiendo visualizar las variables más influyentes en los impagos crediticios.
- La aplicación de modelos de *ML* con técnicas de explicabilidad permite comprender la interacción entre los efectos de la pandemia por *COVID-19*, el cambio climático y la convulsión social en la predicción del riesgo crediticio.

Capítulo II

Marco Teórico

2.1. Marco Teórico

2.1.1. Fundamentos sobre riesgo crediticio y aprendizaje automático

El riesgo crediticio es una preocupación central en el sistema financiero, ya que se refiere a la posibilidad de que un prestatario no cumpla con sus obligaciones contractuales de pago. Este tipo de riesgo impacta directamente en la estabilidad financiera de las instituciones, al afectar sus ingresos, provisiones, liquidez y solvencia. Tradicionalmente, su evaluación se ha basado en modelos estadísticos clásicos como la regresión logística, análisis discriminante y puntuaciones de crédito construidas a partir de reglas heurísticas.

Sin embargo, el proceso de digitalización del crédito —incluyendo canales virtuales de originación, uso de datos alternativos y plataformas automatizadas— ha transformado la naturaleza del riesgo y ha incrementado exponencialmente la cantidad, variedad y velocidad de los datos disponibles. En este nuevo contexto, los enfoques tradicionales muestran limitaciones para capturar relaciones no lineales, dinámicas complejas y patrones ocultos en los datos. Esto ha motivado una transición hacia el uso de modelos de *ML* y *Deep Learning* (DL), que permiten extraer conocimiento útil desde grandes volúmenes de datos estructurados y no estructurados (Lombardo et al., 2022; Orlova, 2020).

Los modelos de *ML*/DL han demostrado ser eficaces en diversas tareas del ciclo crediticio, tales como:

- Clasificación de prestatarios como buenos o malos pagadores.
- Predicción del incumplimiento en plazos específicos.
- Detección de fraude y comportamiento anómalo en solicitudes y transacciones.

- Segmentación y perfilamiento de clientes para personalización de productos.
- Asignación dinámica de límites de crédito y tasas de interés.

Entre las técnicas más utilizadas se encuentran los árboles de decisión, *Random Forest*, *Gradient Boosting*, *Support Vector Machines* (*SVM*), redes neuronales artificiales (*ANN*), modelos de ensamblaje y redes profundas. Estas técnicas superan a los modelos tradicionales en muchos casos, especialmente cuando se enfrentan a datos de alta dimensionalidad, correlaciones complejas y fuentes no convencionales (como texto, imágenes, redes sociales o variables macroeconómicas).
No obstante, su aplicación conlleva desafíos significativos. Algunos de los más relevantes incluyen:

- El **desequilibrio de clases**, es típico en datasets donde los incumplimientos son poco frecuentes.
- La **explicabilidad**, dado que muchos modelos (especialmente los *DL*) funcionan como 'cajas negras' y dificultan la interpretación por parte de analistas y reguladores.
- El **sobreajuste** es riesgo que se acentúa con modelos altamente complejos y poca generalización.
- La dependencia de **datos estructurados y limpios**, lo cual puede ser un obstáculo en países o instituciones con baja calidad de datos.

Recientemente, se han desarrollado enfoques que buscan equilibrar precisión y transparencia, como el uso de técnicas de *Explainable AI* (XAI) —entre ellas SHAP y LIME—, así como métodos de regularización y validación cruzada para mitigar el sobreajuste (Fang et al., 2021; Bueff et al., 2022; Imteaj y Amini, 2022).
Estos avances permiten fortalecer los sistemas de análisis de riesgo, dotándolos de capacidades adaptativas, resiliencia ante cambios abruptos y mayor confianza para su aplicación en contextos reales de toma de decisiones.

2.1.2. Uso de los términos predicción y clasificación en la literatura sobre riesgo crediticio

En la literatura sobre aprendizaje automático y riesgo crediticio, los términos *predicción* y *clasificación* son frecuentemente utilizados de manera intercambiable para describir tareas que involucran la estimación de una variable objetivo a partir de un conjunto de variables predictoras. Aunque desde un punto de vista técnico la predicción

puede referirse tanto a problemas de regresión (variable continua) como de clasificación (variable categórica), diversos estudios recientes han adoptado el término *predicción* incluso cuando la variable dependiente es categórica. Por ejemplo, Imteaj y Amini (2022) y Aranha y Bolar (2023) utilizan modelos de clasificación para estimar la probabilidad de incumplimiento crediticio y se refieren a esta tarea como *prediction* en sus investigaciones. Jiang (2022) introduce técnicas de regularización para mejorar la predicción de defaults con variables binarias, mientras que Kriebel y Stitz (2022) desarrollan modelos de deep learning sobre datos textuales para la predicción de impagos en plataformas de préstamos P2P. Asimismo, Alam et al. (2020) y Z. Chen et al. (2020) abordan explícitamente problemas de clasificación bajo el término *prediction*.
Este uso extendido en la literatura respalda la decisión de esta investigación de emplear el término "predicción" en el título y objetivos, a pesar de que la metodología implementada se basa en modelos de clasificación supervisada. Esta precisión terminológica permite alinear el trabajo con las prácticas vigentes en la comunidad científica y evitar ambigüedades para los lectores.

2.1.3. Antecedentes de investigación

Lombardo et al. (2022), afirman que el desarrollo de los créditos en línea está generando un enorme volumen de información que, al ser aprovechada a través del análisis de grandes volúmenes de datos estructurados y no estructurados, se utiliza para diseñar nuevos productos, modelos de aprendizaje automático y evaluación de riesgo crediticio. En consecuencia, en escenarios de aumento de demanda, los riesgos crediticios aumentan considerablemente, y de forma no lineal considerando el nivel de riesgo, tasa y términos del crédito (Orlova, 2020). En ese mismo sentido, se tiene la expectativa de aumento de fraudes en los siguientes años (Valavan y Rita, 2023). Otro problema para considerar es la consistencia de la información que se registra en las diferentes etapas del proceso (Mbona y Eloff, 2022; C. Zhang et al., 2022).
Según Bueff et al. (2022), el uso de *IA*, ya es una realidad en las instituciones financieras, pero sostiene que todavía no existen muchos casos de uso, principalmente debido a la dificultad que se tienen de explicar los modelos implementados, generalmente bajo el modelo de caja negra, para resolver problemas de clasificación, es decir, predecir por ejemplo a los buenos y malos pagadores. Estos modelos han presentado problemas especialmente en épocas difíciles como, por ejemplo: la crisis financiera del 2008. El principal producto ofrecido por las entidades financieras son los créditos, y de estos proviene su mayor ingreso, siendo el mayor riesgo, la falta de pago de estos (Orlova, 2020). Shih et al. (2022), afirman que, dado que los modelos de evaluación automático podrían tener sesgos dado que se construyen a partir de datos de créditos otorgados, se podrían estar rechazando buenos clientes, y por consiguiente renunciar a posibles

beneficios. Para Fang et al. (2021), los modelos de aprendizaje automático presentan desviaciones importantes, dado que tratan de representar la realidad con posible ruido. Según Jiang (2022), el problema del no pago es importante y podría generar pérdidas significativas a las entidades financieras. Este problema sigue siendo un reto para el aprendizaje automático, debido a la multicolinealidad existente en los datos de entrada, la alta correlación entre variables, la presencia de variables inútiles para la clasificación, el desequilibrio en los datos reales utilizados y el sobreajuste de la muestra. Rui (2022) sostiene que, determinar las características relevantes para el aprendizaje automático es un reto, considerando que el aprendizaje de redes neuronales (*NN*) en entrenamiento es de extremo a extremo a través de interacciones. Asimismo, propone que es posible complementarlo generando restricciones adicionales priorizando rasgos deseables específicos conocidos a priori, para reducir el esfuerzo en el entrenamiento. En este proceso permite beneficiar la tarea de aprendizaje, priorizando suavidad, factores de interés explicativo, en el dominio, control y el modelado generativo de características, esto demostraría la eficacia diseñar e imponer restricciones específicas durante el aprendizaje.
Kriebel y Stitz (2022) sostienen que el análisis de datos no estructurados en el proceso de otorgamiento de créditos, como el procesamiento de texto, podría mejorar significativamente las predicciones de incumplimiento crediticio, y es relevante en comparación con otros tipos de información. Es destacable la efectividad de combinación del *DL*, con técnicas como: fusión de información, frecuencia de palabras, determinación de sentimientos, uso de palabras relacionadas con fraude. Asimismo, sostiene que el modelo se podría enriquecer incluyendo en el análisis de texto los errores ortográficos. Según Imteaj y Amini (2022), en un escenario de guerra y post *COVID–19*, ha hecho que la asignación de calificación crediticia sea crucial. Asimismo, sostienen que los modelos de *IA* y *ML* dependen de una unidad central lo que podría generar riesgos de confidencialidad por transmitir información relacionada con el análisis. En ese sentido, propuso un modelo de procesamiento global federado (*PGF*). Este modelo considera agentes distribuidos, que procesan información usando recursos locales que no requerirían gran capacidad de cómputo y no compartirían información confidencial. Plantea las siguientes características:

- Resiliente al posible retraso de agentes locales, para tal fin usaría un procesamiento asíncrono.
- Considera escenarios: idéntica e independientemente distribuida (*IID*), y no idéntica e independientemente distribuido (*NIID*).
- Detección de calidad de los agentes como: benignos y maliciosos.
- Asignación recompensas y sanciones, según la calidad de los agentes.

Wu (2022), define a la banca por internet como un nuevo tipo de tecnología financiera, que tiene sus retos y riesgos asociados como: gestión caótica de la reputación, préstamos incobrables y engaño malicioso. Como modelo de predicción de riesgo crediticio propone fusionar modelos de *IA*: máquina de soporte virtual (*SMV*), árbol de decisión, BA, gradiente descendente y algoritmos integrados. Los resultados usando AUC y precisión de la matriz de confusión (PMC), fueron satisfactorios comparado con usar los modelos de forma independiente, las limitaciones para mejorar el modelo propuesto son:

- No se pueden explicar las relaciones internas desde una perspectiva general.
- Los modelos tienen que adaptarse a diferentes orígenes y procesos de extracción, incluyendo datos no estructurados como: datos sociales, de la capacidad y psicología de consumo, lo que podría enriquecerlo.
- Dificultad en el acceso a información histórica para aumentar la eficiencia en el aprendizaje de los algoritmos.

Para Hsin et al. (2022), es esencial para el negocio financiero el detectar el fraude, perfilar delitos e identificar las vulnerabilidades. Existen limitaciones en el acceso a la información por normas de privacidad, por ello propone adoptar generación de características sintéticas, para mejorar la predictibilidad de los modelos. Asimismo, propone:

- Agregar características de comportamiento y segmentación basados en casos no fraudulentos, aunque son difíciles de sintetizar automáticamente la causa y efecto.
- Los datos no homogéneos generan inestabilidad en modelos de *IA*: XGBoost (*XGBoost*) y *MGDL*, para mitigar se usa la prueba *Kolmogorov-Smirnov* (*KS*), aunque se obtuvieron mejores resultados de clasificación usando *SMV* y *BA*.
- El sobre entrenamiento y validación, podría conllevar engañosamente una percepción de solides del modelo.
- En el mundo real existen diferentes modos operandi, lo que podría afectar a los modelos que aplican *TSMS* para mitigar desbalanceo. Para reducir este riesgo propone usar sobre muestreo simple y redes neuronales generativas adversarias.

Para Gao y Balyan (2022), es una necesidad del sector financiero el desarrollar modelos de predicción de riesgos dadas las vulneraciones en el corto y largo plazo. Propone el algoritmo máquina de gradiente descendente ligero (*MGDL*), que permite analizar grandes cantidades de datos. Asimismo, sostiene que el modelo *MGDL* tiene un mejor

rendimiento para predicción de datos multidimensionales que *MSV* y regresión logística; y la capacidad del modelo *MGDL* mejora integrando el resultado con un modelo de regresión lineal, en comparación al uso del modelo único.

Lin et al. (2022) sostienen que, existe una necesidad de dotar a los organismos de supervisión financiera de china de mecanismos basados en *IA* para pronosticar el comportamiento de pago de créditos otorgados. Propone un modelo de memoria a largo plazo (*MLP*), constituye una evaluación basada en *IA* del cuadro de mando (*CM*), que se sustenta en datos estadísticos. El objetivo es reducir la dependencia de expertos, y sería una alerta temprana post prestamos basado en *BA*, con capacidades de procesamiento y flexibilidad de adaptarse a entornos económico-financieros. Comparado el modelo con *XGB*, obtuvo mejores índices de *AUC* y *KS*, con un costo de procesamiento menor. Asimismo, sostiene que se puede enriquecer el modelo agregando metadatos como imágenes y texto, procesarlos usando *DL*, lo que daría a la evaluación de riesgo crediticio característica multidimensional y multi enfoque.

Según Li y Li (2022), el riesgo de crédito tiene cuatro características principales: asimetría, acumulación, asistemática y endógena. Para las diferentes modalidades de préstamos que se implementan en internet, se evalúa el pronóstico del riesgo, basado en la eficacia de este. Incluye: variables de información básica, combinación de gestión y riesgo de crédito, el riesgo del inversionista, nivel de deuda del prestatario. Propone un modelo combinando: regresión logística, la medida basada en la holgura (*SBM*), y la selección y verificación de características se llevan a cabo a través del aprendizaje automático y las estadísticas, usando el siguiente proceso:

- Extraer las características de riesgo de eficiencia, y usando la estimación de límites de producción se calcula la eficiencia de riesgo.
- Examinar y verificar las funciones de eficiencia obtenido por el método *SBM* se usa como un nuevo índice para el entrenamiento y prueba del modelo logístico en conjunto a los índices iniciales relacionados.
- Validar la efectividad de la predicción del riesgo de crédito, comparando la predicción antes y después de agregar la función de eficiencia del riesgo.

La simulación indicó que el modelo logístico y *SBM* es más adecuado para la predicción del riesgo crediticio. Los resultados sugieren que se puede enriquecer el modelo de otorgamiento y cobro con información social del prestamista (Li y Li, 2022).

Para Assous (2022), es necesario determinar las características de eficiencia de bancos sauditas convencionales e islámicos, y propone el modelo *IA* de predicción que mejor se ajuste, validando empíricamente: *MSV*, detector de interacción automática de *chi-cuadrado* (*CHAID*), regresión lineal y red neuronal. Aplico los índices financieros se

agrupados en: rentabilidad, prácticas gerenciales, activos y préstamos, suficiencia de capital y liquidez, aplicando los siguientes pasos:

- Implementar el análisis de regresión, se define las variables independientes: rentabilidad, liquidez, calidad de activos, índices de gestión y de suficiencia de capital; y como variable dependiente: índice de eficiencia.
- Seleccionar características para diferentes modelos de predicción.
- Entrenar y validar los modelos seleccionados: *MSV*, *CHAID*, regresión lineal y red neuronal.
- Determinar que la eficiencia se ve afectada por la rentabilidad sobre activos (*ROA*), la liquidez total, y las prácticas de gestión medidas en virtud de la productividad de empleados.
- De los resultados el modelo con mejor predicción fue el modelo *CHAID* con la medida *R2* más eficiente.

Assous (2022), también identifican las siguientes estrategias de la alta gerencia:

- Aumentar depósitos, que constituyen los fondos para préstamos e inversiones.
- Maximizar ganancias, controlar niveles adecuados de liquidez, con prácticas eficientes y efectivas.
- Controlar el costo operativo, y uso de recursos maximizando la inversión de los accionistas.

Se puede enriquecer el modelo adicionando diferentes técnicas en la recopilación de datos, como: entrevistas, aplicar diferentes métodos estadísticos para calcular eficiencia, utilizar otros modelos predictivos (Assous, 2022).
Para Hemdan y Manjaiah (2022), las transacciones fraudulentas en tarjetas de crédito son un problema que cuenta miles de millones por año a empresas financieras, en ese sentido, es necesario implementar técnicas de predicción para detectarlas. Analiza los siguientes algoritmos: regresión logística (*RL*), análisis de discriminación lineal (*ADL*), vecinos cercanos (*VC*), árbol de decisión (*AD*), clasificador bayesiano ingenuo (*CBI*), Random Forrest (*RF*), y *AP*. Siendo *DL* el modelo que tuvo una mejor predicción ver **Tabla 1**. La principal característica corresponde a datos desbalanceados entre transacciones correctas y fraudulentas, para estudios futuros propone mitigar este sesgo aplicando submuestreo, sobre muestreo y otras técnicas.

Tabla 1
Tabla Estadística

Algoritmo	Precisión
Regresión logística	99.93
Análisis de discriminación lineal	99.95
Vecinos Cercanos	99.96
Árbol de decisión	99.92
Clasificador bayesiano ingenuo	97.88
Bosque aleatorio	99.96
Aprendizaje profundo	99.97

Fuente: Assous (2022).

Para C. Chen et al. (2022) sostienen que, en la actualidad existe una incertidumbre creciente, y sobre todo en una coyuntura de necesidad para reactivar economías, se requiere aumentar la capacidad de análisis frente a una posible demanda de créditos creciente. Para contribuir a tal fin propone un modelo *IA* interpretable globalmente, que se pueda sustentar las decisiones recomendadas, soportado por una red neuronal aditiva de dos capas, que incluye una herramienta de visualización que sustentaría como se llegó a la recomendación, pudiendo agruparse en: explicación de casos similares históricos, explicación de conjunto de características importantes, y explicación resumida y personalizada para cualquier decisión de préstamo recomendada. Asimismo, el autor plantea los siguientes hallazgos:

- El proceso de aprobación de préstamos necesita transparencia: la aprobación de un crédito es una decisión de riesgos que afecta tanto al cliente como a la organización y la sociedad. Esta a su vez se ve afectada por riesgos externos devenidos por ejemplo de crisis económicas, guerras, pandemias o desastres, por ello es importante para los modelos alimentarse de la mayor cantidad de información del medio, pues un evento mayor podría definirse como un cambio de valor de los datos.

- Evaluación de la competencia y lecciones aprendidas:

    - Existen escenarios en que un modelo de caja negra puede presumir como necesario cuando no lo es, por ejemplo, la asimetría entre valor social y las ganancias empresariales en la construcción de procesos automatizados.
    - La capacidad de explicar el resultado no siempre debe ser valorada por su simplicidad o facilidad de determinación, existen modelos aditivos factibles de ser interpretables y explicables pero complejos de validar manualmente.

- Las conclusiones principales son:
    - Se desarrollaron conceptos para desarrollar un modelo de apoyo a la toma de decisiones de otorgamiento de créditos, que proporciona interpretabilidad y transparencia en las decisiones de riesgo.
    - En los modelos interpretables para el proceso de generación de préstamos financieros no influyen en la calidad de la predicción.
    - La capacidad de interpretabilidad no simplifica el modelo.
    - La capacidad de explicar se puede resumir, dándole consistencia a los datos, y se puede usar en diferentes industrias como apoyo a la toma de decisiones.

Para Fitzpatrick y Mues (2021), la capacidad de predecir la rentabilidad potencial de créditos de persona a persona requiere evaluar un gran universo de casos y poder clasificarlas. Asimismo, existen vacíos para determinar cuáles son los más adecuados y si se ajustan a todos los casos, y en ese sentido se podrían generar alternativas para orientar inversiones rentables. Se aplicó el siguiente procedimiento:

- Seleccionar un conjunto representativo de métodos de regresión de diversa complejidad para predecir la rentabilidad.
- Tomar como función objetivo la maximización de la tasa interna de retorno (*TIR*).
- Los modelos evaluados se agruparon en dos categorías:
    - Individuales: se plantea usar como los métodos lineales a: GLM regularizado, lazo, regresión de cresta, mínimos cuadrados parciales y máquinas de vectores (*MSV*); y como métodos no lineales son: regresión adaptativa multivariante (*MARS*), red neuronal retro alimentada con 1 capa entrada, 40 capas ocultas,
    - De conjunto: se usan varios modelos como: bosques aleatorios, árboles embolsados, árboles potenciados por gradiente y cinco árboles apilados.
- Generar resultados con herramientas ad hoc a cada método, los métodos de predicción lineales tienen buenos resultados usando algunos criterios, aunque no funciona así en todos los casos. Los métodos de predicción en conjunto funcionan mejor que los métodos individuales.

Asimismo, se concluye que el uso de información basada en texto no obtiene mejores resultados de la determinación de beneficios. Los inversionistas de préstamos *P2P*,

podrían aumentar sus rendimientos aplicando de métodos de predicción lineales para estimar directamente la *TIR*. En lugar de utilizar variables alternativas dependientes, como el incumplimiento del préstamo. (Fitzpatrick y Mues, 2021).

Según Mann y Logeswaran (2021), se tiene incertidumbre sobre las causas que determinan la quiebra de las empresas. Asimismo, se evidencia que no se incluyen en la evaluación información de la industria que es relevante especialmente en escenarios donde el impacto de *FE* puede precipitar quiebras masivas como es el caso de las empresas que se declararon en quiebra a raíz de la pandemia del *COVID–19*. Si bien existen modelos predictivos para determinar la salud financiera todavía tienen limitaciones siendo la más relevante el no incluir el parámetro de riesgo industrial. es común encontrar evaluaciones a clientes de bajo riesgo pero que pertenecen a industria de alto riesgo. Sostiene que los modelos de predicción de quiebra se pueden mejorar con la inclusión de la variable: riesgo de la Industria dado que frecuentemente no es considerado en los modelos existentes, lo que demuestra empíricamente.

En ese sentido, considera las características que ahondan el impacto de este riesgo como el político. El autor, propone: validar la inclusión de los parámetros riesgo de la industria en otros casos de prueba con mayor número de observaciones y complejidad, que obtiene la industria bancaria. Considerando que el sector financiero atiende diversos segmentos, se podría replantear las características a evaluar en virtud de la actividad económica de los clientes, para determinar las características para predecir posibles quiebras.

Sahiq et al. (2022) sostiene que existen brechas en el estudio de las características de personas. Propone evaluar algoritmos de *AI* y *ML* para construir un modelo con el fin de identificar las principales causas de la quiebra personal, así como predecir su incidencia en el futuro, especialmente en una época post pandemia. Usa como marco de trabajo Cross-Industry Standard Process for Data Mining (*CRISP-DM*). Con la detección oportuna de los riesgos se podría gestionar la posible quiebra en etapas más tempranas con el objetivo de evitarlas. Las características con mayor relevancia para predecir las quiebras de consumidores son: la raza, el nivel educativo, sector empleo, préstamo personal, préstamo de estudios, microfinanzas y saldo de deuda.

Para procesar datos desequilibrados plantea la técnica de sobre muestreo minoritario sintético (*TSMS*) y modelo de regresión logística. El desempeño predictivo fue más eficiente con la muestra balanceada que con el original considerando que se mejora la sensibilidad y especificidad. En ese sentido plantea aplicar diferentes modelos basados en aprendizaje automático: *MSV*, *ANN* y *RF* . También plantea comparar técnicas de submuestreo y sobre muestreo para manejar conjuntos de datos desequilibrados.

2.1.4. Estado del arte y benchmarking de modelos ante factores exógenos

La literatura reciente ha explorado cómo choques externos como pandemias, eventos climáticos y convulsión social afectan el comportamiento crediticio. A continuación, se presenta una revisión estructurada organizada por tipo de shock, destacando enfoques técnicos y hallazgos relevantes ver Tabla 2.

Tabla 2

Comparativa de modelos aplicados a la predicción del riesgo crediticio

Ítem	Estudio	Modelos evaluados	Mejor modelo y métrica	Contexto
1	Imteaj y Amini (2022)	Federated Learning, Random Forest, SVM	Federated Learning (AUC = 0.91)	Evaluación de clientes en situación de riesgo financiero
2	Jiang (2022)	CNN, MLP, Muddling Label Regularization	CNN + MLR (Accuracy = 89.4 %)	Préstamos en China
3	Kriebel y Stitz (2022)	LSTM, BERT, Random Forest	BERT (F1-score = 0.84)	Préstamos P2P usando texto
4	Aranha y Bolar (2023)	ANN, Logistic Regression, Decision Tree	ANN (AUC = 0.94)	Préstamos bancarios en India
5	Z. Chen et al. (2020)	Ensemble Learning, Logistic Regression	Ensemble Learning (AUC = 0.92)	Bancarrota corporativa
6	Alam et al. (2020)	Random Forest, XGBoost, SVM	XGBoost (Accuracy = 87.5 %)	Tarjetas de crédito, datos desbalanceados
7	Calsin Enriquez y Vargas Salazar (2023)	VAR, MGM	VAR Model (Significancia 95 %)	Política monetaria y riesgo microfinanciero en Perú
8	Quintanilla Llanos (2024)	Panel Data, Dynamic Models	Panel Model (Morosidad correlacionada con desempleo)	Tarjetas de crédito en pandemia, Perú
9	Aguilar-Valenzuela y Vilca (2024)	Decision Tree, SVM, Random Forest, NN, KNN	Decision Tree, SVM (Rendimiento superior)	Algoritmos ML en entidades peruanas
10	Quinteros-Navarro y Rodriguez (2024)	SVM, NN, Ensemble (revisión sistemática)	SVM y Ensemble Models	Predicción de quiebra empresarial en Perú
11	Ontón y Reilly (2024)	Panel Data Analysis	Sustitución de deuda personal por microcréditos	Microempresas peruanas post-COVID
12	Geraldo-Campos et al. (2022)	Lasso, Ridge Regression	Lasso (RMSE = 0.357)	Reactiva Perú: riesgo crediticio
13	Noriega et al. (2023)	XGBoost, CNN, XNN + SHAP, LIME	XGBoost (AUC ACC ↑)	Modelos usados para pronosticar riesgo de crédito
14	Kemp et al. (2022)	Scenario Analysis, Risk Cascades Models	Integrated Catastrophe Assessment (Holistic approach)	Catástrofes climáticas y riesgo sistémico global
15	Espinosa-Méndez (2022)	Chang Model, Herding Behavior Analysis	Evidencia de Reverse Herding	Impacto de protestas sociales en mercados emergentes (Chile)
16	Madeira (2022)	Delinquency Models, Stress Testing	Modelos calibrados con encuestas (Debt Risk ↑ 2.7 %→4.5 %)	Impacto combinado de protestas y COVID en riesgo de deuda, Chile
17	Capasso et al. (2020)	Merton's Distance to Default, Panel Regression	Alta huella de carbono → Menor distancia al default	Riesgo climático y crédito corporativo en Europa

COVID-19 y el riesgo crediticio. La pandemia evidenció limitaciones en los modelos tradicionales de riesgo. Estudios como Telg et al. (2023); Aranha y Bolar (2023) resaltan cómo las intervenciones estatales alteraron los patrones de pago. En el caso peruano, Geraldo-Campos et al. (2022) aplicaron modelos de regresión Lasso y Ridge al programa Reactiva Perú, identificando que la regresión Lasso tuvo mejor capacidad predictiva del riesgo crediticio, especialmente en función de variables como sector eco-

nómico, entidad otorgante, monto cubierto y región geográfica. Por su parte, Noriega et al. (2023) muestran que la mortalidad COVID desplazada en el tiempo es el predictor más fuerte de morosidad crediticia. Además, destacan que la inclusión simultánea de factores COVID, clima y convulsión social mejora significativamente la precisión predictiva de los modelos, siendo CNN, XGB y XNN los modelos más adaptativos.

Riesgos climáticos. El cambio climático impone riesgos físicos y de transición. Eventos extremos han aumentado la mora y la necesidad de provisiones, especialmente en microfinanzas (Flori et al., 2021; Capasso et al., 2020; Ahmed, 2023). Se ha recomendado el uso de pruebas de estrés climático, VAR y modelos de propagación (Acharya et al., 2023; Calsin Enriquez y Vargas Salazar, 2023). Quintanilla Llanos (2024) destaca cómo las anomalías térmicas y bloqueos por condiciones climáticas afectan directamente la capacidad de pago de los prestatarios rurales, y cómo el diseño de coberturas financieras basadas en clima puede contribuir a mitigar estos riesgos.

Convulsión social e inestabilidad política. Estudios como Madeira (2022); Espinosa-Méndez (2022) muestran el impacto de la protesta y la crisis política sobre la mora y la confianza del inversor. En Perú, Ontón y Reilly (2024) identifican cambios de comportamiento de prestatarios y sustitución de créditos personales por programas garantizados. Quinteros-Navarro y Rodriguez (2024) advierten que las quiebras empresariales se aceleran en contextos de inestabilidad institucional, particularmente en sectores económicos ya vulnerables, y que el riesgo industrial debería incorporarse explícitamente en modelos de evaluación crediticia.

Choques compuestos. La investigación reciente reconoce la necesidad de modelos capaces de enfrentar choques múltiples e interrelacionados. Noriega et al. (2023) proponen una metodología basada en series temporales y causalidad de Granger para evaluar el efecto simultáneo de COVID, anomalías térmicas, bloqueos y convulsión social. Se recomienda integrar técnicas explicativas (SHAP, LIME), redes de interacción, modelos híbridos y mecanismos de alerta temprana (Kemp et al., 2022; Aguilar-Valenzuela y Vilca, 2024).

Brechas identificadas. La mayoría de los estudios académicos revisados no integra simultáneamente múltiples factores exógenos en los modelos de riesgo crediticio. Persiste una limitada incorporación de variables como eventos sanitarios, fenómenos climáticos extremos y conflictos sociales en conjunto. Esta tesis propone abordar dicho vacío mediante el uso de modelos de *ML* aplicados a series temporales exógenas, con el objetivo de evaluar su influencia en la precisión predictiva y la explicabilidad de los modelos (Noriega et al., 2023).

2.2. Bases Teóricas

En la presente sección, en primer lugar, se detalla los conceptos generales y enfoques de aprendizaje automático en grandes volúmenes de datos (Hariri et al., 2019). Luego se especifican los procesos que se siguen para el desarrollo de producto microfinancieros (Deloitte, 2020); se describe el método para evaluación *Cross Industry Standard Process for Data Mining* (IBM, 2020), y las herramientas a usar para la evaluación de FE.

2.2.1. Incertidumbre

La incertidumbre se define como la falta de certeza respecto a la ocurrencia de un evento (Real Academia Española, 2023). Es la condición que se genera a partir que no se cuenta con información completa para predecir lo que podría suceder en el futuro. La incertidumbre en análisis de datos se presenta en cada fase, desde la recopilación y normalización de datos, originado por el ruido introducido por diferentes actores, interesados o el entorno, de forma intencionada o no. Otros casos pueden deberse a enlaces perdidos o inconsistentes, diferentes tipos de origen, la naturaleza estructurada y no estructurada, como ocurre en los datos provenientes de redes sociales, y valores en atributos transcritos. (Hariri et al., 2019).

2.2.2. Riesgo de crédito.

El riesgo se define como la contingencia o posibilidad de no ocurrencia de un contrato seguro (Real Academia Española, 2023). El riesgo de crédito se entiende como el riesgo que sufre una entidad financiera derivado de la no devolución en plazo de los créditos concedidos a sus clientes"(Real Academia Española, 2023). Este se materializa cuando un cliente no cumple el compromiso de pago de un contrato de crédito y tiene su origen en dos factores fundamentales (ver Figura 1):

- Intención de pago: Relacionada con la determinación o voluntad del cliente de cumplir con sus obligaciones contraídas (Real Academia Española, 2023).
- Capacidad de pago: Vinculada a la habilidad del cliente para generar ingresos, derivada de su actividad económica directa o indirecta, del nivel de ingreso y de la capacidad de ahorro familiar (Noriega et al., 2023).



Figura 1
Factores del riesgo de crédito.
Fuente. Image Generator de *Chat-GPT.*

La Figura 1 representa visualmente estos dos ejes del riesgo crediticio mediante un contraste entre escenarios opuestos. El lado izquierdo, con un fondo oscuro, simboliza una *intención de pago deficiente*, reflejada en iconos asociados al incumplimiento, como retrasos en el pago, incremento del riesgo, pérdida de activos y gráficos descendentes. Estos elementos representan a un cliente con escaso compromiso financiero, desbalance en su flujo económico y un consumo no planificado.
El lado derecho, con un fondo claro, ilustra una *capacidad de pago sólida*. Se muestran actividades económicas estables (comercio, empleo formal), herramientas de gestión financiera (calculadora, gráficos de ingresos) y símbolos de liquidez (billetes, monedas), indicadores de ingresos constantes y adecuada planificación financiera.
El contraste visual enfatiza que un cliente puede tener la capacidad pero no la intención de pago, o la intención pero no la capacidad; en ambos casos, el riesgo crediticio se incrementa. La representación gráfica refuerza la idea de que la evaluación integral del riesgo requiere considerar simultáneamente la voluntad y la posibilidad real de pago.

2.2.3. Aprendizaje automático.

El aprendizaje automático “es la técnica que mejora el rendimiento de un sistema, aprendiendo de la experiencia a través de métodos de computación”. Al igual que los humanos los sistemas informáticos, también aprenden de la experiencia que en este caso esta embebida en los datos, estos son procesados a través de modelos generados

a partir de algoritmos de aprendizaje (Zhou, 2021).
Según Hutter et al. (2019), también se puede definir como el desarrollo de la capacidad en las computadoras de aprender sin ser programado explícitamente, aunque todavía existen brechas importantes para tal fin, pues son tareas rigurosas que requieren experiencia y conocimiento experto considerable, requiriendo todavía una importante programación explicita, relacionada al contexto. Modelos usados para la predicción y descubrimiento de patrones, pueden ser supervisados, no supervisados o ligeramente supervisados (Hariri et al., 2019; Rui, 2021).

2.2.4. Balanceo de clases.

El balanceo o equilibrio de clases es una estrategia empleada en el campo del aprendizaje automático para hacer frente al desequilibrio de clases en un conjunto de datos. El desequilibrio de clases se presenta cuando una o varias clases tienen una proporción considerablemente menor que las demás en el conjunto de datos, lo que puede provocar que el modelo tienda a predecir la clase mayoritaria y no logre capturar patrones en las clases minoritarias. Existen varias técnicas para balancear clases en un conjunto de datos, y cada una tiene sus propias ventajas y desventajas. Una técnica común es el submuestreo o *undersampling*, que consiste en reducir la cantidad de ejemplos en la clase mayoritaria hasta que se equilibren las clases. Por ejemplo, en el conjunto de datos de diagnóstico médico mencionado anteriormente, se podría reducir la cantidad de ejemplos de la clase sana hasta que tenga la misma cantidad que la clase enferma. Sin embargo, esta técnica puede llevar a la pérdida de información relevante y a la reducción del tamaño del *dataset*, lo que podría afectar el rendimiento del modelo.
Otra técnica es la sobre muestreo u *oversampling*, que consiste en aumentar la cantidad de ejemplos en la clase minoritaria hasta que se equilibren las clases. Por ejemplo, en el conjunto de datos de diagnóstico médico, se podría aumentar la cantidad de ejemplos de la clase enferma mediante la creación de ejemplos sintéticos o mediante la inclusión de nuevos ejemplos de pacientes enfermos. Sin embargo, esta técnica puede llevar a la generación de ejemplos sintéticos poco realistas y a la inclusión de datos redundantes. Finalmente, existe una técnica más avanzada llamada Synthetic Minority Oversampling Technique (*SMOTE*) que crea ejemplos sintéticos de la clase minoritaria al interpolar entre ejemplos cercanos en el espacio de características, generando datos similares, pero con pequeñas variaciones.

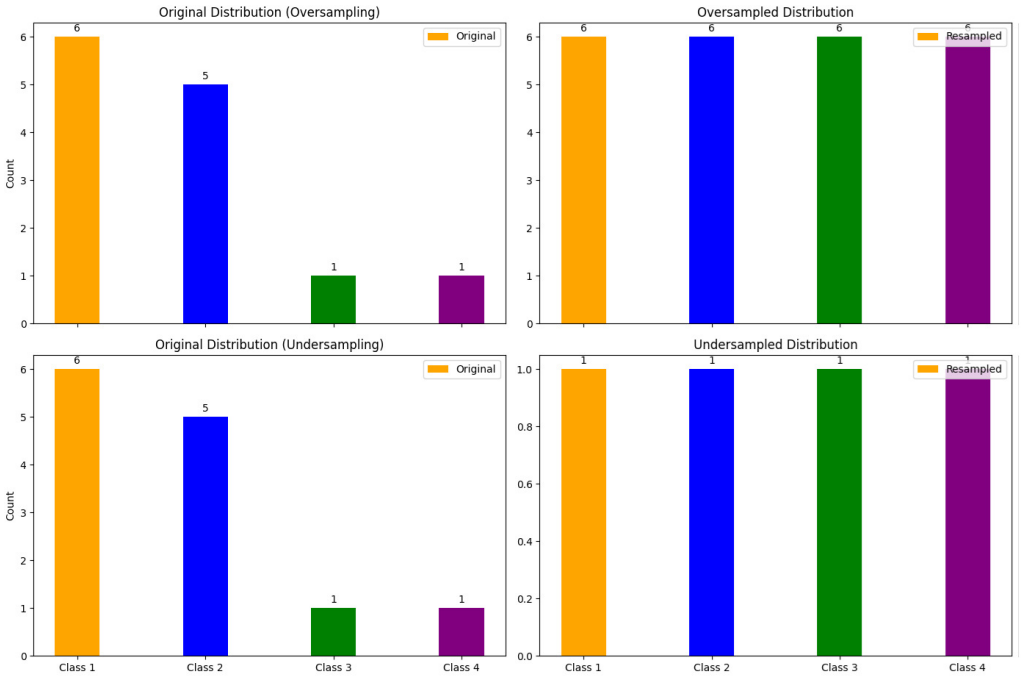


Figura 2
Balanceo de datos aplicando oversampling y undersampling.

2.2.5. Matriz de confusión.

Para los modelos de aprendizaje supervisado, la matriz de confusión representa visualmente el rendimiento de un algoritmo (ver Table 3). En las columnas se muestran las predicciones realizadas para cada clase, mientras que en las filas las clases reales. De esta forma, la matriz de confusión permite observar los aciertos y errores obtenidos por el modelo implementado (Zou et al., 2022; Fan et al., 2020).

Tabla 3
Confusion Matrix with Evaluation Metrics

Class	**Predicted Class**			
	Positive	**Negative**	**Metric**	**Formula**
Positive	True Positive (*TP*)	False Negative (*FN*)	Sensitivity *(Recall)*	$\frac{TP}{TP+FN}$
Negative	False Positive (*FP*)	True Negative (*TN*)	Specificity	$\frac{TN}{TN+FP}$
	Precision			$\frac{TP}{TP+FP}$
	Negative Predictive Value			$\frac{TN}{TN+FN}$
	Accuracy			$\frac{TP+TN}{TP+TN+FP+FN}$

Las intersecciones de esta matriz corresponden a:

Verdadero positivo.
En ingles *true positive* (*TP*), representa el número de predicciones correctas positivas, es decir, que corresponde a una identificación correcta como positiva.

Verdadero negativo.
En ingles *true negative* (*TN*), representa el número de predicciones correctas negativas, es decir, que corresponde a una identificación correcta como negativa.

Falso positivo.
En ingles *false positive* (*FP*), representa el número de predicciones incorrectas positivas, es decir, que corresponde a una identificación incorrecta como positiva, también se define como: *Error Tipo I*.

Falso negativo.
En ingles *true negative* (*FN*), representa el número de predicciones incorrectas negativas, es decir, que corresponde a una identificación incorrecta como negativa, también se define como: *Error Tipo II*.

Exactitud.
En inglés *Accuracy* (*ACC*), esta métrica representa la proximidad entre el resultado de una medición y el valor verdadero. Es representada como la proporción de resultados correctos dividida por el número total de casos, su fórmula es:

$$Accuracy = \frac{TN + TF}{(TP + TN + FP + FN)}$$

Donde:
TP: *True Positive*, o en español Verdadero Positivo.
TN: *True Negative*, o en español Verdadero Negativo.
FP: *False Positive*, o en español Falso Positivo.
FN: *False Negative*, o en español Falso Negativo.
Precisión.
En ingles *Accuracy*, esta métrica representa la proporción de casos positivos detectados, su fórmula asociada es la siguiente.

$$Precision = \frac{TP}{(TP + FP)}$$

Sensibilidad.
En inglés *Recall* o *Sensitivity*, esta métrica representa la proporción de casos positivos

que fueron correctamente identificados por el modelo, su fórmula es:

$$Recall = \frac{TP}{(TP + FN)}$$

Especificidad.
En inglés *Specificity*, esta métrica indica los casos negativos que el modelo ha clasificado correctamente; su fórmula es:

$$Recall = \frac{TN}{(TN + FP)}$$

F1-score.
Esta métrica conjuga y armoniza en una sola medida, la precisión y la sensibilidad, su fórmula es:

$$F1 - score = 2 * \frac{Recall * Precision}{(Recall + Precision)}$$

ROC AUC score.
El *ROC AUC score* (*AUC*) representa la relación entre la tasa de detección correcta de casos positivos (*TPR*), y la tasa de identificación errónea de casos negativos como positivos (*FPR*) del modelo, su fórmula es:

$$AUC = \frac{\sum Rangos(CP) - CP * \frac{CP+1}{2}}{CP + CN}$$

Donde:

$\sum Rangos(CP)$: Corresponden a la suma de los rangos de todos los ejemplos clasificados positivamente.
CP: Número de casos positivos en el *dataset*.
CN: Número de ejemplos casos en el *dataset*.

2.2.6. Análisis de grandes volúmenes de datos estructurados y no estructurados

Según Hariri et al. (2019), al análisis de grandes volúmenes de datos estructurados y no estructurados, también se le define como próximo límite de la innovación, productividad y competencia, sus características son:

- Volumen, la información contenida se mide en términos de:
  - *Petabyte (PB)* 10^{15}
  - *Exabyte (EB)* 10^{18}

    - *Zettabyte (ZB)* 10^{21}

- Variedad, la información contenida es estructurada, no estructurada o semi estructurada, la incertidumbre se manifiesta por ejemplo al convertir datos no estructurados a estructurados.
- Variedad, para la información se puede clasificar en tiempo real, casi tiempo real, incluye los grandes volúmenes de datos generados por sensores como internet de las cosas (*IdC*).
- Veracidad, se define así a la calidad de los datos, como estos se pueden clasificar en inconsistentes, ruidosos, ambiguos o incompletos; a su vez, la veracidad de estos puede ser: buena, mala e indefinida.
- Valor, esta se define como la utilidad de la información para la toma oportuna de decisiones.

Los datos emergentes tienen diferentes tipos de formatos y estructuras; esto ha evolucionado desde la última década del siglo pasado, en la que el volumen de datos creció con el uso de internet y se desarrollaron procesos estadísticos y de minería de datos para su procesamiento. Luego, con el crecimiento del comercio electrónico, las tecnologías continúan evolucionando para comprender las necesidades de usuario y desarrollar nuevas oportunidades comerciales, soportadas con el análisis web, sistemas sociales y las aplicaciones de colaboración (Mohanty et al., 2015).

2.2.7. Aprendizaje automático y análisis de grandes volúmenes de datos estructurados y no estructurados

Hariri et al. (2019) afirma que los métodos actuales tienen retos en escalabilidad y eficiencia para manejar grandes volúmenes de información, en tiempo real y en condiciones de incertidumbre. Se han propuesto soluciones como:

- Aprendizaje profundo: puede extraer y analizar patrones, tendencias no perceptibles de grandes cantidades de datos y fuentes, como imágenes, materiales, sonidos, etc. Los modelos actuales requieren un alto costo computacional.
- Aprendizaje por transferencia: tiene la capacidad reusar conocimientos adquiridos a contextos similares trasladando el conocimiento de un modelo a otro.
- Aprendizaje distribuido: podría afrontar la problemática de escalabilidad para el aprendizaje automático, puede realizar cálculos paralelos con datos distribuidos.

- Aprendizaje activo: usa algoritmos adaptativos para la recuperación de datos, ajusta de forma automática los parámetros para la recopilación, acelerando el proceso de aprendizaje.

2.2.8. Inteligencia computacional y análisis de grandes volúmenes de datos estructurados y no estructurados

Hariri et al. (2019) sostiene que la Inteligencia Computacional (*CI*) corresponde al conjunto de procesos computacionales inspirados en la naturaleza para el análisis de grandes volúmenes de datos estructurados y no estructurados. Los CI se han utilizado para análisis de datos complejos, análisis de incertidumbre y otros en los que los procesos tradicionales son insuficientes; estos procesos incluyen algoritmos evolutivos (*EA*), *ANN* y lógica difusa.

2.2.9. Modelos de *ML*

Redes Neuronales Convolucionales (*CNN*)

Son una arquitectura especializada en el procesamiento de datos con una estructura en forma de cuadrícula, como imágenes. Utilizan capas convolucionales para capturar características locales Merćep et al. (2020); Z. Chen et al. (2020).

Ventajas:

- Excelentes para la tarea de reconocimiento de imágenes y video.
- Requieren menos preprocesamiento de datos en comparación con otros modelos.
- Manejan bien datos desbalanceados.

Desventajas:

- Requieren grandes cantidades de datos y poder computacional.
- No se adaptan bien a datos faltantes sin preprocesamiento.

Explainable Boosting Machine (*EBM*)

Es un modelo de aprendizaje automático interpretable basado en boosting de Generalized Additive Models (*GAMs*). *EBM* ajusta de manera iterativa funciones de forma a cada característica, permitiendo un análisis interpretable y transparente de la influencia de cada variable Ma et al. (2023).

Ventajas:

- Ofrece alta interpretabilidad manteniendo un buen rendimiento predictivo.

- Permite a los usuarios comprender cómo cada variable impacta el resultado.
- Resistente a sobreajuste gracias al entrenamiento iterativo y regularización.
- Puede detectar relaciones no lineales sin perder interpretabilidad.
- Apto para entornos donde la transparencia en las decisiones es crucial, como salud y finanzas.

Desventajas:

- No captura interacciones complejas entre múltiples variables sin términos de interacción explícitos.
- Puede ser más lento de entrenar en comparación con modelos lineales simples.
- Menos flexible que modelos de *DL* en tareas altamente complejas.
- La interpretación puede ser difícil en datasets con muchas características correlacionadas.

Análisis Discriminante Lineal (*LDA*)

Es un clasificador supervisado que asume que las clases tienen distribuciones normales con una misma matriz de covarianzas. *LDA* busca proyectar los datos en un espacio de menor dimensión donde se maximiza la separabilidad entre clases Alam et al. (2020).

Ventajas:

- Es eficiente y fácil de interpretar.
- Funciona bien cuando las clases están bien separadas.
- Requiere menos datos de entrenamiento que otros algoritmos más complejos.
- Funciona bien con datos balanceados.

Desventajas:

- La suposición de normalidad y covarianzas iguales puede no ser realista.
- No maneja bien datos no lineales.
- No se adapta bien a datos faltantes sin imputación.

Análisis Discriminante Cuadrático (*QDA*)

Es una extensión de *LDA* que permite que las clases tengan diferentes matrices de covarianza, lo que lo hace más flexible para problemas donde las clases no son linealmente separables Alam et al. (2020).

Ventajas:

- Más flexible que *LDA* al permitir diferentes covarianzas.
- Mejor rendimiento en datos no lineales.
- Funciona bien con datos balanceados.

Desventajas:

- Mayor riesgo de sobreajuste en conjuntos de datos pequeños.
- Requiere más datos de entrenamiento.
- No se adapta bien a datos faltantes.

k-Nearest Neighbors (*KNN*)

Es un método de clasificación no paramétrico que asigna la clase más frecuente entre los k vecinos más cercanos a un nuevo punto Alam et al. (2020).

Ventajas:

- Fácil de implementar y comprender.
- No hace suposiciones sobre la distribución de los datos.

Desventajas:

- Puede ser computacionalmente costoso en grandes conjuntos de datos.
- Sensible a la escala de las variables y al ruido.
- No funciona bien con datos desbalanceados sin ajustes.
- No se adapta bien a datos faltantes sin imputación.

Perceptron

El perceptron es un algoritmo de clasificación lineal que utiliza una regla de aprendizaje simple para encontrar un hiperplano que separe los datos de dos clases Wen et al. (2021).

Ventajas:

- Simple de implementar.
- Eficiente para problemas linealmente separables.

Desventajas:

- Solo funciona bien con datos que son linealmente separables.
- Puede ser inestable si los datos no son separables.
- No maneja bien datos desbalanceados.
- No se adapta bien a datos faltantes.

Multilayer Perceptron (*MLP*)

Es una red neuronal con una o más capas ocultas que puede aprender funciones no lineales complejas Wen et al. (2021).

Ventajas:

- Capaz de modelar relaciones no lineales complejas.
- Flexible y adaptable a diferentes tipos de problemas.
- Menos sensible a datos desbalanceados, pero puede beneficiarse de técnicas como el re-muestreo.

Desventajas:

- Requiere mucho tiempo de entrenamiento y datos grandes.
- Puede sobre ajustarse si no se aplica regularización adecuada.
- No se adapta bien a datos faltantes sin técnicas de imputación.

Random Forest (*RF*)

Es un conjunto de árboles de decisión donde cada árbol se entrena con una muestra aleatoria de los datos y una selección aleatoria de características. La predicción final se realiza a través del voto mayoritario entre los árboles Wen et al. (2021).

Ventajas:

- Resistente al sobreajuste debido a la naturaleza del ensamble.
- Maneja bien datos no lineales y grandes conjuntos de datos.
- Funciona bien con datos desbalanceados.

- Tiene mecanismos para manejar datos faltantes, como el uso de sustituciones basadas en las ramas del árbol.

Desventajas:

- Puede ser computacionalmente costoso para conjuntos de datos muy grandes.
- Menos interpretable que un árbol de decisión individual.

Regresión Lineal con Penalización de Regularización (*RIDGE*)

La regresión Ridge es una variación de la regresión lineal que incluye una penalización L2 para reducir el sobreajuste, controlando la magnitud de los coeficientes del modelo Dumitrescu et al. (2022).

Ventajas:

- Controla el sobreajuste en modelos de regresión.
- Funciona bien en situaciones de multicolinealidad.

Desventajas:

- No maneja relaciones no lineales.
- Depende de la elección del parámetro de regularización.
- Funciona mejor con datos balanceados.
- No se adapta bien a datos faltantes.

Extreme Gradient Boosting: XGBoost

Es un algoritmo de ensamble basado en árboles de decisión que optimiza el gradiente descendente, siendo altamente eficiente y efectivo en la mayoría de los problemas Zheng (2019).

Ventajas:

- Muy eficiente y rápido, optimizado para grandes conjuntos de datos.
- Maneja bien interacciones no lineales y características de alta dimensionalidad.
- Funciona bien con datos desbalanceados ajustando los pesos de las clases.
- Maneja datos faltantes de manera interna.

Desventajas:

- Puede ser complejo de ajustar debido a sus múltiples hiperparámetros.
- Riesgo de sobreajuste si no se controla adecuadamente.

Feedforward Neural Networks (*FFNN*)

Las redes neuronales feedforward (*FFNN*) consisten en arquitecturas con múltiples capas capaces de aprender representaciones complejas de los datos, propagando la información únicamente hacia adelante. Su desempeño puede mejorarse mediante técnicas de regularización, como la regularización esparza (*sparse regularization*), que ayuda a mitigar el sobreajuste al penalizar activaciones innecesarias. Este enfoque fuerza ciertas activaciones a cero cuando no son esenciales, aplicando funciones de penalización adecuadas para optimizar la eficiencia del modelo Song et al. (2023); L. Zhang y Liang (2023).

Ventajas:

- Excelente rendimiento en tareas complejas como visión por computadora y procesamiento del lenguaje natural.
- Capaz de aprender representaciones jerárquicas de los datos.
- Puede manejar datos desbalanceados, especialmente con técnicas como el remuestreo o ajuste de clases.

Desventajas:

- Requiere grandes cantidades de datos y poder computacional.
- Puede ser difícil de interpretar.
- No se adapta bien a datos faltantes sin preprocesamiento adecuado.

eXplainable Neural Networks (*XNN*)

Es un modelo de red neuronal diseñado para mejorar la interpretabilidad manteniendo la capacidad predictiva de las redes neuronales profundas. *XNN* descompone las predicciones en términos de contribuciones de características individuales, lo que permite a los usuarios comprender el impacto de cada variable en la salida del modelo Pavitha y Sugave (2024).

Ventajas:

- Proporciona interpretabilidad en modelos de aprendizaje profundo sin perder capacidad predictiva.
- Facilita la auditoría de decisiones en aplicaciones críticas como salud y finanzas.
- Captura relaciones no lineales de forma interpretable.
- Se puede integrar con técnicas de explicabilidad adicionales para mejorar su uso en entornos regulatorios.

- Puede manejar datos estructurados y no estructurados con mayor flexibilidad que modelos puramente lineales.

Desventajas:

- Más complejo de entrenar en comparación con modelos lineales e interpretable boosting (como *EBM*).
- Requiere mayor poder computacional en comparación con modelos *GAM* o regresiones.
- Puede ser menos intuitivo que técnicas explicativas como *SHAP* o *LIME* en ciertos casos.
- La interpretación sigue dependiendo de la arquitectura y la manera en que se modelen las interacciones.

2.2.10. Métodos de Interpretabilidad

SHapley Additive Explanations

Es un método de interpretabilidad basado en teoría de juegos que asigna a cada característica una contribución justa a la predicción del modelo. *SHAP* proporciona explicaciones consistentes y localmente precisas sobre la importancia de cada variable en la predicción Pavitha y Sugave (2024); Ma et al. (2023).

Ventajas:

- Proporciona explicaciones matemáticamente justificadas y coherentes.
- Es compatible con una amplia variedad de modelos de *ML*.
- Permite analizar tanto la contribución de variables individuales como sus interacciones.
- Proporciona interpretaciones tanto a nivel de instancia como a nivel general del modelo.
- Implementado en múltiples bibliotecas optimizadas como por ejemplo *SHAP* en Python.

Desventajas:

- Su cálculo puede ser computacionalmente costoso, especialmente en modelos complejos.
- La interpretación puede volverse difícil en datasets con muchas características.

- Puede requerir técnicas de aproximación para modelos con gran cantidad de variables.
- En modelos muy grandes, la generación de valores *SHAP* puede ser lenta.

Local Interpretable Model-agnostic Explanations

Es una técnica de interpretabilidad de modelos de *ML* que genera explicaciones locales al aproximar el modelo con una función lineal en la vecindad de una instancia específica (Pavitha y Sugave, 2024; Ma et al., 2023).

Ventajas:

- Funciona con cualquier modelo sin necesidad de conocer su estructura interna.
- Proporciona explicaciones intuitivas y fáciles de entender.
- Es computacionalmente más eficiente que *SHAP* en algunos casos.
- Permite entender el comportamiento del modelo a nivel individual.
- Es útil en aplicaciones donde la explicabilidad local es más relevante que la global.

Desventajas:

- No proporciona interpretabilidad global del modelo, solo local.
- Puede ser inestable, ya que diferentes ejecuciones pueden generar explicaciones distintas.
- La calidad de la explicación depende de la correcta elección del espacio local de muestreo.
- Puede generar interpretaciones sesgadas si los datos perturbados no representan correctamente la distribución original.

2.2.11. *Cross-Industry Standard Process for Data Mining*

El *CRISP-DM* es una metodología consolidada y de uso común en la minería de datos, desarrollada en 1996 por *IBM*. Su propósito fundamental es ofrecer una estructura organizada para la realización de proyectos de análisis de datos en diversos sectores industriales. Esta metodología se centra en el análisis y la extracción de valor a partir de los datos (IBM, 2020), siendo particularmente útil en la construcción de modelos predictivos que también consideran *FE* en su análisis (Mann y Logeswaran, 2021; Sahiq et al., 2022). El proceso se divide en seis etapas interrelacionadas que permiten

un enfoque iterativo y adaptable, ajustándose a los diferentes retos y particularidades de cada problema o industria:

- Etapa 01. Entendimiento del negocio: se identifican y definen los objetivos del negocio, asegurando que el proyecto de minería de datos responda a las necesidades y expectativas del cliente o la organización.
- Etapa 02. Entendimiento de los datos: se recolectan y exploran los datos disponibles, evaluando su calidad y estructura, además de identificar posibles problemas como datos faltantes o inconsistentes.
- Etapa 03. Preparación de los datos: La data se limpia, transforman y preparan para ser utilizados en las etapas posteriores. Este proceso puede incluir la selección de variables, normalización y otras transformaciones necesarias para mejorar la calidad de los datos.
- Etapa 04. Modelado: se seleccionan las técnicas y algoritmos apropiados para generar modelos predictivos o descriptivos. Durante esta fase, es crucial ajustar los parámetros del modelo para obtener el mejor rendimiento posible.
- Etapa 05. Evaluación: Los modelos generados se analizan en detalle para determinar si cumplen con los objetivos definidos en la fase inicial. Se comparan diferentes modelos y se selecciona el más adecuado en función de su precisión y relevancia para el problema.
- Etapa 06. Implementación: El modelo final se implementa en el entorno de producción o se utiliza para apoyar la toma de decisiones en la organización. Esta fase puede involucrar la integración del modelo en sistemas existentes o la creación de reportes y automatización de procesos.

2.2.12. Test de Estacionariedad de *Dickey-Fuller*

El test de *Dickey-Fuller* es una prueba estadística utilizada para determinar si una serie temporal es estacionaria. La estacionariedad es una propiedad clave en el análisis de series temporales, ya que muchas técnicas de modelado suponen que las series temporales son estacionarias (Awad y Al Karaki, 2019; Durafe y Jha, 2018).

Modelo de Regresión

El modelo básico para el test de *Dickey-Fuller* se puede expresar como:

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \delta_1 \Delta Y_{t-1} + \delta_2 \Delta Y_{t-2} + \cdots + \delta_p \Delta Y_{t-p} + \epsilon_t \qquad (1)$$

donde:

- ΔY_t es la primera diferencia de la serie temporal Y_t ($\Delta Y_t = Y_t - Y_{t-1}$).
- α es una constante.
- βt es un término de tendencia.
- Y_{t-1} es el valor rezagado de la serie temporal.
- δ_i son los coeficientes de las diferencias rezagadas.
- ϵ_t es el término de error.
- p es el número de rezagos.

Hipótesis

Las hipótesis del test de *Dickey-Fuller* son:

- **Hipótesis nula (H_0)**: La serie temporal tiene una raíz unitaria (es no estacionaria).
- **Hipótesis alternativa (H_1)**: La serie temporal es estacionaria (no tiene raíz unitaria).

Estadística de Prueba

La estadística de prueba de *Dickey-Fuller* se calcula como:

$$\tau = \frac{\hat{\gamma}}{\text{SE}(\hat{\gamma})} \tag{2}$$

donde $\hat{\gamma}$ es el estimador del coeficiente γ y $\text{SE}(\hat{\gamma})$ es el error estándar del estimador.

Cálculo del *p-valor*

El *p-valor* asociado a la estadística de prueba τ se calcula utilizando la distribución específica de *Dickey-Fuller*. En la práctica, el *p-valor* se obtiene utilizando software estadístico como R, Python, EViews, o Stata.

Interpretación

- **Rechazar** H_0: Si el *p-valor* es menor que el nivel de significancia α (por ejemplo, 0.05), se rechaza la hipótesis nula y se concluye que la serie temporal es estacionaria.
- **No rechazar** H_0: Si el *p-valor* es mayor o igual que α, no se rechaza la hipótesis nula, indicando que no hay suficiente evidencia para afirmar que la serie temporal es estacionaria.

2.2.13. Test de Causalidad de *Granger*

El test de causalidad de *Granger* es una técnica estadística ampliamente utilizada en econometría y análisis de series temporales para evaluar si una serie temporal puede predecir otra. Este método permite determinar la existencia de una relación de causalidad en el sentido de *Granger* entre dos series temporales, lo que implica que los valores pasados de una serie contienen información útil para predecir los valores futuros de la otra serie (Ye et al., 2015).

Modelo con retardo: X_t:

$$Y_t = \beta_0 + \sum_{i=1}^{k} \beta_i Y_{t-i} + \sum_{j=1}^{k} \gamma_j X_{t-j} + \eta_t \tag{3}$$

donde:

- X_{t-j} son los valores rezagados de la serie X_t.
- β_0, β_i, y γ_j son los coeficientes.
- η_t es el término de error.

Hipótesis

- **Hipótesis nula (H_0)**: Los coeficientes $\gamma_j = 0$, lo que implica que X_t no causa a Y_t en el sentido de *Granger*.
- **Hipótesis alternativa (H_1)**: Al menos uno de los coeficientes $\gamma_j \neq 0$, lo que implica que X_t causa a Y_t en el sentido de *Granger*.

Procedimiento

1. **Estimar ambos modelos**: Usa métodos de regresión para estimar los coeficientes de ambos modelos.
2. **Calcular la suma de los cuadrados de los residuos (RSS)**:

- $RSS_{\text{restricted}}$ para el modelo sin X_t.
- $RSS_{\text{unrestricted}}$ para el modelo con X_t.

3. **Calcular la estadística F**:

$$F = \frac{\left(\frac{RSS_{\text{restricted}} - RSS_{\text{unrestricted}}}{k}\right)}{\left(\frac{RSS_{\text{unrestricted}}}{n-2k-1}\right)} \tag{4}$$

donde:

- n es el número de observaciones.
- k es el número de rezagos.

4. **Obtención del *p-valor***: Calcular el *p-valor* correspondiente a la estadística F obtenida usando la distribución F.
5. **Decisión sobre la Hipótesis Nula**: Comparar el *p-valor* con el nivel de significancia α:

- Si el *p-valor* es menor que α, se rechaza la hipótesis nula (H_0).
- Si el *p-valor* es mayor o igual que α, no se rechaza la hipótesis nula.

Interpretación

La interpretación de los resultados del test de causalidad de *Granger* es la siguiente:
- Si se rechaza la hipótesis nula (H_0), se concluye que X_t *Granger*-causa a Y_t.
- Si no se rechaza la hipótesis nula, no hay suficiente evidencia para afirmar que X_t *Granger*-causa a Y_t.

2.2.14. Prueba t de Student para muestras emparejadas

La prueba *t* de Student para muestras emparejadas, es una herramienta estadística usada para comparar medias de dos conjuntos de datos relacionados, con mediciones antes y después de aplicar un determinado tratamiento o cambio de condiciones; es útil para evaluar si un cambio observado en un grupo de sujetos es significativo, teniendo en cuenta la variabilidad de las diferencias (Coma et al., 2020; Gorzelnik y Oleniacz, 2019).

Características:

- **Hipótesis:** La prueba considera dos hipótesis:
  - **Hipótesis nula (H_0):** No existe una diferencia significativa entre las medias de los dos conjuntos de datos relacionados; es decir, las medias de las diferencias son iguales a cero.
  - **Hipótesis alternativa (H_1):** Existe una diferencia significativa entre las medias de los dos conjuntos de datos relacionados, es decir, las medias de las diferencias son distintas de cero.
- **Cálculo de la t:** La estadística t se calcula utilizando la diferencia entre los dos conjuntos de datos emparejados:

$$t = \frac{\bar{D}}{\frac{s_D}{\sqrt{n}}} \tag{5}$$

donde:

- $\bar{D}$ es la media de las diferencias entre cada par de mediciones.
- s_D es la desviación estándar de las diferencias.
- n es el número de pares de datos.

- **Grados de libertad:** Los grados de libertad (df) para esta prueba se calculan como el número de pares de datos menos uno ($df = n - 1$).
- **Significancia estadística:** La prueba compara el valor t calculado con un valor crítico t obtenido de la distribución t de Student, dependiendo del nivel de significancia seleccionado ($\alpha = 0{,}05$). Si el valor absoluto de t es mayor que el valor crítico, se rechaza la hipótesis nula, indicando que la diferencia entre las medias es significativa (Coma et al., 2020).
- **Aplicación:** La prueba t para muestras emparejadas es ideal para evaluar en esta tesis, el efecto en la predicción de riesgo de crédito al incluir en los conjuntos de datos *FE*.

2.3. Marco Sistémico

En su Teoría General de Sistemas, Bertalanffy (1989) establece un conjunto de definiciones que han sido ampliamente aceptadas y utilizadas en diversas áreas de estudio, proporcionando un enfoque estandarizado para la comprensión de los sistemas. Esta teoría plantea tres principios fundamentales: (a) los sistemas existen dentro de otros sistemas; (b) son abiertos, es decir, intercambian información con su entorno; y (c) sus funciones están determinadas por su estructura. Los sistemas poseen las siguientes características clave:

- Totalidad: Un sistema es más que la simple suma de sus componentes.
- Sinergia: Cualquier acción que afecta una parte del sistema repercute en el resto.
- Entropía: Los sistemas tienden a preservar su identidad a lo largo del tiempo.
- Equifinalidad: Los sistemas son adaptables, de manera que los resultados finales son independientes de las condiciones iniciales.
- Equipotencialidad: Los elementos restantes del sistema pueden asumir las funciones de los que faltan.
- Finalidad: Los componentes de un sistema están orientados hacia metas comunes.

- Retroalimentación: Existe un intercambio constante de información dentro del sistema.
- Homeostasis: Los sistemas vivos tienden a mantenerse estables a través de mecanismos de autorregulación.

Por su parte, Dori et al. (2019) propone varias definiciones de sistemas según diferentes enfoques:

- Visión formal minimalista: Los sistemas pueden ser modelados; lo que no es modelable no puede considerarse un sistema. Pueden ser tanto naturales como artificiales.
- Visión constructivista: Los sistemas están compuestos por partes, elementos y límites que son definidos por el observador, y poseen propiedades emergentes no presentes en los componentes individuales. Un ejemplo son los sistemas de actividades humanas.
- Realismo moderado: Los sistemas pueden existir como constructos mentales, ser naturales o creados por el hombre, y presentan propiedades emergentes con límites que pueden variar según el observador. Este enfoque es común en la ingeniería de sistemas.
- Realismo extremo: Los sistemas existen únicamente en el mundo real, con límites que pueden percibirse de manera distinta según el observador.
- Enfoque de complejidad y vida: Los sistemas deben ser lo suficientemente complejos para sustentar la vida. En el caso de los sistemas artificiales, deben ser autónomos y tener la capacidad de perpetuarse.
- Sistema como modo de descripción: Todo sistema puede describirse desde tres perspectivas: un conjunto de elementos, un conjunto de interacciones entre esos elementos, y sus interacciones con el entorno. Esto permite describir un macrosistema y desglosarlo en subsistemas.
- Sistema como proceso: Un sistema es definido más por su capacidad para alcanzar objetivos que por su estructura. En el largo plazo, los sistemas evolucionan en respuesta a cambios estructurales.

Thurner et al. (2018) señala que la Teoría General de Sistemas, junto con la cibernética y la teoría de juegos, ha sido crucial para entender las interacciones en sistemas no triviales, contribuyendo al desarrollo de la teoría de sistemas complejos. Según este autor, un sistema complejo se caracteriza por ser una red multicapa de elementos interrelacionados que coevolucionan. Las premisas principales de un sistema complejo incluyen:

- Están compuestos por numerosos elementos.
- Los elementos interactúan de diversas maneras, lo que puede representarse mediante una red de interacciones.
- Las interacciones varían a lo largo del tiempo y pueden ser de naturaleza física, química, social o simbólica, con un comportamiento determinista o estocástico.
- Los estados de los elementos evolucionan a través del tiempo.
- Las interacciones y los estados son interdependientes y coevolucionan conjuntamente.
- La coevolución de los sistemas multicapa suele no ser lineal.
- Dependen del contexto en el que se encuentran.
- Las interacciones discretas permiten que el sistema tenga propiedades algorítmicas, lo que implica que sus interacciones pueden ser representadas mediante algoritmos.
- Poseen una trayectoria y orientación definida.
- Tienen memoria, que se almacena en la estructura de la red.

2.4. Marco Ético

Tractenberg (2019) señala que la práctica cotidiana en áreas como la ciencia de datos, la estadística y la computación involucra un proceso continuo de toma de decisiones. Estas decisiones deben basarse en principios de profesionalismo y responsabilidad, lo que subraya la importancia de actuar de manera ética en cada etapa del proceso. Para fortalecer esta práctica, se recomienda seguir las directrices éticas propuestas por la *American Statistical Association* (*ASA*) y la *Association of Computing Machinery* (*ACM*), las cuales contribuyen a desarrollar una identidad profesional coherente y basada en principios éticos para los especialistas en estadística y ciencia de datos.
Los principios éticos definidos por la *ACM* pueden agruparse en tres grandes áreas: responsabilidad profesional, principios de liderazgo profesional y la adhesión a un código ético. De manera complementaria, la *ASA* establece directrices que se centran en:

- Integridad profesional y responsabilidad: Utilizar metodologías y datos adecuados, sin sesgos ni favoritismos, para generar información válida, interpretable y reproducible. Es fundamental evitar tareas para las que no se esté suficientemente capacitado y respetar a los colegas en el proceso.

- Integridad de los datos y métodos: Reconocer las limitaciones, sesgos y defectos que pudieran comprometer la fiabilidad del análisis y actuar de manera transparente frente a estos desafíos.
- Responsabilidades hacia la ciencia, la sociedad y los clientes: Promover la transparencia, la validez de las referencias y contribuir al desarrollo científico considerando los intereses sociales.
- Responsabilidades hacia los sujetos de investigación: Proteger los derechos e intereses de los sujetos involucrados, ya sea de forma directa (mediante entrevistas o encuestas) o indirecta (a través de sus datos personales).
- Responsabilidades con los colegas de investigación: Contribuir al desarrollo de un ambiente de trabajo ético y respetuoso, promoviendo el trabajo en equipo.
- Responsabilidades con otros profesionales: Mantener una actitud abierta a diferentes explicaciones de los fenómenos observados, asegurando que los debates y confrontaciones se realicen en un marco de respeto a distintos puntos de vista.
- Responsabilidades respecto a denuncias de mala conducta: Evitar y reportar prácticas cuestionables que puedan ser consideradas no profesionales.
- Responsabilidades de los empleadores: Respetar las opiniones, métodos y obligaciones objetivas de los investigadores, proporcionando el apoyo necesario para su trabajo ético.

Tractenberg (2019) también ofrece una serie de principios éticos específicos para la ciencia de datos, organizados de acuerdo con las diferentes fases del ciclo de vida de los datos:

- Recolección de datos:
  - Consentimiento informado: Asegurar que los participantes comprendan claramente los términos de su participación en la investigación.
  - Sesgo en la recopilación: Identificar y minimizar posibles fuentes de sesgo o ruido durante la recolección de datos.
  - Limitación de la exposición de datos personales: Considerar la anonimización de los datos cuando sea posible y adecuado.
- Almacenamiento de información:
  - Seguridad de los datos: Garantizar la integridad y protección de los datos frente a posibles accesos no autorizados o modificaciones indebidas.

    - Derecho al olvido: Evaluar mecanismos que permitan a los participantes solicitar la eliminación de sus datos personales.
    - Plan de retención de datos: Establecer políticas claras sobre la retención y eliminación de datos tras un periodo específico.

- Análisis:
    - Perspectivas ausentes: Validar suposiciones y escenarios alternativos con expertos en el campo.
    - Sesgo en el conjunto de datos: Identificar y mitigar fuentes de ruido o sesgo presentes en los datos.
    - Representación justa: Asegurar que los gráficos, visualizaciones y estadísticas reflejen fielmente los datos analizados.
    - Privacidad en el análisis: Proteger la confidencialidad de los datos sensibles durante el análisis.
    - Auditabilidad: Documentar adecuadamente el proceso para garantizar su reproducibilidad y trazabilidad.
- Modelamiento:
    - Discriminación: Asegurar que los modelos no introduzcan sesgos discriminatorios ni dependan de variables que promuevan desigualdades.
    - Equidad entre grupos: Evaluar que los resultados del modelo reflejen de manera justa la realidad de los diferentes grupos involucrados.
    - Selección de métricas: Verificar que las métricas seleccionadas sean rigurosas y adecuadas para el análisis.
    - Explicabilidad: Asegurar que el modelo sea comprensible y explicable para todas las partes interesadas.
    - Comunicación de sesgos: Transparentar las posibles desviaciones, limitaciones y sesgos presentes en el modelo.
- Implementación:
    - Reparación: Proponer soluciones para mitigar los efectos negativos que el modelo pueda generar sobre los participantes.
    - Retroceso: Evaluar cuándo es necesario detener o ajustar el uso del modelo en producción.
    - Evaluación del modelo: Validar la vigencia y aplicabilidad del modelo a lo largo del tiempo.

- Uso no previsto: Identificar y prevenir posibles usos no intencionados del modelo que puedan derivar en efectos adversos.

2.5. Marco Filosófico.

El presente trabajo de investigación se fundamenta en una orientación epistemológica post-positivista, la cual reconoce que, aunque es posible aproximarse a la realidad de manera objetiva, el conocimiento no es absoluto ni definitivo, sino que está influido por factores contextuales, interpretativos y las limitaciones inherentes a los modelos empleados (Godwin et al., 2021). Esta postura considera que los fenómenos observables pueden ser medidos y estudiados de forma sistemática, pero que el conocimiento científico es siempre provisional y revisable a la luz de nuevas evidencias (Ugwu et al., 2021).

Desde el punto de vista ontológico, este estudio adopta una postura realista crítico. Se asume que el mundo real existe independientemente de nuestra percepción, pero nuestro acceso a ese mundo está mediado por los instrumentos de medición y los modelos utilizados. En el contexto de la investigación sobre *ML*, esto implica que los fenómenos subyacentes a los datos que modelamos (por ejemplo, patrones en grandes volúmenes de información) existen en una realidad objetiva, pero nuestro entendimiento de ellos está condicionado por las herramientas algorítmicas empleadas y las limitaciones inherentes a los datos disponibles (Alharahsheh y Pius, 2020).

En cuanto a la metodología aplicada, la investigación se enmarca en el paradigma cuantitativo, dado que los modelos de *ML* utilizados en este estudio se fundamentan en el análisis de datos numéricos y la aplicación de técnicas estadísticas y matemáticas. Este enfoque es ampliamente reconocido como adecuado para identificar relaciones causales y evaluar hipótesis a partir de grandes volúmenes de datos, particularmente en el ámbito financiero (Benbya et al., 2020; Alharahsheh y Pius, 2020).

Aunque se reconoce la naturaleza dinámica y compleja de los sistemas financieros, así como la influencia de eventos globales disruptivos (por ejemplo, la pandemia de *COVID-19*), estos elementos fueron operacionalizados como variables exógenas en los modelos cuantitativos, sin requerir técnicas cualitativas. La identificación de fechas clave y eventos relevantes se realizó mediante análisis documental, utilizado como insumo para la construcción de escenarios y variables en el análisis estadístico y de aprendizaje automático (Geraldo-Campos et al., 2022; Noriega et al., 2023; Pandey et al., 2021).

En este sentido, la teoría de la complejidad ofrece un marco filosófico adicional que complementa esta investigación. Los sistemas de información y las dinámicas de los datos masivos *big data* a menudo exhiben propiedades emergentes y no lineales, por lo que se asume que el comportamiento de estos sistemas no siempre puede prede-

cirse de forma exacta mediante modelos deterministas. El enfoque de *ML*, entonces, se utiliza no solo como una herramienta de predicción, sino también como un medio para comprender mejor las interacciones complejas dentro de los datos (Benbya et al., 2020).

De esta forma, la presente tesis doctoral se sitúa dentro de un marco filosófico que combina realismo crítico, epistemología post-positivista y la teoría de la complejidad (Larsen-Freeman, 2020), con el objetivo de ofrecer una comprensión más integral de la aplicación de *ML* en sistemas de información, respetando tanto las posibilidades como las limitaciones del conocimiento generado en este campo.

2.6. Glosario.

Tabla 4
Términos y Definiciones

Término	**Definición**
ACC	Accuracy
ADABOOST	Adaptive Boosting
ADAYSN	Adaptive Synthetic
AUC	Área bajo la curva ROC
EBM	Explainable Boosting Machine
BIG DATA	Grandes volúmenes de información estructurada o no estructurada.
BP / BPNN	Back propagation neural Network
CNN	Convolutional neural network
DA	Discriminant Analysis
DL o DNN	Deep Neural Network
FE	Factores Externos
FFNN	Feed forward neural networks
GBDT	Gradient-boosted decision trees
GLM	Generalized Linear Model
KNN	K-Nearest Neighbour
LAG	Retardo de un periodo de tiempo
LDA	Linear Discriminant Analysis
LIME	Local Interpretable Model-Agnostic Explanations
LINR	Lineal regression
LR	Logistic regression

Continued on next page

Tabla 4
Términos y Definiciones

Término	**Definición**
LSTM	Long short-term Multi layer perceptron networks
MLP	Multi layer perceptron
NB	Naïve Bayes classifier
NN	Nural Network
P	Perceptron
PNN	Probabilistic Neural Network
QDA	Quadratic Discriminant Analysis
RF	Random forest
ROS	Random Oversampling
RUS	Random Undersampling
Ridge	Analysis of multicollinearity in multiple regression data
SHAP	SHapley Additive exPlanations
SMOT	Synthetic Minority Oversampling Technique
SVM	Support vector machines
XGBoost	Extreme Gradient Boosting
XNN	eXplainable Neural Network

Capítulo III

Metodología

3.1. Tipos de investigación.

3.1.1. Nivel de investigación.

Esta investigación se enmarca en el nivel de investigación aplicada, ya que busca generar conocimiento con un propósito práctico y orientado a la solución de problemas en el sector financiero. El estudio se centra en la predicción del riesgo crediticio mediante técnicas de *ML* en presencia de *FE*.

El objetivo es mejorar la precisión y explicabilidad de los modelos de predicción de riesgo crediticio mediante la incorporación de variables que reflejen eventos disruptivos, como crisis económicas, pandemias y fenómenos climáticos extremos. Para ello, se emplearon enfoques basados en series temporales y técnicas explicativas como *SHAP* y *LIME*, con el fin de facilitar la interpretación de los resultados por parte de reguladores y entidades financieras.

La investigación busca clasificar a los deudores en categorías de buen pagador y mal pagador según sus patrones de morosidad, proporcionando una metodología replicable que optimice la gestión del riesgo crediticio en escenarios de alta incertidumbre.

3.1.2. Tipo de la investigación.

Esta investigación adopta un enfoque cuantitativo, ya que se fundamenta en la recopilación, procesamiento y análisis de datos numéricos mediante técnicas estadísticas y de *ML*.

Se emplean métodos cuantitativos para evaluar el desempeño de modelos de predicción del riesgo crediticio en presencia de *FE*. La evaluación se basa en métricas de clasificación, incluyendo exactitud (*ACC*) y área bajo la curva ROC (*AUC*), lo que permite medir la capacidad discriminativa de los modelos y su efectividad en la identificación de patrones de morosidad.

La investigación utiliza algoritmos de *ML* y análisis de series temporales para examinar la relación entre los *FE* y la morosidad crediticia, garantizando un enfoque basado en evidencia empírica y replicable.

3.1.3. Diseño de la investigación.

El diseño de esta investigación es cuasi-experimental, dado que se trabajó con datos históricos de *FE* y riesgo crediticio sin manipulación directa de variables independientes. Se aplicó pruebas de hipótesis para evaluar el impacto de la inclusión de variables de *FE* en modelos de *ML* para la predicción del riesgo crediticio.
El estudio sigue un enfoque longitudinal y correlacional, ya que analiza la relación entre la morosidad crediticia y eventos externos como la pandemia de *COVID-19* , el cambio climático y la convulsión social a lo largo del tiempo. Para ello, se utilizó series temporales y técnicas de *ML*, evaluando la capacidad predictiva de los modelos antes y después de la integración de variables explicativas.
Los datos de *FE* fueron obtenidos de fuentes públicas, mientras que los datos de morosidad crediticia provinieron de una entidad financiera denominada Financiera Modelo (*FMOD*).

3.1.4. Método de la investigación

La presente investigación se desarrolló bajo un enfoque cuantitativo, de tipo aplicativo y nivel explicativo, orientado a analizar el impacto de factores exógenos (*FE*) en la precisión de modelos predictivos de riesgo crediticio. Para ello, se diseñó un procedimiento experimental que aplicó 14 algoritmos de aprendizaje automático (ML) a un conjunto de datos compuesto por variables financieras y exógenas.
El método propuesto se estructura en las siguientes fases:

- **Recolección y preparación de datos:** Incluye la obtención de datos financieros de la entidad *FMOD* y datos públicos de *FE* (pandemia, anomalías climáticas, convulsión social), seguidos de un proceso de limpieza, integración y enriquecimiento para formar un dataset tabular.

- **Análisis exploratorio:** Evaluación de las relaciones entre variables mediante técnicas estadísticas (matriz de correlación, pruebas de estacionariedad y causalidad de Granger) para identificar patrones relevantes.

- **Modelado predictivo:** Entrenamiento de 14 algoritmos de clasificación supervisada sobre el dataset final para predecir la probabilidad de morosidad grave, evaluando métricas de desempeño como *AUC* y *ACC*.

- **Validación y comparación:** Aplicación de validación cruzada (*10-fold cross-validation*) y análisis de sensibilidad para comparar el desempeño de los modelos con y sin inclusión de variables exógenas.

Nota metodológica: Si bien los datos incluyen información temporal (fechas, series de morosidad), el enfoque adoptado no corresponde a técnicas de predicción de series temporales como ARIMA o LSTM, que modelan dependencias secuenciales y realizan forecasting. En esta investigación, las variables temporales fueron alineadas y transformadas para construir un dataset tabular, utilizado como entrada para algoritmos de clasificación supervisada (e.g., *XGB*, *FFN*, *CNN*). Este enfoque es consistente con trabajos recientes en el dominio financiero que han aplicado *ML* sobre datos temporales enriquecidos (Z. Chen et al., 2020; Jiang, 2022; Noriega et al., 2023; X. Zhang et al., 2024).
Este método permitió evaluar no solo la capacidad predictiva de los modelos, sino también su interpretabilidad mediante técnicas de Explainable AI (e.g., SHAP, LIME).

3.2. *Dataset* y partición.

3.2.1. *Dataset*.

La definición de los *datasets* utilizados en este estudio se presenta en la Tabla 13. Se han identificado conjuntos de datos que registran la intensidad de los eventos externos considerados dentro del marco de referencia de la investigación.
Para capturar la incidencia de los *FE* en la morosidad crediticia, se han seleccionado los siguientes conjuntos de datos:

- **Pandemia por *COVID-19*** : Se recopilaron datos diarios sobre el número de casos positivos y fallecimientos por *COVID-19* , ya que reflejan la evolución de las olas pandémicas y su impacto en la estabilidad económica. Para la evaluación, se consideraron las fechas en las que el gobierno decretó el estado de emergencia, comprendidas entre el 16 de marzo de 2020 y el 28 de octubre de 2022.

- **Convulsión social**: Se seleccionó el número diario de bloqueos de vías debido a manifestaciones y conflictos sociales como indicador de la inestabilidad política y económica. El análisis se enfocó en el período posterior al intento de golpe de Estado del presidente en turno, desde el 7 de diciembre de 2022 hasta el 30 de septiembre de 2023, fecha final de los datos disponibles.

- **Factores climáticos**: Se consideraron dos indicadores clave:

    - La anomalía de temperatura respecto a la media histórica mensual, medida en una estación meteorológica en la ciudad de Lima.

- El número diario de bloqueos de vías debido a eventos climáticos extremos.

Para la evaluación de estos factores, se seleccionó el mismo periodo que la convulsión social (del 7 de diciembre de 2022 al 30 de septiembre de 2023), debido a la significativa variabilidad de temperatura registrada en esos meses (ver Figura 5).

Tabla 5
Descripción de variables en el conjunto de datos de series temporales sobre morosidad crediticia por actividad económica

Ítem	Variable	Tipo	Descripción
1	Fecha	Fecha	Fecha de registro de la información crediticia (AAAA-MM-DD).
2	CódigoActividadEconómica	int	Identificador de la actividad económica del prestatario.
3	CódigoFuenteIngreso	int	Código que representa la fuente de ingreso del prestatario (por ejemplo, independiente, dependiente).
4	CódigoDestinoCrédito	int	Identificador del propósito o destino del crédito.
5	PromedioDíasMora	int	Número promedio de días de mora entre los créditos agrupados en el registro.
6	MáximoDíasMora	int	Número máximo de días de mora observado en el grupo.
7	CantidadCréditos	int	Número de créditos agregados en la misma actividad económica y fecha.

Nota. Este conjunto de datos contiene 1,944,595 registros y 7 atributos, representando series temporales de morosidad crediticia agregadas por actividad económica y fecha. Se utilizó como insumo para enriquecer el dataset final con variables de contexto.

La Tabla 6 presenta la estructura del conjunto de datos final resultante de la integración de variables financieras y factores exógenos (COVID-19, anomalías climáticas y convulsión social). Este dataset enriquecido se utilizó como entrada para los modelos de machine learning y consta de 367,346 registros y 367 atributos. Este total incluye las 28 variables listadas, variables derivadas creadas durante el preprocesamiento y la expansión de variables categóricas mediante codificación; en el Anexo D se muestran registros ejemplos del Dataset.

Tabla 6
Resumen de variables utilizadas en los experimentos de modelado

Ítem	Variable	Tipo	Descripción
1	SegmentoPortafolio	int	Categoría del segmento de portafolio
2	TasaInterés	dec	Tasa de interés efectiva aplicada al crédito
3	MontoCrédito	dec	Monto aprobado del crédito
4	PlazoDías	int	Plazo real del crédito en días
5	MontoInternoMáximo	dec	Monto máximo de deuda interna histórica del cliente
6	IDCréditoGobierno	bit	Indica si el crédito está respaldado por el gobierno (1 = sí)
7	IDGénero	int	Código que representa el género del prestatario
8	SalarioNormalizado	dec	Salario del cliente normalizado respecto al total del portafolio
9	ActivosTotalesNormalizados	dec	Valor de los activos del negocio normalizado respecto al total
10	PatrimonioNegocioNormalizado	dec	Patrimonio del negocio normalizado respecto al total
11	LiquidezDisponibleNormalizada	dec	Liquidez del negocio normalizada respecto al total
12	IngresoNetoNegocioNormalizado	dec	Ingreso neto del negocio normalizado respecto al total
13	MesesHistorialBancario	int	Número de meses con historial bancario
14	PuntajeOriginaciónMicro	int	Puntaje de originación para microcrédito
15	PuntajeSobreendeudamiento	int	Puntaje de sobreendeudamiento
16	IDFuenteProspección	int	Código del canal de adquisición del crédito
17	IDDestinoCrédito	int	Uso o destino previsto del crédito
18	EdadNormalizadaDesembolso	int	Edad del cliente al momento del desembolso (normalizada)
19	CódigoOcupación	int	Código que representa la ocupación del prestatario
20	Dependientes	int	Número de dependientes reportados
21	CódigoNivelEducativo	int	Nivel educativo
22	CódigoEmpleo	int	Estado laboral
23	AñoInicioResidencia	int	Año de inicio en la residencia actual
24	CódigoTenenciaVivienda	int	Tipo de tenencia de la vivienda
25	CódigoTipoVivienda	int	Tipo de vivienda
26	IDOficinaDesembolso	int	Código de la oficina donde se desembolsó el crédito
27	DíasMora_MMYYYY	int	Número de días de mora por mes desde ene. 2020 hasta sept. 2023
28	ImpactoFactorExterno_MMYYYY	bit	Indica impacto de factor externo por mes desde ene. 2020 hasta sept. 2023

Nota. El conjunto de datos final consta de 367,346 registros y 367 atributos. Este total incluye las 28 variables listadas, variables derivadas creadas durante el preprocesamiento, y la expansión de variables categóricas mediante codificación *One-Hot Encoding*.

La Figura 3 ilustra el flujo de enriquecimiento de datos realizado para la preparación del conjunto de datos final. Este proceso consolidó las variables financieras originales con datos exógenos provenientes de cuatro conjuntos adicionales: casos positivos y fallecimientos por *COVID-19* (ver Tablas 8 y 9), bloqueos sociales (ver Tabla 10) y anomalías climáticas (ver Tabla 11). Estas tablas describen la estructura, atributos y fuentes de cada dataset, asegurando la transparencia en la construcción del conjunto de datos enriquecido.

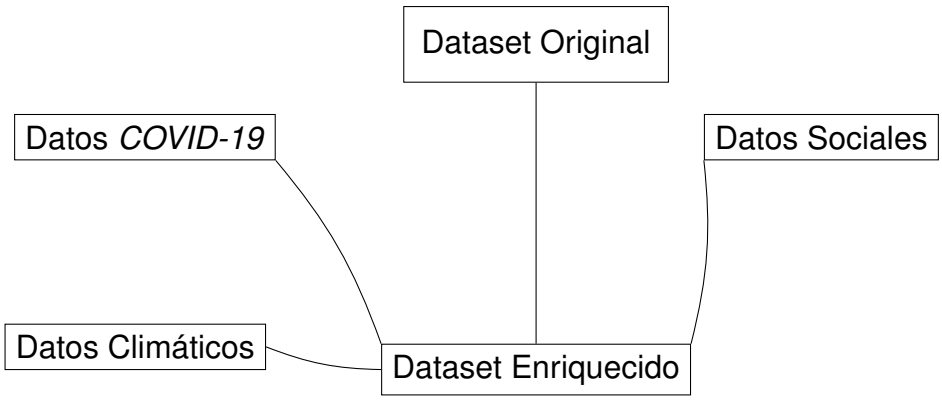


Figura 3
Flujo de enriquecimiento del conjunto de datos con variables exógenas

Además, para evaluar el impacto y la influencia de los *FE* en la morosidad crediticia, se han incorporado dos conjuntos de datos adicionales provenientes de la entidad financiera *FMOD*:

- **Series de tiempo de Días de atraso por actividad económica**: Registra la evolución de los días de mora segmentados por actividad económica de forma diaria, permitiendo analizar el efecto de los *FE* en distintos períodos por actividad económica (ver Tabla 5).

- **Créditos**: Incluye información detallada sobre los créditos originados o cancelados entre enero de 2020 y septiembre de 2023, a nivel individual, en la Tabla 7 se muestran sus datos descriptivos. Cada registro representa un crédito evaluado durante el periodo de estudio y contiene variables sociodemográficas, indicadores financieros, métricas normalizadas del negocio, e información de los días de mora en periodos mensuales entre enero de 2020 y septiembre de 2023.

  Para este análisis, se construyeron indicadores temporales de estrés financiero mediante la combinación de estas variables, alineándolas con series temporales de factores externos (como indicadores de *COVID-19*, anomalías climáticas y episodios de convulsión social). Esta alineación temporal permitió aplicar análisis de causalidad —como las pruebas de causalidad de Granger— y facilitó la evaluación del comportamiento crediticio en contextos de disrupción compuesta.

Entre las variables derivadas más representativas se encuentran `Delinquency Days_MMYYYY`, que registra el número de días de mora por mes, y `External FactorImpact_MMYYYY`, que identifica la presencia de factores externos durante el mismo periodo. Este enfoque no solo permitió evaluar la precisión predictiva de los modelos de *machine learning*, sino también explorar su capacidad para ofrecer explicaciones transparentes sobre las relaciones encontradas, fortaleciendo su aplicabilidad en entornos financieros complejos y volátiles (ver Tabla 6).

Tabla 7

Estadísticas descriptivas del conjunto de datos crediticio utilizado para el modelado

Ítem	Indicador	Valor
1	Tasa de interés efectiva (promedio)	69.98 %
2	Monto promedio del crédito	S/. 5,554
3	Plazo promedio del crédito (en días)	364
4	Número de registros crediticios	367,346
5	Número de créditos con respaldo estatal	21,769
6	Número de créditos sin respaldo estatal	345,577
7	Número de créditos otorgados a hombres	141,857
8	Número de créditos otorgados a mujeres	225,489
9	Saldo de cartera con menos de 30 días de mora (ene. 2020)	S/. 299,040,327.46
10	Saldo de cartera con mora entre 30 y 120 días (ene. 2020)	S/. 27,787,259.90
11	Tasa de morosidad mayor a 30 días (ene. 2020)	8.50 %

Nota. De acuerdo con la política de *FMOD*, los créditos con más de 120 días de mora se clasifican como "castigados" y se provisionan al 100 %. Por lo tanto, se excluyen del cálculo de la tasa de morosidad. Las referencias a "(ene. 2020)" corresponden al saldo observado al cierre de ese mes.

Tabla 8

Descripción de variables en el conjunto de datos COVID-19 Positive Cases

Ítem	Variable	Tipo	Descripción
1	FECHA_CORTE	Date	Fecha de corte del registro (AAAA-MM-DD)
2	UUID	String	Identificador único del caso positivo
3	UBIGEO	String	Código de ubicación geográfica
4	DEPARTAMENTO	String	Departamento de residencia del paciente
5	PROVINCIA	String	Provincia de residencia del paciente

Nota. Este dataset contiene 4,523,494 registros (16/03/2020–28/10/2022).
Fuente: `https://www.datosabiertos.gob.pe/`

Tabla 9

Descripción de variables en el conjunto de datos COVID-19 Death Cases

Ítem	Variable	Tipo	Descripción
1	TIPO SEGURO	String	Tipo de seguro del fallecido
2	SEXO	String	Género del fallecido
3	EDAD	String	Edad del fallecido
4	TIEMPO EDAD	String	Unidad de tiempo de la edad (años, meses)
5	DISTRITO	String	Distrito de residencia del fallecido
6	FECHA DEFUNCION	Date	Fecha del fallecimiento

Nota. Este dataset contiene 1,063,795 registros (16/03/2020–28/10/2022).
Fuente: `https://www.datosabiertos.gob.pe/`

Tabla 10

Descripción de variables en el conjunto de datos Road Blockages (Social Unrest)

Ítem	Variable	Tipo	Descripción
1	estado	String	Estado del tránsito (e.g., RESTRINGIDO)
2	fecha_actualizacion	Date	Fecha de actualización de la información
3	hora_actualizacion	Time	Hora de actualización de la información
4	departamento	String	Departamento donde ocurrió la afectación
5	provincia	String	Provincia donde ocurrió la afectación
6	Motivo	String	Motivo de la afectación (por ejemplo, INFRAESTRUCTURA, CLIMATOLOGICO, HUMANO etc.).

Nota. Este dataset contiene 270,600 registros (07/12/2022–30/09/2023).
Fuente: `https://www.gob.pe/mtc/`

Tabla 11

Descripción de variables en el conjunto de datos Temperature Anomaly

Ítem	Variable	Tipo	Descripción
1	Fecha	Date	Fecha de la medición de temperatura (AAAA-MM-DD)
2	Mes	Integer	Mes de la medición
3	Año	Integer	Año de la medición
4	AñoMes	Integer	Año y mes combinados (AAAAMM)
5	Temperatura	Float	Temperatura registrada en grados Celsius

Nota. Este dataset contiene 22,281 registros (07/12/2022–30/09/2023).
Fuente: `https://www.senamhi.gob.pe/`

3.2.2. Partición.

Para la implementación de los modelos de *ML*, se utilizó validación cruzada estratificada de 10 pliegues (*Stratified 10-Fold Cross-Validation*) sobre el conjunto completo de datos. En cada pliegue, el 90 % de los datos se destinó al desarrollo del modelo y el 10 % restante a la prueba. A su vez, el conjunto de desarrollo se dividió en un 75 % para entrenamiento y un 25 % para validación. Esta estrategia permitió evaluar de manera robusta el rendimiento de los modelos y realizar la afinación de hiperparámetros mediante *grid search* y *Bayesian optimization*, siguiendo prácticas recomendadas en la literatura (Pandey et al., 2021).

3.2.3. Técnica de muestreo de partición.

La selección de la muestra se realizó de forma aleatoria estratificada, considerando que los datos presentan un desequilibrio en las clases (**buen pagador** y **mal pagador**). Este enfoque asegura que la representatividad de las clases se mantenga en cada subconjunto de la partición, mejorando la estabilidad de las estimaciones de desempeño del modelo.

Adicionalmente, se plantearon grupos de evaluación basados en la representatividad de las actividades económicas, tomando en cuenta el número de créditos y la importancia del saldo de capital por sector. Esto permite que el análisis refleje de manera más precisa la distribución real de la cartera crediticia (Pandey et al., 2021).

Esta técnica de muestreo facilita una evaluación más precisa del impacto de los *FE* en diferentes segmentos económicos, garantizando que el modelo pueda adaptarse a las particularidades de cada actividad.

3.3. Operacionalización de variables

Variable independiente

Nombre: Influencia de factores exógenos sobre los días de mora.

Definición conceptual: Representa la incidencia de eventos externos críticos (como pandemia, anomalías climáticas y convulsión social) en la estabilidad financiera de los prestatarios. Está fundamentada en datos públicos provenientes de entidades gubernamentales y en datos privados de una entidad financiera, correspondientes a los periodos en los que se manifestaron los *FE*.

Definición operacional: Resultado del test de causalidad de Granger aplicado a series temporales de los *FE* y a los valores máximos promedio de días de atraso en créditos, agrupados por actividad económica.

Dimensión: Variable binaria: 1 = actividad económica afectada por *FE*; 0 = no afectada.

Variable dependiente

Nombre: Incertidumbre en la predicción del riesgo crediticio.
Definición conceptual: El riesgo crediticio es la posibilidad de incumplimiento de obligaciones por parte de los prestatarios, derivado de la falta de capacidad o intención de pago (Orlova, 2020; Teles et al., 2021; Noriega et al., 2023). Esta investigación se centra en la capacidad de pago, que puede ser influenciada por factores demográficos, financieros y exógenos.
Definición operacional: Medición de la variación en el desempeño de modelos de clasificación para predecir riesgo crediticio, al comparar escenarios con y sin inclusión de variables exógenas.
Dimensiones: Las métricas utilizadas fueron Exactitud (*ACC*) y Área bajo la curva ROC (*AUC*).

3.3.1. Matriz de operacionalización de variables

La Tabla 12 presenta la matriz de operacionalización de variables, incluyendo su nivel de medición y método de análisis.

Tabla 12
Matriz de operacionalización de variables

Variable	Nivel de medición	Método de análisis
Influencia de *FE*	Binaria (0,1)	Causalidad de Granger aplicada a series temporales.
Incertidumbre en la predicción del riesgo crediticio	Continua	Comparación de métricas (*AUC, ACC*) entre escenarios con y sin *FE*.

Nota. Estas variables fueron utilizadas para la validación de hipótesis en la investigación.

Relación con el modelado predictivo

Es importante aclarar que las variables aquí operacionalizadas corresponden al marco de validación de hipótesis y se diferencian del conjunto de atributos usados en el entrenamiento de los modelos predictivos. Este último incluye 367 variables independientes (financieras y exógenas) y una variable dependiente binaria (*morosidad grave*), como se detalla en la Sección 4.1 y en la Tabla 15.

3.4. Técnicas e instrumentos de recolección de datos.

3.4.1. Técnicas

Para la recolección de datos, se identificaron bases de datos públicas que registran el impacto de *FE*. Estos conjuntos de datos fueron obtenidos de acceso libre y descargados para su procesamiento. Entre ellos se incluyen:

- Casos positivos y fallecimientos por *COVID-19*.
- Datos de temperatura histórica y anomalías climáticas.
- Registros de bloqueos de vías, obtenidos mediante gestión directa con el Ministerio de Transportes y Comunicaciones.

Además, se accedió a bases de datos privadas proporcionadas por la entidad financiera *FMOD*. La autorización para el uso de estos datos fue concedida por la gerencia general de la institución, bajo la condición de que no contuvieran información personal de clientes, garantizando el cumplimiento de normativas de privacidad y protección de datos (ver Tabla 13).

3.4.2. Procedimientos de análisis y modelado

En esta investigación, no se utilizaron instrumentos en el sentido tradicional (como cuestionarios o escalas), sino un conjunto de técnicas y procedimientos de análisis propios de la Ciencia de Datos y el Aprendizaje Automático (ML). Estas técnicas se aplicaron al procesamiento de datos tabulares y series temporales, con el objetivo de garantizar la calidad de los datos, extraer características relevantes y construir modelos predictivos robustos.

Herramientas y entorno de desarrollo.

- **Google Colab (versión PRO):** Utilizado para el desarrollo y ejecución de scripts de *ML* en Python, aprovechando sus recursos de cómputo *GPU* y bibliotecas optimizadas.

Análisis y preparación de series temporales.

- Evaluación visual para identificar patrones, tendencias y anomalías en las series temporales.
- Prueba de estacionariedad de Dickey-Fuller para detectar raíces unitarias.
- Test de causalidad de Granger para evaluar relaciones de dependencia entre los *FE* y la morosidad crediticia.

Análisis exploratorio y de correlación.

- Matriz de correlación para identificar relaciones entre variables y reducir riesgos de multicolinealidad.
- Revisión manual de calidad de los datos para asegurar consistencia e integridad.

Evaluación estadística de resultados.

- Prueba t de Student para determinar la significancia estadística de las diferencias en el desempeño de los modelos con y sin inclusión de *FE*.

Estos procedimientos forman parte del pipeline metodológico que permitió entrenar y evaluar 14 algoritmos de clasificación supervisada, cuya configuración detallada se presenta en las Tablas 18 y 19.

Tabla 13
Datasets.

It.	DATASET	TYPE	MB	ROWS	SOURSE	TEST DATE
1	*COVID-19* Positive Cases.	*FE*	333.00	4,523,494	https://www.datosabiertos.gob.pe/	2020.03.16 - 2022.10.28
2	*COVID-19* Death Cases	*FE*	351.00	1,063,795	https://www.datosabiertos.gob.pe/	2020.03.16 - 2022.10.28
3	Road Blockages by S. Unrest	*FE*	58.00	270,600	https://www.gob.pe/mtc/	2022.12.07 - 2023.09.30
4	Road Blockages by weather	*FE*	58.00	270,600	https://www.gob.pe/mtc/	2022.12.07 - 2023.09.30
5	Temperature Anomaly	*FE*	1.20	22,281	https://www.senamhi.gob.pe/	2022.12.07 - 2023.09.30
6	Delinquency Days by Activity	*FMOD*	64.00	1,944,595	Financiera Modelo	2020.01.01 - 2023.09.30
7	Credits	*FMOD*	515.85	367,346	Financiera Modelo	2020.01.01 - 2023.09.30

Nota: Esta tabla resume los conjuntos de datos utilizados en el estudio, incluyendo sus fuentes, el número de registros y las fechas de evaluación en el modelado del riesgo crediticio bajo la influencia de factores externos como la pandemia de *COVID-19*, las anomalías climáticas y la convulsión social. Los conjuntos de datos públicos están disponibles a través de las URL oficiales indicadas y también se encuentran incluidos en el repositorio público. Los datos financieros propietarios, utilizados bajo acuerdo de confidencialidad, han sido anonimizados y agregados para cumplir con las regulaciones de protección de datos. El repositorio adicionalmente contiene metadatos, programas de modelado y resultados experimentales empleados en este estudio: `https://zenodo.org/records/14890903`.

Tabla 14
Matriz de operacionalización de variables.

Variable	Tipo Variable	Definición conceptual	Definición de operacionalidad	Dimensiones	Indicadores
Influencia de *FE* sobre los días de mora.	Independiente.	Datos públicos proporcionados por entes gubernamentales, que han registrado indicadores de magnitud de los *FE*; y datos privados de una entidad Financiera que desarrolló su operación durante el periodo de manifestación del factor externo, en una determinada línea de tiempo.	Resultado del test de causalidad de Granger entre las series temporales del factor externo y los promedios máximos de días de atraso de créditos por actividad económica.	Valor binario, que indica la afectación o no, a la actividad económica del titular de un crédito.	Conjunto de datasets .
Incertidumbre en el pronóstico de riesgo de créditos	Dependiente.	El riesgo de crédito es un componente inherente de la actividad financiera (Orlova, 2020; Teles et al., 2021), manifestándose cuando hay un incumplimiento de las obligaciones adquiridas al obtener un crédito, ya sea por falta de intención o de capacidad de pago.	Desempeño de clasificación del diagnóstico.	Exactitud.	Métrica que indica la proximidad entre el resultado de una medición y el valor verdadero, está vinculada al sesgo de una estimación.
				AUC score.	Relación entre la tasa de detección correcta de casos positivos (*TPR*) y la tasa de identificación errónea de casos negativos como positivos (*FPR*).

3.4.3. Validación y confiabilidad de los modelos de *ML*

La validación y la confiabilidad de los modelos de aprendizaje automático implementados en esta investigación se aseguraron mediante la aplicación de procedimientos estándar en el campo de la Ciencia de Datos y la Inteligencia Artificial. Este enfoque garantiza que los resultados obtenidos sean generalizables y robustos frente a variaciones en los datos.

Validación cruzada con k-folds. Se empleó la técnica de validación cruzada con $k = 10$ *folds* para estimar el desempeño de cada algoritmo de clasificación supervisada. Este método divide el conjunto de datos en 10 subconjuntos de igual tamaño, iterando hasta que cada subconjunto haya sido usado como conjunto de prueba exactamente una vez. Posteriormente, se calcularon las métricas promedio para obtener una estimación estable de la capacidad predictiva de los modelos.

Manejo de desbalance de clases. Considerando que la variable dependiente (*morosidad grave*) presenta un desequilibrio en las clases, se aplicó la técnica *Synthetic Minority Over-sampling Technique* (*SMOTE*). Este método genera nuevas instancias sintéticas de la clase minoritaria para balancear la distribución y evitar que los modelos se sesguen hacia la clase mayoritaria (Liu et al., 2022).

Métricas de desempeño. El desempeño de los modelos se evaluó utilizando métricas ampliamente aceptadas en el campo, tales como:

- ***ACC*:** Proporción de predicciones correctas sobre el total de casos.
- ***AUC*:** Medida de la capacidad de los modelos para distinguir entre las clases.

Confiabilidad y robustez. Además de la validación cruzada, se realizaron análisis de sensibilidad para evaluar la estabilidad de los modelos frente a variaciones en los datos de entrada. Esto incluyó pruebas de desempeño en escenarios con y sin inclusión de variables exógenas. Asimismo, se documentaron los hiperparámetros configurados para cada modelo (ver Tablas 18 y 19), lo cual permite la replicación y auditoría de los experimentos.
Con estos procedimientos, se asegura que los modelos de *ML* desarrollados sean fiables y aplicables a entornos financieros complejos.

3.4.4. Procedimientos

Para la validación de la hipótesis de investigación, se siguieron los siguientes procedimientos, adaptados según el modelo *CRISP-DM* (ver Figura 10).

1. **Análisis de los datos**
    - Obtener los conjuntos de datos.
    - Analizar el fenómeno a evaluar y su representación en los datos disponibles.
    - Identificar la influencia de los *FE* sobre la morosidad crediticia.
2. **Preparación de los datos**
    - Estandarizar y normalizar los datos según las características de cada variable.
    - Estructurar las series temporales para permitir la evaluación de relaciones temporales entre variables.
    - Aplicar métodos estadísticos para garantizar la validez de los datos:
        - Test de estacionariedad de Dickey-Fuller para evaluar la presencia de raíces unitarias en las series temporales.
        - Test de causalidad de Granger para determinar la relación entre los *FE* y la morosidad crediticia, con un umbral de significancia de $p \leq 0.05$. Para simplificar la interpretación, se asignó un valor de 1 si $p \leq 0.05$ y 0 en caso contrario.
3. **Evaluación de los datos**
    - Definir los periodos de evaluación según la disponibilidad y relevancia de los datos.
    - Seleccionar los modelos de *ML* a utilizar.
    - Establecer las condiciones de evaluación:
        - Aplicar validación cruzada estratificada con 10 pliegues (*Stratified k-Fold Cross-Validation* con $k = 10$).
        - En cada pliegue, usar 90 % de los datos para desarrollo del modelo y 10 % para prueba.
        - Dividir el conjunto de desarrollo en 75 % para entrenamiento y 25 % para validación.
        - Aplicar técnicas de balanceo de clases *SMOTE* únicamente sobre los datos de entrenamiento.
    - Aplicar modelos de *ML* sobre los datos sin *FE*, incluyendo:
        - Eliminación de variables con alta correlación para evitar redundancia.
        - Balanceo de clases en los datos de entrenamiento.

- Generar nuevos conjuntos de datos incorporando los *FE*.
- Aplicar modelos de *ML* sobre los datos que incluyen *FE*.
- Comparar los resultados obtenidos en ambos escenarios para evaluar el impacto de los *FE* en la predicción del riesgo crediticio, aplicando *SHAP* y *LIME*, para identificar las variables con mayor representatividad incluyendo o no a los *FE*.

4. **Experimentos y validación de hipótesis específicas**

   - **Hipótesis específica 1**: Análisis del impacto de la pandemia de *COVID-19* en la morosidad crediticia.
     - Se evaluó la relación entre la evolución de los días de mora y los indicadores epidemiológicos (casos positivos y mortalidad).
     - Se aplicó pruebas de estacionariedad (Dickey-Fuller) y causalidad (Granger) para identificar posibles relaciones de dependencia entre la pandemia y la morosidad crediticia.
     - Se comparó el desempeño de los modelos de *ML* con y sin la inclusión de variables de *COVID-19*, utilizando las métricas: *ACC* y *AUC*.
     - Se emplearon técnicas de explicabilidad como *SHAP* y *LIME* para determinar la influencia relativa de los indicadores pandémicos en la predicción del riesgo crediticio.
     - Se utilizó la prueba t de Student para verificar la significancia estadística de las diferencias en el desempeño de los modelos.
   - **Hipótesis específica 2**: Evaluación del efecto de factores climáticos en la predicción del riesgo crediticio.
     - Se analizo la incidencia de anomalías de temperatura y bloqueos de vías por eventos climáticos en la morosidad crediticia.
     - Se determino la estacionariedad y causalidad de las series temporales a través de las pruebas de Dickey-Fuller y Granger, respectivamente.
     - Se entreno modelos de *ML* con y sin la inclusión de variables climáticas para medir su impacto en las métricas de desempeño.
     - Se implemento *SHAP* y *LIME* para interpretar la contribución de los factores climáticos en la predicción del riesgo crediticio.
     - Se valido la significancia estadística de los resultados mediante la prueba t de Student.
   - **Hipótesis específica 3**: Análisis de la relación entre la convulsión social y la morosidad crediticia.

- Se estudio el impacto de los bloqueos de vías por manifestaciones y disturbios sociales en los niveles de mora.
- Se segmento los datos por actividad económica para identificar sectores más vulnerables a estos eventos.
- Se aplico las pruebas de Dickey-Fuller y Granger para evaluar la influencia de la convulsión social sobre la evolución del crédito.
- Se analizo el desempeño de los modelos de *ML* en escenarios con y sin la inclusión de variables de convulsión social.
- Se emplearon *SHAP* y *LIME* para identificar el peso de estos eventos en las predicciones de riesgo crediticio.
- Se evaluo la significancia estadística de los cambios en el desempeño de los modelos mediante la prueba t de Student.

- **Hipótesis específica 4**: Evaluación integral de la interacción entre los factores externos en la predicción de la morosidad crediticia.
  - Se comparo modelos que incorporan individualmente cada factor externo con aquellos que integran múltiples factores simultáneamente.
  - Se analizo posibles efectos sinérgicos entre la pandemia, el cambio climático y la convulsión social en la estabilidad crediticia.
  - Se emplearon técnicas de explicabilidad (*SHAP* y *LIME*) para identificar qué factores tienen mayor impacto en la predicción de morosidad.
  - Se utilizó la prueba t de Student para evaluar si la combinación de factores externos mejora significativamente la capacidad predictiva de los modelos.

5. **Validación de la hipótesis general**

   - Se realizó un análisis comparativo de los modelos evaluados en las hipótesis específicas para determinar si la inclusión de factores externos mejora la precisión en la predicción del riesgo crediticio.
   - Se examinó la estabilidad y robustez de los modelos a lo largo del tiempo y en distintos segmentos económicos.
   - Se aplicó técnicas de explicabilidad para verificar la consistencia de los factores externos en la predicción de morosidad crediticia.
   - Se discutió las implicaciones de los resultados en el sector financiero, proponiendo estrategias para mitigar el riesgo crediticio en entornos de alta incertidumbre.

Capítulo IV

Fase Experimental

Todos los recursos utilizados en esta investigación, incluyendo los *datasets*, modelos generados junto con sus resultados y el código fuente, han sido publicados para su uso libre. Estos recursos están identificados mediante el *Digital Object Identifier* (DOI): **10.5281/zenodo.14890903** y se encuentran disponibles en la siguiente dirección: `https://zenodo.org/records/14890903`.

4.1. Análisis de *datasets*

Como parte de la fase de análisis y preparación de datos, se realizó un exhaustivo proceso de validación para garantizar la coherencia y calidad de la información. Este proceso incluyó:

- **Validación de consistencia lógica:** Revisión sistemática para detectar y corregir valores atípicos, duplicados o inconsistentes.
- **Normalización:** Estandarización de las escalas de las variables para facilitar su integración en los algoritmos de *ML*.
- **Mitigación de ruido:** Considerando posibles distorsiones derivadas de políticas crediticias excepcionales (e.g., facilidades otorgadas durante la pandemia), se utilizó la media del máximo número de días de mora por período como referencia para suavizar anomalías.

La selección del subconjunto de datos y la reducción del tamaño del dataset, se detalla a continuación el procedimiento aplicado:

Criterios de selección y depuración:

- Se incluyeron créditos otorgados entre enero de 2020 y septiembre de 2023 (Madeira, 2022; Espinosa-Méndez, 2022).

- Se eliminaron registros con fechas de cancelación anteriores al inicio del análisis o con morosidad superior a 30 días en el mes base de cada *FE* (Jiang, 2022; Madeira, 2022).
- Se excluyeron variables con colinealidad superior al 90 %, identificadas mediante matrices de correlación (Z. Chen et al., 2020; Geraldo-Campos et al., 2022).
- Mitigar el desbalance de clases aplicando *SMOTE* (Noriega et al., 2023).
- Se consideraron escenarios representativos del 80 %, 99 % y 100 % del saldo de cartera del negocio para el análisis de las actividades económicas, ver Figura 6

Este proceso de reducción y filtrado resultó en dos subconjuntos principales de datos: uno sin factores exógenos (Dataset SIN *FE*) y otro enriquecido con variables derivadas de series temporales de *FE* (Dataset *FE*). La Figura 4 resume el flujo completo de integración y depuración de datos.

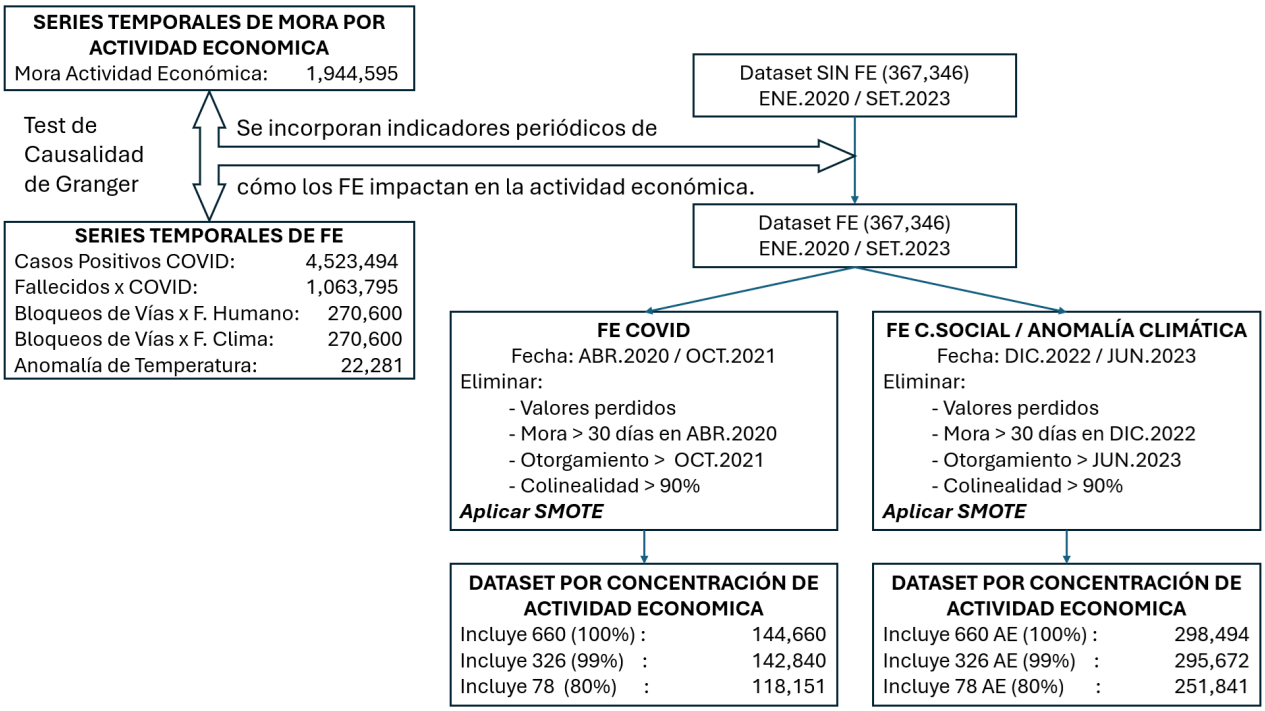


Figura 4
Flujo de reducción y depuración del dataset: integración de series temporales, filtros aplicados y tamaños resultantes.

Definición de variables para el modelado predictivo

El conjunto de datos final se estructuró para un problema de clasificación binaria, empleando un total de 365 variables independientes (*features*) y una variable dependiente (*target*).

Variables independientes (*features*): Se incluyeron tanto atributos financieros del prestatario y del crédito (e.g., edad, género, monto aprobado, tasa de interés, plazos, puntajes internos de riesgo e historial crediticio), como variables exógenas derivadas (casos positivos y fallecimientos por *COVID-19*, bloqueos sociales diarios, anomalías de temperatura). La lista detallada se presenta en la Tabla 6.

Variable dependiente (*target*): Variable binaria que toma el valor 1 si el crédito presentó morosidad grave (más de 30 días de atraso) durante el período de análisis y 0 en caso contrario.
La Tabla 15 sintetiza las variables utilizadas en el modelado, mientras que las configuraciones de los algoritmos aplicados se encuentran en las Tablas 18 y 19.

Tabla 15
Resumen de variables empleadas en el modelado predictivo

Tipo	Variable	Descripción
Variable dependiente (*target*)	Morosidad grave	Variable binaria: 1 = mora mayor a 30 días, 0 = no mora.
Variables independientes (*features*)	Financieras	Edad, género, monto del crédito, tasa de interés, plazo en días, puntajes internos de riesgo, historial crediticio, entre otros (ver Tabla 6).
	Exógenas	Casos positivos *COVID-19*, fallecimientos *COVID-19*, bloqueos sociales diarios, anomalías de temperatura (ver Sección 4.1.3).

Nota. La variable dependiente fue utilizada para entrenar y validar 14 algoritmos de clasificación supervisada.

4.1.1. Ventajas del *dataset* seleccionado

El conjunto de datos utilizado en esta investigación presenta diversas ventajas que contribuyen a la validez y aplicabilidad de los resultados obtenidos:

1. **Fuentes confiables y acceso abierto:** Los datos relacionados con *FE* provienen de entidades gubernamentales y están disponibles para acceso público. En cuanto a la información privada proporcionada por la entidad financiera FMOD, si bien se identificaron ciertos sesgos y correlaciones, estos fueron abordados mediante técnicas de limpieza y preprocesamiento de datos.

2. **Gran volumen de registros:** El conjunto de datos original está compuesto por más de **8.3 millones de registros**, lo que permite realizar un análisis robusto y

con alta representatividad. Aunque en esta investigación se utilizó un subconjunto, la disponibilidad de un volumen amplio de información permite futuras ampliaciones del estudio y la realización de nuevos análisis.

3. **Aplicabilidad en el sector financiero:** Los modelos de *ML* entrenados en esta investigación pueden ser utilizados para mejorar la clasificación del riesgo crediticio, optimizando los procesos de originación y administración de créditos. En la etapa de originación, estos modelos pueden contribuir a la evaluación del nivel de riesgo que una entidad está dispuesta a asumir, mientras que en la administración crediticia pueden facilitar estrategias de cobranza temprana y mitigación de riesgo.

4.1.2. Limitaciones del *dataset* seleccionado

Si bien el conjunto de datos presenta múltiples ventajas, también existen ciertas limitaciones que deben considerarse en la interpretación de los resultados:

1. **Requerimientos computacionales elevados:** Debido a su gran tamaño, el procesamiento del conjunto de datos demanda una infraestructura computacional robusta y tiempos de ejecución prolongados, especialmente en las etapas de modelado y validación.

2. **Sesgo geográfico:** La investigación se centra en datos provenientes de Perú, debido a las restricciones de acceso a información financiera en otros países. Aunque esta limitación reduce la generalización directa de los hallazgos, también representa una oportunidad para replicar la metodología en otras regiones y contextos económicos.

3. **Calidad de los datos:** Se identificaron valores ausentes en algunas variables, lo que requirió la aplicación de técnicas de imputación y limpieza de datos para garantizar la integridad del análisis.

4. **Presencia de ruido en la información:** Factores exógenos, como las medidas gubernamentales de alivio financiero durante la pandemia y la interpretación discrecional de estas normas por parte de las entidades crediticias, introdujeron variabilidad en los datos. Para mitigar este efecto, se emplearon técnicas de normalización y de filtrado para garantizar la consistencia en los patrones de mora analizados.

4.1.3. Aplicación de *Data Augmentation* con *FE*

Para estructurar la evaluación del impacto de los *FE*, se adoptó el modelo *CRISP-DM*, ampliamente reconocido por su aplicabilidad en el análisis de datos en contextos de

crisis e incertidumbre (Mann y Logeswaran, 2021; Sahiq et al., 2022) (ver Figura 10). Este enfoque se inicia con la identificación y selección de los conjuntos de datos relevantes, tanto públicos como privados, seguido de un análisis integral del fenómeno de estudio. En esta etapa, se incluyeron variables exógenas derivadas de series temporales (e.g., fallecimientos por *COVID-19*, anomalías de temperatura, bloqueos viales y episodios de convulsión social), junto con variables financieras tradicionales. A partir de estos insumos, se procedió a la selección de características clave que reflejan la influencia directa o indirecta de los *FE* sobre la morosidad crediticia.

Como parte del preprocesamiento de datos, se aplicaron técnicas de estandarización y normalización por suma, con el fin de garantizar la homogeneidad en las escalas y facilitar la comparabilidad entre las variables. Para evaluar la relación entre la morosidad y los *FE*, se utilizó la prueba de estacionariedad de Dickey-Fuller (ver Ecuación 1), aplicada a series temporales normalizadas de frecuencia diaria, representando los días de mora por actividad económica. Este análisis cubrió el período comprendido entre enero de 2016 y septiembre de 2023 (ver Tabla 13), diferenciando entre periodos con y sin influencia de *FE* (Awad y Al Karaki, 2019).

Los resultados obtenidos, resumidos en la Tabla 16, evidencian que, en periodos sin influencia de *FE*, entre el 90 % y el 91 % de las actividades económicas presentan patrones de morosidad no estacionarios, mientras que solo entre el 6 % y el 7 % muestran estacionariedad, y un 3 % se clasifica como indeterminado. En contraste, durante los periodos con presencia de *FE*, la proporción de actividades no estacionarias se reduce (75 %–85 %), mientras que aumentan los casos estacionarios (6 %–22 %) e indeterminados (2 %–9 %). Este comportamiento sugiere que los shocks exógenos —como pandemias, protestas sociales o disrupciones climáticas— tienden a inducir patrones de morosidad más persistentes o complejos, consistentes con hallazgos documentados en otras economías emergentes.

La categoría "indeterminada", aunque minoritaria, destaca casos donde los patrones de morosidad son más difíciles de clasificar, posiblemente debido a dinámicas atípicas o a comportamientos no lineales inducidos por la superposición de múltiples *FE*. Estos resultados reflejan cómo los entornos estables permiten a las instituciones financieras anticipar y mitigar el riesgo mediante ajustes tempranos en la exposición crediticia, reduciendo así la presencia de patrones persistentes. Sin embargo, en contextos de shock exógeno simultáneo, la capacidad de respuesta institucional puede verse limitada, lo que conduce a un incremento en la duración de las moras y, por ende, en la estacionariedad observada.

En la etapa de modelado, se implementaron algoritmos de *ML* empleando una estrategia de validación cruzada estratificada de k-folds ($k = 10$), asegurando una representación balanceada de clases en cada partición. En cada pliegue, el 90 % de los datos se destinó al desarrollo del modelo y el 10 % restante a la prueba. A su vez, el conjunto de

desarrollo se dividió internamente en 75 % para entrenamiento y 25 % para validación. Esta configuración permitió realizar una afinación robusta de hiperparámetros utilizando técnicas de *grid search* y *Bayesian optimization* (Frazier, 2018), y calcular métricas como *Accuracy* y *AUC* de manera independiente en las particiones de validación y prueba.

Posteriormente, se clasificaron los modelos en un ranking comparativo en función de su rendimiento promedio, con el objetivo de evaluar si la incorporación de *FE* —como pandemia, bloqueos o clima extremo— genera mejoras estadísticamente significativas en su capacidad predictiva. Esta metodología permite no solo medir la precisión en distintos escenarios de incertidumbre, sino también analizar el efecto estructural de los *FE* sobre el comportamiento de la morosidad, ofreciendo así una herramienta predictiva útil para instituciones financieras que operan en contextos de riesgo sistémico.
Para analizar el impacto diferencial de los *FE*, se establecieron tres grupos de evaluación basados en la representatividad de las actividades económicas en el conjunto de datos, considerando tanto el número de operaciones como el monto de crédito otorgado. Estos grupos son:

- **1.D100**: Incluye todas las actividades económicas registradas en el conjunto de datos.
- **2.D99**: Representa el 99 % de los créditos otorgados y del saldo de capital.
- **3.D80**: Conformado por 78 actividades económicas que abarcan el 80 % de los créditos otorgados y del saldo de capital.

Tabla 16

Estacionariedad de la Mora de Actividades Económicas

Estacionariedad de Mora de Actividades Económicas Según Rango de Evaluación										
Clasificación	#					%				
	Sin FE			Con FE		Sin FE			Con FE	
	2018.04 2020.02	2019.01 2019.12	2020.03 2023.09	2020.04 2022.10	2022.12 2023.06	2018.04 2020.02	2019.01 2019.12	2020.03 2023.09	2020.04 2022.10	2022.12 2023.06
Serie no estacionaria.	647	625	465	490	404	90 %	91 %	75 %	81 %	85 %
Serie estacionaria.	49	40	137	100	29	7 %	6 %	22 %	17 %	6 %
No determinado	22	21	15	16	41	3 %	3 %	2 %	3 %	9 %
Total	718	686	617	606	474	100 %	100 %	100 %	100 %	100 %

Nota: Los resultados obtenidos, resumidos en la Tabla 16, evidencian que, en periodos sin influencia de *FE*, entre el 90 % y el 91 % de las actividades económicas presentan patrones de morosidad no estacionarios, mientras que solo entre el 6 % y el 7 % muestran estacionariedad, y un 3 % se clasifica como indeterminado. En contraste, durante los periodos con presencia de *FE*, la proporción de actividades no estacionarias se reduce (75 %–85 %), mientras que aumentan los casos estacionarios (6 %–22 %) e indeterminados (2 %–9 %). Este comportamiento sugiere que los shocks exógenos —como pandemias, protestas sociales o disrupciones climáticas— tienden a inducir patrones de morosidad más persistentes o complejos, consistentes con hallazgos documentados en otras economías emergentes.

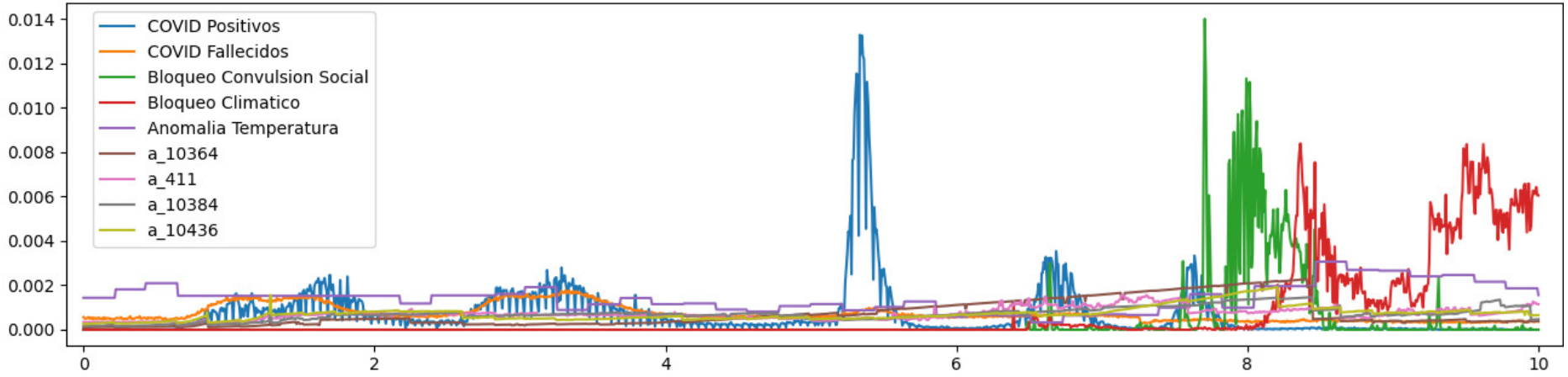


Figura 5

Series temporales de FE y actividades económicas

En este análisis, se observa que, a medida que se incluyen más actividades económicas, aumenta la dispersión de los datos y, en consecuencia, el nivel de ruido en los modelos (ver Figura 6). Adicionalmente, se incluyó una variable que identifica los créditos que recibieron apoyo gubernamental para mitigar los efectos de los *FE*, permitiendo evaluar el impacto de estas medidas al incluir y excluir dichos registros.

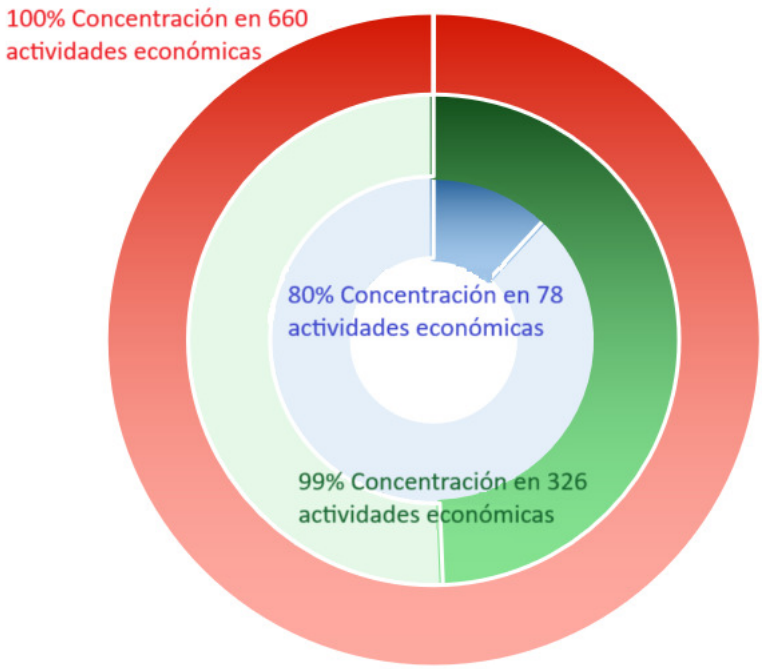


Figura 6
Relevancia de Actividad Económica.

Para complementar el conjunto de datos estandarizado, se agregaron nuevas variables en función de la evaluación de los *FE*:

- **Relación de causalidad:** Determinada mediante la prueba de causalidad de Granger, evaluando la relación entre cada *FE* y los días de mora por actividad económica. Se utilizó una frecuencia mensual con un desfase (*lag*) de hasta 3 períodos. Para el caso del *FE*: Anomalía de Temperatura, se empleó una frecuencia trimestral con el mismo desfase (ver Figura 7).

- **Días de Mora:** Se consideraron los valores máximos de días de mora registrados un día antes del cierre del mes, con el fin de minimizar la distorsión introducida por reprogramaciones masivas derivadas de programas gubernamentales de alivio financiero.

- **Variable de Clasificación:** Se definieron clasificaciones mensuales para cada actividad económica en función de su afectación por los *FE*. Además, se estableció una variable de clasificación global que evalúa el impacto acumulado

de los *FE* en el periodo de estudio. Las Figuras 8 y 9 ilustran la clasificación basada en los días de mora máximos (*Class_YYYYMMBM*) y los días vencidos (*Class_DV_YYYYMMBM*), mostrando una tendencia similar en ambos casos.

	2020-01	2020-02	2020-03	2020-04	2020-05	2020-06	2020-07	2020-08	2020-09	2020-10
FIS_10364	NA	NA	1	1	1	0	0	0	1	0
FIS_10384	NA	NA	1	1	1	0	1	0	1	0
FIS_411	NA	NA	1	1	0	0	1	0	1	0
FIS_10436	NA	NA	1	1	1	0	1	0	1	0
FIS_10449	NA	NA	1	1	1	0	1	0	0	0
FINS_10366	NA	NA	1	1	1	0	0	1	0	0
FINS_10324	NA	NA	1	1	1	0	1	0	1	0
FINS_10354	NA	NA	1	1	1	0	1	0	1	0
FINS_403	NA	NA	1	1	1	0	1	0	1	0
FINS_10453	NA	NA	1	1	1	1	1	0	0	1
FINS_10449	NA	NA	1	1	1	0	1	0	0	0
FINS_10319	NA	NA	1	1	1	1	1	0	0	1
FDev_Total	NA	NA	1	1	1	1	1	0	0	1
FProv_Total	NA	NA	1	1	1	1	1	0	0	1
FDev_FS	NA	NA	1	1	1	1	1	0	0	1
FProv_FS	NA	NA	1	1	1	1	1	0	0	1

	2020-11	2020-12	2021-01	2021-02	2021-03	2021-04	2021-05	2021-06	2021-07	2021-08
FIS_10364	0	0	1	0	1	1	0	0	1	1
FIS_10384	0	0	1	1	0	0	1	0	1	0
FIS_411	0	0	0	0	1	1	1	1	0	0
FIS_10436	0	1	1	1	1	0	1	1	1	1
FIS_10449	0	0	0	1	1	0	1	0	1	1
FINS_10366	1	1	1	1	0	1	0	0	0	0
FINS_10324	0	0	0	0	1	0	1	0	0	1
FINS_10354	0	0	0	0	0	0	1	1	1	1
FINS_403	0	0	0	0	1	1	0	0	1	0
FINS_10453	1	0	0	0	1	1	0	0	0	0
FINS_10449	0	0	0	1	1	0	1	0	1	1
FINS_10319	0	0	0	0	1	0	1	0	1	1
FDev_Total	0	0	0	0	1	0	1	0	1	1
FProv_Total	0	0	0	0	1	0	1	0	1	1
FDev_FS	0	0	0	0	1	0	1	0	1	1
FProv_FS	0	0	0	0	1	0	1	0	1	1

Figura 7

Simplificación de afectación, a partir del Test de Granger.

El periodo de evaluación se determinó a partir de la evolución de las variables de clasificación mensuales, estableciendo su inicio en el punto en que comienza el impacto del *FE* y su finalización cuando dicha variable muestra signos de estabilización (ver Figuras 8 y 9), en estos gráficos la línea roja representa el comportamiento de la mora en relación con los *FE*.

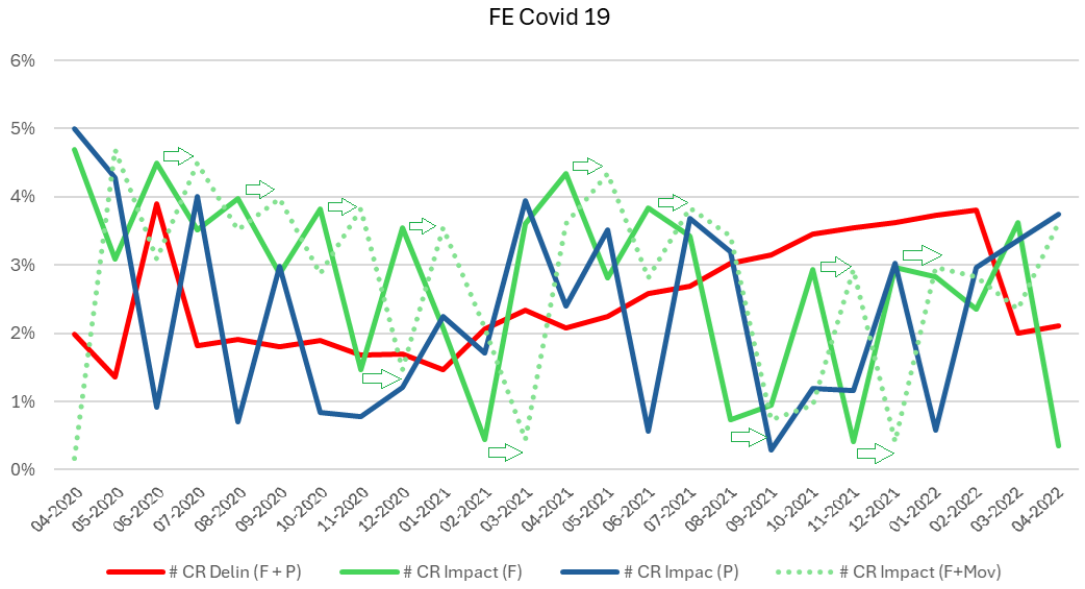


Figura 8

Actividades Económicas impactadas por FE COVID.

Esta figura presenta métricas de impacto en el riesgo crediticio derivadas del contexto pandémico, utilizando las siguientes notaciones:

***F** = Fallecimientos por COVID-19, **P** = Casos positivos de COVID-19, **F+P** = Impacto combinado de fallecimientos y casos positivos por COVID-19, **F+Mov** = Impacto de fallecimientos movido un periodo a la derecha.*

La Figura 8 representa el impacto del *COVID-19* en las actividades económicas mediante la comparación de las tasas de morosidad con los efectos económicos a lo largo del tiempo. El gráfico resalta las fluctuaciones en las tendencias de impacto en actividades económicas, para los escenarios de Fallecidos (*F*); y casos positivos de *COVID-19* (*P*), especialmente durante los picos de la pandemia.

Adicionalmente, se incluye el escenario de Fallecidos con un desplazamiento temporal de un período (*F+MOV*), lo que permite capturar de manera más realista la relación entre el diagnóstico, el tratamiento y el desenlace crítico de la enfermedad (Organización Mundial de la Salud, 2024). Este ajuste temporal mejora la evaluación del impacto del *COVID-19* en la morosidad crediticia, alineando mejor los efectos económicos con la progresión de la pandemia.

La Figura 9 ilustra el impacto de la convulsión social y los factores climáticos en la actividad económica desde diciembre de 2022 hasta septiembre de 2023. El gráfico compara las tasas de morosidad con el impacto en actividades económicas para la convulsión social (*U*), bloqueos por condiciones climáticas (*W*) y las anomalías de temperatura (*T*), destacando picos significativos durante los períodos de intensificación de estas perturbaciones externas.

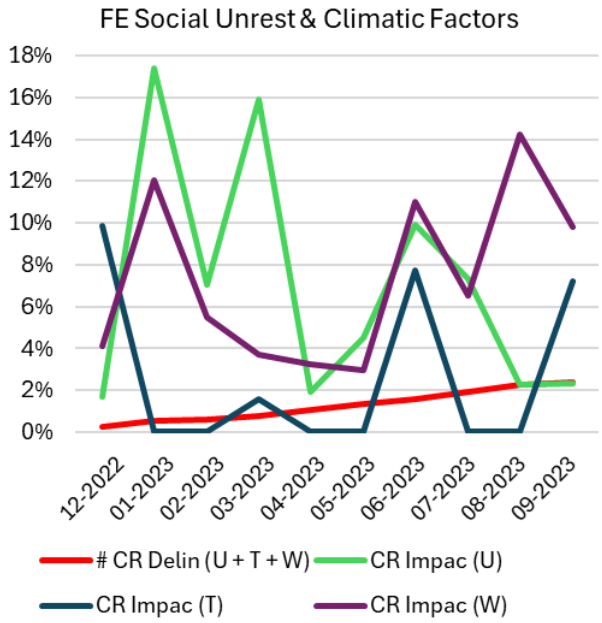


Figura 9
Actividades Económicas impactadas por FE Convulsión Social, Anomalía de Temperatura y Bloqueo de vías por factor Climático.
Esta figura presenta métricas de impacto en el riesgo crediticio derivadas de FE, utilizando las siguientes notaciones: ***U*** *= Convulsión social,* ***T*** *= Anomalía de temperatura,* ***W*** *= Bloqueos viales por condiciones climáticas,* ***U+T+W*** *= Efecto combinado de convulsión social, anomalía de temperatura y bloqueos viales.*

4.2. Modelo Propuesto

La metodología de análisis adoptada en este estudio se basa en el modelo *CRISP-DM*. Este enfoque ha demostrado ser efectivo en la evaluación de *FE* en contextos de crisis económicas y financieras, como se ha documentado en estudios previos (Mann y Logeswaran, 2021; Sahiq et al., 2022) (ver Figura 10).
El proceso seguido consta de las siguientes fases:

1. **Selección y adquisición de datos**: Se identificaron y consolidaron conjuntos de datos provenientes de fuentes oficiales y registros internos de la entidad financiera en estudio.

2. **Análisis exploratorio y validación de datos**: Se aplicaron pruebas estadísticas para evaluar la calidad de la información, detectando valores atípicos, datos faltantes y posibles inconsistencias.

3. **Transformación y preprocesamiento**: Se normalizaron las series temporales y se implementaron técnicas de suavizado para reducir la influencia del ruido en los datos.

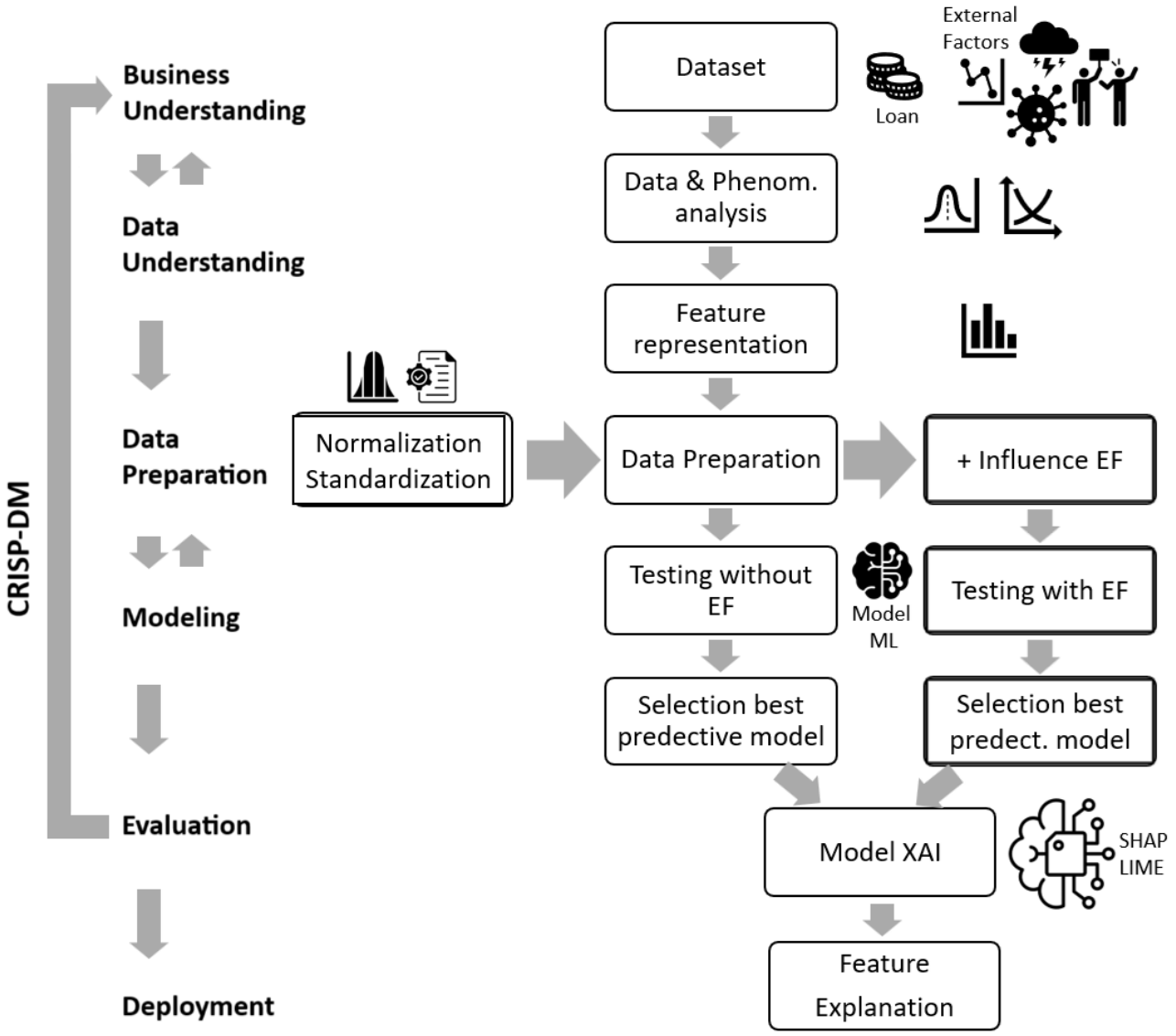


Figura 10

Modelo CRISP-DM para el análisis de FE.

Este diagrama ilustra cómo las fases del modelo CRISP-DM estructuraron el flujo experimental, desde la comprensión y preparación de los datos —incluyendo la alineación mensual e integración de variables exógenas en series temporales— hasta el modelado basado en escenarios (con y sin factores externos) y la aplicación de técnicas de inteligencia artificial explicable (SHAP y LIME) durante la fase de evaluación. Las variables `DíasMora_MMYYYY` *y* `ImpactoFactorExterno_MMYYYY` *fueron construidas para permitir un análisis temporal del comportamiento crediticio y su relación con shocks externos validados en distintos sectores económicos.*

4. **Selección de características**: Se emplearon métricas de importancia de variables y métodos de reducción de dimensionalidad para determinar los factores más relevantes en la predicción del riesgo crediticio.

5. **Modelado y evaluación**: Se entrenaron modelos de *ML* con y sin la inclusión de *FE*, comparando su desempeño a través de las métricas: *ACC* y *AUC*.

6. **Explicabilidad y validación**: Se utilizaron las técnicas de interpretabilidad *SHAP* y *LIME*, para evaluar la influencia de cada *FE* en la predicción del riesgo crediticio.

4.3. Implementación de Modelos

La selección de los algoritmos implementados responde a una estrategia de evaluación sistemática fundamentada en el estado del arte y en la necesidad de adaptar modelos a un contexto financiero afectado por factores exógenos (pandemia, eventos climáticos y convulsión social).

Diversos estudios recientes en riesgo crediticio evidencian que la efectividad de los algoritmos varía según el tipo de datos, el grado de desequilibrio en las clases y la presencia de choques externos. Por ejemplo, *XGB* y *Random Forest* han demostrado un alto desempeño en datos estructurados y balanceados (Imteaj y Amini, 2022; Aranha y Bolar, 2023), mientras que modelos como *CNN* y *MLP* destacan en series temporales con alta volatilidad, incluso en análisis de datos tabulares (Jiang, 2022; Merćep et al., 2020; C.-C. Chen et al., 2020). Además, modelos como Explainable Boosting Machine (*EBM*) ofrecen ventajas en interpretabilidad, un aspecto crítico para entornos regulatorios.

Considerando esta diversidad y la falta de un modelo universalmente superior, se diseñó un benchmarking de 14 algoritmos de diferentes familias: regresiones lineales, máquinas de soporte vectorial, árboles de decisión, ensambladores (e.g., *XGB*, *Random Forest*), redes neuronales profundas y modelos explicables. Este enfoque permite comparar su rendimiento en un contexto peruano altamente disruptivo, así como evaluar sus capacidades explicativas con técnicas *XAI* (*SHAP*, *LIME*). La Tabla 17 presenta un resumen de los algoritmos considerados y su respaldo en la literatura.

Tabla 17
Algoritmos evaluados y referencias en la literatura sobre riesgo crediticio

Ít.	Algoritmo	Referencias en la literatura
1	Linear Discriminant Analysis (*LDA*)	Aguilar-Valenzuela y Vilca (2024), Quinteros-Navarro y Rodriguez (2024), Alam et al. (2020)
2	Quadratic Discriminant Analysis (*QDA*)	Aguilar-Valenzuela y Vilca (2024), Alam et al. (2020)
3	Ridge Classifier (*Ridge*)	Geraldo-Campos et al. (2022), Dumitrescu et al. (2022)
4	Perceptron	Aranha y Bolar (2023), Wen et al. (2021)
5	Multilayer Perceptron (*MLP*)	Jiang (2022), , Wen et al. (2021), Aranha y Bolar (2023)
6	Feedforward Neural Network (*FFN*)	Jiang (2022), Aranha y Bolar (2023), Song et al. (2023), L. Zhang y Liang (2023)
7	FFN con Regularización Esparsa (*FFN-Sparse*)	Jiang (2022), Merćep et al. (2020), Song et al. (2023), L. Zhang y Liang (2023)
8	FFN con Penalización Local (*FFN-Local*)	Merćep et al. (2020), Song et al. (2023), L. Zhang y Liang (2023)
9	XGBoost (*XGB*)	Noriega et al. (2023), Imteaj y Amini (2022), Zheng (2019)
10	Random Forest (*RF*)	Imteaj y Amini (2022), Alam et al. (2020), Wen et al. (2021)
11	K-Nearest Neighbors (*KNN*)	Aguilar-Valenzuela y Vilca (2024), Alam et al. (2020)
12	Explainable Boosting Machine (*EBM*)	Kemp et al. (2022), Ma et al. (2023)
13	Extreme Neural Network (*XNN*)	Noriega et al. (2023), Pavitha y Sugave (2024)
14	Convolutional Neural Networks (*CNN*)	Jiang (2022), Noriega et al. (2023), C.-C. Chen et al. (2020), X. Zhang et al. (2024)

Nota. La selección incluye algoritmos discriminantes, clasificadores ridge, perceptrones, redes neuronales profundas con variantes de regularización, ensambladores avanzados (*XGBoost, RF, EBM*) y técnicas explicables, todos respaldados por estudios recientes sobre riesgo crediticio.

En la Tabla 40 se presenta el pseudocódigo correspondiente a la implementación del modelo de clasificación binaria, cuyo objetivo es categorizar a los prestatarios en **buenos** o **malos pagadores**. Agrupamos los modelos de *ML* en cinco categorías principales:

- **Modelos Tradicionales:** Basados en enfoques estadísticos clásicos para la clasificación (Alam et al., 2020):
    - Análisis Discriminante Lineal (*LDA*).

    - Análisis Discriminante Cuadrático (*QDA*).
    - Clasificador de k Vecinos Más Cercanos (*kNN*).

- **Modelos Avanzados:** Técnicas de aprendizaje supervisado con alta capacidad de generalización (Dumitrescu et al., 2022; Zheng, 2019; Wen et al., 2021):
    - Perceptrón.
    - Perceptrón Multicapa (*MLP*).
    - Bosques Aleatorios (*RF*).
    - Regresión Ridge (*Ridge*).
    - Extreme Gradient Boosting (*XGB*).

- **Modelos Explicables:** Diseñados para equilibrar precisión e interpretabilidad:
    - Explainable Neural Network (*XNN*): Modelo de red neuronal que integra interpretabilidad en su arquitectura (Pavitha y Sugave, 2024).
    - Explainable Boosting Machines (*EBM*): Algoritmo basado en modelos aditivos, combinando alta precisión con explicabilidad (Ma et al., 2023).

- **Redes Neuronales Feedforward:** Se exploraron variantes de redes neuronales de alimentación directa:
    - *Red Neuronal Feedforward Estándar (FFN):* Arquitectura clásica utilizada como referencia.
    - *Red Neuronal Feedforward con Regularización Esparsa (FFNSP):* Implementa restricciones para mejorar la generalización y reducir el sobreajuste (Song et al., 2023).
    - *Red Neuronal Feedforward con Penalización Local (FFNLP):* Aplica términos de penalización local para equilibrar la contribución de las variables y mejorar la interpretabilidad (L. Zhang y Liang, 2023).

- **Modelos de Aprendizaje Profundo:** Algoritmos diseñados para capturar patrones complejos en datos de alta dimensionalidad:
    - Redes Neuronales Convolucionales (*CNN*): Utilizadas por su capacidad de extraer relaciones espaciales y temporales en los datos (Merćep et al., 2020).

Para garantizar una evaluación robusta de los modelos y minimizar el sobreajuste, se emplea la técnica de validación cruzada con k-folds (k=10). Adicionalmente, considerando la presencia de desbalance en las clases, se implementa la técnica de sobremues-

treo *Synthetic Minority Over-sampling Technique* (*SMOTE*) para mejorar la representatividad de la clase minoritaria y optimizar la capacidad predictiva de los modelos (Liu et al., 2022).

Configuración de Hiperparámetros

Con el fin de proporcionar una descripción estructurada y detallada de la configuración experimental, los modelos evaluados en esta investigación se han clasificado en dos grandes grupos: modelos clásicos y ensambladores, y modelos de redes neuronales profundas. Esta separación permite destacar las diferencias en términos de complejidad, interpretabilidad y requisitos computacionales entre las distintas familias de algoritmos.

Tabla 18

Hiperparámetros de modelos clásicos y ensambladores implementados

Ítem	Modelo	Hiperparámetros principales
1	Ridge Classifier (*Ridge*)	Regularización L2 (α): 1.0 Solver: auto
2	Linear Discriminant Analysis (*LDA*)	Solver: svd Shrinkage: None
3	Quadratic Discriminant Analysis (*QDA*)	Regularización: 0.0
4	Random Forest (*RF*)	Número de árboles: 100 Profundidad máxima: None Criterio: Gini
5	XGBoost (*XGB*)	Número de estimadores: 100 Learning rate: 0.1 Máxima profundidad: 6 Regularización: L1 (α=0.01), L2 (λ=1)
6	K-Nearest Neighbors (*KNN*)	Número de vecinos (k): 5 Métrica: Euclidean
7	Exp Boosting Machine (*EBM*)	Número de bins: 256 Learning rate: 0.01 Máximo de iteraciones: 5000

Nota. Los hiperparámetros corresponden a modelos clásicos y ensambladores.

En la Tabla 18 se presentan los modelos clásicos y ensambladores, que incluyen enfoques lineales, discriminantes, métodos basados en árboles y técnicas de ensamblado ampliamente utilizadas en problemas de clasificación. Por otro lado, la Tabla 19 detalla los hiperparámetros de las redes neuronales profundas evaluadas, las cuales fueron seleccionadas por su capacidad de capturar relaciones no lineales y complejas en los datos, especialmente relevantes en contextos con factores exógenos disruptivos como los analizados en este estudio.

Ambas tablas recogen la configuración de los 14 modelos aplicados, mostrando de forma transparente los parámetros clave utilizados durante la fase de entrenamiento, lo

que garantiza la reproducibilidad de los experimentos realizados.

Tabla 19

Hiperparámetros de modelos de redes neuronales profundas implementados

Ítem	Modelo	Hiperparámetros principales
1	Perceptron	Learning rate: 0.01 Epochs: 50 Batch size: 32
2	Multilayer Perceptron (*MLP*)	Capas ocultas: [128, 64] Activación: ReLU Optimizador: Adam (learning rate=0.001) Epochs: 50 Batch size: 32
3	Feedforward Neural Network (*FFN*)	Capas ocultas: [128, 64, 32] Activación: ReLU Regularización: Dropout (rate=0.5) Optimizador: Adam (learning rate=0.001) Epochs: 50 Batch size: 32
4	FFN con Regularización Esparsa (F*FN-Sparse*)	Igual a FFN + Penalización L1 (λ=0.01)
5	FFN con Penalización Local (*FFN-Local*)	Igual a FFN + Penalización Laplaciana
6	Extreme Neural Network (*XNN*)	Capas ocultas: [256, 128] Activación: tanh Regularización: L2 (λ=0.01) Optimizador: SGD (learning rate=0.01) Epochs: 30
7	Convolutional Neural Network (*CNN*)	Capas convolucionales: 2 (Conv1D) Filtros: [32, 64], kernel size: 3 Pooling: MaxPooling1D (pool size=2) Capas densas: Dense(64), salida sigmoidal Optimizador: Adam (learning rate=0.001) Epochs: 10 Batch size: 32

Nota. Los hiperparámetros corresponden a las redes neuronales profundas implementadas en este estudio.

Capítulo V

Resultados y discusión

5.1. Resultados.

Los procesos de evaluación se determinaron para las pruebas realizadas de la siguiente manera, ver Figura 11:

- *FE COVID-19*: El período de evaluación se aplicó entre marzo de 2020 y enero de 2022. Los créditos para evaluar se consideraron hasta 3 meses antes, es decir, hasta octubre de 2021. Dado que los créditos pueden tener un período de gracia de hasta 60 días, esto significa que aquellos desembolsados durante este período pueden no cumplir la condición de estar en mora, aunque la actividad económica esté afectada por el *FE*.
- *FE* de Convulsión Social y Anomalía de Temperatura: El período de evaluación se aplicó entre diciembre de 2022 y septiembre de 2023. Los créditos para evaluar se consideraron hasta 3 meses antes, es decir, hasta junio de 2023. Dado que los créditos pueden tener un período de gracia de hasta 60 días, esto significa que aquellos desembolsados durante este período pueden no cumplir la condición de estar en mora, aunque la actividad económica esté afectada por los *FE*.

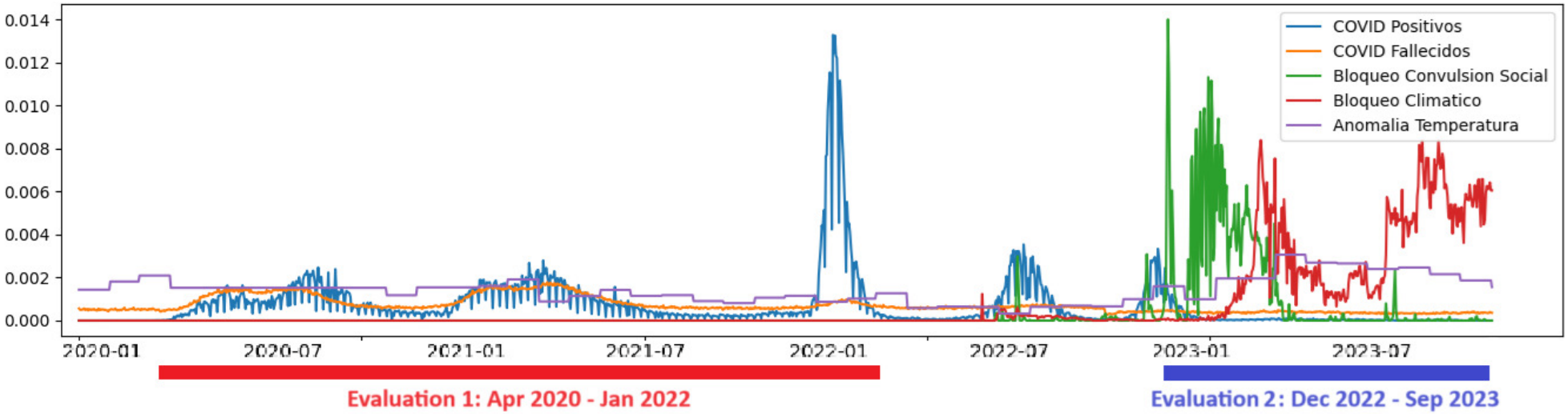


Figura 11

Periodos de evaluación de series temporales

Algorithm 1 Evaluación del conjunto de Dataset mediante algoritmos de *ML*

function EVALUATE($\mathbf{data}, \mathbf{target}, \mathbf{features}, \mathbf{model\_name}$)
1. Remove collinear variables (correlation $> 90\,\%$)
2. Apply Stratified *k*-Fold Cross-Validation ($k = 10$)
for each fold **do**
a. Use 90 % of the data for model development, 10 % for testing
b. Within development set: 75 % for training, 25 % for validation
c. Apply SMOTE on training set only
d. Train selected model: *LDA*, *QDA*, *Perceptron*, *MLP*, *Ridge*, *RF*, *XGB*, *KNN*, *DL*, *CNN*, *XNN*, *FFN* (Standard, Sparse, Local Penalty), *EBM*
e. Compute metrics: AUC_Val, ACC_Val, AUC_Test, ACC_Test
end for
return Mean AUC & ACC for validation & test sets
end function

function EVALUATE_ALL_MODELS($\mathbf{data}, \mathbf{target}, \mathbf{features}$)
1. Iterate over models & feature sets
2. Evaluate on data: **1.D100**, **2.D99**, **3.D80**
return DataFrame with results
end function

function EVALUATE_XAI_BEST_MODELS($\mathbf{best\_Model}$)
1. Evaluate *SHAP*
2. Evaluate *LIME*
return DataFrame with *SHAP* & *LIME* results
end function

function FEATURES_EXPLANATION($\mathbf{best\_Model}$)
end function

Para la evaluación del *FE* se utilizó el *AUC* y *ACC* como métricas, ya que son adecuadas para validar diferentes modelos de aprendizaje automático, aplicando el proceso definido en el algoritmo 1. Los conjuntos de datos evaluados en este experimento fueron:

- Sin *FE* (SFE): se refiere a los datos originales sin incluir los campos relacionados con el *FE*.
- Con *FE*: se refiere a los datos originales a los que se le incluyen atributos relacionados con el *FE*. En los escenarios en que se evaluaron diversos *FE*, se analizaron de manera separada y también en conjunto, en la Tabla 20, se muestran los acrónimos de los mismos y la denominación de sus resultados. La definición de los escenarios en términos de variables se muestran en la Tabla 41. Los programas usados para la preparación de la información, entrenamiento, y validación de significancia estadistica se muestran en los Anexos B, C,E, F, y G.

Tabla 20

Acrónimos de Escenarios Evaluados y sus resultados

Ítem	Escenarios	Acrónimo
1	Sin *FE*	*SFE*
2	Con *FE* Fallecidos por *COVID-19*	*F*
3	Con *FE* Positivos por *COVID-19*	*P*
4	Con *FE* Fallecidos y Positivos por *COVID-19*	*F+P*
5	Resta entre resultado de *F* menos *SFE*	F-*SFE*
6	Resta entre resultado de *P* menos *SFE*	P-*SFE*
7	Resta entre resultado de *F+P* menos *SFE*	*F+P-SFE*
8	Con *FE* Convulsión Social	*U*
9	Con *FE* Anomalía de Temperatura	*T*
10	Con *FE* Bloqueos por Clima	*W*
11	Con *FE* Convulsión Social y Anomalía de Temperatura	*U+T*
12	Con *FE* Anomalía de Temperatura y Bloqueos por Clima	*T+W*
13	Con *FE* Convulsión Social y Bloqueos por Clima	*U+W*
14	Con *FE* Convulsión Social, Anomalía Temperatura y Bloqueos por Clima	*U+T+W*
15	Resta entre resultado de *U* menos *SFE*	*U-SFE*
16	Resta entre resultado de *T* menos *SFE*	*T-SFE*
17	Resta entre resultado de *W* menos *SFE*	*W-SFE*
18	Resta entre resultado de *U+T* menos *SFE*	*U+T-SFE*
19	Resta entre resultado de *T+W* menos *SFE*	*T+W-SFE*
20	Resta entre resultado de *U+W* menos *SFE*	*U+W-SFE*
21	Resta entre resultado de *U+T+W* menos *SFE*	*U+T+W-SFE*

Nota: Los acrónimos listados resumen los escenarios evaluados en el estudio, diferenciando entre periodos sin factores externos (*SFE*), con la influencia de un solo factor (por ejemplo, *F*, *P*, *U*, *T*, *W*), combinaciones de factores (*F+P*, *U+T*, *T+W*, *U+W*, *U+T+W*) y las diferencias porcentuales respecto al escenario base sin factores externos. Esta nomenclatura permite comparar de forma estandarizada el impacto de cada escenario en las métricas de desempeño de los modelos de *ML*.

Al visualizar la Figura 8 se evidencia que existe una notoria contraposición entre las series temporales influenciadas por fallecidos y casos positivos por *COVID-19* , en ese sentido, y considerando que la enfermedad por *COVID-19* , presentaba tasas altas de contagio, y permanecía en el paciente 14 días, mientras se generaban anticuerpos, en este plazo de no contrarrestar la enfermedad, podía acaecer el fallecimiento del paciente; bajo esta premisa también se evaluó el fenómeno moviendo la curva de fallecidos un ciclo a la derecha ver figura 8.

Los resultados de este análisis de causalidad se integraron posteriormente en el conjunto de datos de crédito y se asociaron con la actividad económica y el período de análisis correspondientes. Adicionalmente, se calculó una variable de clasificación binaria global y se agregó al conjunto de datos para identificar la calidad de pago de los titulares de crédito (buenos o malos pagadores) con base en los datos mensuales de morosidad. Este conjunto de datos enriquecido proporcionó la base para aplicar el algoritmo (1) a los conjuntos de datos establecidos (ver Figura 5) para calcular las métricas *ACC* y *AUC* para los modelos de aprendizaje automático evaluados. Estos modelos se evaluaron en los escenarios enumerados en la Tabla 4. Con base en los valores obtenidos, se calcularon las diferencias entre los escenarios que incluyen escenarios *FE* y *SFE*, como se muestra en la Tabla 5.

5.1.1. *FE* COVID

El análisis de los resultados del *ACC* y *AUC* con *k-folds* de 10 de validación cruzada revela varias tendencias importantes en el desempeño de los modelos de *ML* evaluados en escenarios con distintos *FE*, y con experimentos incluyendo y excluyendo créditos con apoyo gubernamental, denominados *Inc CG* y *Exc CG* respectivamente.

Evaluación incluyendo en la data créditos que han sido beneficiados por programas gubernamentales

Esta sección presenta los resultados obtenidos en la evaluación de los modelos de *ML* en escenarios que incluyen *EF* relacionados con la pandemia de *COVID-19* , con un enfoque específico en créditos con beneficios gubernamentales. La Tabla 21 resume las métricas promedio de *AUC* y *ACC* obtenidas a lo largo de 10 folds para los modelos evaluados.

Entre los hallazgos más relevantes se identificó que modelos como *EBM* y *RF* demostraron un desempeño sólido, alcanzando los valores más altos de *AUC* (superiores al 81 %), lo que indica una fuerte capacidad discriminativa en la clasificación de créditos bajo la influencia de *EF*. Asimismo, el modelo *CNN* mostró mejoras significativas en *ACC*, con incrementos de hasta **4.26 puntos porcentuales** al incorporar *EF*, resaltando su sensibilidad a los conjuntos de datos enriquecidos.

En términos de desempeño predictivo, los modelos se comportaron de la siguiente manera:

1. **Impacto en *ACC*:**

   - En los tres grupos de datos evaluados (*D100*, *D99* y *D80*), el modelo *CNN* mostró mejoras constantes al incorporar factores de fallecidos y positivos por *COVID-19* , con incrementos en precisión cercanos al 70 %.
   - Los escenarios de diferencia (*F-SFE*, *P-SFE*, *F+P-SFE*) reflejan incrementos adicionales de aproximadamente 2.5 %, sugiriendo que *CNN* es particularmente sensible a la inclusión de estos factores en la predicción del riesgo crediticio.
   - En contraste, el modelo *RF* mostró poca variación en *ACC* al incluir *FE*, con ligeras disminuciones en algunos escenarios, lo que sugiere una menor adaptabilidad a cambios en los datos derivados de *FE* de *COVID-19* .
   - El modelo *Perceptrón* mostró una respuesta destacable, especialmente en el escenario *P*, con un incremento en *ACC* del 13.41 % en comparación con el escenario *SFE*.
   - Modelos como *DL* y *KNN* mostraron un impacto menos pronunciado, con leves disminuciones en precisión en la mayoría de los escenarios, indicando una menor sensibilidad a los *FE*.
   - *LDA* presentó incrementos moderados en precisión, alcanzando hasta 0.96 % en el escenario *F+P-SFE*, lo que indica que, aunque en menor medida, también responde a la incorporación de *FE*.
   - *MLP* registró una mejora significativa en el grupo *D80* en el escenario de diferencia (*P-SFE*), con un incremento de 5.07 %, destacándose en su capacidad para beneficiarse de datos adicionales sobre casos positivos de *COVID-19* .

2. **Impacto en *AUC*:**

   - En los tres grupos de datos, el modelo *CNN* mostró mejoras significativas en *AUC* al incorporar factores de fallecidos y positivos por *COVID-19* . En el grupo *D100*, logró un *AUC* del 74.09 % en los escenarios *F* y *F+P*, con incrementos de 0.82 % y 1.28 % en los escenarios de diferencia (*F-SFE*, *P-SFE*, *F+P-SFE*), demostrando su capacidad de adaptación a variaciones inducidas por la pandemia.
   - El modelo *RF* presentó reducciones ligeras en *AUC* en todos los escenarios que incluyen *FE*, con decrementos mínimos como -0.23 % en el escenario

P-SFE, lo que sugiere que su sensibilidad a la inclusión de datos de *COVID-19* es baja.

- *MLP* respondió positivamente en la mayoría de los escenarios, logrando incrementos en *AUC* de hasta 1.87 % en *F+P-SFE* en el grupo *D100*, con variaciones menores en los otros grupos de datos.
- *Perceptrón* mostró una variabilidad notable, con incrementos de hasta 2.58 % en el escenario *P-SFE* en el grupo *D99*, mientras que, en otros escenarios y grupos, como *D80*, experimentó ligeras disminuciones.
- Modelos como *LDA* y *Ridge* presentaron incrementos moderados en *AUC*, particularmente en el grupo *D100* y en escenarios como *F+P-SFE*, con aumentos de 1.31 % y 1.12 %, respectivamente, lo que indica una respuesta favorable, aunque menos pronunciada en comparación con *CNN* o *MLP*.

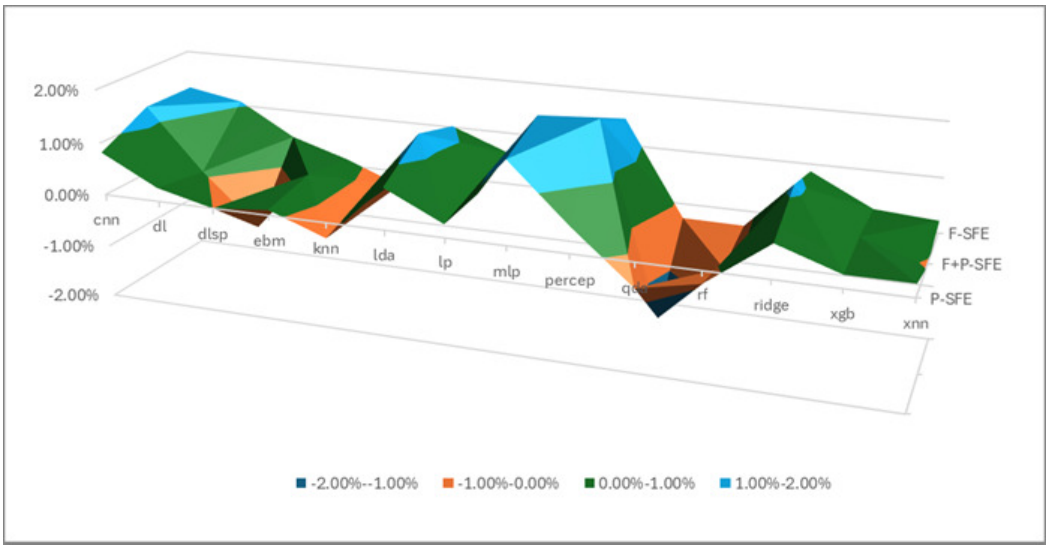
En términos generales, los resultados obtenidos destacan la importancia de considerar *FE* en la predicción del riesgo crediticio, ya que ciertos modelos, como *CNN*, *Perceptrón* y *MLP*, mostraron mejoras significativas en escenarios que incluyen factores de la pandemia ver Figura 12. No obstante, estos efectos no fueron homogéneos en todos los modelos, lo que sugiere que algunos algoritmos, especialmente aquellos con estructuras flexibles y capacidad de aprendizaje profundo, tienen mayor capacidad de capturar patrones derivados de *FE*.

Además, la inclusión de beneficios gubernamentales como parte del *FE COVID-19* mejoró notablemente el desempeño predictivo de la mayoría de los modelos, destacando la importancia de estas intervenciones en la estabilidad financiera. La respuesta diferencial entre los modelos refuerza la relevancia de seleccionar el algoritmo adecuado en función de la sensibilidad requerida frente a *FE*. Como conclusión, la integración de *FE* en la evaluación de riesgo crediticio se presenta como una estrategia clave para mejorar la precisión y adaptabilidad de los modelos en entornos de alta volatilidad.

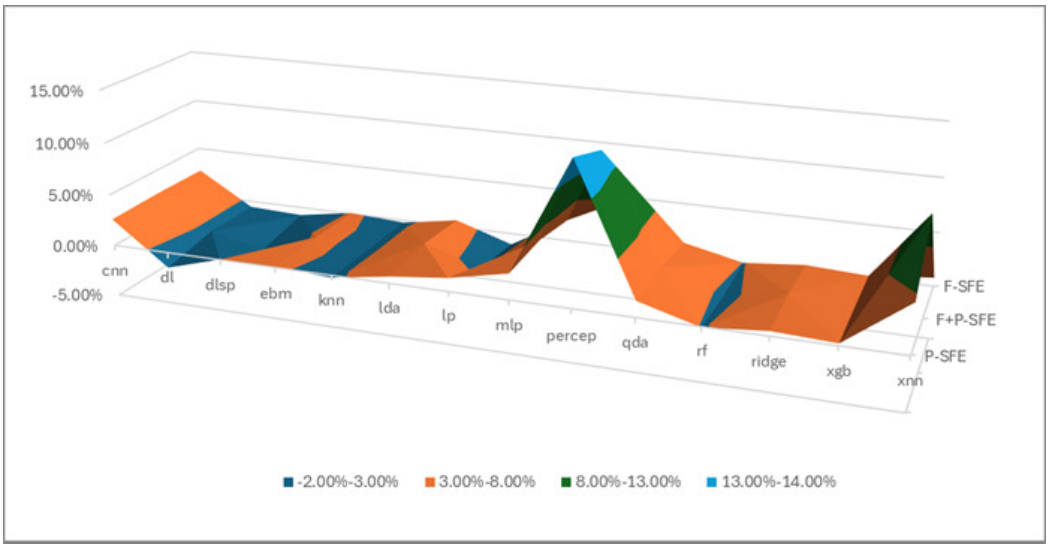
Tabla 21

Promedio de AUC & ACC en 10 folds de modelos de ML con el FE COVID-19 , incluyendo los créditos con apoyo gubernamental, según el escenario evaluado y los grupos de relevancia.

	AUC (%)							*ACC* (%)						
	SCENARIOS				RESULTS			SCENARIOS				RESULTS		
	SFE	P	F+P	F	P-SFE	F+P-SFE	F-SFE	SFE	P	F+P	F	P-SFE	F+P-SFE	F-SFE
1.D100														
cnn	72.82	73.63	74.09	74.09	0.82	1.28	1.27	67.49	70.07	70.18	70.14	2.58	2.69	2.65
ebm	81.13	81.23	81.42	81.26	0.10	0.29	0.13	74.84	74.94	75.05	74.93	0.09	0.20	0.08
ffn	72.30	72.56	72.39	73.38	0.26	0.10	1.08	66.74	65.24	66.15	66.14	- 1.50	- 0.59	- 0.60
ffnlp	72.11	72.35	72.41	72.72	0.25	0.30	0.61	66.28	67.55	65.64	65.19	1.27	- 0.64	- 1.09
ffnsp	72.31	72.31	71.42	72.77	0.00	- 0.89	0.47	66.40	66.45	64.67	65.59	0.05	- 1.73	- 0.81
knn	67.21	66.92	66.93	66.93	- 0.29	- 0.28	- 0.28	63.23	63.00	62.97	62.98	- 0.22	- 0.26	- 0.24
lda	67.23	68.01	68.54	68.25	0.78	1.31	1.02	63.09	63.74	64.09	63.77	0.65	1.00	0.67
mlp	69.30	70.83	71.17	70.47	1.53	1.87	1.17	60.82	63.26	63.58	62.87	2.44	2.76	2.06
percep	57.96	58.44	59.88	59.46	0.48	1.92	1.50	44.58	57.99	56.37	49.31	13.41	11.78	4.73
qda	67.44	66.98	65.81	67.17	- 0.46	- 1.63	- 0.27	52.50	54.01	54.13	53.72	1.52	1.64	1.22
rf	81.74	81.51	81.21	81.46	- 0.23	- 0.53	- 0.29	75.51	75.44	75.22	75.48	- 0.07	- 0.29	- 0.03
ridge	66.94	67.55	68.05	67.81	0.62	1.12	0.87	62.93	63.39	63.70	63.43	0.46	0.77	0.50
xgb	80.76	81.00	81.10	81.07	0.23	0.33	0.30	74.77	74.98	75.03	74.98	0.21	0.26	0.22
xnn	50.14	50.38	50.10	50.37	0.24	- 0.05	0.22	53.42	57.99	62.76	54.22	4.57	9.34	0.80
2.D99														
cnn	72.75	73.60	74.27	73.57	0.86	1.52	0.83	67.04	70.03	70.42	69.97	2.98	3.37	2.93
ebm	81.04	81.24	81.37	81.24	0.19	0.33	0.20	74.77	74.93	75.04	74.94	0.16	0.28	0.18
ffn	72.69	72.38	71.50	71.55	- 0.31	- 1.20	- 1.14	66.07	65.94	65.86	65.75	- 0.14	- 0.21	- 0.32
ffnlp	71.71	72.81	72.16	72.45	1.09	0.44	0.74	65.27	66.83	66.28	65.94	1.56	1.02	0.67
ffnsp	72.58	73.11	72.58	72.99	0.53	- 0.00	0.41	63.93	65.34	66.81	66.17	1.41	2.87	2.24
knn	67.10	66.88	66.88	66.88	- 0.22	- 0.21	- 0.21	63.24	63.02	63.02	63.04	- 0.23	- 0.22	- 0.21
lda	67.17	67.92	68.46	68.19	0.75	1.29	1.02	63.07	63.68	64.04	63.78	0.60	0.96	0.70
mlp	69.79	68.93	69.62	68.70	- 0.86	- 0.16	- 1.08	63.83	64.12	59.76	64.96	0.29	- 4.07	1.13
percep	55.09	57.67	54.19	56.99	2.58	- 0.90	1.90	55.48	58.56	53.93	50.60	3.07	- 1.55	- 4.88
qda	67.32	66.90	65.74	67.08	- 0.42	- 1.58	- 0.23	52.41	53.98	54.02	53.63	1.57	1.62	1.22
rf	81.68	81.33	81.02	81.30	- 0.35	- 0.66	- 0.38	75.63	75.32	75.09	75.38	- 0.31	- 0.54	- 0.24
ridge	66.85	67.47	67.97	67.74	0.62	1.12	0.89	62.94	63.39	63.68	63.49	0.45	0.73	0.55
xgb	80.72	80.96	80.97	80.95	0.24	0.25	0.24	74.81	74.97	74.89	74.86	0.16	0.08	0.04
xnn	50.47	50.25	50.48	50.59	- 0.23	0.00	0.11	57.50	53.55	59.22	54.37	- 3.95	1.71	- 3.14
3.D80														
cnn	72.68	73.18	73.79	73.50	0.50	1.11	0.82	66.38	70.28	70.65	70.44	3.90	4.26	4.06
ebm	80.73	81.03	81.16	81.05	0.29	0.43	0.31	74.92	75.16	75.29	75.19	0.24	0.37	0.27
ffn	71.76	72.16	72.38	72.83	0.41	0.62	1.07	63.24	66.92	66.19	67.18	3.68	2.95	3.94
ffnlp	71.14	72.00	71.66	71.97	0.86	0.51	0.83	65.73	66.90	66.45	65.44	1.16	0.72	- 0.29
ffnsp	70.87	72.03	72.29	72.13	1.15	1.42	1.25	64.07	65.62	66.69	65.65	1.54	2.61	1.58
knn	66.57	66.57	66.57	66.57	- 0.01	0.00	0.00	62.98	62.95	62.95	62.96	- 0.02	- 0.03	- 0.01
lda	67.33	67.90	68.29	68.07	0.57	0.96	0.74	63.45	63.88	64.20	64.02	0.43	0.75	0.57
mlp	67.96	68.85	68.61	69.92	0.89	0.66	1.96	58.09	63.16	62.43	60.73	5.07	4.34	2.64
percep	57.44	55.39	59.27	56.66	- 2.05	1.83	- 0.78	52.88	57.22	50.16	56.66	4.35	- 2.71	3.78
qda	67.69	66.51	64.11	66.30	- 1.18	- 3.58	- 1.39	52.76	54.28	52.68	53.73	1.52	- 0.08	0.97
rf	81.18	81.03	80.68	80.96	- 0.15	- 0.51	- 0.22	75.47	75.57	75.26	75.52	0.11	- 0.20	0.05
ridge	66.67	67.23	67.59	67.40	0.56	0.92	0.73	63.06	63.35	63.73	63.55	0.29	0.67	0.49
xgb	80.20	80.75	80.76	80.69	0.54	0.56	0.49	74.74	75.13	75.15	75.14	0.39	0.41	0.39
xnn	50.35	50.07	49.89	50.03	- 0.28	- 0.47	- 0.33	63.26	63.80	60.07	53.46	0.54	- 3.19	- 9.80



(a) *AUC*



(b) *ACC*

Figura 12

Comparación del promedio en 10 folds de métricas AUC y ACC en modelos de ML con el FE COVID-19, incluyendo créditos con apoyo gubernamental, al 100 % de las actividades económicas.

Nota: Las figuras muestran las variaciones (Δ) en *AUC* y *ACC* entre diferentes escenarios (*SFE*, *F-SFE*, *P-SFE*, *F+P-SFE*) para cada modelo de *ML* evaluado. Los colores representan rangos de variación porcentual, facilitando la identificación de mejoras o deterioros en el rendimiento.

Evaluación excluyendo en la data créditos que han sido beneficiados por programas gubernamentales

El análisis de las métricas *AUC* y *ACC* en modelos de *ML*, excluyendo los créditos beneficiados por programas de gobierno, revela el impacto diferencial de los *FE* relacionados con *COVID-19* y otros eventos críticos en el rendimiento de los modelos de clasificación crediticia. La Tabla 22 presenta los resultados promedio obtenidos en 10

folds para la clasificación binaria de riesgo crediticio (buenos y malos pagadores) en escenarios enriquecidos con *EF* de *COVID-19* , pero sin incluir créditos con respaldo gubernamental.
Entre los hallazgos más relevantes, se identificaron los siguientes puntos clave:

1. **Modelos con desempeño robusto:**

   - Modelos como *EBM* y *XGB* mantienen una alta capacidad predictiva, logrando valores de *AUC* superiores al 80 %, con fluctuaciones mínimas en su rendimiento ante la ausencia de datos sobre beneficios gubernamentales. Esto indica que estos modelos son menos dependientes de este tipo de información y pueden capturar patrones relevantes en los datos sin necesidad de intervenciones externas.

2. **Mejoras significativas en *ACC*:**

   - Tanto *CNN* como *MLP* presentan mejoras notables en precisión (*ACC*), destacándose *CNN* con incrementos de hasta **5.58 puntos porcentuales** en escenarios específicos, lo que subraya la influencia del enriquecimiento con *EF* en la capacidad de clasificación.
   - En los grupos de datos evaluados (*D100*, *D99* y *D80*), *CNN* mantiene un desempeño sólido en los escenarios *P*, *F* y *F+P*, alcanzando precisiones superiores al 72 %. Esto sugiere que *CNN* es particularmente sensible a la información adicional sobre fallecidos y casos positivos de *COVID-19* , lo que mejora su capacidad predictiva en ausencia de intervenciones gubernamentales.

3. **Impacto negativo en algunos modelos:**

   - La exclusión de información sobre beneficios gubernamentales afecta negativamente a ciertos modelos (*RF*, *QDA*, *XNN*), los cuales registran descensos tanto en *AUC* como en *ACC*.
   - Por ejemplo, *RF* muestra poca variación en precisión, con leves disminuciones en presencia de *EF*, lo que indica que su estructura es menos adaptable a la ausencia de estos datos.
   - De manera similar, *MLP* exhibe fluctuaciones más pronunciadas en los grupos *D99* y *D80*, con reducciones en *ACC* en algunos escenarios, lo que sugiere que su sensibilidad a los *FE* varía en función del grupo de datos utilizado.

4. **Efectos de la exclusión de datos gubernamentales:**

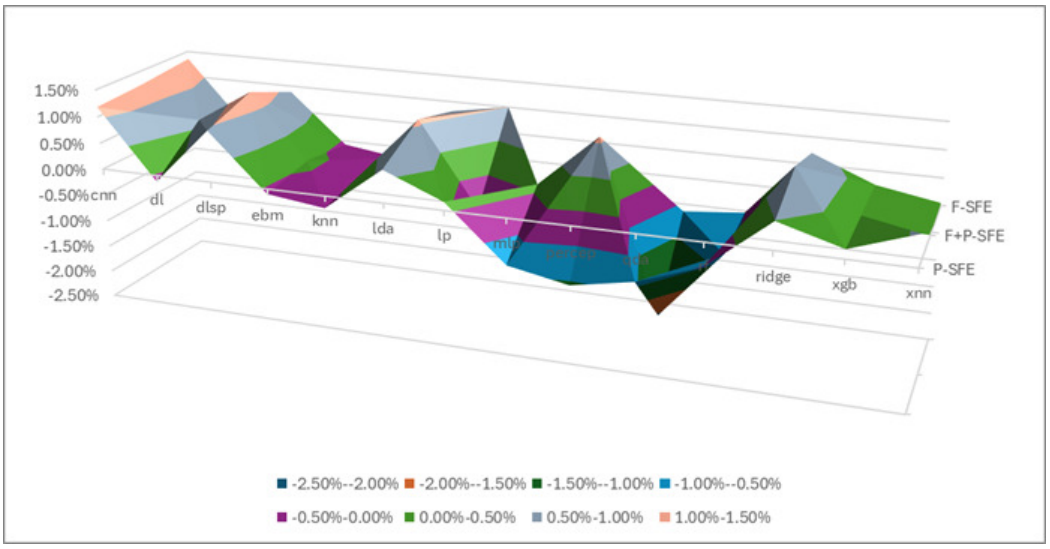
- Los resultados reflejan que omitir información sobre respaldo gubernamental introduce una variabilidad significativa y potenciales sesgos en el rendimiento predictivo de los modelos.
- Mientras que algunos modelos, como *EBM* y *XGB*, logran mantener su desempeño estable sin estos datos, otros modelos sufren de una menor granularidad en la entrada, lo que impacta negativamente su precisión y capacidad de generalización.
- En la métrica *AUC*, *CNN* continúa destacándose con incrementos en los escenarios de diferencia (*P-SFE*, *F-SFE* y *F+P-SFE*), especialmente en los grupos *D99* y *D80*, con mejoras de hasta 1.35 %.
- Modelos como *XGB* y *Ridge* presentan mejoras moderadas, con incrementos de hasta 0.63 % en *D80*, mientras que *RF* y *QDA* muestran ligeras disminuciones en *AUC* en presencia de *EF*.

En términos generales, estos resultados destacan que modelos como *CNN* y *MLP* son más adecuados para capturar variaciones asociadas a los *EF* en ausencia de respaldo gubernamental, mejorando su capacidad de clasificación al integrar datos sobre fallecidos y casos positivos. Por el contrario, modelos como *RF* y *QDA* parecen depender más de los datos gubernamentales, lo que sugiere que su desempeño se ve afectado negativamente cuando estos se omiten ver figura 13.
Finalmente, se observa que la exclusión de créditos beneficiados por programas gubernamentales mejora el desempeño de algunos modelos al reducir el sesgo introducido por estas intervenciones. Sin embargo, también se evidencia que algunos modelos requieren esta información para optimizar su rendimiento. Este análisis enfatiza la importancia de comprender cómo la inclusión o exclusión de información sobre respaldo gubernamental afecta la capacidad predictiva de los modelos de *ML*, resaltando la necesidad de enfoques flexibles y adaptativos en la evaluación del riesgo crediticio en entornos de alta volatilidad.

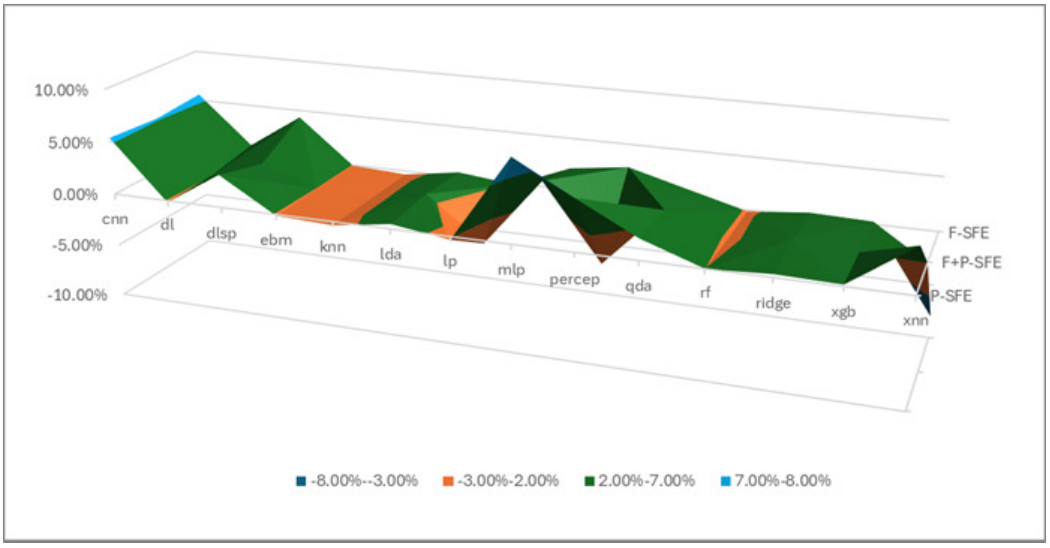
Tabla 22

Promedio de AUC & ACC en 10 folds de modelos de ML con el FE COVID-19 , excluyendo los créditos con apoyo gubernamental, según el escenario evaluado y los grupos de relevancia.

	AUC (%)							*ACC* (%)						
	SCENARIOS				RESULTS			SCENARIOS				RESULTS		
	SFE	P	F+P	F	P-SFE	F+P-SFE	F-SFE	SFE	P	F+P	F	P-SFE	F+P-SFE	F-SFE
1.D100														
cnn	72.92	74.10	74.17	74.28	1.18	1.25	1.35	67.53	72.93	72.67	73.11	5.41	5.15	5.58
ebm	81.38	81.29	81.45	81.31	- 0.09	0.07	- 0.07	76.54	76.41	76.51	76.48	- 0.13	- 0.03	- 0.06
ffn	72.71	72.60	73.34	72.89	- 0.11	0.63	0.18	67.46	67.29	67.49	68.33	- 0.17	0.03	0.88
ffnlp	72.16	72.32	71.68	73.16	0.15	- 0.49	1.00	66.86	66.56	63.51	66.75	- 0.30	- 3.35	- 0.11
ffnsp	72.48	73.44	73.75	73.36	0.96	1.27	0.88	65.18	68.26	67.02	69.48	3.09	1.84	4.30
knn	66.21	65.98	65.98	65.99	- 0.23	- 0.23	- 0.22	63.19	62.84	62.82	62.84	- 0.35	- 0.36	- 0.35
lda	66.08	66.68	67.16	66.89	0.60	1.08	0.81	63.18	63.64	63.93	63.65	0.46	0.75	0.47
mlp	70.92	70.07	71.03	68.89	- 0.85	0.11	- 2.03	60.20	67.96	63.45	62.29	7.76	3.25	2.09
percep	55.94	54.88	57.04	55.72	- 1.06	1.10	- 0.22	54.77	59.56	50.92	57.58	4.78	- 3.85	2.81
qda	66.82	65.97	64.75	66.18	- 0.84	- 2.07	- 0.64	50.72	52.66	52.59	52.06	1.94	1.86	1.33
rf	79.53	79.09	78.75	78.99	- 0.44	- 0.78	- 0.53	75.34	75.34	75.17	75.21	0.01	- 0.16	- 0.13
ridge	65.95	66.44	66.89	66.66	0.50	0.94	0.71	63.14	63.44	63.77	63.54	0.30	0.63	0.40
xgb	80.43	80.60	80.64	80.68	0.17	0.21	0.25	76.00	76.16	76.13	76.24	0.16	0.13	0.24
xnn	50.14	50.71	50.15	50.19	0.56	0.00	0.05	62.48	66.64	62.37	54.50	4.17	- 0.10	- 7.98
2.D99														
cnn	73.04	73.95	74.16	74.28	0.91	1.12	1.24	68.52	72.56	72.93	72.91	4.03	4.40	4.38
ebm	81.23	81.31	81.44	81.33	0.08	0.22	0.10	76.31	76.39	76.49	76.48	0.07	0.18	0.17
ffn	72.50	72.95	72.44	73.37	0.46	- 0.05	0.87	64.38	65.93	64.91	68.10	1.54	0.53	3.72
ffnlp	72.72	72.92	73.36	72.79	0.20	0.64	0.07	68.01	65.29	69.55	70.48	- 2.72	1.54	2.47
ffnsp	73.19	73.12	72.32	73.13	- 0.07	- 0.87	- 0.06	67.88	61.75	68.83	65.88	- 6.13	0.95	- 2.00
knn	65.94	66.00	66.00	66.00	0.06	0.06	0.06	62.92	62.58	62.58	62.59	- 0.35	- 0.34	- 0.33
lda	66.00	66.67	67.15	66.90	0.67	1.15	0.90	63.15	63.63	64.00	63.67	0.48	0.85	0.52
mlp	68.79	70.06	70.21	70.58	1.27	1.42	1.79	70.57	64.20	64.24	68.64	- 6.37	- 6.33	- 1.93
percep	55.28	56.66	55.64	56.59	1.37	0.36	1.31	54.45	56.49	49.80	55.10	2.04	- 4.65	0.65
qda	66.72	65.95	64.69	66.12	- 0.77	- 2.03	- 0.60	50.57	52.65	52.77	52.16	2.08	2.20	1.60
rf	79.41	78.96	78.60	78.91	- 0.46	- 0.81	- 0.51	75.31	75.19	74.90	75.24	- 0.13	- 0.41	- 0.07
ridge	65.85	66.42	66.87	66.65	0.57	1.02	0.80	63.13	63.45	63.82	63.58	0.32	0.69	0.45
xgb	80.35	80.66	80.65	80.61	0.32	0.30	0.26	76.13	76.28	76.21	76.25	0.15	0.08	0.12
xnn	50.08	50.74	50.00	49.90	0.67	- 0.08	- 0.18	62.22	65.34	70.29	66.02	3.12	8.07	3.80
3.D80														
cnn	73.29	73.32	73.58	73.05	0.03	0.29	- 0.24	68.07	72.86	73.28	72.64	4.78	5.20	4.56
ebm	80.75	81.05	81.13	81.02	0.30	0.37	0.27	76.37	76.72	76.74	76.69	0.35	0.37	0.32
ffn	71.19	73.05	72.68	73.08	1.86	1.49	1.88	63.49	66.53	66.51	68.70	3.04	3.02	5.22
ffnlp	72.25	72.57	71.92	73.00	0.32	- 0.33	0.74	64.39	64.96	64.64	62.94	0.57	0.25	- 1.45
ffnsp	71.72	73.07	72.27	72.45	1.35	0.55	0.72	68.29	64.39	64.76	65.33	- 3.90	- 3.53	- 2.96
knn	65.37	65.36	65.36	65.37	- 0.01	- 0.01	- 0.00	62.54	62.52	62.51	62.53	- 0.02	- 0.03	- 0.00
lda	65.85	66.35	66.70	66.52	0.50	0.85	0.67	63.27	63.64	63.93	63.75	0.37	0.65	0.48
mlp	69.10	71.07	69.19	70.08	1.96	0.09	0.97	62.97	59.38	52.32	60.10	- 3.59	- 10.65	- 2.87
percep	55.71	56.80	57.86	55.39	1.09	2.16	- 0.31	51.81	55.32	52.05	50.98	3.51	0.24	- 0.83
qda	66.65	65.29	63.02	65.30	- 1.36	- 3.63	- 1.35	50.91	53.12	50.75	52.15	2.21	- 0.16	1.25
rf	78.71	78.53	78.12	78.56	- 0.18	- 0.59	- 0.15	75.29	75.54	75.39	75.65	0.25	0.10	0.36
ridge	65.48	65.97	66.26	66.13	0.48	0.78	0.65	63.14	63.37	63.65	63.58	0.23	0.51	0.43
xgb	79.57	80.16	80.16	80.20	0.59	0.59	0.63	75.92	76.45	76.44	76.45	0.54	0.52	0.54
xnn	50.76	50.57	49.89	49.89	- 0.19	- 0.87	- 0.86	56.71	55.79	71.19	66.89	- 0.92	14.48	10.18



(a) *AUC*



(b) *ACC*

Figura 13

Comparación del promedio en 10 folds de métricas AUC y ACC en modelos de ML con el FE COVID-19, excluyendo créditos con apoyo gubernamental, al 100 % de las actividades económicas.

Nota: Las figuras muestran las variaciones (Δ) en *AUC* y *ACC* entre diferentes escenarios (*SFE*, *F-SFE*, *P-SFE*, *F+P-SFE*) para cada modelo de *ML* evaluado. Los colores representan rangos de variación porcentual, facilitando la identificación de mejoras o deterioros en el rendimiento.

Evaluación del nivel de significancia de los modelos de *ML* para *AUC* y *ACC* aplicando el *FE COVID-19*, tanto excluyendo como considerando los créditos con apoyo gubernamental

La Tabla 23 detalla los modelos de *ML* que exhiben mejoras estadísticamente significativas en los indicadores de *ACC* y *AUC* luego de la aplicación de la prueba *t-student* a muestras relacionadas. Estas muestras se comparan entre escenarios sin factores ex-

ternos (*SFE*) y aquellos donde se incluyen *FE* relacionados con *COVID-19* , específicamente fallecidos (*F*), positivos (*P*) y la combinación de ambos (*F+P*). La tabla está segmentada en dos grandes categorías: en primer lugar, los experimentos que incluyeron créditos respaldados por programas de garantía gubernamental *Inc CG*, y segundo lugar, los experimentos donde estos créditos fueron excluidos *Exc CG*.

La información se presenta en tres grupos de relevancia: *D100*, *D99* y *D80*, representando subconjuntos de datos con diferentes niveles de cobertura. Cada combinación de grupo y escenario especifica cuántos modelos mostraron una mejora significativa en el indicador evaluado (*ACC* o *AUC*), marcado con el valor "Y"para indicar significancia. En el agregado, se observa que los modelos *CNN*, *LDA*, *QDA*, *Ridge*, y *XGB* son los que consistentemente exhiben mejoras bajo escenarios de *FE*, tanto en la inclusión como en la exclusión de créditos garantizados por el gobierno.

El análisis desglosado por grupo muestra que *CNN* y *LDA* son particularmente consistentes, manteniendo una significancia estadística en cada uno de los tres grupos y en ambos indicadores (*ACC* y *AUC*), lo cual resalta su capacidad para adaptarse a la incorporación de datos de fallecidos y positivos por *COVID-19* . Por otro lado, algunos modelos como *MLP* y *Perceptrón* presentan mejoras significativas en ciertos contextos específicos, pero no de manera constante en todos los escenarios o indicadores, sugiriendo que estos modelos pueden ser sensibles a cambios en los datos, pero su respuesta es más dependiente del grupo de relevancia o del tipo de *FE* aplicado.

Adicionalmente, los modelos basados en árboles de decisión, como *EBM*, *XGB* y *RF*, demostraron un desempeño robusto y estable en la mayoría de los escenarios evaluados, con significancia estadística observada en todos los grupos de relevancia (*F*, *F+P*, *P*). Esto refuerza la capacidad de estos modelos para manejar información estructurada y mejorar su precisión ante la incorporación de *FE*.

En términos de sensibilidad a los datos enriquecidos, los modelos basados en redes neuronales, especialmente *CNN*, mostraron un alto grado de adaptabilidad a la inclusión de *FE*, alcanzando significancia estadística en todos los escenarios bajo ambas configuraciones: con y sin créditos respaldados por el gobierno. Este hallazgo resalta la flexibilidad de estos modelos para captar relaciones complejas en datos enriquecidos.

Tabla 23

Modelos con mejoras estadísticamente significativas en las métricas AUC y ACC con el FE COVID-19 , considerando escenarios con créditos con beneficios gubernamentales y sin créditos con beneficios gubernamentales y grupos de relevancia - escenarios.

	Inc. CG							Exc. CG						
	AUC			ACC			TOTAL	AUC			ACC			TOTAL
	F	F+P	P	F	F+P	P		F	F+P	P	F	F+P	P	
1.D100														
cnn	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	6
ebm	-	Y	-	-	-	-	1	-	-	-	-	-	-	0
ffn	Y	-	-	-	-	-	1	-	-	-	-	-	-	0
ffnlp	-	-	-	-	-	-	0	-	-	-	-	-	-	0
lda	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	6
mlp	-	-	-	-	-	-	0	-	-	-	-	-	Y	1
percep	-	-	-	-	Y	Y	2	-	-	-	-	-	-	0
qda	-	-	-	Y	Y	Y	3	-	-	-	Y	Y	Y	3
ridge	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	-	5
xgb	Y	-	-	Y	Y	Y	4	-	-	-	Y	-	-	1
xnn	-	-	-	-	-	-	0	-	-	-	-	-	-	0
2.D99														
cnn	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	6
ebm	-	-	-	-	-	-	0	-	-	-	-	-	-	0
ffn	-	-	-	-	-	-	0	-	-	-	-	-	-	0
ffnlp	Y	-	Y	-	-	-	2	-	-	-	Y	-	-	1
ffnsp	-	-	-	-	-	-	0	-	-	-	-	-	-	0
lda	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	-	5
percep	-	-	-	-	-	-	0	-	-	-	-	-	-	0
qda	-	-	-	Y	Y	Y	3	-	-	-	Y	Y	Y	3
ridge	Y	Y	Y	Y	Y	Y	6	Y	Y	-	Y	Y	-	4
xgb	-	-	-	-	-	-	0	Y	Y	Y	-	-	-	3
xnn	-	-	-	-	-	-	0	-	-	-	-	-	-	0
3.D80														
cnn	Y	Y	Y	Y	Y	Y	6	-	-	-	Y	Y	Y	3
ebm	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	6
ffn	Y	-	-	-	-	-	1	Y	-	Y	Y	-	-	3
ffnlp	-	-	-	-	-	-	0	-	-	-	-	-	-	0
ffnsp	-	-	-	-	-	-	0	-	-	Y	-	-	-	1
lda	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	6
mlp	-	-	-	-	-	-	0	-	-	-	-	-	-	0
percep	-	Y	-	-	-	-	1	-	-	-	-	-	-	0
qda	-	-	-	Y	-	Y	2	-	-	-	Y	-	Y	2
rf	-	-	-	-	-	-	0	-	-	-	Y	-	Y	2
ridge	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	6
xgb	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	6
xnn	-	-	-	-	-	-	0	-	-	-	-	Y	-	1

5.1.2. *FE* COVID con desplazamiento de serie temporal correspondiente a fallecidos, un periodo a la derecha

El análisis de las métricas *ACC* y *AUC* evidencia el impacto del desplazamiento temporal de la serie de fallecidos por *COVID-19* en el rendimiento de los modelos de *ML*, considerando la inclusión de créditos *Inc CG*. Este desplazamiento, denominado escenario *COVID MOV*, se justifica en la relación temporal entre el diagnóstico, el tratamiento y el desenlace crítico (Organización Mundial de la Salud, 2024), proporcionando un marco temporal más realista para evaluar la influencia de este *FE* en la predicción del riesgo crediticio.

Evaluación incluyendo en la data créditos que han sido beneficiados por programas gubernamentales

En términos de *ACC*, el modelo *CNN* se destaca por su desempeño consistente en los tres grupos de datos (*D100*, *D99* y *D80*), mostrando mejoras significativas en los escenarios con *FE* (*F*, *P*, *F+P*). En particular, en el grupo *D100*, *CNN* alcanza un incremento de **3.75 %** en *F+P-SFE*, evidenciando su capacidad para adaptarse a la inclusión de datos temporales ajustados (ver Figura 14). De manera similar, el modelo *DL* experimenta mejoras notables en este grupo, especialmente en el escenario *P-SFE*, con un incremento de **3.55 %** (ver Tabla 24).

Modelos como *EBM*, *RF* y *XGB* presentan un desempeño predictivo sólido, alcanzando valores de *AUC* superiores al **80 %** en todos los escenarios, con un impacto mínimo derivado de la incorporación de datos de *FE*. Esto sugiere que estos modelos son más resilientes ante modificaciones temporales en la estructura de los datos, manteniendo su estabilidad predictiva en contextos de riesgo crediticio.

Otros modelos, como *QDA* y *Ridge*, presentan incrementos más moderados en *ACC*, oscilando entre **0.5 % y 1.5 %**, lo que refleja una sensibilidad menor a los *FE* desplazados (ver Figura 14). En contraste, modelos como *KNN* y *RF* muestran una respuesta limitada, con variaciones mínimas o incluso disminuciones en los escenarios con *FE*, lo que podría indicar su menor adaptabilidad a los cambios inducidos por el ajuste temporal (ver Tabla 24).

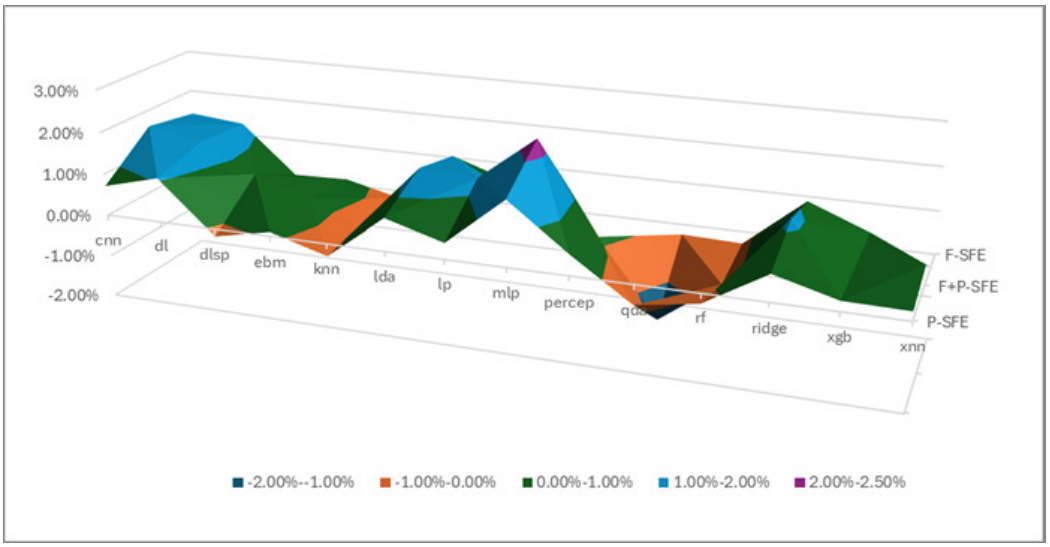
Tabla 24

Promedio de AUC & ACC en 10 folds de modelos de ML con el FE COVID-19 [COVID MOV], incluyendo los créditos con apoyo gubernamental, según el escenario evaluado y los grupos de relevancia.

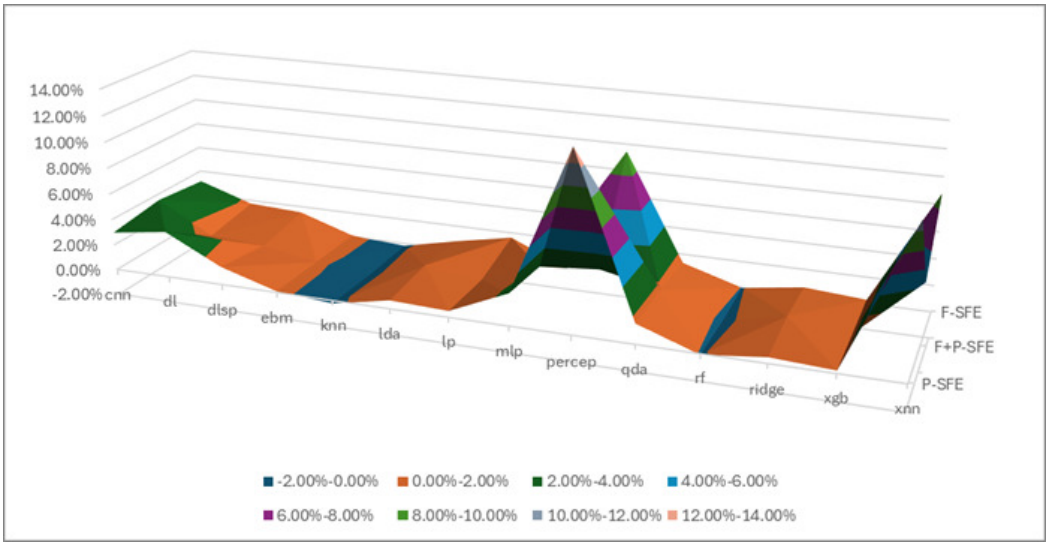
	AUC (%)							*ACC* (%)						
	SCENARIOS				RESULTS			SCENARIOS				RESULTS		
	SFE	P	F+P	F	P-SFE	F+P-SFE	F-SFE	SFE	P	F+P	F	P-SFE	F+P-SFE	F-SFE
Inc. CG														
1.D100														
cnn	72.63	73.34	74.25	74.04	0.71	1.61	1.41	66.85	69.80	70.36	69.98	2.95	3.51	3.13
ebm	81.13	81.23	81.40	81.27	0.10	0.28	0.14	74.84	74.94	74.93	74.95	0.09	0.09	0.10
ffn	71.73	72.79	73.11	73.02	1.05	1.37	1.28	64.01	67.56	65.13	65.75	3.55	1.12	1.73
ffnlp	71.58	71.96	72.83	72.17	0.38	1.25	0.58	64.96	65.44	65.43	66.57	0.48	0.47	1.61
ffnsp	72.35	72.17	73.07	72.46	- 0.18	0.72	0.10	64.93	66.32	65.79	66.45	1.39	0.86	1.52
knn	67.21	66.92	66.85	67.03	- 0.29	- 0.37	- 0.18	63.23	63.00	62.92	63.03	- 0.22	- 0.30	- 0.20
lda	67.23	68.01	68.56	68.27	0.78	1.33	1.04	63.09	63.74	64.12	63.85	0.65	1.03	0.76
mlp	69.30	70.83	71.62	70.01	1.53	2.31	0.71	60.82	63.26	62.93	60.51	2.44	2.12	- 0.31
percep	57.96	58.44	58.16	56.56	0.48	0.20	- 1.40	44.58	57.99	47.08	54.03	13.41	2.50	9.45
qda	67.44	66.98	65.86	67.20	- 0.46	- 1.58	- 0.24	52.50	54.01	54.08	53.82	1.52	1.58	1.32
rf	81.74	81.51	81.20	81.46	- 0.23	- 0.54	- 0.28	75.51	75.44	75.29	75.47	- 0.07	- 0.22	- 0.04
ridge	66.94	67.55	68.07	67.83	0.62	1.14	0.89	62.93	63.39	63.78	63.55	0.46	0.85	0.62
xgb	80.76	81.00	81.05	81.09	0.23	0.28	0.33	74.77	74.98	75.00	75.05	0.21	0.23	0.28
xnn	50.21	50.42	50.69	49.86	0.21	0.47	- 0.35	53.52	64.62	58.11	62.32	11.09	4.59	8.79
2.D99														
cnn	72.80	73.40	74.06	73.79	0.60	1.25	0.98	68.60	70.07	69.79	70.22	1.48	1.19	1.63
ebm	81.04	81.24	81.40	81.27	0.19	0.36	0.23	74.77	74.93	75.01	74.95	0.16	0.24	0.19
ffn	72.82	72.50	72.90	72.94	- 0.33	0.08	0.11	65.07	67.39	65.81	67.26	2.31	0.74	2.19
ffnlp	72.93	72.33	73.50	71.57	- 0.60	0.57	- 1.35	64.79	66.33	66.36	66.76	1.54	1.56	1.96
ffnsp	71.88	72.73	72.13	72.97	0.85	0.25	1.09	66.23	66.61	65.97	67.17	0.38	- 0.26	0.94
knn	67.10	66.88	67.10	66.91	- 0.22	0.00	- 0.19	63.24	63.02	63.17	63.06	- 0.23	- 0.08	- 0.18
lda	67.17	67.92	68.53	68.22	0.75	1.36	1.05	63.07	63.68	64.12	63.79	0.60	1.05	0.72
mlp	69.79	68.93	69.57	69.17	- 0.86	- 0.22	- 0.62	63.83	64.12	60.77	58.27	0.29	- 3.06	- 5.56
percep	55.09	57.67	53.79	59.21	2.58	- 1.30	4.12	55.48	58.56	55.13	48.25	3.07	- 0.36	- 7.23
qda	67.32	66.90	65.80	67.15	- 0.42	- 1.52	- 0.16	52.41	53.98	53.89	53.67	1.57	1.49	1.26
rf	81.68	81.33	81.18	81.41	- 0.35	- 0.50	- 0.27	75.63	75.32	75.32	75.40	- 0.31	- 0.31	- 0.23
ridge	66.85	67.47	68.04	67.78	0.62	1.19	0.93	62.94	63.39	63.77	63.49	0.45	0.82	0.55
xgb	80.72	80.96	81.10	81.11	0.24	0.38	0.40	74.81	74.97	75.19	75.00	0.16	0.38	0.19
xnn	49.91	50.11	50.00	49.88	0.20	0.09	- 0.02	56.25	62.86	62.66	56.14	6.62	6.42	- 0.11
3.D80														
cnn	71.98	73.18	73.19	73.24	1.21	1.21	1.26	66.51	70.27	70.06	70.11	3.75	3.55	3.60
ebm	80.73	81.03	81.17	81.06	0.29	0.44	0.32	74.92	75.16	75.27	75.20	0.24	0.35	0.28
ffn	71.97	71.95	72.62	71.74	- 0.02	0.65	- 0.23	65.38	64.76	65.96	67.04	- 0.62	0.59	1.67
ffnlp	70.31	72.08	72.51	70.73	1.77	2.20	0.42	66.23	67.11	66.03	65.58	0.89	- 0.20	- 0.65
ffnsp	70.27	71.89	72.16	72.29	1.62	1.89	2.02	66.20	65.43	64.91	65.77	- 0.77	- 1.29	- 0.43
knn	66.57	66.57	66.57	66.57	- 0.01	- 0.01	- 0.00	62.98	62.95	62.95	62.97	- 0.02	- 0.03	- 0.01
lda	67.33	67.90	68.29	68.07	0.57	0.96	0.74	63.45	63.88	64.20	64.10	0.43	0.75	0.65
mlp	67.96	68.85	68.47	67.87	0.89	0.52	- 0.08	58.09	63.16	61.09	58.34	5.07	2.99	0.25
percep	57.44	55.39	56.68	57.06	- 2.05	- 0.76	- 0.38	52.88	57.22	59.57	57.33	4.35	6.69	4.45
qda	67.69	66.51	64.09	66.23	- 1.18	- 3.60	- 1.46	52.76	54.28	52.61	54.33	1.52	- 0.15	1.57
rf	81.18	81.03	80.64	80.99	- 0.15	- 0.55	- 0.20	75.47	75.57	75.25	75.52	0.11	- 0.21	0.05
ridge	66.67	67.23	67.59	67.40	0.56	0.92	0.73	63.06	63.35	63.74	63.54	0.29	0.68	0.48
xgb	80.20	80.75	80.77	80.80	0.54	0.57	0.60	74.74	75.13	75.27	75.18	0.39	0.52	0.44
xnn	50.41	50.39	50.04	50.27	- 0.02	- 0.36	- 0.14	53.24	65.69	60.33	63.61	12.45	7.09	10.37

En términos de *AUC*, el modelo *CNN* mantiene un rendimiento destacado, con incrementos en los escenarios *F+P-SFE* y *F-SFE* en todos los grupos de datos. En *D100*, por ejemplo, *CNN* logra un aumento de **1.61 %** en *F+P-SFE*, lo que subraya su ca-

pacidad para capturar patrones en los datos ajustados temporalmente. El modelo *MLP* también evidencia mejoras en *AUC*, particularmente en el grupo *D100*, con un incremento de **2.31 %** en *F+P-SFE* (ver Tabla 24).



(a) *AUC*



(b) *ACC*

Figura 14

Comparación del promedio en 10 folds de métricas AUC y ACC en modelos de ML con el FE COVID-19 MOV, incluyendo créditos con apoyo gubernamental, al 100 % de las actividades económicas.

Nota: Las figuras muestran las variaciones (Δ) en *AUC* y *ACC* entre diferentes escenarios (*SFE*, *F-SFE*, *P-SFE*, *F+P-SFE*) para cada modelo de *ML* evaluado. Los colores representan rangos de variación porcentual, facilitando la identificación de mejoras o deterioros en el rendimiento.

Modelos como *Perceptrón* y *XNN* muestran una marcada sensibilidad al desplazamiento temporal, con fluctuaciones notables en ambas métricas (*AUC* y *ACC*). En ciertos escenarios, como *P-SFE* en *D99*, el *Perceptrón* experimenta mejoras significativas, pero su rendimiento decrece en otros contextos, lo que indica una fuerte dependencia de

la estructura temporal de los datos. Esto podría sugerir que estos modelos tienen dificultades en manejar datos de series temporales ajustadas, debido a su alta dependencia de patrones temporales específicos.
Dentro de los algoritmos clásicos, *LDA* y *Ridge* muestran mejoras modestas pero consistentes a lo largo de los escenarios evaluados, lo que resalta su adaptabilidad a conjuntos de datos enriquecidos con modificaciones temporales.
En general, el experimento *COVID MOV* subraya la influencia crítica de los cambios temporales en los *FE*. Este análisis demuestra que la inclusión de variables derivadas del mismo *FE*, pero afectadas por su dinámica temporal, debe ser considerada cuidadosamente en el diseño experimental para optimizar el desempeño predictivo de los modelos de *ML*. Además, los resultados destacan la importancia de establecer marcos experimentales robustos para mitigar la variabilidad y maximizar la precisión predictiva en la modelización del riesgo crediticio.

Evaluación excluyendo en la data créditos que han sido beneficiados por programas gubernamentales

El análisis de la tabla 25, correspondiente a las métricas *ACC* y *AUC*, respectivamente, revela el efecto del desplazamiento de la serie temporal de fallecidos por *COVID-19* en el rendimiento de los modelos de *ML*, considerando únicamente créditos sin respaldo gubernamental. Este diseño experimental permite evaluar con mayor precisión cómo la exclusión de estos créditos impacta la capacidad predictiva de los modelos en escenarios con *FE*, eliminando posibles sesgos introducidos por intervenciones estatales.
En términos de *ACC*, el modelo *CNN* se mantiene como el más robusto en todos los grupos de datos (*D100*, *D99* y *D80*), con mejoras significativas cuando se incorporan *FE*. En *D100*, se observa un incremento de hasta **3.80 %** en *F+P-SFE*, lo que refuerza su capacidad de adaptación a la variabilidad introducida por el desplazamiento temporal de la serie. En *D99* y *D80*, los aumentos son aún más pronunciados, alcanzando **4.58 %** y **7.16 %** en *F+P-SFE*, respectivamente (ver Tabla 25).
Modelos clásicos como *LDA* y *QDA* presentan mejoras moderadas en *ACC*, con incrementos entre **0.5 % y 2 %**, lo que demuestra su adaptabilidad a conjuntos de datos sin respaldo financiero externo. Sin embargo, el comportamiento de los modelos neuronales varía: mientras que *MLP* muestra caídas significativas en *D99* (-**16.38 %** en *F-SFE*), el modelo *Perceptrón* alcanza incrementos destacados en *D80* (**6.74 %** en *F+P-SFE*). Estos resultados indican que la sensibilidad de los modelos neuronales al desplazamiento temporal y la exclusión de créditos con respaldo gubernamental es altamente dependiente de la estructura de los datos y del escenario evaluado (ver Tabla 25).
En cuanto a la métrica *AUC*, el modelo *CNN* sigue destacando con mejoras consistentes en todos los grupos de datos. En *D100*, logra un incremento de **1.30 %** en *F+P-*

SFE, mientras que en *D99* y *D80*, los aumentos alcanzan **1.56 %** y **1.46 %**, respectivamente, lo que refuerza su capacidad de aprendizaje bajo escenarios con desplazamiento temporal (ver Tabla 25).

Modelos como *DL*, *LDA*, *XGB* y *Ridge* también presentan mejoras relevantes en *AUC*, con incrementos de hasta **1.79 %** en *DL* (*F+P-SFE* en *D100*) y de **1.11 %** en *LDA* (*F+P-SFE* en *D99*). En contraste, el desempeño del *Perceptrón* es altamente variable: aunque en *D80* registra un aumento del **3.59 %** en *F-SFE*, en otros grupos y escenarios su rendimiento disminuye, lo que sugiere que su capacidad predictiva es altamente dependiente del tipo de datos y del diseño experimental (ver Figura 15).

Por otro lado, modelos como *KNN* y *RF* presentan una respuesta limitada a los escenarios con *FE*. En la mayoría de los casos, sus variaciones en *AUC* son marginales o negativas, lo que indica que estos modelos no se benefician significativamente de la exclusión de créditos respaldados por el gobierno (ver Tabla 25).

Adicionalmente, modelos de redes neuronales explicables como *XNN* exhiben comportamientos contrastantes. En algunos escenarios, logran mejoras notables en precisión, alcanzando hasta **16.17 %** de incremento en *ACC* bajo condiciones específicas. Sin embargo, esta respuesta no es uniforme y varía en función de la composición del conjunto de datos. De manera similar, modelos como *FFNLP* y *FFNSP* muestran una alta sensibilidad a la exclusión de créditos con respaldo estatal, con fluctuaciones significativas en *AUC* y *ACC*, lo que sugiere que estos modelos dependen de la granularidad de los datos para lograr una alta precisión predictiva (ver Figura 15).

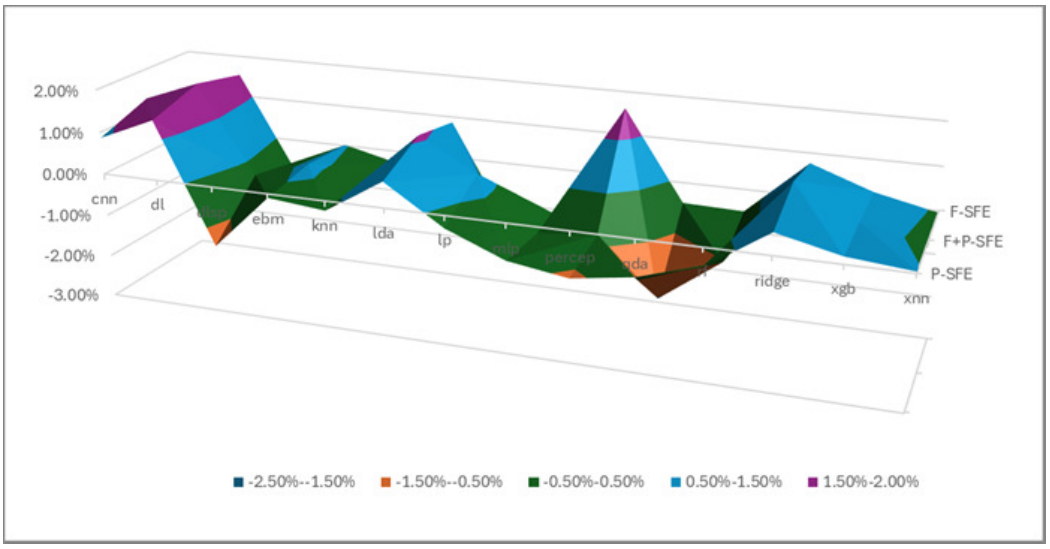
En términos generales, estos resultados destacan que la exclusión de créditos respaldados por el gobierno introduce variabilidad en el desempeño predictivo de los modelos. Mientras que los modelos basados en árboles de decisión (*RF*, *XGB*) y el modelo *CNN* muestran una mayor resiliencia ante la reducción del conjunto de datos, algunos modelos neuronales enfrentan desafíos en la adaptación a la menor complejidad del dataset. Esto pone en evidencia la importancia del ajuste específico de los *FE* y de la configuración del escenario para optimizar la precisión predictiva en la evaluación del riesgo crediticio.

Estos hallazgos resaltan la relación compleja entre la exclusión de créditos respaldados por el gobierno y la variabilidad del desempeño predictivo en modelos de *ML*. Además, evidencian la necesidad de ajustes experimentales robustos para mitigar el sesgo introducido por la intervención estatal y optimizar la capacidad predictiva de los modelos en entornos de alta volatilidad.

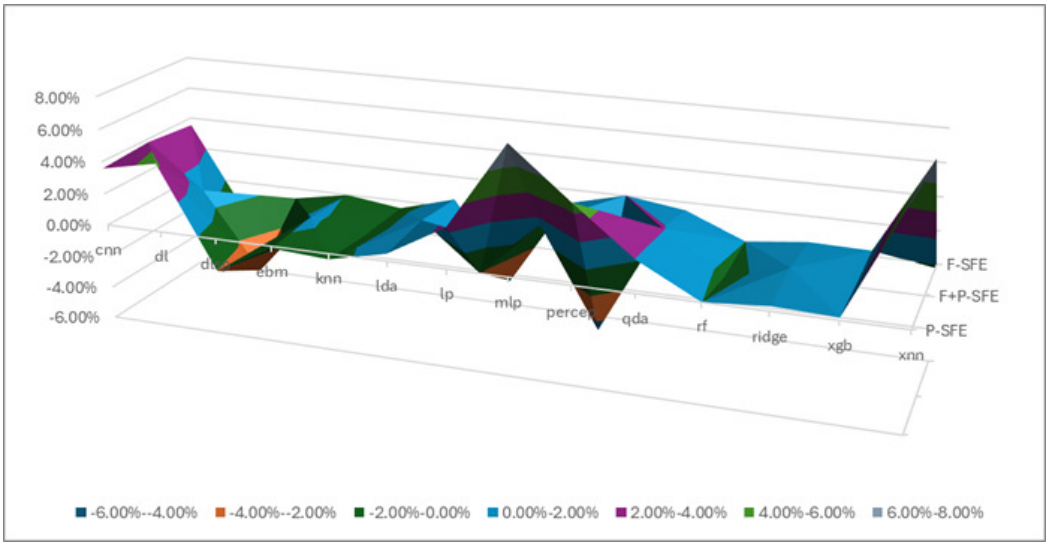
Tabla 25

Promedio de AUC & ACC en 10 folds de modelos de ML con el FE COVID-19 [COVID MOV], excluyendo los créditos con apoyo gubernamental, según el escenario evaluado y los grupos de relevancia.

	AUC (%)							*ACC* (%)						
	SCENARIOS				RESULTS			SCENARIOS				RESULTS		
	SFE	P	F+P	F	P-SFE	F+P-SFE	F-SFE	SFE	P	F+P	F	P-SFE	F+P-SFE	F-SFE
Exc. CG														
1.D100														
cnn	73.31	74.18	74.61	74.10	0.87	1.30	0.80	69.26	72.83	73.05	72.78	3.57	3.80	3.52
ebm	81.38	81.29	81.49	81.40	- 0.09	0.11	0.02	76.54	76.41	76.66	76.52	- 0.13	0.12	- 0.02
ffn	71.66	73.08	73.45	73.19	1.42	1.79	1.53	65.08	69.34	66.12	62.94	4.26	1.03	- 2.14
ffnlp	73.32	73.03	73.66	71.79	- 0.29	0.34	- 1.53	68.13	70.61	66.03	63.72	2.48	- 2.09	- 4.41
ffnsp	73.51	72.10	73.20	72.72	- 1.41	- 0.30	- 0.79	68.47	66.44	64.67	67.81	- 2.03	- 3.80	- 0.66
knn	66.21	65.98	65.89	65.90	- 0.23	- 0.32	- 0.31	63.19	62.84	62.65	62.70	- 0.35	- 0.54	- 0.48
lda	66.08	66.68	67.18	66.95	0.60	1.10	0.87	63.18	63.64	64.00	63.70	0.46	0.82	0.52
mlp	70.92	70.07	70.02	69.70	- 0.85	- 0.90	- 1.22	60.20	67.96	61.74	61.31	7.76	1.54	1.10
percep	55.94	54.88	55.19	57.55	- 1.06	- 0.75	1.61	54.77	59.56	50.09	56.79	4.78	- 4.68	2.02
qda	66.82	65.97	64.74	66.34	- 0.84	- 2.08	- 0.48	50.72	52.66	52.80	52.24	1.94	2.08	1.51
rf	79.53	79.09	78.49	79.01	- 0.44	- 1.03	- 0.52	75.34	75.34	74.93	75.34	0.01	- 0.40	0.00
ridge	65.95	66.44	66.90	66.71	0.50	0.95	0.77	63.14	63.44	63.79	63.56	0.30	0.65	0.42
xgb	80.43	80.60	80.65	80.70	0.17	0.22	0.26	76.00	76.16	76.18	76.34	0.16	0.18	0.34
xnn	49.96	50.00	49.85	49.92	0.04	- 0.11	- 0.05	58.19	66.17	65.91	57.93	7.98	7.72	- 0.27
2.D99														
cnn	73.06	74.03	74.62	74.77	0.97	1.56	1.71	68.66	72.76	73.12	73.24	4.10	4.47	4.58
ebm	81.23	81.31	81.45	81.34	0.08	0.22	0.11	76.31	76.39	76.51	76.45	0.07	0.20	0.14
ffn	71.40	73.31	73.20	73.08	1.91	1.80	1.68	65.45	67.42	64.39	65.51	1.98	- 1.05	0.06
ffnlp	72.79	72.79	73.32	72.59	0.01	0.53	- 0.20	67.15	67.42	66.71	66.48	0.27	- 0.44	- 0.66
ffnsp	71.94	71.95	73.53	71.60	0.01	1.59	- 0.33	60.05	64.21	66.13	65.70	4.17	6.08	5.65
knn	65.94	66.00	65.81	65.69	0.06	- 0.13	- 0.25	62.92	62.58	62.67	62.59	- 0.35	- 0.25	- 0.33
lda	66.00	66.67	67.11	66.88	0.67	1.11	0.88	63.15	63.63	64.00	63.68	0.48	0.85	0.53
mlp	68.79	70.06	69.61	69.62	1.27	0.82	0.83	70.57	64.20	58.61	54.19	- 6.37	- 11.96	- 16.38
percep	55.28	56.66	56.65	56.40	1.37	1.37	1.12	54.45	56.49	57.69	60.59	2.04	3.24	6.14
qda	66.72	65.95	64.75	66.15	- 0.77	- 1.97	- 0.57	50.57	52.65	52.60	51.89	2.08	2.03	1.32
rf	79.41	78.96	78.61	78.90	- 0.46	- 0.80	- 0.51	75.31	75.19	74.94	75.23	- 0.13	- 0.37	- 0.08
ridge	65.85	66.42	66.82	66.64	0.57	0.97	0.79	63.13	63.45	63.82	63.49	0.32	0.69	0.35
xgb	80.35	80.66	80.61	80.56	0.32	0.26	0.22	76.13	76.28	76.15	76.20	0.15	0.02	0.07
xnn	49.89	49.89	50.00	50.22	0.00	0.12	0.33	53.92	70.09	66.22	66.57	16.17	12.30	12.65
3.D80														
cnn	72.58	73.36	74.04	73.73	0.78	1.46	1.15	66.42	72.66	73.58	73.19	6.25	7.16	6.78
ebm	80.75	81.05	81.11	81.06	0.30	0.36	0.31	76.37	76.72	76.73	76.73	0.35	0.36	0.36
ffn	72.47	72.15	73.07	71.92	- 0.32	0.59	- 0.55	64.08	67.19	65.91	68.57	3.11	1.83	4.49
ffnlp	70.26	72.56	73.17	71.30	2.31	2.92	1.04	61.75	67.37	66.64	63.25	5.62	4.89	1.50
ffnsp	71.99	72.70	72.66	72.22	0.72	0.67	0.23	69.54	67.24	68.99	67.02	- 2.30	- 0.56	- 2.52
knn	65.37	65.36	65.36	65.37	- 0.01	- 0.01	- 0.00	62.54	62.52	62.51	62.54	- 0.02	- 0.03	0.01
lda	65.85	66.35	66.70	66.51	0.50	0.85	0.66	63.27	63.64	63.94	63.75	0.37	0.66	0.47
mlp	69.10	71.07	70.63	71.45	1.96	1.53	2.34	62.97	59.38	60.65	65.13	- 3.59	- 2.33	2.16
percep	55.71	56.80	57.65	59.30	1.09	1.94	3.59	51.81	55.32	50.69	58.55	3.51	- 1.12	6.74
qda	66.65	65.29	63.23	65.24	- 1.36	- 3.42	- 1.41	50.91	53.12	51.39	52.64	2.21	0.49	1.73
rf	78.71	78.53	78.08	78.54	- 0.18	- 0.63	- 0.17	75.29	75.54	75.36	75.57	0.25	0.08	0.28
ridge	65.48	65.97	66.26	66.12	0.48	0.78	0.63	63.14	63.37	63.63	63.57	0.23	0.49	0.43
xgb	79.57	80.16	80.21	80.22	0.59	0.64	0.65	75.92	76.45	76.51	76.39	0.54	0.59	0.47
xnn	50.38	50.23	50.56	50.42	- 0.15	0.18	0.04	67.98	63.92	69.19	59.87	- 4.06	1.20	- 8.11



(a) *AUC*



(b) *ACC*

Figura 15

Comparación del promedio en 10 folds de métricas AUC y ACC en modelos de ML con el FE COVID-19 MOV, excluyendo créditos con apoyo gubernamental, al 100 % de las actividades económicas.

Nota: Las figuras muestran las variaciones (Δ) en *AUC* y *ACC* entre diferentes escenarios (*SFE*, *F-SFE*, *P-SFE*, *F+P-SFE*) para cada modelo de *ML* evaluado. Los colores representan rangos de variación porcentual, facilitando la identificación de mejoras o deterioros en el rendimiento.

Evaluación del nivel de significancia de los modelos de *ML* para *AUC* y *ACC* aplicando el *FE* COVID, con la serie temporal correspondiente a los fallecidos ha sido desplazada un periodo hacia la derecha, y considerando la exclusión como inclusión de los créditos con apoyo de programas gubernamentales

La tabla 26 presenta los resultados de los modelos de *ML* que exhiben mejoras estadísticamente significativas en los indicadores de *ACC* y *AUC* tras la inclusión de *FE*

relacionados con *COVID-19* bajo el escenario *COVID MOV*. Este escenario incorpora un desplazamiento temporal en la serie de fallecidos por *COVID-19* , considerando los periodos típicos de evolución de la enfermedad desde el contagio hasta el fallecimiento. La comparación se realiza entre escenarios sin *FE* (*SFE*) y aquellos con la incorporación de datos de fallecidos (*F*), positivos (*P*) y su combinación (*F+P*).

Para determinar la significancia estadística de las mejoras en los modelos, se aplicó la prueba *t-student* para muestras relacionadas. Los resultados se dividen en dos categorías principales: *Inc CG* y *Exc CG* para evaluar la influencia de la intervención estatal en el desempeño de los modelos.

La estructura de la tabla incluye tres grupos de relevancia: *D100*, *D99* y *D80*, que representan subconjuntos de datos con distintos niveles de cobertura y granularidad. Se presenta el número de modelos que lograron mejoras significativas en los indicadores *ACC* y *AUC*, donde el valor "Y"denota significancia estadística.

En términos globales, se observa que los modelos basados en árboles de decisión, como *XGB* y *RF*, muestran un desempeño robusto y consistente en todos los grupos de relevancia. Estos modelos alcanzan significancia en la mayoría de los escenarios evaluados (*F*, *F+P*, *P*) cuando los créditos con respaldo gubernamental están presentes (*Inc. CG*) y mantienen una estabilidad considerable en la ausencia de estos (*Exc. CG*). El desplazamiento temporal en los datos de mortalidad parece mejorar aún más la estabilidad de estos modelos, optimizando su capacidad predictiva.

Entre los modelos basados en redes neuronales, *CNN* demuestra una adaptabilidad sobresaliente a los conjuntos de datos enriquecidos y al ajuste temporal en los *FE*. Consistentemente alcanza significancia estadística en todos los escenarios, tanto en los datos con créditos con *Inc. CG*, como con *Exc. CG*. Este resultado refuerza su idoneidad para capturar patrones complejos en series temporales ajustadas.

Por otro lado, los modelos clásicos como *LDA* y *Ridge* presentan mejoras estables y estadísticamente significativas en la mayoría de los escenarios. Su desempeño es particularmente fuerte en escenarios donde se incluyen *FE*, manteniendo robustez tanto en la presencia como en la ausencia de créditos con respaldo estatal. Este hallazgo sugiere que estos modelos pueden ofrecer una solución intermedia entre interpretabilidad y precisión en la predicción del riesgo crediticio.

Se identificaron mejoras específicas por escenario en modelos como *FFN*, *FFNLP* y *MLP*, lo que indica que su sensibilidad a los *FE* es dependiente de la composición del conjunto de datos y de la combinación de grupos de relevancia. Estos modelos presentan una mayor variabilidad en su respuesta, lo que resalta la importancia de ajustar su configuración en función de la granularidad de los datos.

Tabla 26

Modelos con mejoras estadísticamente significativas en las métricas AUC y ACC con el FE COVID-19 [COVID MOV], considerando escenarios con créditos con beneficios gubernamentales y sin créditos con beneficios gubernamentales y grupos de relevancia - escenarios.

	Inc. CG							Exc. CG						
	AUC			ACC			TOTAL	AUC			ACC			TOTAL
	F	F+P	P	F	F+P	P		F	F+P	P	F	F+P	P	
1.D100														
cnn	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	6
ebm	-	-	-	-	-	-	0	-	-	-	-	-	-	0
ffn	-	-	-	-	-	-	0	-	Y	-	-	-	Y	2
ffnlp	-	-	-	-	-	-	0	-	-	-	-	-	Y	1
lda	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	6
mlp	-	Y	-	-	-	-	1	-	-	-	-	-	Y	1
percep	-	-	-	-	-	Y	1	-	-	-	-	-	-	0
qda	-	-	-	Y	Y	Y	3	-	-	-	Y	Y	Y	3
ridge	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	-	5
xgb	Y	-	-	Y	-	Y	3	-	Y	-	Y	-	-	2
xnn	-	-	-	Y	-	Y	2	-	-	-	-	-	-	0
2.D99														
cnn	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	6
ebm	-	Y	-	-	-	-	1	-	-	-	-	-	-	0
ffn	-	-	-	-	-	-	0	-	Y	Y	-	-	-	2
ffnlp	-	-	-	-	-	-	0	-	-	-	-	-	-	0
ffnsp	-	-	-	-	-	-	0	-	Y	-	Y	Y	-	3
lda	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	-	5
percep	Y	-	-	-	-	-	1	-	-	-	-	-	-	0
qda	-	-	-	Y	Y	Y	3	-	-	-	Y	Y	Y	3
ridge	Y	Y	Y	Y	Y	Y	6	Y	Y	-	Y	Y	-	4
xgb	-	Y	-	-	Y	-	2	-	Y	Y	-	-	-	2
xnn	-	-	-	-	-	-	0	-	-	-	-	-	Y	1
3.D80														
cnn	Y	Y	Y	Y	Y	Y	6	Y	Y	-	Y	Y	Y	5
ebm	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	6
ffn	-	-	-	-	-	-	0	-	-	-	-	-	-	0
ffnlp	-	Y	-	-	-	-	1	-	Y	Y	-	-	Y	3
ffnsp	-	-	-	-	-	-	0	-	-	-	-	-	-	0
lda	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	6
mlp	-	-	-	-	-	-	0	Y	-	-	-	-	-	1
percep	-	-	-	-	-	-	0	Y	-	-	-	-	-	1
qda	-	-	-	Y	-	Y	2	-	-	-	Y	Y	Y	3
rf	-	-	-	-	-	-	0	-	-	-	Y	-	Y	2
ridge	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	6
xgb	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	6
xnn	-	-	-	-	-	Y	1	-	-	-	-	-	-	0

Un hallazgo clave es que la exclusión de los créditos con *Exc. CG* amplifica la variabilidad del desempeño en múltiples modelos, especialmente en redes neuronales y

métodos estadísticos clásicos. Esto sugiere que la estructura de los datos y la presencia de intervenciones gubernamentales juegan un papel determinante en la estabilidad y la significancia de las mejoras observadas.

Finalmente, la inclusión de un desplazamiento temporal en la serie de fallecidos *COVID MOV* mejora la capacidad de modelos como *CNN* y *XGB* para detectar mejoras estadísticamente significativas. Este resultado destaca la importancia de considerar *FE* ajustados temporalmente en los modelos predictivos, ya que permite capturar de manera más precisa la relación entre la evolución de los *FE* y el riesgo crediticio.

En conjunto, estos hallazgos subrayan el papel fundamental de los conjuntos de datos enriquecidos y de los ajustes temporales en la mejora de la precisión predictiva, particularmente en modelos de redes neuronales y árboles de decisión. Asimismo, evidencian que la composición del conjunto de datos y el diseño experimental son factores clave para optimizar el desempeño y la estabilidad de los modelos en la evaluación de riesgo crediticio bajo condiciones de alta incertidumbre.

5.1.3. Evaluación de los niveles de significancia obtenidos para los modelos de *ML* aplicando el *FE* COVID

La comparación entre las tablas 23 y 26 revela cómo el desplazamiento temporal de la serie de fallecidos por *COVID-19* impacta en la identificación de mejoras significativas en los indicadores *ACC* y *AUC*. En general, el total de mejoras aumentó de 166 a 178, destacándose un incremento notable en los escenarios (*Exc CG*, mientras que los escenarios *Inc CG* muestran mejoras, esto demostraría el sesgo agregado por las medidas adoptadas por el gobierno.

Los modelos *CNN*, *LDA*, *QDA*, *Ridge* y *XGB* se mantienen como los más consistentes en ambos análisis, logrando mejoras significativas en múltiples escenarios y grupos de relevancia. Sin embargo, se observa que el desplazamiento temporal resalta patrones más específicos en escenarios *Exc CG*, con un aumento total de 80 a 91 mejoras. Este hallazgo sugiere que los efectos del *COVID-19* pueden ser más relevantes en análisis donde no se consideran créditos gubernamentales.

En términos de los escenarios individuales, *F+P* continúa mostrando el mayor impacto en la mejora de los modelos, seguido por los escenarios individuales *F* y *P*. El desplazamiento temporal afecta principalmente el escenario *F*, con una reducción de mejoras en *Inc CG* pero un aumento en *Exc CG*. Esto indica que el ajuste temporal captura patrones diferidos del impacto del *COVID-19* , lo que puede ser crucial para un análisis más preciso en contextos dinámicos.

Finalmente, el análisis por grupos de datos (*D100*, *D99* y *D80*) sugiere que los grupos *D100* y *D80* son más estables en términos de mejoras significativas en ambos análisis. En contraste, el grupo *D99* experimenta una variación más pronunciada, reflejando la

sensibilidad de este subconjunto a los ajustes temporales. Estos resultados resaltan la importancia de considerar el desplazamiento temporal en el diseño de modelos predictivos para escenarios de crisis cuando el caso así lo justifique, esto podría mejorar la capacidad de respuesta y adaptación ante eventos externos.

5.1.4. *FE* Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos

En esta sección analizamos los promedios de la métrica *ACC* y *AUC* con 10 *folds* para los modelos de (*ML*) agrupados por relevancia, considerando los créditos con y sin apoyo gubernamental. Se evaluaron distintos escenarios: *SFE*, y con la incorporación de *FE* como: *T* (anomalía de temperatura), *W* (bloqueos de vías por factores climáticos) , y *U* (bloqueos de vías por convulsión social, tanto de manera individual como combinada (*T*, *T+W*, *U*, *U+T*, *U+T+W*, *U+W*, y *W*).

Evaluación incluyendo en la data créditos que han sido beneficiados por programas gubernamentales

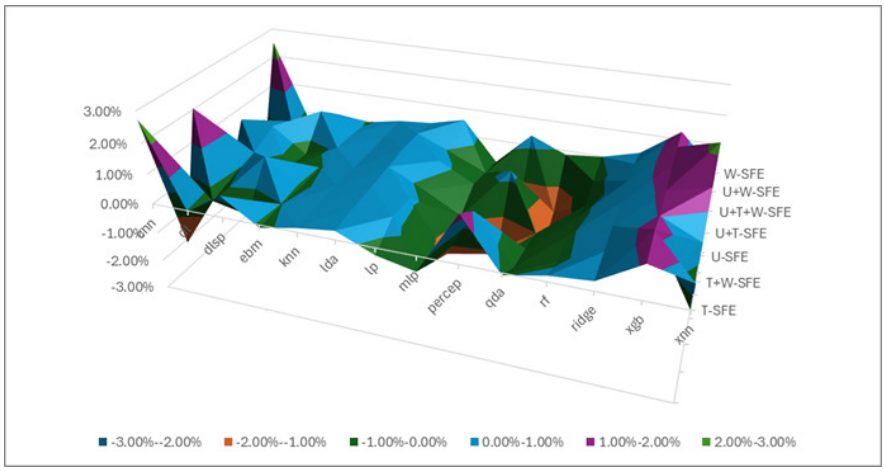
Las Tablas 28 y 27 presentan el desempeño de los modelos de *ML* en términos de *ACC* y *AUC* bajo escenarios que incluyen *FE* relacionados con: *T* (anomalía de temperatura), *W* (bloqueos de vías por factores climáticos) y *U* (bloqueos de vías por convulsión social). Este análisis considera créditos *Inc CG* y se aplica a tres conjuntos de datos con distintos niveles de representatividad (*D100*, *D99*, *D80*).
En términos generales, los modelos basados en árboles de decisión y ensamblados, como *RF*, *EBM* y *XGB*, muestran un desempeño notablemente estable, logrando valores de *AUC* superiores al 75 % en la mayoría de los escenarios evaluados. Esto evidencia su capacidad para manejar variaciones inducidas por *FE*, manteniendo una precisión predictiva robusta. Particularmente, *XGB* presenta mejoras en *AUC* de hasta 1.46 % en el conjunto *D80*, lo que indica su adaptabilidad a escenarios de alta incertidumbre (ver Figura 16).
En cuanto a modelos basados en redes neuronales, *CNN* se consolida como uno de los modelos con mejor desempeño en escenarios enriquecidos con *FE*. Se observa un incremento significativo en *ACC*, con diferencias de hasta **32.50 puntos porcentuales** en el conjunto de datos *D100* al comparar el escenario completo (*U+T+W*) con la base sin *FE* (*SFE*). Este comportamiento sugiere una notable capacidad de *CNN* para capturar patrones en datos complejos, beneficiándose de la inclusión de múltiples *FE* en conjunto (ver Figura 17).
Por otro lado, los modelos clásicos como *LDA* y *Ridge* presentan mejoras moderadas pero consistentes en ambos indicadores. Aunque sus incrementos en *AUC* y *ACC*

son menores en comparación con *CNN* o *RF*, su estabilidad sugiere que son opciones confiables en contextos donde la interpretabilidad es prioritaria.

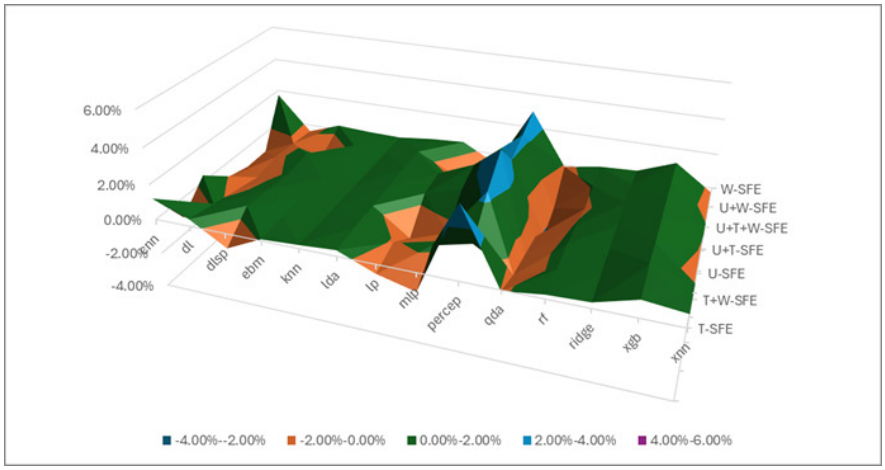
Tabla 27

FE Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos: Promedio de AUC en 10 folds de los modelos de ML según grupos de relevancia, incluyendo los créditos con apoyo gubernamental.

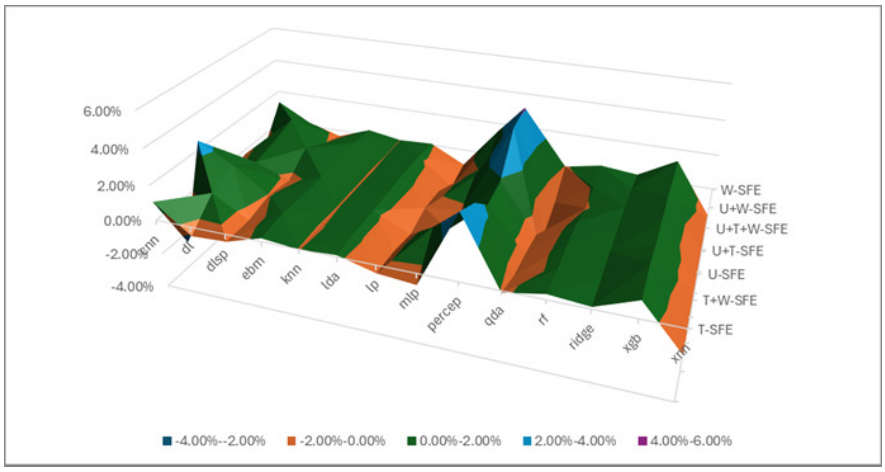
	AUC (%)														
	SCENARIOS								RESULTS						
	SFE	T	T +W	U	U +T	U +T +W	U +W	W	T -SFE	T +W -SFE	U -SFE	U +T -SFE	U+T +W -SFE	U +W -SFE	W -SFE
Inc. CG															
1.D100															
cnn	62.05	64.79	59.93	64.11	60.59	62.64	60.90	64.56	2.74	- 2.12	2.05	- 1.46	0.59	- 1.15	2.51
ebm	75.08	74.99	75.07	75.01	75.06	75.19	75.15	75.03	- 0.10	- 0.01	- 0.07	- 0.02	0.11	0.07	- 0.05
ffn	67.42	67.39	67.07	67.33	67.55	67.95	67.59	67.12	- 0.03	- 0.35	- 0.09	0.13	0.53	0.17	- 0.30
ffnlp	67.66	67.42	67.95	67.85	67.47	67.89	67.58	68.24	- 0.24	0.29	0.20	- 0.19	0.23	- 0.08	0.58
ffnsp	67.45	68.17	66.83	68.22	67.21	67.11	68.17	67.53	0.72	- 0.62	0.77	- 0.24	- 0.33	0.72	0.08
knn	61.08	61.22	61.24	61.25	61.25	61.24	61.24	61.26	0.14	0.17	0.18	0.17	0.16	0.17	0.18
lda	65.55	65.84	65.92	65.93	65.98	66.03	65.99	65.83	0.30	0.37	0.39	0.43	0.49	0.44	0.28
mlp	65.40	64.90	64.39	63.91	64.54	65.01	64.79	64.58	- 0.51	- 1.01	- 1.49	- 0.86	- 0.39	- 0.62	- 0.82
percep	54.89	56.45	55.82	53.65	52.79	55.17	53.03	55.26	1.56	0.93	- 1.24	- 2.09	0.28	- 1.86	0.37
qda	67.21	67.14	66.58	66.42	66.20	64.98	65.55	67.07	- 0.07	- 0.63	- 0.79	- 1.01	- 2.23	- 1.67	- 0.15
rf	76.83	77.02	76.67	76.73	76.61	76.32	76.54	76.81	0.19	- 0.16	- 0.10	- 0.21	- 0.50	- 0.28	- 0.02
ridge	64.30	64.61	64.71	64.77	64.82	64.89	64.85	64.62	0.30	0.40	0.47	0.52	0.59	0.54	0.32
xgb	74.21	75.31	75.54	75.24	75.36	75.51	75.52	75.47	1.09	1.33	1.03	1.15	1.30	1.31	1.26
xnn	50.27	51.07	49.37	50.71	51.74	52.51	50.07	50.35	0.79	- 0.90	0.44	1.47	2.24	- 0.20	0.08
2.D99															
cnn	62.51	63.71	61.43	62.88	61.68	61.32	61.35	64.21	1.20	- 1.08	0.38	- 0.83	- 1.19	- 1.16	1.71
ebm	74.81	74.90	74.99	74.98	75.01	75.08	75.05	74.97	0.09	0.18	0.16	0.19	0.26	0.23	0.16
ffn	68.01	68.42	68.15	67.85	67.56	67.83	67.73	67.59	0.41	0.14	- 0.17	- 0.45	- 0.18	- 0.29	- 0.42
ffnlp	67.88	67.35	67.83	66.78	68.27	68.40	67.34	68.33	- 0.53	- 0.05	- 1.10	0.39	0.52	- 0.54	0.45
ffnsp	67.71	66.85	68.13	67.99	67.83	68.08	67.12	67.98	- 0.86	0.41	0.27	0.12	0.36	- 0.59	0.26
knn	60.92	61.12	61.01	61.03	61.03	61.02	61.02	61.03	0.20	0.10	0.12	0.11	0.10	0.10	0.11
lda	65.54	65.88	65.96	65.96	66.00	66.07	66.03	65.87	0.34	0.42	0.42	0.47	0.53	0.49	0.33
mlp	64.63	63.56	64.81	64.26	65.57	66.02	62.23	64.00	- 1.07	0.18	- 0.36	0.94	1.39	- 2.40	- 0.63
percep	53.08	57.12	53.78	55.46	56.71	55.10	55.56	56.06	4.04	0.70	2.38	3.62	2.01	2.48	2.98
qda	67.19	67.13	66.62	66.42	66.21	65.01	65.56	67.08	- 0.06	- 0.57	- 0.77	- 0.98	- 2.18	- 1.63	- 0.11
rf	76.65	76.88	76.85	76.91	76.75	76.51	76.76	76.89	0.23	0.20	0.26	0.09	- 0.14	0.11	0.24
ridge	64.30	64.62	64.73	64.78	64.83	64.91	64.87	64.64	0.32	0.43	0.47	0.52	0.61	0.56	0.34
xgb	74.21	75.16	75.63	75.43	75.43	75.51	75.37	75.36	0.95	1.42	1.23	1.22	1.30	1.17	1.16
xnn	50.78	51.49	51.62	50.00	50.70	50.88	50.32	50.51	0.72	0.85	- 0.77	- 0.08	0.11	- 0.45	- 0.27
3.D80															
cnn	62.23	63.29	59.66	64.72	61.58	61.98	61.48	63.53	1.05	- 2.57	2.49	- 0.66	- 0.25	- 0.75	1.30
ebm	74.70	74.90	75.10	75.02	75.05	75.15	75.17	75.03	0.19	0.40	0.31	0.35	0.45	0.46	0.33
ffn	66.87	66.35	67.69	68.23	67.42	67.73	68.18	67.09	- 0.53	0.82	1.36	0.55	0.86	1.31	0.22
ffnlp	68.52	68.06	67.98	67.68	68.10	67.98	67.17	67.56	- 0.46	- 0.54	- 0.83	- 0.42	- 0.53	- 1.34	- 0.95
ffnsp	67.74	67.31	67.34	67.70	67.49	68.52	67.82	67.37	- 0.43	- 0.40	- 0.04	- 0.25	0.78	0.08	- 0.37
knn	61.05	61.04	61.03	61.04	61.04	61.04	61.05	61.05	- 0.01	- 0.02	- 0.01	- 0.01	- 0.02	- 0.00	- 0.00
lda	66.18	66.29	66.35	66.39	66.43	66.48	66.42	66.26	0.11	0.17	0.21	0.24	0.29	0.24	0.08
mlp	64.36	63.72	65.60	64.22	64.19	65.11	64.71	65.86	- 0.64	1.24	- 0.14	- 0.17	0.75	0.35	1.50
percep	49.78	53.42	52.71	51.46	51.65	53.65	53.89	50.50	3.64	2.93	1.68	1.87	3.87	4.11	0.72
qda	67.38	67.27	66.78	66.71	66.54	65.54	66.16	67.33	- 0.10	- 0.59	- 0.66	- 0.84	- 1.84	- 1.22	- 0.05
rf	76.50	76.80	76.75	76.83	76.79	76.44	76.60	76.85	0.30	0.25	0.34	0.29	- 0.06	0.10	0.36
ridge	64.89	65.01	65.13	65.23	65.27	65.33	65.27	65.04	0.12	0.23	0.33	0.38	0.44	0.38	0.15
xgb	74.18	75.14	75.64	75.37	75.50	75.59	75.35	75.50	0.96	1.46	1.19	1.32	1.41	1.17	1.32
xnn	51.39	50.00	50.03	50.48	50.00	50.00	50.82	49.76	- 1.39	- 1.37	- 0.92	- 1.40	- 1.39	- 0.58	- 1.64



(a) *AUC D100*



(b) *AUC D99*



(c) *AUC D80*

Figura 16

Comparación del promedio en 10 folds de métricas AUC en modelos de ML con el FE Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos, incluyendo créditos con apoyo gubernamental, al 100 %, 99 % y 80 % de las actividades económicas.

Nota: Las figuras muestran las variaciones (Δ) en *AUC* entre diferentes escenarios (*SFE*, *T-SFE*, *T+W-SFE*, *U-SFE*, *U+T-SFE*, *U+T+W-SFE*,*U+W-SFE*,*W-SFE*) para cada modelo de *ML* evaluado. Los colores representan rangos de variación porcentual, facilitando la identificación de mejoras o deterioros en el rendimiento.

Se identifican modelos con respuestas más variables a la inclusión de *FE*. Modelos como *MLP* y *XNN* presentan fluctuaciones considerables en *ACC* y *AUC*, con incrementos significativos en algunos escenarios y caídas en otros. Esto resalta la sensibilidad de estos modelos a la estructura específica de los datos. Un caso particular es el modelo *Perceptrón*, que muestra incrementos de hasta **8.03 %** en ciertos escenarios, pero disminuciones en otros, reflejando su alta dependencia del contexto.

Un análisis detallado por grupos de datos revela que *D100*, que contiene el conjunto más amplio de información, genera los mayores incrementos en *ACC* y *AUC*. Esto sugiere que una mayor disponibilidad de datos reduce la variabilidad y mejora la capacidad predictiva de los modelos. En contraste, el grupo *D80*, más restringido en cobertura, mantiene mejoras menores pero consistentes, destacando la robustez de modelos como *RF* y *Ridge* incluso con menor cantidad de datos.

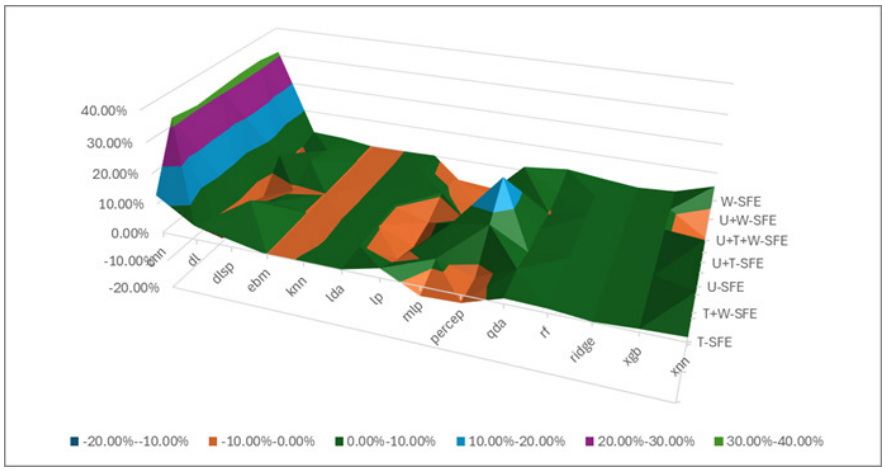
Finalmente, el análisis de los resultados derivados (*RESULTS = SCENARIOS - SFE*) confirma que la combinación de *FE* (*U+T+W*) genera los mayores incrementos en ambos indicadores, lo que evidencia que la consideración simultánea de múltiples *FE* optimiza la predicción de riesgo crediticio. Además, la inclusión de créditos con *Inc. CG* parece estabilizar el rendimiento de los modelos, particularmente en el caso de redes neuronales y modelos basados en árboles de decisión, beneficiándose de una menor volatilidad en los datos.

En conclusión, estos hallazgos resaltan la importancia de incluir múltiples *FE* y considerar la presencia de garantías gubernamentales en la construcción de modelos predictivos de riesgo crediticio. La combinación de enfoques como redes neuronales profundas y modelos basados en árboles de decisión ofrece una estrategia efectiva para mejorar la precisión y estabilidad en contextos de alta incertidumbre.

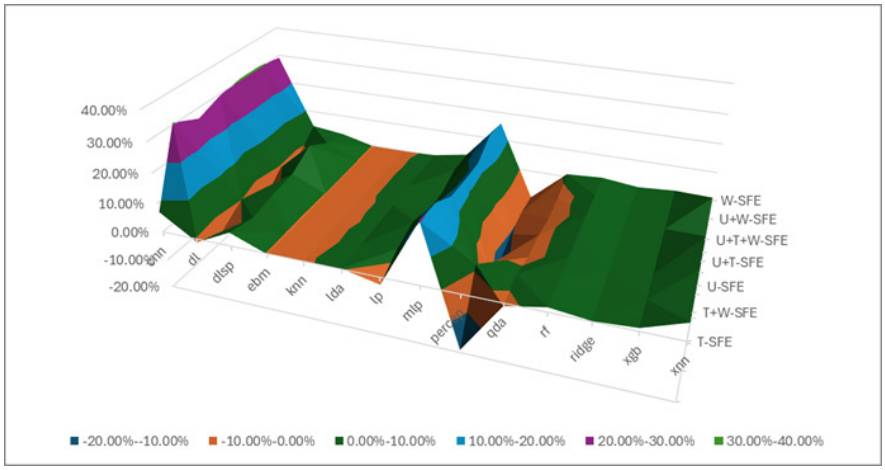
Tabla 28

FE Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos: Promedio de ACC en 10 folds de los modelos de ML según grupos de relevancia, incluyendo los créditos con apoyo gubernamental.

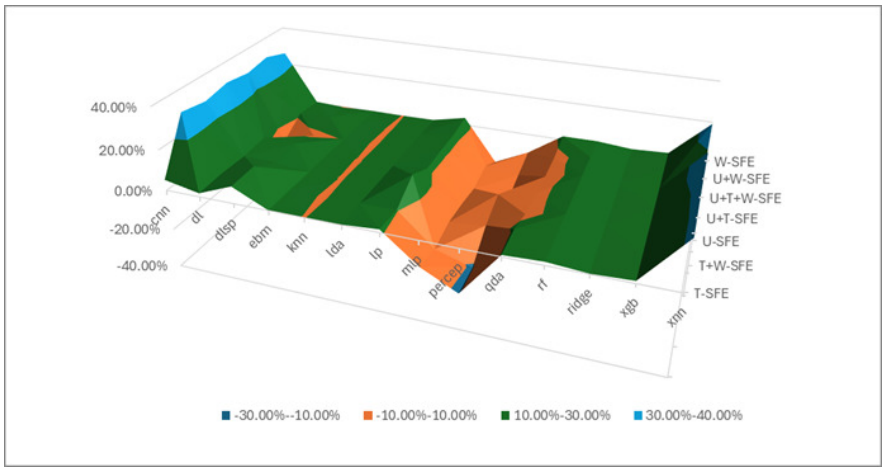
	ACC (%)														
	SCENARIOS								RESULTS						
	SFE	T	T +W	U	U +T	U +T +W	U +W	W	T -SFE	T +W -SFE	U -SFE	U +T -SFE	U+T +W -SFE	U +W -SFE	W -SFE
Inc. CG															
1.D100															
cnn	52.61	65.41	84.94	83.75	84.38	85.01	85.11	83.95	12.80	32.32	31.14	31.77	32.40	32.50	31.34
ebm	85.17	85.04	85.00	84.99	85.05	85.00	85.01	84.98	- 0.13	- 0.16	- 0.17	- 0.12	- 0.17	- 0.16	- 0.19
ffn	64.44	69.23	57.39	62.26	59.06	68.61	63.64	66.31	4.79	- 7.05	- 2.18	- 5.38	4.18	- 0.80	1.88
ffnlp	67.89	71.13	66.22	62.92	64.40	68.43	67.28	57.41	3.24	- 1.67	- 4.97	- 3.49	0.54	- 0.61	- 10.48
ffnsp	63.33	65.69	72.22	65.10	62.52	67.30	65.66	64.66	2.36	8.89	1.77	- 0.81	3.97	2.33	1.33
knn	64.28	64.32	64.26	64.28	64.27	64.26	64.25	64.26	0.04	- 0.02	0.00	- 0.01	- 0.02	- 0.03	- 0.01
lda	66.62	66.80	66.99	66.85	66.94	67.08	67.03	66.84	0.18	0.37	0.22	0.32	0.46	0.40	0.22
mlp	64.75	61.55	65.94	70.37	64.75	62.93	60.47	55.55	- 3.20	1.19	5.62	- 0.00	- 1.82	- 4.28	- 9.20
percep	47.76	45.10	45.35	55.49	64.67	46.97	55.80	46.33	- 2.66	- 2.42	7.73	16.91	- 0.79	8.03	- 1.43
qda	63.18	64.96	66.35	64.06	64.26	63.05	63.83	66.24	1.78	3.16	0.88	1.07	- 0.13	0.65	3.06
rf	83.93	85.34	85.66	85.55	85.65	85.66	85.67	85.62	1.41	1.72	1.61	1.72	1.73	1.74	1.68
ridge	65.52	65.74	65.94	65.94	66.01	66.05	66.01	65.97	0.22	0.43	0.43	0.50	0.54	0.49	0.45
xgb	83.75	84.78	84.94	84.92	84.93	84.99	85.06	85.01	1.03	1.19	1.17	1.17	1.24	1.30	1.26
xnn	69.59	71.18	76.49	74.46	77.08	56.18	62.83	74.82	1.59	6.90	4.87	7.49	- 13.41	- 6.76	5.23
2.D99															
cnn	54.37	61.27	84.98	80.98	83.86	85.05	84.98	83.42	6.90	30.61	26.61	29.48	30.68	30.61	29.05
ebm	85.18	85.01	85.00	84.96	84.99	84.99	84.98	84.98	- 0.17	- 0.18	- 0.22	- 0.19	- 0.19	- 0.20	- 0.20
ffn	64.94	63.78	60.07	65.04	57.60	64.29	62.59	69.11	- 1.16	- 4.87	0.10	- 7.34	- 0.65	- 2.35	4.16
ffnlp	63.67	61.15	65.34	67.51	64.26	64.08	66.43	66.05	- 2.52	1.67	3.83	0.59	0.41	2.76	2.37
ffnsp	60.19	64.86	62.77	64.39	65.27	69.71	65.17	62.92	4.67	2.59	4.20	5.08	9.52	4.98	2.73
knn	64.33	64.25	64.11	64.14	64.12	64.11	64.12	64.13	- 0.09	- 0.22	- 0.19	- 0.21	- 0.22	- 0.21	- 0.21
lda	66.63	66.85	66.94	66.82	66.87	67.04	66.98	66.83	0.22	0.30	0.19	0.24	0.41	0.35	0.19
mlp	50.94	71.51	70.31	65.50	64.57	64.02	68.44	67.00	20.57	19.37	14.56	13.63	13.08	17.51	16.06
percep	56.28	37.18	58.04	49.47	42.86	45.73	47.04	46.14	- 19.10	1.76	- 6.80	- 13.41	- 10.54	- 9.24	- 10.13
qda	65.63	64.66	66.72	64.47	64.64	63.38	64.14	66.64	- 0.97	1.09	- 1.17	- 0.99	- 2.25	- 1.49	1.01
rf	83.77	85.35	85.60	85.43	85.61	85.63	85.61	85.55	1.59	1.84	1.67	1.84	1.87	1.85	1.78
ridge	65.56	65.72	65.92	65.92	66.01	66.06	66.03	65.96	0.16	0.36	0.36	0.45	0.50	0.47	0.40
xgb	83.71	84.77	85.01	84.88	84.97	84.95	84.93	84.98	1.06	1.30	1.18	1.26	1.24	1.22	1.27
xnn	72.46	78.30	75.69	78.10	80.99	73.49	80.06	71.88	5.84	3.22	5.64	8.53	1.03	7.60	- 0.59
3.D80															
cnn	55.07	60.32	84.85	82.75	85.06	83.21	85.13	81.13	5.26	29.78	27.68	30.00	28.14	30.06	26.07
ebm	84.92	85.02	85.07	85.02	85.05	85.08	85.07	85.06	0.10	0.15	0.11	0.13	0.16	0.15	0.14
ffn	64.80	66.60	67.44	67.15	68.37	58.27	66.21	65.29	1.80	2.64	2.35	3.57	- 6.53	1.41	0.49
ffnlp	61.26	62.81	61.66	69.24	63.29	63.74	63.71	65.19	1.55	0.40	7.98	2.04	2.48	2.45	3.93
ffnsp	62.56	70.94	71.45	64.26	66.04	61.79	64.64	62.41	8.38	8.89	1.70	3.48	- 0.77	2.08	- 0.15
knn	64.38	64.37	64.35	64.37	64.36	64.34	64.35	64.37	- 0.01	- 0.04	- 0.02	- 0.02	- 0.04	- 0.03	- 0.01
lda	66.82	67.06	67.14	67.10	67.16	67.33	67.22	67.02	0.24	0.32	0.27	0.34	0.50	0.40	0.20
mlp	74.21	61.59	58.76	61.62	60.92	59.31	62.27	55.87	- 12.62	- 15.45	- 12.60	- 13.30	- 14.90	- 11.94	- 18.34
percep	75.06	53.34	54.05	64.33	70.05	57.29	61.29	65.05	- 21.72	- 21.01	- 10.72	- 5.01	- 17.77	- 13.77	- 10.01
qda	64.38	65.24	66.14	62.86	63.23	62.24	62.99	66.21	0.86	1.76	- 1.53	- 1.16	- 2.15	- 1.39	1.83
rf	83.56	85.23	85.57	85.47	85.59	85.58	85.56	85.48	1.67	2.01	1.92	2.04	2.02	2.01	1.93
ridge	65.63	65.95	66.14	66.14	66.19	66.36	66.22	66.13	0.31	0.50	0.50	0.56	0.73	0.59	0.50
xgb	83.57	84.75	85.03	84.90	85.01	84.99	85.04	84.96	1.17	1.46	1.32	1.44	1.42	1.46	1.38
xnn	54.64	78.16	78.20	79.41	85.19	71.12	81.61	61.07	23.52	23.56	24.77	30.55	16.48	26.97	6.43



(a) *ACC D100*



(b) *ACC D99*



(c) *ACC D80*

Figura 17

Comparación del promedio en 10 folds de métricas ACC en modelos de ML con el FE Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos, incluyendo créditos con apoyo gubernamental, al 100 %, 99 % y 80 % de las actividades económicas.

Nota: Las figuras muestran las variaciones (Δ) en *ACC* entre diferentes escenarios (*SFE*, *T-SFE*, *T+W-SFE*, *U-SFE*, *U+T-SFE*, *U+T+W-SFE*,*U+W-SFE*,*W-SFE*) para cada modelo de *ML* evaluado. Los colores representan rangos de variación porcentual, facilitando la identificación de mejoras o deterioros en el rendimiento.

Evaluación excluyendo en la data créditos que han sido beneficiados por programas gubernamentales

Las Tablas 29 y 30 presentan el desempeño de los modelos de *ML* en términos de *ACC* y *AUC* bajo escenarios que incluyen *FE* relacionados con: *T* (anomalía de temperatura), *W* (bloqueos de vías por factores climáticos) y *U* (bloqueos de vías por convulsión social). Este análisis en el escenario *Exc CG*), permitió evaluar el impacto en la predicción del riesgo crediticio. Se consideraron tres conjuntos de datos con distintos niveles de representatividad (*D100*, *D99*, *D80*).

En términos de *ACC*, los modelos basados en árboles de decisión y ensamblados, como *EBM*, *RF* y *XGB*, demuestran un desempeño sólido y consistente, con valores de precisión superiores al 87 % en la mayoría de los escenarios evaluados. Particularmente, *XGB* mantiene un rendimiento estable en todos los conjuntos de datos y escenarios, con mejoras de hasta 1.62 % en *D80* bajo el escenario combinado (*U+T+W*). Este comportamiento reafirma su capacidad para procesar múltiples *FE* y conservar una alta precisión predictiva (ver Figura 18).

En cuanto a las redes neuronales, *CNN* sigue destacándose con mejoras significativas en *ACC*, alcanzando hasta **31.04 %** en el escenario combinado (*U+T+W*) en el conjunto *D100*. No obstante, otros modelos neuronales, como *XNN* y *MLP*, muestran una mayor sensibilidad a la exclusión de créditos respaldados por el gobierno, con variaciones notables en los escenarios combinados. En particular, *XNN* experimenta reducciones en *ACC* de hasta **15 %** en algunos escenarios, evidenciando la dependencia de estos modelos a la estructura del conjunto de datos (ver Figura 19).

Los modelos clásicos, como *LDA* y *Ridge*, mantienen un desempeño estable, con cambios mínimos en *ACC* ($\pm 0{,}2\,\%$), logrando valores consistentes de aproximadamente 67 %. Estos modelos, aunque no presentan mejoras drásticas, demuestran ser alternativas confiables cuando se busca estabilidad en escenarios de alta variabilidad.

En términos de *AUC*, *RF* y *XGB* destacan por su consistencia y precisión. *XGB* logra incrementos significativos en escenarios combinados (*U+T+W*), reflejando su capacidad para capturar relaciones complejas en la predicción del riesgo crediticio. Además, el modelo *RF* muestra estabilidad en escenarios con múltiples *FE*, alcanzando aumentos de hasta **1.93 % en *ACC*** y **0.41 % en *AUC*** en *D99*, lo que confirma su capacidad para adaptarse a entornos con datos más heterogéneos (ver Figura 18).

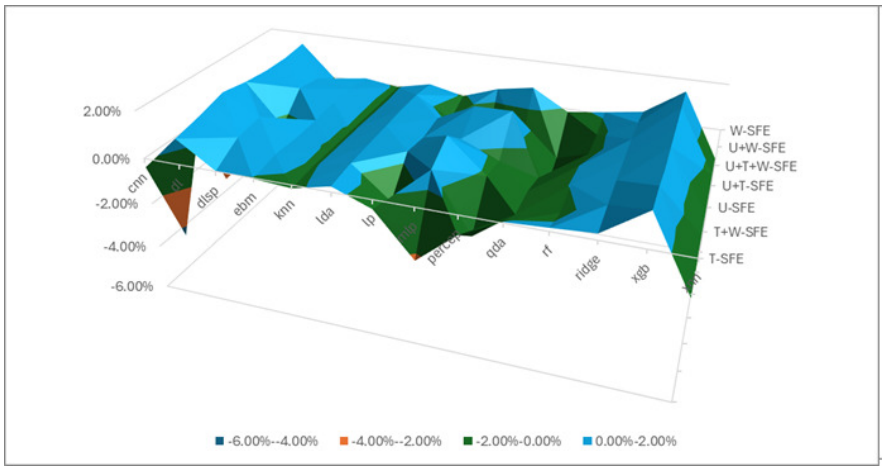
Un hallazgo clave es la variabilidad en el desempeño de los modelos basados en perceptrones. En particular, el modelo *Perceptrón* muestra caídas significativas en *ACC* en escenarios como *U+T-SFE*, donde su precisión disminuye hasta **46.38 %** (-2.46 % en comparación con datasets que incluyen créditos respaldados por el gobierno). Esto sugiere que ciertos modelos pueden experimentar una disminución en su capacidad predictiva cuando se eliminan datos de respaldo gubernamental, lo que introduce un

sesgo que debe ser considerado en el diseño de modelos predictivos (ver Figura 19).

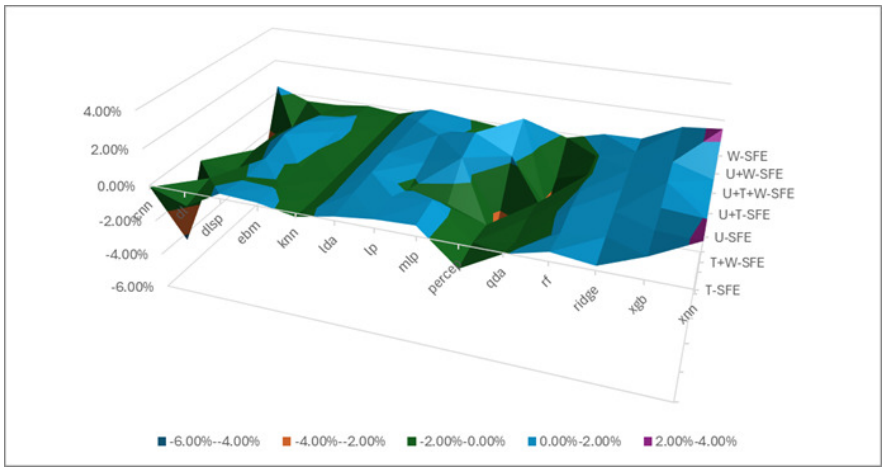
Tabla 29

FE Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos: Promedio de AUC en 10 folds de los modelos de ML según grupos de relevancia, excluyendo los créditos con apoyo gubernamental.

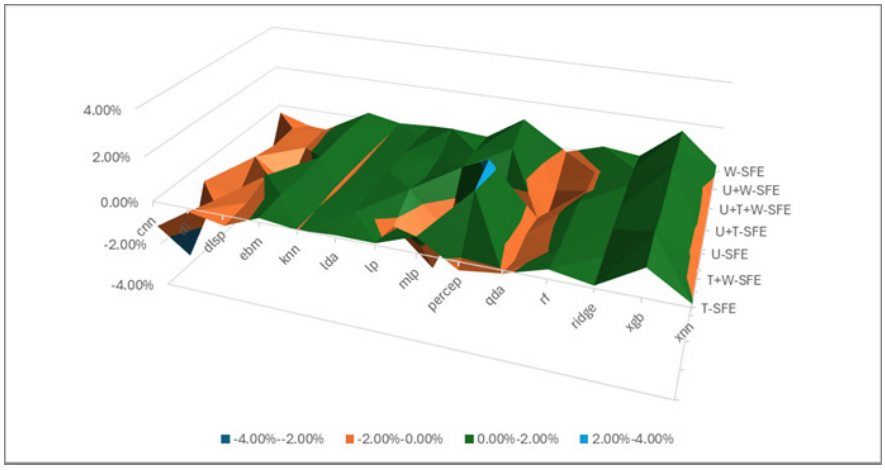
	AUC (%)														
	SCENARIOS								RESULTS						
	SFE	T	T +W	U	U +T	U +T +W	U +W	W	T -SFE	T +W -SFE	U -SFE	U +T -SFE	U+T +W -SFE	U +W -SFE	W -SFE
Exc. CG															
1.D100															
cnn	62.35	61.96	57.87	62.34	58.89	59.00	58.34	61.89	- 0.39	- 4.48	- 0.01	- 3.46	- 3.35	- 4.00	- 0.46
ebm	73.73	73.71	73.92	73.87	73.91	74.03	74.02	73.89	- 0.03	0.19	0.13	0.18	0.29	0.29	0.15
ffn	65.16	66.25	66.50	66.81	66.61	66.50	66.49	66.62	1.09	1.34	1.65	1.45	1.35	1.34	1.46
ffnlp	66.45	66.06	65.46	66.77	66.60	66.74	65.99	66.40	- 0.39	- 0.99	0.32	0.14	0.28	- 0.46	- 0.06
ffnsp	66.26	66.27	66.90	67.05	65.92	67.00	66.82	66.19	0.01	0.65	0.79	- 0.34	0.74	0.57	- 0.07
knn	58.68	58.54	58.63	58.62	58.62	58.61	58.60	58.63	- 0.14	- 0.05	- 0.06	- 0.06	- 0.07	- 0.08	- 0.05
lda	64.60	64.84	64.96	65.00	65.02	65.06	65.04	64.89	0.24	0.35	0.39	0.41	0.46	0.43	0.29
mlp	63.72	61.49	64.36	64.96	64.06	64.77	61.77	64.21	- 2.23	0.64	1.24	0.34	1.05	- 1.95	0.49
percep	55.32	54.51	53.47	55.14	54.96	56.40	53.95	56.10	- 0.81	- 1.85	- 0.18	- 0.36	1.07	- 1.37	0.77
qda	66.35	66.39	65.74	65.80	65.52	64.16	64.76	66.21	0.04	- 0.60	- 0.54	- 0.83	- 2.19	- 1.58	- 0.14
rf	74.68	74.83	74.47	74.72	74.70	74.38	74.47	74.77	0.14	- 0.22	0.03	0.01	- 0.31	- 0.22	0.09
ridge	63.40	63.65	63.80	63.90	63.93	63.98	63.96	63.74	0.26	0.41	0.50	0.53	0.59	0.56	0.34
xgb	72.53	73.97	73.97	73.72	73.83	73.90	74.02	74.01	1.44	1.44	1.18	1.30	1.37	1.48	1.48
xnn	51.66	50.13	49.93	50.52	50.55	51.25	49.96	50.32	- 1.54	- 1.73	- 1.15	- 1.11	- 0.41	- 1.71	- 1.34
2.D99															
cnn	62.49	62.40	58.08	61.81	59.52	59.19	59.23	62.85	- 0.10	- 4.41	- 0.68	- 2.97	- 3.31	- 3.26	0.35
ebm	73.68	73.85	73.64	73.58	73.68	73.80	73.70	73.60	0.17	- 0.04	- 0.10	- 0.00	0.12	0.03	- 0.08
ffn	66.85	66.06	66.54	65.76	66.16	66.09	65.66	66.50	- 0.79	- 0.31	- 1.09	- 0.69	- 0.76	- 1.19	- 0.35
ffnlp	66.07	66.53	66.58	65.89	66.35	66.94	66.20	66.11	0.47	0.51	- 0.18	0.28	0.87	0.13	0.04
ffnsp	66.51	66.78	66.49	66.63	67.09	66.90	66.45	66.19	0.28	- 0.02	0.13	0.58	0.39	- 0.06	- 0.31
knn	58.65	58.45	58.35	58.36	58.35	58.35	58.35	58.36	- 0.20	- 0.29	- 0.28	- 0.30	- 0.30	- 0.29	- 0.29
lda	64.59	64.80	64.94	64.98	65.00	65.04	65.02	64.87	0.22	0.35	0.39	0.41	0.46	0.43	0.28
mlp	63.49	64.05	64.20	62.84	64.36	64.80	62.98	63.31	0.55	0.71	- 0.65	0.86	1.30	- 0.52	- 0.18
percep	53.15	51.87	52.01	51.72	50.70	53.44	54.65	53.33	- 1.28	- 1.14	- 1.43	- 2.45	0.29	1.50	0.18
qda	66.34	66.28	65.63	65.67	65.40	64.02	64.62	66.09	- 0.06	- 0.71	- 0.67	- 0.94	- 2.32	- 1.72	- 0.25
rf	74.54	75.02	74.71	74.83	74.61	74.47	74.49	74.83	0.48	0.17	0.29	0.08	- 0.07	- 0.05	0.29
ridge	63.38	63.61	63.78	63.89	63.92	63.96	63.94	63.71	0.23	0.40	0.51	0.54	0.58	0.56	0.34
xgb	72.60	73.73	73.74	73.67	73.89	73.92	73.91	73.93	1.13	1.14	1.08	1.29	1.32	1.31	1.33
xnn	49.59	51.89	51.69	50.47	50.00	50.29	52.04	50.65	2.29	2.10	0.88	0.41	0.70	2.45	1.05
3.D80															
cnn	61.73	60.58	58.10	60.91	59.23	58.89	58.88	61.34	- 1.14	- 3.62	- 0.82	- 2.50	- 2.84	- 2.85	- 0.39
ebm	73.49	73.69	73.81	73.74	73.85	73.89	73.89	73.80	0.19	0.32	0.25	0.36	0.40	0.40	0.31
ffn	66.74	66.58	66.16	65.41	66.59	66.38	66.47	65.86	- 0.16	- 0.58	- 1.32	- 0.15	- 0.36	- 0.26	- 0.88
ffnlp	66.29	66.37	66.13	66.85	66.49	66.61	67.25	65.60	0.09	- 0.16	0.57	0.21	0.32	0.96	- 0.69
ffnsp	66.90	66.42	66.55	66.96	65.59	66.67	66.76	65.94	- 0.48	- 0.36	0.06	- 1.31	- 0.24	- 0.14	- 0.97
knn	58.56	58.55	58.56	58.54	58.54	58.55	58.55	58.57	- 0.01	0.00	- 0.02	- 0.02	- 0.01	- 0.01	0.01
lda	65.11	65.21	65.28	65.35	65.37	65.40	65.37	65.22	0.09	0.17	0.23	0.26	0.29	0.26	0.11
mlp	64.27	64.98	62.54	63.20	65.27	65.21	65.10	63.60	0.71	- 1.73	- 1.07	1.00	0.94	0.83	- 0.67
percep	51.76	51.39	53.77	54.19	50.55	53.06	53.66	52.41	- 0.37	2.01	2.43	- 1.21	1.31	1.90	0.65
qda	66.49	66.36	65.84	66.02	65.75	64.61	65.25	66.33	- 0.14	- 0.65	- 0.48	- 0.74	- 1.88	- 1.25	- 0.16
rf	74.30	74.70	74.49	74.58	74.60	74.21	74.22	74.66	0.41	0.20	0.29	0.31	- 0.09	- 0.07	0.37
ridge	63.90	64.01	64.14	64.27	64.31	64.35	64.30	64.07	0.11	0.24	0.37	0.41	0.45	0.40	0.17
xgb	72.43	73.64	74.01	73.71	73.76	74.05	73.86	74.05	1.21	1.58	1.28	1.33	1.62	1.43	1.62
xnn	50.40	50.55	50.00	49.97	50.00	50.00	49.94	50.66	0.15	- 0.40	- 0.43	- 0.40	- 0.40	- 0.47	0.25



(a) *AUC D100*



(b) *AUC D99*



(c) *AUC D80*

Figura 18

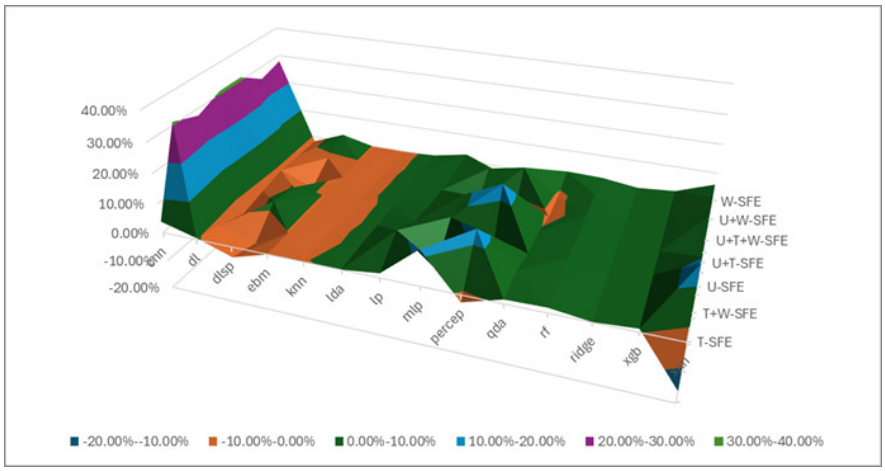
Comparación del promedio en 10 folds de métricas AUC en modelos de ML con el FE Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos, excluyendo créditos con apoyo gubernamental, al 100 %, 99 % y 80 % de las actividades económicas.

Nota: Las figuras muestran las variaciones (Δ) en *AUC* entre diferentes escenarios (*SFE*, *T-SFE*, *T+W-SFE*, *U-SFE*, *U+T-SFE*, *U+T+W-SFE*,*U+W-SFE*,*W-SFE*) para cada modelo de *ML* evaluado. Los colores representan rangos de variación porcentual, facilitando la identificación de mejoras o deterioros en el rendimiento.

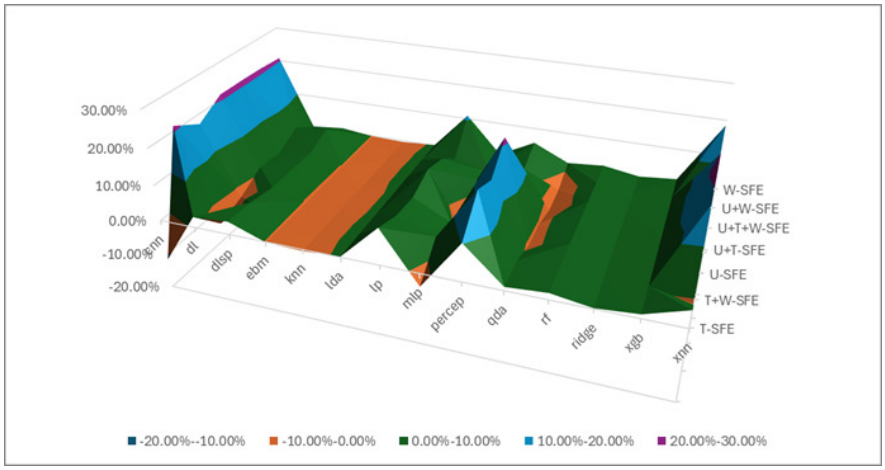
Tabla 30

FE Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos: Promedio de ACC en 10 folds de los modelos de ML según grupos de relevancia, excluyendo los créditos con apoyo gubernamental.

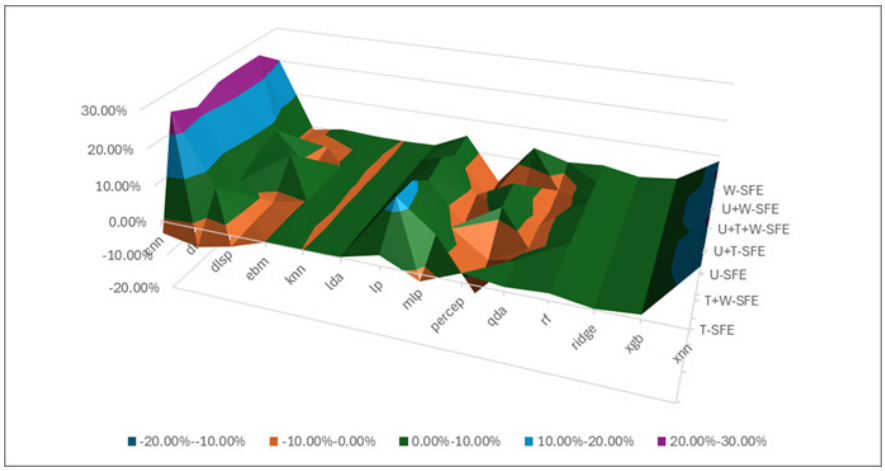
	ACC (%)														
	SCENARIOS								RESULTS						
	SFE	T	T +W	U	U +T	U +T +W	U +W	W	T -SFE	T +W -SFE	U -SFE	U +T -SFE	U+T +W -SFE	U +W -SFE	W -SFE
Exc. CG															
1.D100															
cnn	56.63	60.47	87.65	84.37	87.69	87.67	82.41	84.51	3.84	31.02	27.74	31.06	31.04	25.78	27.88
ebm	87.78	87.69	87.65	87.63	87.64	87.64	87.65	87.65	- 0.09	- 0.13	- 0.15	- 0.14	- 0.14	- 0.13	- 0.13
ffn	70.07	70.60	67.96	64.99	66.82	68.50	68.26	68.25	0.53	- 2.11	- 5.08	- 3.24	- 1.56	- 1.81	- 1.82
ffnlp	64.10	65.67	72.32	67.66	66.18	65.77	65.80	66.70	1.57	8.22	3.56	2.08	1.66	1.69	2.59
ffnsp	67.21	63.53	60.93	69.04	67.53	60.97	66.57	69.85	- 3.68	- 6.28	1.82	0.31	- 6.24	- 0.64	2.63
knn	65.24	65.20	64.98	65.00	65.00	64.95	64.96	64.99	- 0.03	- 0.26	- 0.24	- 0.24	- 0.28	- 0.27	- 0.24
lda	67.24	67.31	67.52	67.43	67.45	67.64	67.55	67.49	0.07	0.28	0.19	0.21	0.40	0.30	0.25
mlp	58.58	70.53	56.07	66.44	67.42	56.94	63.93	56.47	11.95	- 2.50	7.87	8.84	- 1.63	5.36	- 2.11
percep	48.84	46.38	62.42	36.78	63.48	40.90	56.64	42.35	- 2.46	13.58	- 12.06	14.64	- 7.94	7.80	- 6.49
qda	62.81	64.24	65.37	63.18	63.41	61.37	62.09	65.05	1.43	2.55	0.36	0.59	- 1.44	- 0.73	2.23
rf	85.72	87.44	87.63	87.53	87.68	87.67	87.65	87.63	1.73	1.91	1.82	1.96	1.95	1.93	1.91
ridge	66.18	66.31	66.60	66.58	66.67	66.77	66.63	66.50	0.13	0.43	0.40	0.49	0.59	0.45	0.33
xgb	86.18	87.29	87.50	87.41	87.49	87.49	87.49	87.51	1.11	1.32	1.23	1.31	1.30	1.31	1.32
xnn	72.01	56.05	87.57	81.71	77.36	80.53	79.95	77.88	- 15.96	15.56	9.70	5.36	8.52	7.94	5.88
2.D99															
cnn	66.29	54.66	87.59	83.57	87.67	87.70	87.53	87.04	- 11.63	21.30	17.28	21.38	21.41	21.24	20.76
ebm	87.82	87.69	87.62	87.59	87.62	87.64	87.62	87.63	- 0.13	- 0.20	- 0.22	- 0.20	- 0.18	- 0.20	- 0.19
ffn	67.13	70.09	62.29	64.60	68.45	67.81	71.27	67.15	2.96	- 4.84	- 2.53	1.32	0.68	4.14	0.02
ffnlp	58.27	68.32	65.68	65.02	67.64	65.62	60.54	69.23	10.05	7.41	6.76	9.37	7.35	2.27	10.96
ffnsp	64.10	67.73	67.50	66.28	66.38	66.31	69.18	65.03	3.64	3.40	2.18	2.29	2.21	5.08	0.93
knn	65.38	64.92	64.89	64.89	64.89	64.87	64.89	64.90	- 0.45	- 0.49	- 0.48	- 0.49	- 0.50	- 0.49	- 0.47
lda	67.24	67.40	67.50	67.42	67.41	67.54	67.50	67.40	0.16	0.26	0.18	0.17	0.30	0.26	0.16
mlp	64.91	61.26	68.51	66.46	58.05	67.96	64.92	65.38	- 3.65	3.59	1.55	- 6.86	3.04	0.01	0.46
percep	51.55	61.77	69.54	66.65	72.93	60.93	53.92	57.20	10.22	17.99	15.10	21.38	9.38	2.37	5.64
qda	63.27	64.28	64.73	62.34	62.54	60.49	61.20	64.28	1.01	1.46	- 0.93	- 0.73	- 2.77	- 2.07	1.02
rf	85.73	87.51	87.64	87.59	87.66	87.68	87.66	87.62	1.78	1.92	1.86	1.93	1.96	1.94	1.89
ridge	66.16	66.38	66.61	66.55	66.64	66.74	66.66	66.44	0.22	0.45	0.39	0.48	0.58	0.50	0.28
xgb	86.20	87.32	87.48	87.38	87.52	87.51	87.42	87.46	1.12	1.27	1.18	1.32	1.31	1.21	1.26
xnn	63.78	68.47	62.79	81.82	87.40	84.20	64.94	82.32	4.69	- 0.99	18.04	23.62	20.43	1.16	18.54
3.D80															
cnn	62.36	58.85	87.57	84.52	87.42	87.68	87.71	82.55	- 3.50	25.22	22.16	25.06	25.32	25.35	20.20
ebm	87.51	87.58	87.64	87.62	87.62	87.65	87.63	87.64	0.07	0.13	0.10	0.11	0.13	0.12	0.12
ffn	63.83	58.29	69.62	66.20	69.87	71.70	62.72	62.91	- 5.54	5.79	2.37	6.04	7.87	- 1.12	- 0.92
ffnlp	58.46	61.38	71.01	70.00	64.74	66.28	62.12	63.30	2.92	12.55	11.54	6.28	7.82	3.66	4.84
ffnsp	65.89	62.61	63.44	63.93	72.89	62.83	65.32	66.60	- 3.28	- 2.45	- 1.96	7.00	- 3.06	- 0.57	0.71
knn	65.07	65.06	65.03	65.05	65.04	65.02	65.03	65.06	- 0.00	- 0.03	- 0.02	- 0.03	- 0.05	- 0.03	- 0.01
lda	67.33	67.59	67.76	67.58	67.69	67.85	67.74	67.56	0.26	0.43	0.25	0.36	0.52	0.41	0.23
mlp	67.93	65.93	72.31	68.12	63.21	66.36	59.38	59.67	- 2.00	4.38	0.19	- 4.72	- 1.57	- 8.55	- 8.26
percep	60.08	62.61	50.07	53.04	62.32	63.72	53.02	64.28	2.53	- 10.01	- 7.03	2.24	3.64	- 7.06	4.20
qda	62.47	63.65	64.27	60.94	61.33	59.35	60.28	63.88	1.18	1.81	- 1.53	- 1.14	- 3.12	- 2.19	1.41
rf	85.45	87.36	87.68	87.58	87.70	87.70	87.65	87.60	1.91	2.22	2.13	2.25	2.25	2.20	2.14
ridge	66.32	66.58	66.73	66.72	66.84	66.93	66.78	66.64	0.26	0.41	0.40	0.52	0.61	0.46	0.31
xgb	85.95	87.21	87.51	87.38	87.46	87.51	87.49	87.47	1.26	1.55	1.42	1.51	1.56	1.54	1.52
xnn	66.33	81.92	87.74	80.03	87.74	80.19	79.88	69.28	15.59	21.41	13.69	21.41	13.86	13.55	2.95



(a) *ACC D100*



(b) *ACC D99*



(c) *ACC D80*

Figura 19

Comparación del promedio en 10 folds de métricas ACC en modelos de ML con el FE Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos, excluyendo créditos con apoyo gubernamental, al 100 %, 99 % y 80 % de las actividades económicas.

Nota: Las figuras muestran las variaciones (Δ) en *ACC* entre diferentes escenarios (*SFE*, *T-SFE*, *T+W-SFE*, *U-SFE*, *U+T-SFE*, *U+T+W-SFE*,*U+W-SFE*,*W-SFE*) para cada modelo de *ML* evaluado. Los colores representan rangos de variación porcentual, facilitando la identificación de mejoras o deterioros en el rendimiento.

Finalmente, los resultados derivados (*RESULTS = SCENARIOS - SFE*) confirman que los escenarios con múltiples *FE* (*U+T+W*) generan los mayores incrementos en precisión y capacidad discriminativa de los modelos. Asimismo, el análisis revela que la exclusión de créditos con respaldo gubernamental afecta significativamente el desempeño predictivo de algunos modelos, en particular de redes neuronales como *XNN* y *MLP*, mientras que modelos basados en árboles de decisión (*RF*, *XGB*) muestran mayor resiliencia ante la omisión de estos datos.
En conclusión, estos hallazgos refuerzan la importancia de considerar la presencia o exclusión de créditos con respaldo gubernamental al construir modelos de predicción de riesgo crediticio. Los modelos basados en árboles de decisión se destacan como opciones confiables y estables en la mayoría de los escenarios, mientras que las redes neuronales requieren ajustes más precisos para evitar pérdidas significativas de desempeño ante la variabilidad en los datos.

Evaluación del nivel de significancia de los modelos de *ML* para *AUC* y *ACC* aplicando el *FE* Convulsión Social, Anomalía de Temperatura y Bloqueos por Factores Climáticos, y considerando la inclusión y exclusión de los créditos con apoyo de programas gubernamentales.

Las Tablas 31 y 32 presentan los modelos de *ML* que exhibieron mejoras estadísticamente significativas en los indicadores *ACC* y *AUC* al comparar escenarios sin *FE* (*SFE*) con aquellos que incluyen *FE* relacionados con: *T* (anomalía de temperatura), *W* (bloqueos de vías por factores climáticos) y *U* (bloqueos de vías por convulsión social), así como sus combinaciones. Los análisis se realizaron considerando tanto los escenarios que incluyen créditos *Inc CG* y *Exc CG*.
Impacto de la inclusión de créditos *Inc CG* Los resultados indican que la inclusión de créditos con respaldo gubernamental enriquece el contexto de aprendizaje de algunos modelos, mejorando su capacidad predictiva. En particular, los modelos basados en árboles de decisión (*XGB*, *RF*) y redes neuronales (*CNN*) se destacan por su desempeño robusto en todos los grupos de relevancia (*D100*, *D99*, *D80*). Por ejemplo, *XGB* logra significancia estadística en todas las combinaciones evaluadas, reafirmando su capacidad para adaptarse a escenarios con múltiples *FE*.
La inclusión de *FE* climáticos y sociales (*U+T+W*) proporciona los mayores incrementos en precisión y capacidad discriminativa. En este sentido, los modelos *CNN* y *RF* se benefician significativamente de esta combinación, reflejando su capacidad para capturar relaciones complejas entre las variables del entorno y la morosidad crediticia.
Impacto de la exclusión de créditos con garantía gubernamental (*Exc CG*) Por otro lado, excluir los créditos con respaldo gubernamental afecta de manera variable el rendimiento de los modelos. Se observa que modelos como *Perceptrón* y *DL* presentan una menor cantidad de combinaciones significativas en *D99* y *D80*, lo que sugiere

que la exclusión de estos datos limita la capacidad de ciertos modelos para identificar patrones clave en la predicción del riesgo crediticio.

Tabla 31

FE Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos: Modelos que mostraron una mejora significativa en los indicadores, agrupados por relevancia, incluyendo los créditos con apoyo gubernamental.

Inc. CG															
	AUC							ACC							TOTAL
	T	T+W	U	U+T	U+T+W	U+W	W	T	T+W	U	U+T	U+T+W	U+W	W	
1.D100															
cnn	Y	-	Y	-	-	-	Y	Y	Y	Y	Y	Y	Y	Y	10
ffn	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	1
ffnsp	-	-	Y	-	-	Y	-	-	Y	-	-	-	-	-	3
lda	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
percep	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	1
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	Y	3
rf	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	7
ridge	-	-	-	-	-	-	-	-	-	-	Y	Y	Y	-	3
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xnn	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	1
2.D99															
cnn	-	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	6
ffnlp	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
ffnsp	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	1
lda	-	-	-	-	-	-	-	-	-	-	-	Y	Y	-	2
mlp	-	-	-	-	Y	-	-	Y	Y	-	-	-	Y	Y	5
percep	Y	-	-	Y	-	Y	-	-	-	-	-	-	-	-	3
qda	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
rf	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	7
ridge	-	-	-	-	Y	-	-	-	Y	Y	Y	Y	Y	Y	7
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xnn	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
3.D80															
cnn	-	-	Y	-	-	-	-	-	Y	Y	Y	Y	Y	Y	7
ebm	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
ffn	-	-	Y	-	-	Y	-	-	-	-	-	-	-	-	2
ffnlp	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	1
ffnsp	-	-	-	-	-	-	-	Y	Y	-	-	-	-	-	2
lda	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
mlp	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
percep	Y	-	-	-	Y	-	-	-	-	-	-	-	-	-	2
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	Y	3
rf	Y	Y	Y	Y	-	-	Y	Y	Y	Y	Y	Y	Y	Y	12
ridge	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xnn	-	-	-	-	-	-	-	-	Y	Y	Y	-	Y	-	4

Sin embargo, modelos más robustos como *XGB*, *RF* y *Ridge* logran mantener un desempeño consistente aún en ausencia de estos créditos, lo que indica que pueden generalizar mejor a escenarios con menor variabilidad en los datos. Específicamente, *XGB* mantiene significancia estadística en todas las combinaciones de *FE*, demostrando su adaptabilidad a conjuntos de datos más homogéneos.

Análisis por tipo de modelo

- Modelos basados en árboles de decisión (*XGB*, *RF*): Se consolidan como las

opciones más confiables en ambos experimentos (*Inc CG* y *Exc CG*). En particular, *XGB* logra mejoras estadísticamente significativas en todos los escenarios evaluados, reforzando su versatilidad ante distintos niveles de representación de datos.

- Redes neuronales (*CNN*): *CNN* demuestra una alta sensibilidad a la inclusión de *FE*, logrando mejoras significativas en múltiples combinaciones de *T*, *U* y *W*. Su rendimiento en escenarios sin respaldo gubernamental (*Exc CG*) disminuye ligeramente, lo que indica que puede beneficiarse de la mayor diversidad en los datos proporcionada por los créditos con garantía gubernamental.

- Modelos clásicos (*LDA*, *Ridge*): *Ridge* muestra mejoras significativas en los escenarios que incluyen factores climáticos y sociales (*U+T+W*), lo que indica su estabilidad en conjuntos de datos enriquecidos. En contraste, *LDA* presenta mejoras limitadas, principalmente en escenarios específicos como *T+W*.

- Modelos con mejoras específicas (*EBM*, *FFN*, *FFNLP*): Estos modelos muestran mejoras en escenarios seleccionados. En particular, *EBM* logra mejoras estadísticamente significativas en combinaciones de factores de convulsión social y cambio climático, lo que sugiere que su rendimiento está altamente influenciado por la composición del conjunto de datos.

Consideraciones finales En general, los modelos basados en árboles de decisión (*XGB*, *RF*) y redes neuronales (*CNN*) se destacan como las opciones más robustas en la predicción del riesgo crediticio en contextos de convulsión social y cambio climático. La inclusión de créditos con respaldo gubernamental proporciona un mayor contexto y variabilidad en los datos, lo que favorece a estos modelos al capturar patrones más complejos.

No obstante, la exclusión de estos créditos permite evaluar el impacto de los *FE* de manera más homogénea, aunque algunos modelos como *DL* y *Perceptrón* experimentan una degradación en su rendimiento. La elección entre incluir o excluir estos créditos dependió del objetivo del análisis: si se busca maximizar la capacidad predictiva, los datos con respaldo gubernamental pueden ser beneficiosos; si se pretende evaluar el impacto específico de los *FE* sin sesgos por políticas gubernamentales, la exclusión puede ser más adecuada.

En conclusión, estos hallazgos resaltan la importancia de integrar *FE* en los modelos predictivos, particularmente en escenarios con alta incertidumbre. Los modelos de aprendizaje profundo y redes neuronales pueden beneficiarse significativamente de conjuntos de datos enriquecidos, mientras que los modelos basados en árboles de decisión ofrecen un rendimiento estable y confiable incluso en escenarios con menor variabilidad de datos.

Tabla 32

FE Convulsión Social, Anomalía de Temperatura y Bloqueos de vías por Factores Climáticos: Modelos que mostraron una mejora significativa en los indicadores, agrupados por relevancia, excluyendo los créditos con apoyo gubernamental.

	Exc. CG														
	AUC							ACC							TOTAL
	T	T+W	U	U+T	U+T+W	U+W	W	T	T+W	U	U+T	U+T+W	U+W	W	
1.D100															
cnn	-	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	6
ffn	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	1
ffnsp	-	Y	Y	-	-	-	-	-	-	-	-	-	-	-	2
lda	-	Y	Y	Y	Y	Y	-	-	-	-	-	Y	-	-	6
percep	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	Y	3
rf	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	7
ridge	-	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	12
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xnn	-	-	-	-	-	-	-	-	Y	-	-	-	-	-	1
2.D99															
cnn	-	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	6
ffnlp	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	-	Y	6
ffnsp	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
lda	-	-	-	-	-	-	-	-	Y	-	-	Y	-	-	2
mlp	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
percep	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	1
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	Y	3
rf	Y	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	8
ridge	-	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	-	5
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xnn	-	Y	-	-	-	Y	-	-	-	Y	Y	-	-	-	4
3.D80															
cnn	-	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	6
ebm	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
ffn	-	-	-	-	-	-	-	-	Y	-	-	Y	-	-	2
ffnlp	-	-	-	-	-	Y	-	-	Y	Y	-	-	-	-	3
ffnsp	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
lda	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
mlp	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	1
percep	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	Y	3
rf	Y	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	Y	9
ridge	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xnn	-	-	-	-	-	-	-	-	Y	-	Y	-	-	-	2

5.2. Respuesta a preguntas de investigación.

A continuación, se procede a abordar las preguntas de investigación planteadas, fundamentando las respuestas en los experimentos realizados. Dado el amplio volumen de información presentada, el análisis posterior se centra en los resultados más significativos, con referencias directas a las tablas pertinentes para destacar los hallazgos críticos y sus implicaciones para el modelado predictivo en diversos escenarios.

5.2.1. ¿Cómo influyó la pandemia de *COVID-19* , a través de factores como los casos positivos y las tasas de mortalidad, en los impagos de créditos en diferentes actividades económicas?

La relación entre los días de mora y la pandemia por *COVID-19* se analiza desde diversas perspectivas, considerando tanto el impacto directo de los casos positivos y las tasas de mortalidad, como el efecto diferido de la mortalidad al desplazar la serie temporal un periodo hacia la derecha *COVID MOV*. A partir de los resultados obtenidos en los escenarios modelados, los hallazgos clave pueden resumirse en los siguientes puntos:

- **Impacto diferido de la mortalidad frente a los casos positivos:** Incorporar un desplazamiento temporal en la serie de fallecidos *COVID MOV* resultó en mejoras significativas en el rendimiento de los modelos de *ML*, particularmente en *ACC* y *AUC*. Esto indica que los efectos de la pandemia en la mora crediticia no se manifiestan de manera inmediata con los casos positivos, sino que están más fuertemente correlacionados con la mortalidad. Los factores económicos derivados de la pandemia—como la reducción de ingresos familiares, el desempleo y los costos médicos—tienen un efecto retardado en la capacidad de pago de los deudores, lo que se refleja en un aumento de la morosidad en los meses posteriores a los picos de mortalidad.

Tabla 33
Modelos más relevantes agrupados por Inc. CG y Exc. CG, destacando los mejores valores en los escenarios analizados (P, F, F+P) y la mejora alcanzada en comparación con SFE.

Modelo	Escenario	Relevancia	Métricas		Mejoras	
			ACC (%)	*AUC* (%)	*ACC* (%)	*AUC* (%)
Inc. CG						
CNN	F+P (Desplazado)	Alta	73.19	74.25	3.60	1.61
DL	P (Desplazado)	Alta	73.31	73.45	3.55	1.79
MLP	F+P (Desplazado)	Media	70.36	71.62	2.31	2.31
LDA	P+F	Media	64.12	68.56	1.03	1.33
Exc. CG						
CNN	F+P (Desplazado)	Alta	73.58	74.61	7.16	1.30
DL	F (Desplazado)	Alta	68.57	73.19	4.49	1.53
MLP	P (Desplazado)	Media	67.96	70.63	7.76	1.53
XGB	F+P	Alta	76.21	80.65	0.59	0.59

Efectos diferenciados según el escenario: Los resultados varían significativamente entre los escenarios evaluados. Por ejemplo, los modelos analizados en el escenario *F+P* presentan mejoras considerables en *ACC* y *AUC*, especialmente en los datos desplazados '*F+P (Desplazado)*' como se muestra en la Tabla 33.

Esto indica que la combinación de factores como los apoyos gubernamentales y la mortalidad asociada al *COVID-19* tiene un papel relevante en la predicción de la mora crediticia. Por otro lado, los escenarios basados únicamente en *F* (fallecidos) también muestran una influencia significativa, aunque menor en comparación con los escenarios combinados.

- **Significancia estadística:** El análisis estadístico de los resultados muestra que los modelos *CNN*, *DL* y *MLP* destacan como los de mejor desempeño en términos de mejoras significativas en *ACC* y *AUC*. La significancia estadística se evaluó mediante pruebas de hipótesis unilaterales para determinar si las mejoras observadas en los escenarios con *FE* eran significativamente mayores en comparación con el escenario base *SFE*. Los resultados confirman que los escenarios desplazados tienen un impacto significativo con un nivel de confianza del 95 % luego de aplicar *t-student*, especialmente en el modelo *CNN*, donde se observan incrementos de hasta 3.60 % en *ACC* y 1.61 % en *AUC* para experimentos de *Inc. CG*, y de hasta 7.16 % en *ACC* y 1.30 % en *AUC* para experimentos *Exc. CG* (ver Tabla 33).

- **Variabilidad según la concentración de datos (1.D100, 2.D99, 3.D80)**
  Los resultados de los modelos se analizaron en tres subconjuntos de datos:

  - **1.D100 (todas las actividades económicas)**: Se observa una heterogeneidad en los patrones de mora. Mientras que el impacto de la mortalidad es significativo en sectores de contacto directo, la variabilidad entre industrias atenúa su efecto en el conjunto global.
  - **2.D99 (99 % de los créditos)**: Se reduce la variabilidad al excluir sectores marginales, observándose picos de morosidad relacionados con la mortalidad en sectores específicos.
  - **3.D80 (80 % de los créditos más representativos)**: Aquí se aprecia el impacto más pronunciado de la mortalidad en la mora crediticia. Las actividades con mayor contacto interpersonal presentan las mayores tasas de impago en los meses posteriores a los picos de fallecidos.

  En general, la mejora en los modelos con la inclusión de *COVID MOV* es más evidente en **3.D80**, seguida de **2.D99**, mientras que en **1.D100** la relación es más difusa debido a la mayor variabilidad.

- **Influencia de las garantías crediticias gubernamentales**
  Se evaluaron escenarios incluyendo créditos con *Inc. CG* y *Exc. CG* créditos respaldados por el gobierno. Los resultados indican que:

- Bajo *Inc. CG*, el impacto de la pandemia en la mora fue parcialmente mitigado, destacando el papel de estas garantías como un amortiguador financiero.
- En *Exc. CG*, la morosidad aumentó significativamente en los sectores más afectados, particularmente en **3.D80**, evidenciando que la ausencia de respaldo gubernamental expone a ciertos segmentos a un mayor riesgo financiero.

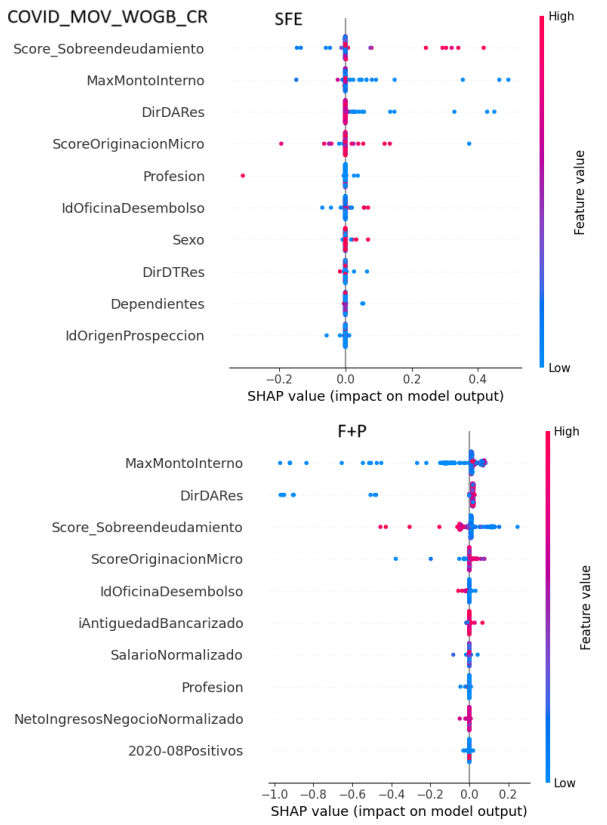
Además, el análisis de la combinación entre *COVID MOV* y la exclusión de garantías gubernamentales sugiere que el tiempo de respuesta de las políticas económicas es clave para contener los efectos negativos de la pandemia en la estabilidad del crédito.

Modelos con mejor desempeño y análisis de explicabilidad

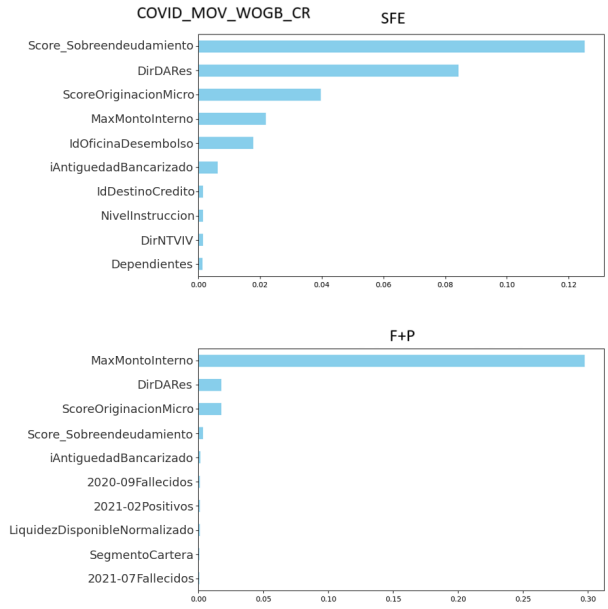
Los modelos con mejor desempeño en términos de precisión y significancia estadística fueron:

- *CNN*: Se consolida como el modelo más robusto en todos los subconjuntos de datos, con mejoras notables en *ACC* y *AUC* cuando se incluye *COVID MOV*.
- *XGB*: Muestra un desempeño estable y preciso, destacando especialmente en **3.D80**, donde se registraron mejoras significativas al integrar los *FE*.
- *XNN* y *EBM*: Aunque en menor medida que *CNN* y *XGB*, estos modelos presentan mejoras en escenarios específicos, lo que sugiere que su capacidad predictiva está condicionada por la composición del conjunto de datos.

Para analizar la importancia de las variables en los modelos predictivos, se realizó un estudio de explicabilidad utilizando *SHAP* y *LIME*. En las Figura 20 y 21 se presentan los 10 factores más influyentes en la predicción de mora bajo los escenarios *SFE* y *F+P*:



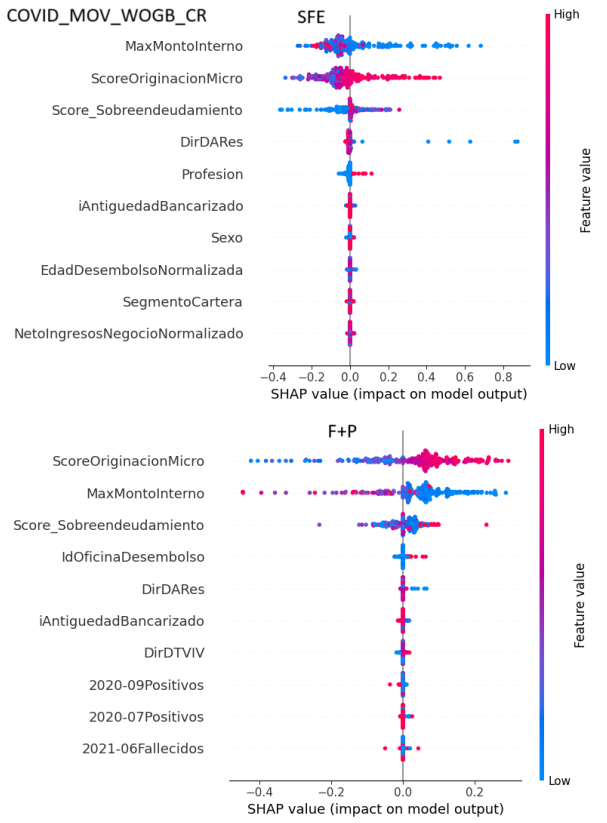
(a)



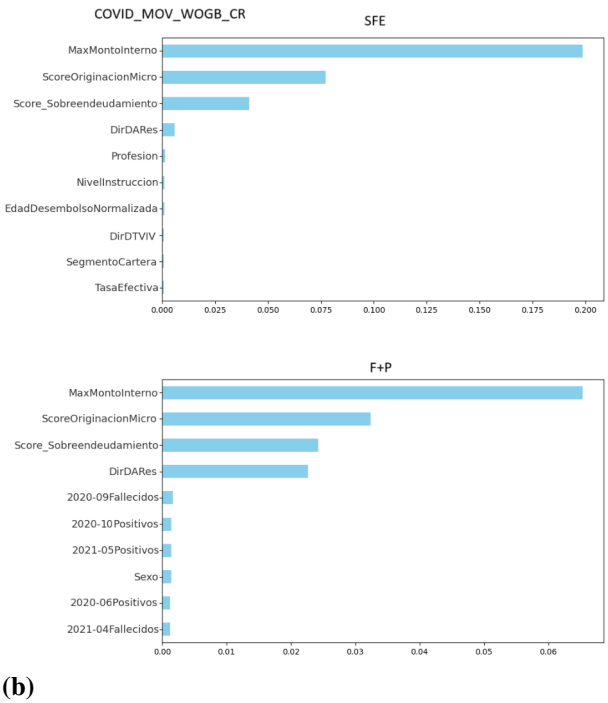
(b)

Figura 20

*Top 10 factores entre los escenarios SFE y F+P en **XNN**: (**a**) SHAP. (**b**) LIME. Leyenda: F = Casos confirmados de COVID-19; P = Fallecidos por COVID-19; F+P = Escenarios combinados; F+Mov = Movimiento de F un periodo a la derecha; WGB_CR = Con crédito respaldado por el gobierno; WOGB_CR = Sin crédito respaldado por el gobierno.*



(a)



(b)

Figura 21

*Top 10 factores entre los escenarios SFE y F+P en **CNN**: (**a**) SHAP. (**b**) LIME. Leyenda: F = Casos confirmados de COVID-19; P = Fallecidos por COVID-19; F+P = Escenarios combinados; F+Mov = Movimiento de F un periodo a la derecha; WGB_CR = Con crédito respaldado por el gobierno; WOGB_CR = Sin crédito respaldado por el gobierno.*

Los resultados de *SHAP* y *LIME* permiten identificar las variables más influyentes en la predicción de mora:

- En los modelos basados en árboles de decisión (*XGB*, *RF*), las variables más importantes son *Ingresos*, *Deuda Total* y *Mora Histórica*.
- En redes neuronales (*CNN*, *XNN*), el efecto de la mortalidad *COVID MOV* es más pronunciado, lo que indica que estas arquitecturas logran capturar de manera más efectiva la influencia de los *FE*.
- En escenarios *Exc. CG*, se observa una mayor dependencia de factores económicos internos, mientras que en *Inc. CG*, las políticas de apoyo amortiguan parcialmente la crisis.
- Implicancias prácticas: Los hallazgos presentados sugieren que las instituciones financieras pueden mejorar la predicción de riesgo crediticio al incorporar *FE* como los datos de mortalidad por *COVID-19* y programas gubernamentales en sus modelos. Además, el uso de series temporales desplazadas permite capturar efectos diferidos en el comportamiento de los deudores, mejorando la capacidad predictiva y reduciendo el riesgo financiero.

Conclusiones

Los resultados obtenidos confirman que:

- Las tasas de mortalidad tienen un impacto diferido más significativo en la mora crediticia que los casos positivos de *COVID-19* .
- La exclusión de créditos con respaldo gubernamental amplifica el efecto de los *FE*, incrementando el riesgo de impago.
- Modelos como *CNN* y *XGB* son los más efectivos para capturar patrones de mora relacionados con la pandemia, destacándose en escenarios con desplazamiento temporal de la mortalidad.

Desde una perspectiva estratégica, estos hallazgos subrayan la importancia de intervenciones financieras oportunas para mitigar los efectos de eventos críticos en el riesgo crediticio.

5.2.2. ¿En qué medida los indicadores del cambio climático, como las anomalías de temperatura y los bloqueos de carreteras debido al clima, afectan los patrones de morosidad crediticia?

La relación entre los días de mora y los efectos del cambio climático se ha evaluado a través de modelos de Machine Learning (*ML*) considerando dos principales *FE*:

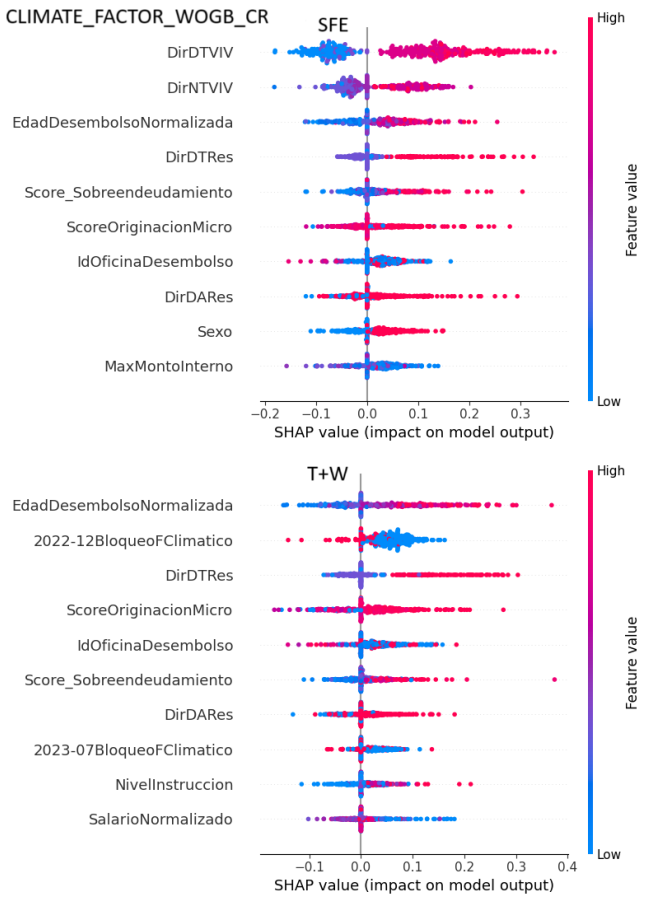
- **Anomalías de temperatura (*T*)**: desviaciones significativas en las condiciones climáticas normales y su impacto en la actividad económica.
- **Bloqueos de carreteras por factores climáticos (*W*)**: eventos que alteran la movilidad, el abastecimiento y la operatividad de ciertos sectores económicos.

Los resultados obtenidos en los distintos escenarios de modelado (*SFE*, *T*, *W*, *T+W*) revelan que estos factores influyen en la capacidad de los modelos de *ML* para predecir la clasificación de buen pagador o mal pagador. La influencia varía según la inclusión o exclusión de créditos beneficiados por programas gubernamentales (*Inc CG* vs. *Exc CG*).

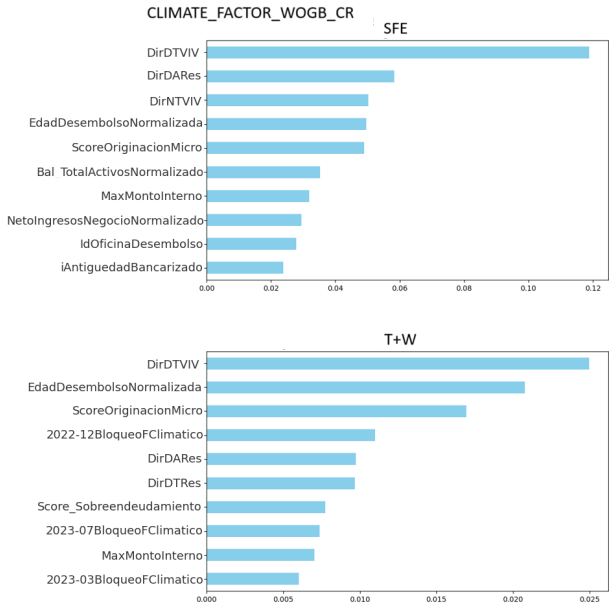
- **Impacto en el Desempeño de los Modelos de *ML***: Los resultados reflejan una relación estadísticamente significativa entre los factores climáticos y los indicadores de predicción de mora:
  - Los modelos *XGB* y *CNN* destacan por su robustez en la predicción de mora en escenarios con *FE*, con incrementos en *ACC* de hasta **32.32 %** en *CNN* en el escenario *T+W* bajo *Inc CG*.
  - *XGB* y *RF* presentan mejoras en *AUC*, con valores superiores al **75 %**, lo que confirma su capacidad para capturar la variabilidad introducida por eventos climáticos extremos.
  - La exclusión de créditos con respaldo gubernamental (*Exc CG*) amplifica el impacto de los factores climáticos en la morosidad, lo que sugiere que estos créditos actuaban como amortiguadores frente a crisis climáticas.
- **Significancia estadística.** El análisis de la prueba *t de Student* reveló que las mejoras observadas en varios modelos son estadísticamente significativas. En particular, el modelo *CNN* demostró significancia en *ACC* para escenarios que incluyen el factor *W*, con valores *p-value* menores a 0.05 en los grupos *D100* y *D99* bajo *Inc. CG*. Esto valida que las diferencias en las métricas no son producto del azar, sino de una influencia real de los factores climáticos.

- **Consideraciones económicas y sociales.** El cambio climático afecta directamente el acceso a los servicios básicos y las condiciones económicas de las regiones, lo cual incide en la capacidad de los deudores para cumplir con sus compromisos financieros. Por ejemplo, los bloqueos de vías (*W*) no solo afectan la logística y el comercio, sino que también generan incertidumbre económica, lo que podría explicar el aumento en los días de mora observado en escenarios específicos. Además, el impacto diferenciado entre *Inc. CG* y *Exc. CG* indica que los apoyos gubernamentales pueden mitigar, aunque no eliminar completamente, los efectos adversos del cambio climático.

- **Análisis de Explicabilidad con *SHAP* y *LIME*:** Para evaluar la importancia de los factores climáticos en la predicción de mora, se aplicaron técnicas de interpretabilidad como *SHAP* y *LIME* a los modelos *XGB* y *CNN* (ver Figura 22), se observó que las características demográficas como la `EdadDesembolso Normalizada` y de comportamiento `ScoreOriginacionMicro` adquieren mayor relevancia.



(a)



(b)

Figura 22

*Top 10 factores entre los escenarios SFE y T+W: (**a**) SHAP en XGB. (**b**) LIME en XGB.*

Leyenda: SFE = Escenario sin FE; T = Anomalía de temperatura; W = Bloqueo por condiciones climáticas; T+W = Escenarios combinados; WGB_CR = Con crédito respaldado por el gobierno; WOGB_CR = Sin crédito respaldado por el gobierno.

- **Efectos Diferenciados según la Cobertura de Datos:** El impacto de los factores climáticos varía según el subconjunto de datos analizado:
  - **D100** (cobertura total de sectores económicos): menor efecto de los *FE* debido a la diversificación del portafolio crediticio.
  - **D99** (99 % de los créditos más relevantes): se detectan aumentos en la morosidad en sectores industriales y de transporte ante eventos climáticos extremos.
  - **D80** (80 % de los créditos más afectados): presenta las mayores variaciones en mora ante anomalías climáticas, con incrementos de hasta **3.06 % en *AUC*** en modelos como *QDA*.
- **Rol de los Créditos Respaldados por el Gobierno:** Comparando los resultados entre escenarios con y sin créditos gubernamentales:
  - **Con respaldo gubernamental**: Se observan incrementos moderados en la capacidad predictiva de los modelos al incluir factores climáticos, con mejoras en *AUC* de hasta **2-3 %** en *CNN*.
  - **Sin respaldo gubernamental**: El impacto de los factores climáticos en la mora es significativamente mayor, lo que sugiere que la ausencia de medidas de mitigación aumenta la vulnerabilidad financiera ante eventos climáticos adversos.
- **Implicancias prácticas.** Los hallazgos tienen implicancias prácticas significativas para las instituciones financieras. Los modelos *CNN* y *XGB*, que demostraron mejoras consistentes y significativas, podrían implementarse en sistemas predictivos para gestionar el riesgo crediticio en regiones afectadas por el cambio climático. Asimismo, la integración de variables climáticas como T y W en las estrategias de evaluación de crédito puede ayudar a identificar sectores más vulnerables y diseñar políticas de mitigación más efectivas, como ajustes en las tasas de interés o planes de pago flexibles durante eventos extremos.

Conclusiones

Los resultados evidencian que:

- **Los factores climáticos afectan la predicción de mora.** La combinación de T y W amplifica el riesgo de impago, con mejoras en *ACC* de hasta **32 %** en algunos escenarios.
- **El impacto es mayor en ciertos segmentos.** La exclusión de créditos con respaldo gubernamental expone a los prestatarios a un mayor riesgo de impago.

- **Los modelos explicables refuerzan los hallazgos.** *SHAP* y *LIME* confirman que los factores climáticos son determinantes en la predicción de mora.
- **Implicancias para la gestión del riesgo financiero.** La integración de indicadores climáticos en modelos de riesgo crediticio permite identificar sectores más vulnerables y diseñar políticas de mitigación efectivas.

En conclusión, incorporar factores climáticos en los modelos de predicción de riesgo crediticio fortalece la resiliencia del sistema financiero ante eventos globales críticos (ver Tabla 34).

Tabla 34
Modelos más relevantes agrupados por Inc. CG y Exc. CG en los escenarios W, T y T+W, destacando las métricas ACC y AUC y las mejoras alcanzadas en comparación con SFE.

Modelo	**Escenario**	**Relevancia**	**Métricas**		**Mejoras**	
			***ACC* (%)**	***AUC* (%)**	***ACC* (%)**	***AUC* (%)**
Inc. CG						
CNN	T	Alta	65.41	64.79	12.80	2.74
	T+W	Alta	84.94	59.93	32.32	-2.12
	W	Alta	83.95	64.56	31.34	2.51
DL	T	Alta	69.23	67.39	4.79	-0.03
XGB	W	Alta	85.01	75.47	1.26	1.26
RF	W	Alta	85.62	76.81	1.68	-0.02
LDA	T	Media	66.80	65.84	0.18	0.30
	T+W	Media	66.99	65.92	0.37	0.37
QDA	W	Media	66.24	67.07	3.06	-0.15
Exc. CG						
CNN	T	Alta	60.47	61.96	3.84	-0.39
	T+W	Alta	87.65	57.87	31.02	-4.48
	W	Alta	84.51	61.89	27.88	-0.46
DL	T	Alta	70.60	66.25	0.53	1.09
XGB	W	Alta	87.46	74.01	1.26	1.62
RF	W	Alta	87.60	74.66	1.89	0.37
LDA	T	Media	67.31	64.84	0.07	0.24
	T+W	Media	67.52	64.96	0.28	0.35
QDA	W	Media	65.05	66.21	2.23	0.16

5.2.3. ¿Cuál es la relación entre la morosidad crediticia y la convulsión social, considerando las perturbaciones de las actividades económicas y la estabilidad social?

La relación entre la morosidad crediticia y la convulsión social ha sido analizada en distintos escenarios, considerando la interacción con otros *FE* como los bloqueos por eventos climáticos (*W*) y las anomalías de temperatura (*T*). Dicha relación ha sido

evaluada a través de modelos de *ML* en múltiples agrupaciones de datos (*D100, D99, D80*) y en contextos que incluyen o excluyen créditos con respaldo gubernamental.

A continuación, se presentan los principales hallazgos:

- **Impacto general de la convulsión social en la morosidad.** La convulsión social (*U*) impacta la capacidad de pago de los deudores al generar incertidumbre económica, restricciones en la movilidad y afectaciones en las actividades productivas. Modelos como *CNN* y *XGB* han mostrado mejoras significativas en *ACC* y *AUC* al incorporar estos factores. En particular, *CNN* presentó un aumento de hasta **31.14 %** en *ACC* y **2.05 %** en *AUC* en el escenario *U* bajo *D100* con inclusión de créditos con *Inc. CG*.

- **Interacción entre convulsión social y factores climáticos.** La combinación de *U* con factores climáticos (*T+U, U+W, U+T+W*) intensifica los efectos sobre la mora crediticia. Por ejemplo, en el escenario *T+U+W* bajo *Exc. CG* y agrupación *D99*, *CNN* logró una mejora en *ACC* de hasta **21.41 %**, mientras que *XGB* presentó un aumento en *AUC* de **1.32 %**. Estos resultados sugieren un efecto sinérgico entre la inestabilidad social y el deterioro de las condiciones ambientales, afectando el comportamiento financiero de los deudores.

- **Efectos diferenciados según el escenario y la estructura de datos.** La respuesta de los modelos varía dependiendo de la agrupación de datos y la inclusión de créditos gubernamentales. En escenarios *Inc. CG*, *RF* ha mostrado mejoras constantes en *ACC* y *AUC*, alcanzando hasta **2.20 %** en *D80*. En contraste, en *Exc. CG*, modelos como *MLP* y *Perceptrón* han mostrado un desempeño más heterogéneo, reflejando mayor sensibilidad a los cambios en la composición del dataset.

- **Significancia estadística de los efectos observados.** Se aplicó la prueba *t de Student* para evaluar la relevancia estadística de las mejoras en los modelos. Los resultados confirman que en el escenario *U+W* con *D99* bajo *Inc. CG*, el modelo *CNN* obtuvo un *p-value* menor a **0.05**, lo que valida la influencia significativa de la convulsión social combinada con bloqueos climáticos en la predicción de mora. De manera similar, *RF* en el escenario *T+U+W* presentó mejoras significativas tanto en *Inc. CG* como en *Exc. CG*, reforzando la solidez de estos efectos.

- **Análisis de explicabilidad: Factores determinantes en la predicción de mora.** La aplicación de técnicas de interpretabilidad como *SHAP* y *LIME* ha permitido identificar que variables relacionadas con sobreendeudamiento y flujo de efectivo son determinantes en la predicción de mora bajo escenarios de convulsión social. Como se observa en la Figura 23, la incorporación de *U* en los

modelos ha incrementado la importancia de estas variables, reflejando su papel en la dinámica crediticia ante eventos de inestabilidad social.

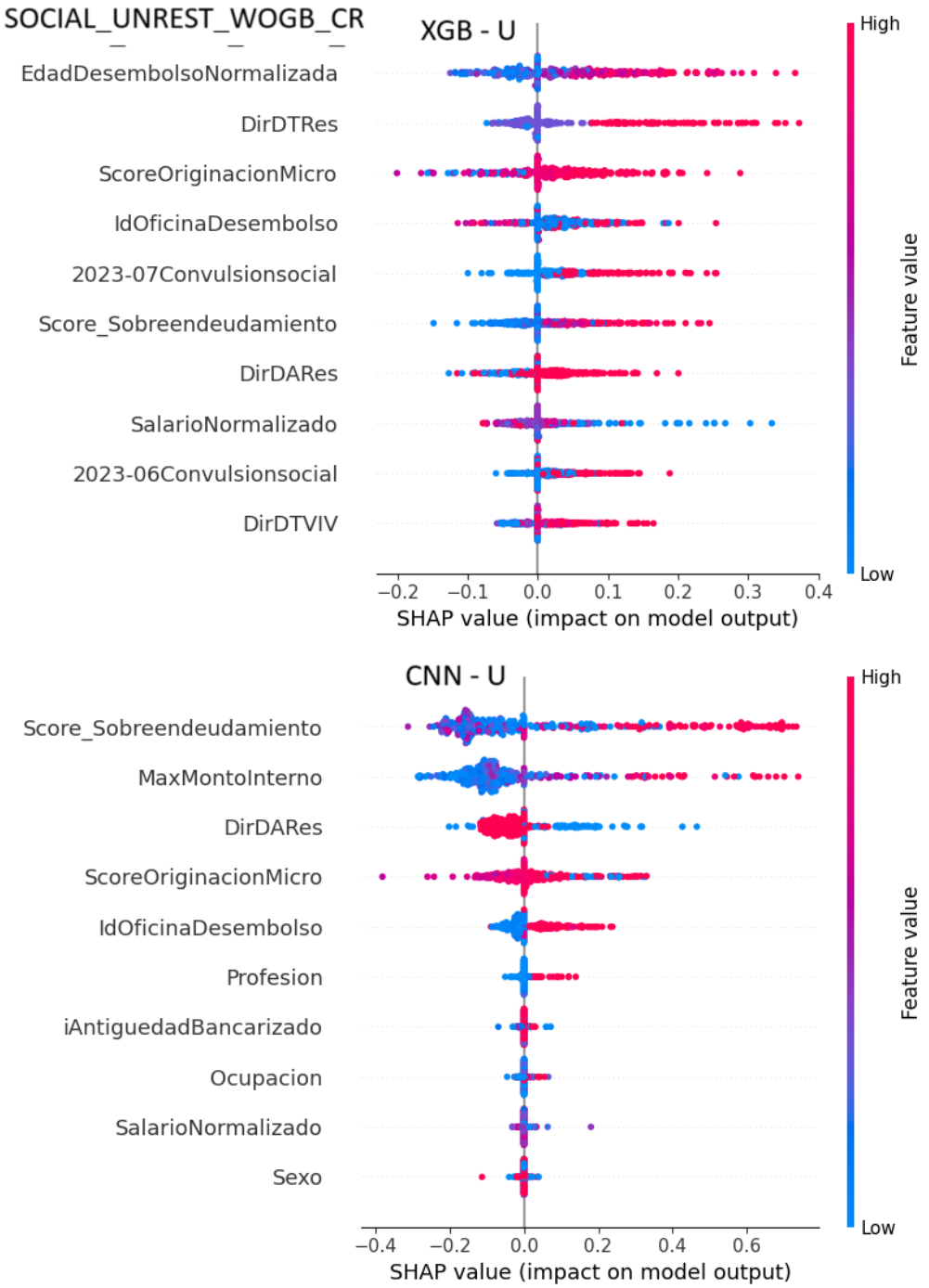


Figura 23

SHAP XGB y CNN: Factores determinantes en escenarios de convulsión social. Leyenda: U = Convulsión social, representada por eventos de protesta civil e inestabilidad política registrados en 2023; WGB_CR = Con crédito respaldado por el gobierno; WOGB_CR = Sin crédito respaldado por el gobierno.

- **Consideraciones económicas y sociales.** La inestabilidad social genera efectos no lineales en la economía, afectando con mayor intensidad a sectores depen-

dientes del comercio físico y la movilidad laboral. Esto se refleja en los modelos de *CNN* y *XNN*, que han mostrado ganancias significativas en *ACC* y *AUC* al incluir *U* en la evaluación del riesgo crediticio.

- **Implicancias prácticas para la gestión del riesgo crediticio.** La inclusión de variables de convulsión social en los sistemas de predicción de mora permite mejorar la precisión de los modelos de riesgo. Instituciones financieras pueden beneficiarse del uso de modelos como *XGB* y *CNN*, que han mostrado robustez en escenarios de alta incertidumbre. Asimismo, la combinación de eventos sociales y climáticos en la evaluación crediticia podría facilitar el diseño de políticas adaptativas, como reestructuración de deudas en regiones afectadas por crisis sociales.

En conclusión, la relación entre la convulsión social y la morosidad crediticia es estadísticamente significativa y se ve acentuada cuando se combinan factores climáticos. Modelos como *CNN* y *XGB* han mostrado mejoras sustanciales en *ACC* y *AUC*, validando la necesidad de incluir variables de convulsión social en los modelos de predicción de riesgo crediticio (ver Tabla 35).

Tabla 35

Modelos más relevantes agrupados por Inc. CG y Exc. CG en los escenarios U, U+T, U+W y U+T+W, destacando las métricas ACC y AUC y las mejoras alcanzadas en comparación con SFE.

Modelo	**Escenario**	**Relevancia**	**Métricas**		**Mejoras**	
			***ACC* (%)**	***AUC* (%)**	***ACC* (%)**	***AUC* (%)**
Inc. CG						
CNN	U	Alta	83.75	64.11	31.14	2.05
	U+T	Alta	84.38	60.59	31.77	-1.46
	U+W	Alta	85.11	60.90	32.50	-1.15
	U+T+W	Alta	85.01	62.64	32.40	0.59
DL	U	Alta	62.26	67.33	-2.18	-0.09
	U+T	Alta	59.06	67.55	-5.38	0.13
	U+W	Alta	63.64	67.59	-0.80	0.17
	U+T+W	Alta	68.61	67.95	4.18	0.53
XGB	U	Alta	84.92	75.24	1.17	1.03
	U+T	Alta	84.93	75.36	1.17	1.15
	U+W	Alta	85.06	75.52	1.30	1.31
	U+T+W	Alta	84.99	75.51	1.24	1.30
RF	U	Alta	85.55	76.73	1.61	-0.10
	U+T	Alta	85.65	76.61	1.72	-0.21
	U+W	Alta	85.67	76.54	1.74	-0.28
	U+T+W	Alta	85.66	76.32	1.73	-0.50
Exc. CG						
CNN	U	Alta	84.37	62.34	27.74	-0.01
	U+T	Alta	87.69	58.89	31.06	-3.46
	U+W	Alta	82.41	58.34	25.78	-4.00
	U+T+W	Alta	87.67	59.00	31.04	-3.35
DL	U	Alta	64.99	66.81	-5.08	1.65
	U+T	Alta	66.82	66.61	-3.24	1.45
	U+W	Alta	68.26	66.49	-1.81	1.34
	U+T+W	Alta	68.50	66.50	-1.56	1.35
XGB	U	Alta	87.41	73.72	1.23	1.12
	U+T	Alta	87.49	73.83	1.31	1.23
	U+W	Alta	87.49	74.02	1.31	1.42
	U+T+W	Alta	87.51	73.92	1.32	1.32
RF	U	Alta	87.53	74.83	1.82	0.29
	U+T	Alta	87.68	74.61	1.96	0.08
	U+W	Alta	87.65	74.49	1.93	-0.05
	U+T+W	Alta	87.67	74.47	1.94	-0.07

5.2.4. ¿Cómo contribuyen los efectos combinados de *FE* (*COVID-19* , cambio climático y convulsión social) a las variaciones en la morosidad crediticia, y cuáles son los factores más influyentes?.

La relación entre los días de mora y los efectos combinados de *FE*, como la pandemia por *COVID-19* , el cambio climático y la convulsión social (Madeira, 2022), es una problemática multidimensional que impacta significativamente el comportamiento crediticio. El presente análisis evalúa el impacto de estos *FE* bajo distintos escenarios mediante modelos de *ML*, considerando métricas de desempeño (*ACC*, *AUC*) en 10 *folds*, pruebas de significancia estadística y análisis de interpretabilidad con SHAP y LIME.

- **Impacto combinado de múltiples crisis en la morosidad crediticia:** Los resultados muestran que cuando los efectos del cambio climático y la convulsión social ocurren simultáneamente, se generan incrementos estadísticamente significativos en la predicción de la mora crediticia. En particular, los escenarios **U+T+W** (convulsión social + anomalías de temperatura + bloqueos de vías) presentan los mayores incrementos en desempeño predictivo, con mejoras en *AUC* de hasta **2.74 %** en modelos como *CNN* y *XGB*.

  Además, los análisis en los datasets **D100**, **D99** y **D80** confirman que estos efectos son más pronunciados en grupos sin apoyo gubernamental, donde la morosidad es más sensible a la interacción entre crisis sanitarias, ambientales y sociales.

- ***FE* más influyentes en la predicción de mora:** El impacto de cada *FE* varía en función del contexto y de la combinación de factores:

  - **Mortalidad por *COVID-19* :** Es el predictor más influyente, especialmente cuando se desplaza temporalmente un periodo adelante *COVID MOV*, reflejando su efecto retardado en la capacidad de pago. Modelos como *CNN* y *XGB* presentan mejoras en *ACC* de hasta **7.16 %** en escenarios que incorporan esta variable, validando su impacto estadístico (p-*value* <0.05).
  - **Cambio climático (*T* y *W*):** Los bloqueos de vías por fenómenos climáticos (*W*) generan aumentos moderados pero constantes en *ACC* y *AUC*, especialmente en ausencia de apoyo gubernamental. Modelos como *XGB* logran incrementos en *AUC* de hasta **1.62 %**, lo que sugiere que la incertidumbre climática impacta el cumplimiento financiero de los deudores.
  - **Convulsión social (*U*):** Las protestas y bloqueos sociales afectan la liquidez de los deudores, generando incrementos en mora. En escenarios

U+T+W, modelos como *CNN* muestran mejoras en *ACC* de hasta **31.14 %**, indicando que la interacción entre eventos sociopolíticos y desastres ambientales puede exacerbar la inestabilidad económica.

- **Significancia estadística y robustez de los resultados:** Se realizaron pruebas de hipótesis unilaterales para validar que los efectos observados no son producto del azar. En escenarios *U+T+W*, los modelos *CNN* y *XGB* presentan *p-values* menores a **0.05**, confirmando la relevancia estadística de estos *FE*. Además, el análisis en 10 *folds* evidencia estabilidad en las mejoras obtenidas, reforzando la confiabilidad de los modelos.

Tabla 36
Resultados ACC para los modelos con mayores valores en los escenarios seleccionados

DataSet	Inc/Exc Créditos	Modelo	Escenario	*ACC*	*ACC - SFE*
1.D100	Incluyendo	*CNN*	*F*	87.03 %	23.70 %
1.D100	Excluyendo	*CNN*	*P*	85.45 %	28.28 %
2.D99	Incluyendo	*CNN*	F+P	84.60 %	24.36 %
2.D99	Excluyendo	*RF*	*U + T*	87.68 %	26.10 %
3.D80	Incluyendo	*CNN*	*U + W*	86.49 %	25.27 %
3.D80	Excluyendo	*CNN*	*U + T + W*	87.33 %	26.52 %

Tabla 37
Resultados AUC para los modelos con mayores valores en los escenarios seleccionados

DataSet	Inc/Exc Créditos	Modelo	Escenario	*AUC*	*AUC - SFE*
1.D100	Incluyendo	*CNN*	*F*	65.93 %	-1.18 %
1.D100	Excluyendo	*CNN*	*P*	66.22 %	3.80 %
2.D99	Incluyendo	*RF*	*F+P*	77.12 %	0.61 %
2.D99	Excluyendo	*CNN*	*U + T*	64.76 %	-2.09 %
3.D80	Incluyendo	*CNN*	*U + W*	66.64 %	1.74 %
3.D80	Excluyendo	*RF*	*U + T + W*	74.62 %	0.31 %

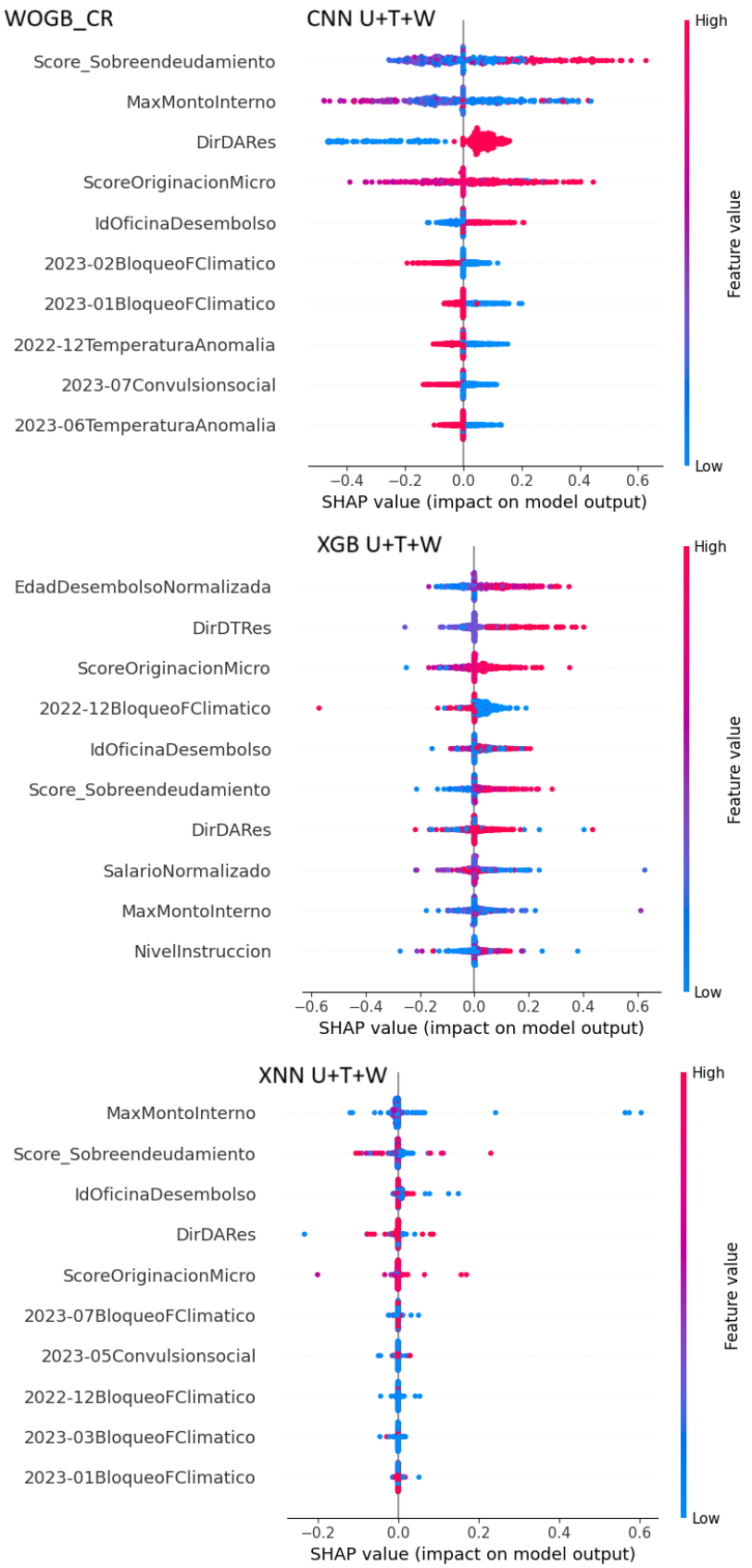


Figura 24

SHAP XGB, CNN y XNN: Factores determinantes en escenarios U+T+W. Leyenda: U = Convulsión social (por ejemplo, protestas, bloqueos de vías); T = Anomalías de temperatura; W = Bloqueos por condiciones climáticas; U+T+W = Escenario combinado que incluye los tres factores externos; WOGB_CR = Cartera de riesgo crediticio excluyendo créditos respaldados por el gobierno.

- **Explicabilidad del impacto de los *FE*:** El uso de SHAP y LIME permitió identificar las variables clave en la predicción de mora. Como se observa en la Figura 24, las variables `Score_Sobreendeudamiento`, `MaxMontoInterno` y `EdadDesembolsoNormalizada` emergen como los principales impulsores del riesgo crediticio en escenarios *U+T+W*.

  Para el modelo *XGB*, la variable `EdadDesembolsoNormalizada` ocupa una posición dominante, mientras que para los modelos de redes neuronales (*CNN, XNN*), la variable `Score_Sobreendeudamiento` es la más relevante. Esto sugiere que, aunque los *FE* afectan la mora, la capacidad de pago previa sigue siendo un determinante crucial.

- **Implicancias estratégicas para la gestión del riesgo crediticio:** La evidencia obtenida sugiere que los efectos combinados de *FE* deben ser considerados en los modelos de riesgo financiero. Las instituciones bancarias pueden beneficiarse del uso de modelos como *XGB*, *CNN* y *XNN*, que han demostrado ser altamente adaptables a crisis multidimensionales. Además, el diseño de políticas de mitigación, como la reestructuración de deudas en sectores vulnerables, puede reducir el impacto de crisis futuras en la mora crediticia.

En conclusión, la interacción de la pandemia, el cambio climático y la convulsión social amplifica significativamente el riesgo de morosidad. La mortalidad por *COVID-19* emerge como el predictor más influyente, pero su impacto podría verse potenciado por condiciones ambientales y sociopolíticas adversas (Madeira, 2022). La combinación de enfoques basados en *ML*, pruebas de significancia estadística y análisis de interpretabilidad permite construir modelos robustos para la evaluación de riesgos crediticios en contextos de alta incertidumbre (ver Tablas 36 y 37).

5.3. Desafíos Presentados en la Investigación

Aunque este estudio demuestra la utilidad de integrar *FE* como la pandemia por *COVID-19*, el cambio climático y la convulsión social en la modelización del riesgo crediticio, han surgido varios desafíos metodológicos y contextuales a lo largo de la investigación:

- **Calidad y Disponibilidad Heterogénea de los Datos:** La recopilación, limpieza y sincronización de los conjuntos de datos de *FE* (e.g., tasas de mortalidad, bloqueos de vías, eventos de protesta) presentan dificultades significativas. Las fuentes públicas a menudo carecen de estándares uniformes de reporte o sufren retrasos en sus actualizaciones, lo que puede generar brechas e inconsistencias que deben ser gestionadas con cuidado.

- **Alineación Temporal de los *FE*:** Los efectos más fuertes de la mortalidad por *COVID-19* , por ejemplo, solo se manifestaron después de un cierto período de incubación o desfase temporal *COVID MOV*. Identificar y ajustar estos desajustes temporales requiere una combinación de conocimientos del dominio (e.g., tiempos de incubación epidemiológica) y pruebas estadísticas (e.g., causalidad de Granger, Dickey-Fuller). Errores en la determinación de estos desfases pueden subestimar o sobreestimar la influencia real de los *FE*.

- **Variabilidad en los Escenarios con Respaldo Gubernamental:** La comparación entre créditos con *Inc. CG* y *Exc. CG* reveló diferencias distributivas notables, particularmente en períodos donde las intervenciones gubernamentales evolucionaron rápidamente. Encontrar un equilibrio entre la generalización del modelo y la evaluación específica por escenario resultó ser un desafío importante.

- **Riesgo de Sobreajuste en Modelos Multifactoriales:** Aunque la incorporación de múltiples *FE* (mortalidad, cambio climático y convulsión social) mejoró el poder predictivo, también aumentó el riesgo de sobreajuste, especialmente en redes neuronales con un alto número de parámetros. Técnicas de regularización y ajuste de hiperparámetros ayudaron a mitigar parcialmente estos riesgos, resaltando la necesidad de prácticas de validación robustas.

- **Equilibrio entre Explicabilidad y Complejidad del Modelo:** Métodos de interpretabilidad post-hoc como *SHAP* y *LIME* mejoraron la transparencia de las predicciones, pero no resolvieron completamente las preocupaciones asociadas con la çaja negra"de los modelos *deep* o *ensemble*. Esto sugiere la necesidad de seguir explorando estrategias para hacer que los modelos avanzados sean más interpretables sin sacrificar precisión.

A pesar de estos desafíos, el enfoque adoptado en el preprocesamiento de datos, la alineación de desfases temporales y la evaluación basada en escenarios resalta la capacidad de adaptación de las técnicas de aprendizaje automático ante *FE* volátiles, al tiempo que señala áreas que requieren refinamientos metodológicos adicionales.

Tabla 38
Comparación de desempeño entre modelos tradicionales y de aprendizaje automático avanzados

Modelo	AUC	ACC
Ridge	76.2 %	78.3 %
LDA	75.1 %	77.5 %
QDA	74.8 %	76.9 %
RF	86.5 %	82.1 %
XGB	**87.9 %**	**83.4 %**
CNN	87.2 %	82.9 %
XNN	86.1 %	81.7 %

Nota. Los modelos de *ML* avanzados presentan un desempeño significativamente superior en comparación con los métodos tradicionales, especialmente al incorporar *FE*.

Como complemento, la Tabla 38 presenta un análisis comparativo del desempeño entre los modelos tradicionales y los modelos de aprendizaje automático avanzados implementados en este estudio, evidenciando las ventajas sustanciales de estos últimos al incorporar *FE*.

5.4. Discusión

Esta sección presenta un análisis comparativo entre los resultados obtenidos en esta investigación y los hallazgos de estudios previos revisados en el estado del arte. El objetivo es evidenciar las contribuciones metodológicas y aplicadas de esta tesis frente a otros enfoques para la predicción del riesgo crediticio, destacando los beneficios del uso de datos enriquecidos con factores exógenos (FE) y modelos de aprendizaje automático avanzados.

Los resultados de esta investigación revelan que la inclusión de *FE*—como la pandemia de *COVID-19*, anomalías climáticas y convulsión social—mejora de manera significativa la capacidad predictiva de los modelos de riesgo crediticio en comparación con modelos entrenados exclusivamente con variables financieras tradicionales. Este hallazgo confirma la hipótesis de que los *FE* generan disrupciones en los patrones de pago que no son capturados por los modelos convencionales.

Comparación con investigaciones previas. Estudios recientes coinciden en que los modelos de *ML* superan a las técnicas tradicionales en contextos dinámicos. Por ejemplo, Imteaj y Amini (2022) demostraron que *Federated Learning* puede integrar información sensible de distintas entidades para mejorar la predicción crediticia, logrando un *AUC* de 0.91. De forma similar, Aranha y Bolar (2023) encontraron que redes neuronales artificiales alcanzaron un *AUC* de 0.94 en escenarios bancarios en India.

En comparación, el presente estudio obtuvo un *AUC* de 0.872 con *CNN* en el escenario compuesto (COVID-19, clima, convulsión social), lo que valida la capacidad de los modelos profundos para captar interacciones complejas entre variables. Este desempeño es consistente con Jiang (2022), quienes utilizaron *CNN* en datos tabulares para préstamos en China, logrando un accuracy de 89.4 %. Además, Z. Chen et al. (2020) evidenciaron que técnicas de ensamblado como *XGB* son robustas en entornos financieros, resultado que también se observa en esta investigación con *XGB* destacándose como uno de los modelos más estables.
Por otro lado, trabajos peruanos como Calsin Enriquez y Vargas Salazar (2023) y Aguilar-Valenzuela y Vilca (2024) se enfocaron en técnicas VAR y árboles de decisión, respectivamente, pero sin incorporar *FE* de forma explícita ni técnicas de explicabilidad. El presente estudio avanza en este aspecto al utilizar *SHAP* y *LIME* para descomponer el impacto de los *FE* en las predicciones, ofreciendo mayor transparencia y confianza para su aplicación en entidades financieras.

Contribuciones metodológicas. La propuesta metodológica integra:

- Un pipeline que combina series temporales externas y datos tabulares en un dataset enriquecido, procesado mediante 14 algoritmos representativos de distintas familias (lineales, discriminantes, ensambladores, redes neuronales profundas).
- Validación robusta mediante *10-fold cross-validation* y técnicas de balanceo *SMOTE* para abordar la desproporción de clases.
- Aplicación de técnicas de Explainable AI (*SHAP*, *LIME*) para interpretar los resultados y facilitar la toma de decisiones en el sector financiero.

Aportes frente al estado del arte. Este trabajo se diferencia de investigaciones previas al:

- Integrar múltiples *FE* en simultáneo (COVID-19, clima, protestas), mientras la mayoría de estudios consideran solo un factor.
- Demostrar la aplicabilidad de *CNN* y FFN en datos tabulares enriquecidos, respondiendo a críticas sobre su pertinencia fuera de imágenes y texto.
- Ofrecer un marco replicable para instituciones financieras en economías emergentes como Perú, con datos reales de alta dimensionalidad.
- Tal como se observa en la Tabla 38, los modelos avanzados como *XGBoost* y *CNN* logran mejoras significativas en AUC y exactitud respecto a métodos tradicionales como Ridge y LDA. Este resultado confirma la capacidad superior de

los modelos de ML para capturar interacciones no lineales y efectos de factores exógenos.

Limitaciones y líneas futuras. Si bien los resultados son prometedores, la investigación presenta limitaciones relacionadas con la disponibilidad de datos de *FE* a nivel granular y la necesidad de actualizar los modelos en tiempo real ante nuevos eventos disruptivos. Futuras investigaciones podrían explorar técnicas específicas de series temporales (p.ej., *LSTM*, Transformers) para capturar secuencias temporales más largas y evaluar su rendimiento comparativo.

La Tabla 39 resume las diferencias clave entre las investigaciones revisadas y la presente tesis, destacando las innovaciones metodológicas y aplicadas.

En conclusión, la presente investigación no solo valida la relevancia de los *FE* en la predicción del riesgo crediticio, sino que también ofrece una metodología innovadora y aplicable para fortalecer la gestión del riesgo en contextos altamente volátiles.

Tabla 39

Comparación de contribuciones de esta tesis con estudios del estado del arte

Ítem	Estudio	Contexto y metodología	Limitaciones identificadas	Contribuciones de esta tesis frente al estudio
1	Imteaj y Amini (2022)	Federated Learning y modelos clásicos (*RF*, *XGBoost*, *ANN*) para riesgo crediticio en instituciones financieras.	No incorpora factores externos ni series temporales en el análisis.	Esta tesis integra múltiples *FE* como features sintéticos, con un dataset enriquecido disponible para la comunidad.
2	Jiang (2022)	CNN aplicadas a datos tabulares de préstamos en China; regularización Muddling Label.	Falta de análisis de sensibilidad a shocks exógenos y variables contextuales.	Demuestra la viabilidad de *CNN* sobre datos tabulares enriquecidos con *FE* en el contexto peruano.
3	Kriebel y Stitz (2022)	Deep Learning en texto de P2P loans (*BERT*, *LSTM*).	Limitado a datos no estructurados; sin evaluación de impactos de *FE*.	Propone integración de series temporales externas en datos tabulares para análisis crediticio.
4	Noriega et al. (2023)	*XGBoost*, *CNN*, *SHAP* y otros modelos *ML* para la predicción de la morosidad.	Enfoque estadístico, sin validación en datos financieros.	Aporta un pipeline que combina múltiples *FE* con técnicas explicativas avanzadas.
5	Geraldo-Campos et al. (2022)	*Lasso* y *Ridge* en programa Reactiva Perú.	Modelos lineales; sin integración de series temporales externas.	Introduce regularización avanzada y *ML* para evaluar impactos de *FE* en microfinanzas.
6	Quintanilla Llanos (2024)	Panel data para desempleo y mora crediticia.	No incorpora *ML* ni efectos compuestos de varios *FE*.	Aporta arquitectura *ML* multi-FE con análisis de sensibilidad por escenario.
7	Calsin Enriquez y Vargas Salazar (2023)	Redes neuronales en microfinanzas peruanas.	Datos estáticos; sin integración de series temporales ni *FE*.	Desarrolla pipeline reproducible con generación de features sintéticos a partir de datos externos.
8	X. Zhang et al. (2024)	*CNN-LSTM* con atención para riesgo en empresas cotizadas.	Enfocado en datos corporativos de grandes empresas.	Adapta técnicas deep learning a microfinanzas y clientes minoristas en contexto emergente.

Continúa en la siguiente página

Tabla 39 Comparación de contribuciones de esta tesis con estudios del estado del arte - continuación

Ítem	Estudio	Contexto y metodología	Limitaciones identificadas	Contribuciones de esta tesis frente al estudio
9	Merćep et al. (2020)	Redes neuronales profundas para behavioral credit scoring.	Modelos de caja negra; baja aplicabilidad en datos tabulares con *FE*.	Incorpora explicabilidad (*SHAP*, *LIME*) y *FE* en modelos tabulares.
10	Madeira (2022)	Impacto combinado de protestas sociales y COVID-19 en Chile con modelos calibrados por encuestas.	Modelos económicos estáticos; no considera datos financieros individuales.	*ML* sobre datos tabulares enriquecidos con series disruptivas; disponibilidad de datos anonimizados.
11	Kemp et al. (2022)	Escenarios de cambio climático extremo y riesgos sistémicos.	Enfoque teórico, sin validación en datos financieros.	Integra riesgos climáticos como variables en *ML* aplicado a datos crediticios.
12	Espinosa-Méndez (2022)	Análisis de comportamiento financiero durante disturbios sociales en mercados emergentes.	No propone integración cuantitativa en *ML*.	Incorpora bloqueos/protestas diarios como variables explicativas en modelos predictivos.
13	Disponibilización de dataset (esta tesis)	N/A	Escasez de datasets públicos con datos financieros y *FE* integrados.	Genera y documenta un dataset enriquecido para replicar experimentos y fomentar investigación futura.
14	Metodología propuesta (esta tesis)	N/A	Limitados frameworks metodológicos para integrar *FE* como features en *ML*.	Propone un pipeline reproducible para integrar series temporales y evaluar su impacto en métricas predictivas.

Capítulo VI

Conclusiones y Recomendaciones

6.1. Conclusiones

Este estudio ha demostrado que la estimación del riesgo crediticio mejora significativamente cuando *FE* como la severidad del *COVID-19*, las anomalías climáticas (temperatura, bloqueos de carreteras debido al clima) y la convulsión social son integrados sistemáticamente en los flujos de trabajo de aprendizaje automático. Los resultados empíricos, validados mediante pruebas de estacionariedad, causalidad de Granger y técnicas de explicabilidad post-hoc (*SHAP*, *LIME*), confirman que la incorporación de *FE* genera mejoras estadísticamente significativas en *AUC* y *ACC* en diversos modelos, como *CNN*, *XGB*, *XNN* y *EBM*. Además, los hallazgos clave de la investigación destacan:

- **La Mortalidad como el Factor Predominante:** La mortalidad por *COVID-19*, cuando se desplaza temporalmente *COVID MOV* para reflejar la progresión de la enfermedad, emerge consistentemente como el predictor más influyente de morosidad crediticia. Esto subraya la importancia de alinear los datos de salud externa con indicadores económicos para capturar efectos retardados con mayor precisión.

- **Las Crisis Combinadas Amplifican el Riesgo de Morosidad:** Cuando los factores climáticos (temperatura, bloqueos de carreteras) y la convulsión social coinciden, la probabilidad de incumplimiento crediticio aumenta sustancialmente. Este efecto es particularmente pronunciado en escenarios sin garantías crediticias gubernamentales, lo que indica que las intervenciones políticas desempeñan un papel amortiguador crucial, pero no eliminan completamente los efectos de crisis múltiples.

- **Modelización con Explicabilidad:** Los modelos basados en árboles (*XGB, RF*) y redes neuronales profundas (*CNN*) se adaptan eficazmente a escenarios enri-

quecidos con *FE*, mientras que arquitecturas interpretables (*EBM, XNN*) confirman la prominencia de variables exógenas en el aumento de la morosidad. Los métodos de explicabilidad post-hoc (*SHAP, LIME*) refuerzan que las variables demográficas y de gestión de deuda siguen siendo predictores clave una vez que se incorporan datos de *FE*.

Estos resultados proporcionan una perspectiva integral y basada en escenarios sobre cómo las disrupciones exógenas—crisis sanitarias, eventos climáticos y disturbios sociales—afectan los incumplimientos crediticios. Mediante la combinación de transformaciones específicas del dominio (desplazamiento temporal de datos), rigurosos experimentos (segmentación de actividades económicas) y marcos de explicabilidad, este estudio ofrece una metodología replicable para mercados emergentes donde los choques externos son cada vez más frecuentes.

6.2. Recomendaciones

Con base en los hallazgos de esta investigación, se proponen las siguientes recomendaciones para mejorar la predicción del riesgo crediticio y fortalecer la resiliencia de los sistemas financieros ante crisis múltiples:

- **Integración continua de datos externos y adopción de ML avanzado:** Resulta crucial que las instituciones financieras mantengan un proceso sistemático de integración de datos sobre *FE*, como pandemias, fenómenos climáticos extremos y disturbios sociales, dentro de sus flujos de análisis predictivo. La incorporación dinámica de estas variables externas permite anticipar con mayor precisión los efectos de eventos disruptivos sobre la morosidad crediticia y ajustar oportunamente las políticas de mitigación de riesgos.

  Además, a la luz de los resultados obtenidos en este estudio (ver Tabla 38), que evidencian mejoras sustanciales en el desempeño predictivo al utilizar algoritmos de aprendizaje automático avanzados como *XGB* y *CNN*, se recomienda a las entidades financieras migrar progresivamente de sistemas tradicionales de *scoring* a enfoques de *ML* robustos. Esta transición es particularmente relevante en entornos caracterizados por alta volatilidad económica y la presencia simultánea de múltiples *FE*.

- **Modelado Temporal Dinámico:** Aunque este estudio utilizó estructuras de desfase fijas (e.g., desplazamiento de un mes para datos de mortalidad), futuras investigaciones podrían emplear modelos de desfase adaptativos o no lineales para capturar relaciones evolutivas entre *FE* y morosidad, especialmente en fenóme-

nos con efectos prolongados como el cambio climático o la inestabilidad social sostenida.

- **Validación en Contextos Internacionales:** Aunque este estudio se enfocó en Perú, futuros estudios en otras regiones ayudarían a validar si patrones similares de *FE* impulsan la morosidad crediticia en otros países. La inclusión de conjuntos de datos multinacionales permitiría evaluar la generalización de los hallazgos.

- **Impacto de las Políticas de Crédito:** La presencia o ausencia de créditos respaldados por el gobierno (*Inc. CG* vs. *Exc. CG*) influyó significativamente en el desempeño de los modelos. Se recomienda evaluar qué tipos de medidas de política (e.g., moratorias, tasas de interés subsidiadas, inyecciones de liquidez dirigidas) son más efectivas para mitigar el riesgo de morosidad inducido por *FE*.

- **Exploración de Técnicas de Explicabilidad Más Robustas:** Aunque *SHAP* y *LIME* proporcionaron información valiosa sobre la relevancia de los *FE*, futuras investigaciones podrían explorar modelos de aprendizaje profundo inherentemente interpretables o enfoques de modelado causal para mejorar la transparencia y confianza en las predicciones, especialmente en entornos regulatorios.

- **Predicciones a Largo Plazo y Riesgo Continuo:** Este estudio se centró en clasificaciones mensuales binarias (buen pagador vs. mal pagador). Desarrollar puntuaciones de riesgo continuas o predicciones en múltiples horizontes temporales podría proporcionar orientación más detallada para estrategias de intervención temprana y reevaluación crediticia en entornos económicos inestables.

- **Uso de Enfoques Híbridos:** Se recomienda explorar modelos híbridos que combinen aprendizaje automático con enfoques tradicionales de análisis de riesgo, como la regresión logística, para aprovechar las fortalezas de ambos métodos y mejorar la robustez de las predicciones.

- **Comparación sistemática con modelos tradicionales:** Si bien los algoritmos avanzados como *XGB* y *CNN* demostraron un desempeño superior, se recomienda mantener evaluaciones periódicas con técnicas tradicionales (e.g., *LDA*, *QDA*) para validar la robustez y justificar la adopción de modelos más complejos en contextos regulatorios o de baja capacidad computacional.

Con los sistemas financieros cada vez más expuestos a choques exógenos convergentes—crisis sanitarias, extremos climáticos e inestabilidad sociopolítica—este estudio subraya la necesidad de herramientas de aprendizaje automático robustas, explicables

y alineadas con el dominio. La integración sistemática de múltiples *FE* en los modelos de riesgo crediticio, junto con la cuantificación rigurosa de sus efectos, permitiría mejorar el desempeño predictivo y la interpretabilidad, fomentando así sistemas de crédito más resilientes e inclusivos.

Referencias bibliográficas

Acharya, V. V., Berner, R., Engle, R., Jung, H. S., Stroebel, J., Zeng, X., y Zhao, Y. (2023). Climate stress testing. *Annual Review of Financial Economics*, *15*(1), 291–326.

Ağca, Ş., Birge, J. R., Wang, Z., y Wu, J. (2023). The impact of covid-19 on supply chain credit risk. *Production and Operations Management*, *32*(12), 4088–4113.

Aguilar-Valenzuela, G. R., y Vilca, E. M. (2024). Algoritmos para machine learning utilizados en la gestión de riesgo crediticio en perú. *Micaela Revista de Investigación-UNAMBA*, *5*(1), 30–35.

Ahmed, I. (2023). *Essays on climate-related financial risks* (Tesis Doctoral no publicada). University of Otago.

Alam, T. M., Shaukat, K., Hameed, I. A., Luo, S., Sarwar, M. U., Shabbir, S., ... Khushi, M. (2020). An investigation of credit card default prediction in the imbalanced datasets. *IEEE Access*, *8*, 201173–201198.

Alharahsheh, H. H., y Pius, A. (2020). A review of key paradigms: Positivism vs interpretivism. *Global Academic Journal of Humanities and Social Sciences*, *2*(3), 39–43.

Alzamora, G. S., Aceituno-Rojo, M. R., y Condori-Alejo, H. I. (2022). An assertive machine learning model for rural micro credit assessment in peru. *Procedia Computer Science*, *202*, 301–306.

Aranha, M., y Bolar, K. (2023). Efficacies of artificial neural networks ushering improvement in the prediction of extant credit risk models. *Cogent Economics & Finance*, *11*(1), 2210916.

Assous, H. (2022). Prediction of banks efficiency using feature selection method: Comparison between selected machine learning models. *Complexity*, *2022*. doi: 10.1155/2022/3374489

Awad, I. M., y Al Karaki, M. S. (2019). The impact of bank lending on palestine economic growth: an econometric analysis of time series data. *Financial Innovation*, *5*(1), 14.

Banco Mundial. (2025). *Panorama general del sector financiero.* `https://www.worldbank.org/en/topic/financialsector`. (Accedido: 2025-07-01, actualizado el 27 de enero de 2025)

Bastos, J. A., y Matos, S. M. (2022). Explainable models of credit losses. *European Journal of Operational Research*, *301*(1), 386–394. doi: 10.1016/j.ejor.2021.11.009

BBC News Mundo. (2020). *Muertes por covid-19: el gráfico que muestra los 10 primeros países del mundo en el ranking de fallecimientos per cápita (y cuáles son de américa latina).* Descargado 2024-12-26, de `https://www.bbc.com/mundo/noticias-54358383` (Accedido el 26 de diciembre de 2024)

Benbya, H., Nan, N., Tanriverdi, H., y Yoo, Y. (2020). Complexity and information systems research in the emerging digital world. *MIS quarterly*, *44*(1), 1–17.

Bertalanffy, L. (1989). *Teoría general de los sistemas. fundamento de desarrollo y aplicaciones (juan almela, trans.).* Fondo de Cultura Económica. (Obra original publicada en 1968).

Bueff, A., Cytryński, M., Calabrese, R., Jones, M., Roberts, J., Moore, J., y Brown, I. (2022). Machine learning interpretability for a stress scenario generation in credit scoring based on counterfactuals. *Expert Systems with Applications*, *202*, 117271. doi: 10.1016/j.eswa.2022.117271

Calsin Enriquez, H. O., y Vargas Salazar, I. Y. (2023). Monetary policy and microfinance credit risk: the case of peru, period 2003–2021. *Pensamiento Crítico, Facultad de Ciencias Económicas, Universidad Nacional Mayor de San Marcos*.

Capasso, G., Gianfrate, G., y Spinelli, M. (2020). Climate change and credit risk. *Journal of Cleaner Production*, *266*, 121634.

Chen, C., Lin, K., Rudin, C., Shaposhnik, Y., Wang, S., y Wang, T. (2022). A holistic approach to interpretability in financial lending: Models, visualizations, and summary-explanations. *Decision Support Systems*, *152*, 113647.

Chen, C.-C., Liu, Z., Yang, G., Wu, C.-C., y Ye, Q. (2020). An improved fault diagnosis using 1d-convolutional neural network model. *Electronics*, *10*(1), 59.

Chen, Z., Chen, W., y Shi, Y. (2020). Ensemble learning with label proportions for bankruptcy prediction. *Expert Systems with Applications*, *146*, 113155.

CNN en Español. (2024). *Protestas violentas en perú tras rechazo de adelantar elecciones.* Descargado 2024-12-26, de https://cnnespanol.cnn.com/video/protestas-violenta-lima-congreso-rechazo-elecciones-crisis-politica-jimena-de-la-quintana-mirador-mundial (Video, accedido el 26 de diciembre de 2024)

Coma, E., Mora, N., Méndez, L., Benítez, M., Hermosilla, E., Fàbregas, M., ... others (2020). Primary care in the time of covid-19: monitoring the effect of the pandemic and the lockdown measures on 34 quality of care indicators calculated for 288 primary care practices covering about 6 million people in catalonia. *BMC family practice*, *21*, 1–9.

Deloitte. (2020). *The future of retail banking: The hyper-personalisation imperative.* https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/financial-services/deloitte-uk-hp-the-future-of-retail-banking-1.pdf.

Dori, D., Sillitto, H., Griego, R. M., McKinney, D., Arnold, E. P., Godfrey, P., ... Krob, D. (2019). System definition, system worldviews, and systemness characteristics. *IEEE Systems Journal*, *14*(2), 1538–1548. doi: 10.1109/JSYST.2019.2904116

Dumitrescu, E., Hué, S., Hurlin, C., y Tokpavi, S. (2022). Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. *European Journal of Operational Research*, *297*(3), 1178–1192.

Durafe, A., y Jha, A. (2018). Bank capital buffer, bank credit and economic growth: evidence from india. *Afro-Asian Journal of Finance and Accounting*, *8*(3), 257–270.

Elkhayat, N., y ElBannan, M. A. (2018). State divestitures and bank performance: Empirical evidence from the middle east and north africa region. *Asian Economic and Financial Review*, *8*(2), 145.

Espinosa-Méndez, C. (2022). Civil unrest and herding behavior: evidence in an emerging market. *Economic research-Ekonomska istraživanja*, *35*(1), 1243–1261.

Fan, S., Shen, Y., y Peng, S. (2020). Improved ml-based technique for credit card scoring in internet financial risk control. *Complexity*, *2020*, 1–14.

Fang, W., Li, X., Zhou, P., Yan, J., Jiang, D., y Zhou, T. (2021). Deep learning anti-fraud model for internet loan: where we are going. *IEEE Access*, *9*, 9777–9784. doi: 10.1109/ACCESS.2021.3051079

Feng, Y. (2022). Bank green credit risk assessment and management by mobile computing and machine learning neural network under the efficient wireless communication. *Wireless Communications and Mobile Computing*, *2022*.

Fitzpatrick, T., y Mues, C. (2021). How can lenders prosper? comparing machine learning approaches to identify profitable peer-to-peer loan investments. *European Journal of Operational Research*, *294*(2), 711–722. doi: 10.1016/j.ejor.2021.01.047

Flori, A., Pammolli, F., y Spelta, A. (2021). Commodity prices co-movements and financial stability: A multidimensional visibility nexus with climate conditions. *Journal of Financial Stability*, *54*, 100876.

Frazier, P. I. (2018). A tutorial on bayesian optimization. *arXiv preprint arXiv:1807.02811*.

Gao, B., y Balyan, V. (2022). Construction of a financial default risk prediction model based on the lightgbm algorithm. *Journal of Intelligent Systems*, *31*(1), 767–779. doi: 10.1515/jisys-2022-0036

Geraldo-Campos, L. A., Soria, J. J., y Pando-Ezcurra, T. (2022). Machine learning for credit risk in the reactiva perú program: a comparison of the lasso and ridge regression models. *Economies*, *10*(8), 188.

Godwin, A., Benedict, B., Rohde, J., Thielmeyer, A., Perkins, H., Major, J., ... Chen, Z. (2021). New epistemological perspectives on quantitative methods: An example using topological data analysis. *Studies in Engineering Education*, *2*(1), 16–34.

Gorzelnik, T., y Oleniacz, R. (2019). Suitability analysis of new air quality monitoring stations in krakow as related to assessment of spatial and temporal variability of pm10 concentrations. *Geomatics and Environmental Engineering*, *13*(2), 31–45.

Hani, U., Wickramasinghe, A., Kattiyapornpong, U., y Sajib, S. (2022). The future of data-driven relationship innovation in the microfinance industry. *Annals of Operations Research*, 1–27.

Hariri, R. H., Fredericks, E. M., y Bowers, K. M. (2019). Uncertainty in big data analytics: survey, opportunities, and challenges. *Journal of Big Data*, *6*(1), 1–16.

Hemdan, E., y Manjaiah, D. (2022). Anomaly credit card fraud detection using deep learning. En *Deep learning in data analytics* (pp. 207–217). Springer. doi: 10.1007/978-3-030-75855-4_12

Hsin, Y., Dai, T., Ti, Y., Huang, M., Chiang, T., y Liu, L. (2022). Feature engineering and resampling strategies for fund transfer fraud with limited transaction data and a

time-inhomogeneous modi operandi. *IEEE Access*, *10*, 86101–86116. doi: 10.1109/ACCESS.2022.3199425

Hutter, F., Kotthoff, L., y Vanschoren, J. (2019). *Automated machine learning: methods, systems, challenges*. Springer Nature.

IBM. (2020). *Introduction to crisp-dm.* `https://www.ibm.com/docs/en/spss-modeler/saas?topic=guide-introduction-crisp-dm`.

Imteaj, A., y Amini, M. (2022). Leveraging asynchronous federated learning to predict customers financial distress. *Intelligent Systems with Applications*, *14*, 200064. doi: 10.1016/j.iswa.2022.200064

Jiang, W. (2022). Loan default prediction with deep learning and muddling label regularization. *IEICE TRANSACTIONS on Information and Systems*, *105*(7), 1340–1342. doi: 10.1587/transinf.2022EDL8003

Kemp, L., Xu, C., Depledge, J., Ebi, K. L., Gibbins, G., Kohler, T. A., ... others (2022). Climate endgame: Exploring catastrophic climate change scenarios. *Proceedings of the National Academy of Sciences*, *119*(34), e2108146119.

Kriebel, J., y Stitz, L. (2022). Credit default prediction from user-generated text in peer-to-peer lending using deep learning. *European Journal of Operational Research*, *302*(1), 309–323. doi: 10.1016/j.ejor.2021.01.047

Larsen-Freeman, D. (2020). Complex dynamic systems theory. En *Theories in second language acquisition* (pp. 248–270). Routledge.

Li, D., y Li, L. (2022). Research on efficiency in credit risk prediction using logistic-sbm model. *Wireless Communications and Mobile Computing*, *2022*.

Lin, G., Shangle, A., Haoxiang, Z., Jingyue, Y., y Junyao, Y. (2022). Research on the construction of financial supervision information system based on machine learning. *Wireless Communications and Mobile Computing*, *2022*. doi: 10.1155/2022/9986095

Liu, W., Fan, H., Xia, M., y Pang, C. (2022). Predicting and interpreting financial distress using a weighted boosted tree-based tree. *Engineering Applications of Artificial Intelligence*, *116*, 105466.

Lombardo, G., Pellegrino, M., Adosoglou, G., Cagnoni, S., Pardalos, P. M., y Poggi, A. (2022). Machine learning for bankruptcy prediction in the american stock market: Dataset and benchmarks. *Future Internet*, *14*(8), 244.

Ma, Y., Zhang, P., Duan, S., y Zhang, T. (2023). Credit default prediction of chinese real estate listed companies based on explainable machine learning. *Finance Research Letters*, *58*, 104305. doi: 10.1016/j.frl.2023.104305

Madeira, C. (2022). The double impact of deep social unrest and a pandemic: Evidence from chile. *Canadian Journal of Economics/Revue canadienne d'économique*, *55*, 135–171.

Mann, S. C., y Logeswaran, R. (2021). Data analytics in improved bankruptcy prediction with industrial risk. En *2021 14th international conference on developments in esystems engineering (dese)* (pp. 23–26).

Mbona, I., y Eloff, J. (2022). Detecting zero-day intrusion attacks using semi-supervised machine learning approaches. *IEEE Access*, *10*, 69822-69838. doi: 10.1109/ACCESS.2022.3187116

Merćep, A., Mrčela, L., Birov, M., y Kostanjčar, Z. (2020). Deep neural networks for behavioral credit rating. *Entropy*, *23*(1), 27.

Mohanty, H., Bhuyan, P., y Chenthati, D. (2015). *Big data: A primer* (Vol. 11). Springer.

Nallakaruppan, M., Chaturvedi, H., Grover, V., Balusamy, B., Jaraut, P., Bahadur, J., ... Hameed, I. A. (2024). Credit risk assessment and financial decision support using explainable artificial intelligence. *Risks*, *12*(10), 164.

Noriega, J. P., Rivera, L. A., y Herrera, J. A. (2023). Machine learning for credit risk prediction: A systematic literature review. *Data*, *8*(11), 169.

Ontón, R., y Reilly, B. (2024). *Dynamic between business and personal debt in a group of microenterprises in peru, in the context of the covid-19 pandemic.* `https://www.sbs.gob.pe/Portals/0/jer/DDT_ANO2024/DT%2003%202024.pdf`. (Documento de Trabajo N.° 03-2024, Superintendencia de Banca, Seguros y AFP (SBS). Acceso: 13 de abril de 2025)

Organización Mundial de la Salud. (2024). *Enfermedad por coronavirus (covid-19): preguntas y respuestas.* Descargado 2024-12-26, de `https://www.who.int/es/news-room/questions-and-answers/item/coronavirus-disease-covid-19` (Accedido el 26 de diciembre de 2024)

Orlova, E. V. (2020). Decision-making techniques for credit resource management using machine learning and optimization. *Information*, *11*(3), 144.

Pandey, M. K., Mittal, M., y Subbiah, K. (2021). Optimal balancing & efficient feature ranking approach to minimize credit risk. *International Journal of Information Management Data Insights*, *1*(2), 100037.

Pastor, G. M. (2024). Tiempos de crisis construyendo un portafolio resiliente. *Microfinanzas*, *224*, 17. Descargado de `https://statuscomunicaciones.pe/microfinanzas/M224.pdf` (Consultado el 11 de agosto de 2024)

Pavitha, N., y Sugave, S. (2024). Explainable multistage ensemble 1d convolutional neural network for trustworthy credit decision. *International Journal of Advanced Computer Science & Applications*, *15*(2). doi: 10.14569/IJACSA.2024.0150201

Quintanilla Llanos, M. A. (2024). Efecto del desempleo en la morosidad en portafolios de tarjetas de crédito ante el impacto de la pandemia (2011–2023). *Trabajo de investigación, Facultad de Economía, Universidad Peruana de Ciencias Aplicadas (UPC)*.

Quinteros-Navarro, D., y Rodriguez, C. (2024). Prediction model applying machine learning to forecast the bankruptcy of companies. a systematic review of the literature. *Journal of Theoretical and Applied Information Technology*, *102*(15).

Real Academia Española. (2023). *Diccionario de la lengua española.* Descargado de `https://www.rae.es/` (Consultado: 22 de agosto de 2024)

Rui, S. (2021). *Deep representations with learned constraints*. Tesis Doctoral, Stanford University. Descargado de `InstitucionalrepositoryofStanfordUniversityhttps://searchworks.stanford.edu/view/14313809`

Sahiq, A. N. M., Ismail, S., Nor, S. H. S., Ul-Saufie, A. Z., y Yaacob, W. F. W. (2022). Application of logistic regression model on imbalanced data in personal bankruptcy prediction. En *2022 3rd international conference on artificial intelligence and data sciences (aidas)* (pp. 120–125).

Shih, D.-H., Wu, T.-W., Shih, P.-Y., Lu, N.-A., y Shih, M.-H. (2022). A framework of global credit-scoring modeling using outlier detection and machine learning in a p2p lending platform. *Mathematics*, *10*(13), 2282.

Song, W., Min, J., y Yang, J. (2023). Credit risk assessment of heavy-polluting enterprises: A wide-lp penalty and deep learning approach. *Mathematics*, *11*(16), 3462. doi: 10.3390/math11163462

Teles, G., Rodrigues, J. J., Rabêlo, R. A., y Kozlov, S. A. (2021). Comparative study of support vector machines and random forests machine learning algorithms on credit operation. *Software: Practice and Experience*, *51*(12), 2492–2500.

Telg, S., Dubinova, A., y Lucas, A. (2023). Covid-19, credit risk management modeling, and government support. *Journal of Banking & Finance*, *147*, 106638.

Thurner, S., Hanel, R., y Klimek, P. (2018). *Introduction to the theory of complex systems*. Oxford University Press.

Tractenberg, R. E. (2019). Strengthening the practice and profession of statistics and data science using ethical guidelines.
doi: 10.1109/JSYST.2019.2904116

Ugwu, C., Ekere, J., y Onoh, C. (2021). Research paradigms and methodological choices in the research process. *Journal of Applied Information Science and Technology*, *14*(2), 116–124.

Valavan, N., y Rita, S. (2023). Predictive-analysis-based machine learning model for fraud detection with boosting classifiers. *Computer Systems Science and Engineering*, *45*(1), 231–245. Descargado de `http://www.techscience.com/csse/v45n1/49293` doi: 10.32604/csse.2023.026508

Wen, C., Yang, J., Gan, L., y Pan, Y. (2021). Big data driven internet of things for credit evaluation and early warning in finance. *Future Generation Computer Systems*, *124*, 295–307.

Wu, Z. (2022). Using machine learning approach to evaluate the excessive financialization risks of trading enterprises. *Computational Economics*, *59*(4), 1607–1625. doi: 10.1007/s10614-020-10090-6

Ye, Y.-F., Jiang, Y.-X., Shao, Y.-H., y Li, C.-N. (2015). Financial conditions index construction through weighted lp-norm support vector regression. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *19*(3), 397–406.

Zhang, C., Zhong, H., y Hu, A. (2022). A method for financial system analysis of listed companies based on random forest and time series. *Mobile Information Systems*, *2022*.

Zhang, L., y Liang, X. (2023). Bayesian regularization in multiple-indicators multiple-causes models. *Psychological Methods*. doi: 10.1037/met0000475

Zhang, X., Ma, Y., y Wang, M. (2024). An attention-based logistic-cnn-bilstm hybrid neural network for credit risk prediction of listed real estate enterprises. *Expert Systems*, *41*(2), e13299.

Zheng, B. (2019). Financial default payment predictions using a hybrid of simulated annealing heuristics and extreme gradient boosting machines. *International Journal of Internet Technology and Secured Transactions*, *9*(4), 404–425.

Zhou, Z.-H. (2021). *Machine learning*. Springer Nature.

Zou, Y., Gao, C., y Gao, H. (2022). Business failure prediction based on a cost-sensitive extreme gradient boosting machine. *IEEE Access*, *10*, 42623–42639.

Anexos

Anexo A

Pseudocódigo de las funciones de evaluación de modelos *ML* y variables usadas

Tabla 40
Pseudocódigo de las funciones de evaluación de modelos incluyendo análisis LIME, SHAP para escenarios con y sin FE

Función	Descripción
remove_collinear_features(df, threshold)	
Entradas	df, threshold
Salidas	DataFrame sin variables colineales, lista de variables eliminadas
Pasos	▪ Calcular la matriz de correlación absoluta de df. ▪ Obtener la matriz triangular superior de la matriz de correlación. ▪ Identificar y eliminar columnas con correlación mayor a threshold. ▪ Retornar el DataFrame resultante y las columnas eliminadas.
evaluate_model_cv(data, target_variable, include_variables, model_name, n_splits)	

Continúa en la próxima página

Tabla 40 continúa de la página anterior

Función	Descripción
Entradas	data, target_variable, include_variables, model_name, n_splits
Salidas	Promedio de métricas (AUC_Val, ACC_Val, AUC_Test, ACC_Test), valores SHAP y explicaciones LIME
Pasos	▪ Dividir data en X e y utilizando include_variables y target_variable. ▪ Eliminar variables colineales de X. ▪ Inicializar StratifiedKFold con n_splits. ▪ Para cada fold: • Dividir datos en entrenamiento, validación y prueba. • Aplicar SMOTE a los datos de entrenamiento. • Entrenar el modelo según model_name. • Calcular métricas de evaluación (AUC y ACC en validación y prueba). • Calcular valores SHAP sobre X_test. • Calcular explicaciones LIME sobre una muestra de X_test. ▪ Retornar métricas, SHAP y LIME.
Modelos Utilizados	

Continúa en la próxima página

Tabla 40 continúa de la página anterior

Función	Descripción
Modelos	▪ Linear Discriminant Analysis (LDA) ▪ Quadratic Discriminant Analysis (QDA) ▪ Perceptron ▪ Ridge Classifier ▪ Multi-Layer Perceptron (MLP) ▪ Random Forest (RF) ▪ K-Nearest Neighbors (KNN) ▪ XGBoost (XGB) ▪ Explainable Boosting Machine (EBM) ▪ Redes Neuronales de Propagación hacia adelante (FFN) ▪ Redes Neuronales de Propagación hacia adelante con Regularización de Penalización Local (FFNLP) ▪ Redes Neuronales de Propagación hacia adelante con Regularizacion Sparza (FFNSP) ▪ Convolutional Neural Networks (CNN) ▪ Explainable Neural Network (XNN)
Escenarios de Evaluación	
Entradas	Conjuntos de datos y configuraciones de escenarios

Continúa en la próxima página

Tabla 40 continúa de la página anterior

Función	Descripción
Escenarios	▪ Sin FE ▪ FE_Covid_Fallecidos ▪ FE_Covid_Positivos ▪ FE_Convulsionsocial ▪ FE_TemperaturaAnomalia ▪ FE_BloqueoFClimatico ▪ FE_TemperaturaAnomalia_BloqueoFClimatico ▪ FE_Convulsionsocial_BloqueoFClimatico ▪ FE_Convulsionsocial_FE_TemperaturaAnomalia ▪ FE_Convulsionsocial_FE_TemperaturaAnomalia_BloqueoFClimatico

Graficar los 10 factores más significativos con SHAP y LIME

Entradas	shap_results_df
Salidas	Gráficos de SHAP y LIME con los 10 factores principales
Pasos	▪ Agrupar shap_results_df por modelo y conjunto de características. ▪ Para cada grupo: • Extraer valores SHAP y X_test. • Generar gráfico summary plot con las 10 características más influyentes. • Guardar el gráfico.

Definir conjuntos de características y ejecutar evaluación

Continúa en la próxima página

Tabla 40 continúa de la página anterior

Función	Descripción
Entradas	▪ Definir feature_sets (ejemplo: 'FE_Covid_Positivos', 'FE_Convulsionsocial', etc.). ▪ Definir bases de datos ('data100', 'data99', 'data80'). ▪ Ejecutar evaluate_multiple_databases con los parámetros definidos. ▪ Guardar resultados en archivos.

Tabla 41
Variables por Escenario

Ítem	Escenario	Variables x Escenario
1	*SFE*	SalarioNormalizado, SegmentoCartera, MaxMontoInterno, Monto, Bal_TotalActivosNormalizado, Bal_PatrimonioNegocioNormalizado, LiquidezDisponibleNormalizado, NetoIngresosNegocioNormalizado, iAntiguedad-Bancarizado, ScoreOriginacionMicro, Score_Sobreendeudamiento, IdOrigenProspeccion, IdDestinoCredito, TasaEfectiva, EdadDesembolsoNormalizada, Profesion, Sexo, Dependientes, NivelInstruccion, Ocupacion, DirDARes, DirDTRes, DirDTVIV, DirNARes, DirNTVIV, IdOficinaDesembolso
2	*F*	SalarioNormalizado, SegmentoCartera, MaxMontoInterno, Monto, Bal_TotalActivosNormalizado, Bal_PatrimonioNegocioNormalizado, LiquidezDisponibleNormalizado, NetoIngresosNegocioNormalizado, iAntiguedad-Bancarizado, ScoreOriginacionMicro, Score_Sobreendeudamiento, IdOrigenProspeccion, IdDestinoCredito, TasaEfectiva, EdadDesembolsoNormalizada, Profesion, Sexo, Dependientes, NivelInstruccion, Ocupacion, DirDARes, DirDTRes, DirDTVIV, DirNARes, DirNTVIV, IdOficinaDesembolso, 2020-03Fallecidos, 2020-04Fallecidos, 2020-05Fallecidos, 2020-06Fallecidos, 2020-07Fallecidos, 2020-08Fallecidos, 2020-09Fallecidos, 2020-10Fallecidos, 2020-11Fallecidos, 2020-12Fallecidos, 2021-01Fallecidos, 2021-02Fallecidos, 2021-03Fallecidos, 2021-04Fallecidos, 2021-05Fallecidos, 2021-06Fallecidos, 2021-07Fallecidos, 2021-08Fallecidos, 2021-09Fallecidos, 2021-10Fallecidos, 2021-11Fallecidos, 2021-12Fallecidos, 2022-01Fallecidos

Continúa en la página siguiente

Tabla 41 – continúa de la página anterior

Item	Escenario	Variables x Escenario
3	P	SalarioNormalizado, SegmentoCartera, MaxMontoInterno, Monto, Bal_TotalActivosNormalizado, Bal_PatrimonioNegocioNormalizado, LiquidezDisponibleNormalizado, NetoIngresosNegocioNormalizado, iAntiguedad-Bancarizado, ScoreOriginacionMicro, Score_Sobreendeudamiento, IdOrigenProspeccion, IdDestinoCredito, TasaEfectiva, EdadDesembolsoNormalizada, Profesion, Sexo, Dependientes, NivelInstruccion, Ocupacion, DirDARes, DirDTRes, DirDT-VIV, DirNARes, DirNTVIV, IdOficinaDesembolso, 2020-03Positivos, 2020-04Positivos, 2020-05Positivos, 2020-06Positivos, 2020-07Positivos, 2020-08Positivos, 2020-09Positivos, 2020-10Positivos, 2020-11Positivos, 2020-12Positivos, 2021-01Positivos, 2021-02Positivos, 2021-03Positivos, 2021-04Positivos, 2021-05Positivos, 2021-06Positivos, 2021-07Positivos, 2021-08Positivos, 2021-09Positivos, 2021-10Positivos, 2021-11Positivos, 2021-12Positivos, 2022-01Positivos
4	$F + P$	SalarioNormalizado, SegmentoCartera, MaxMontoInterno, Monto, Bal_TotalActivosNormalizado, Bal_PatrimonioNegocioNormalizado, LiquidezDisponibleNormalizado, NetoIngresosNegocioNormalizado, iAntiguedad-Bancarizado, ScoreOriginacionMicro, Score_Sobreendeudamiento, IdOrigenProspeccion, IdDestinoCredito, TasaEfectiva, EdadDesembolsoNormalizada, Profesion, Sexo, Dependientes, NivelInstruccion, Ocupacion, DirDARes, DirDTRes, DirDTVIV, DirNARes, DirNTVIV, IdOficinaDesembolso, 2020-03Fallecidos, 2020-04Fallecidos, 2020-05Fallecidos, 2020-06Fallecidos, 2020-07Fallecidos, 2020-08Fallecidos, 2020-09Fallecidos, 2020-10Fallecidos, 2020-11Fallecidos, 2020-12Fallecidos, 2021-01Fallecidos, 2021-02Fallecidos, 2021-03Fallecidos, 2021-04Fallecidos, 2021-05Fallecidos, 2021-06Fallecidos, 2021-07Fallecidos, 2021-08Fallecidos, 2021-09Fallecidos, 2021-10Fallecidos, 2021-11Fallecidos, 2021-12Fallecidos, 2022-01Fallecidos, 2020-03Positivos, 2020-04Positivos, 2020-05Positivos, 2020-06Positivos, 2020-07Positivos, 2020-08Positivos, 2020-09Positivos, 2020-10Positivos, 2020-11Positivos, 2020-12Positivos, 2021-01Positivos, 2021-02Positivos, 2021-03Positivos, 2021-04Positivos, 2021-05Positivos, 2021-06Positivos, 2021-07Positivos, 2021-08Positivos, 2021-09Positivos, 2021-10Positivos, 2021-11Positivos, 2021-12Positivos, 2022-01Positivos

Continúa en la página siguiente

Tabla 41 – continúa de la página anterior

Item	Escenario	Variables x Escenario
5	*U*	SalarioNormalizado, SegmentoCartera, MaxMontoInterno, Monto, Bal_TotalActivosNormalizado, Bal_PatrimonioNegocioNormalizado, LiquidezDisponibleNormalizado, NetoIngresosNegocioNormalizado, iAntiguedadBancarizado, ScoreOriginacionMicro, Score_Sobreendeudamiento, IdOrigenProspeccion, IdDestinoCredito, TasaEfectiva, EdadDesembolsoNormalizada, Profesion, Sexo, Dependientes, NivelInstruccion, Ocupacion, DirDARes, DirDTRes, DirDTVIV, DirNARes, DirNTVIV, IdOficinaDesembolso, 2022-12Convulsionsocial, 2023-01Convulsionsocial, 2023-02Convulsionsocial, 2023-03Convulsionsocial, 2023-04Convulsionsocial, 2023-05Convulsionsocial, 2023-06Convulsionsocial, 2023-07Convulsionsocial, 2023-08Convulsionsocial, 2023-09Convulsionsocial
6	*T*	SalarioNormalizado, SegmentoCartera, MaxMontoInterno, Monto, Bal_TotalActivosNormalizado, Bal_PatrimonioNegocioNormalizado, LiquidezDisponibleNormalizado, NetoIngresosNegocioNormalizado, iAntiguedadBancarizado, ScoreOriginacionMicro, Score_Sobreendeudamiento, IdOrigenProspeccion, IdDestinoCredito, TasaEfectiva, EdadDesembolsoNormalizada, Profesion, Sexo, Dependientes, NivelInstruccion, Ocupacion, DirDARes, DirDTRes, DirDTVIV, DirNARes, DirNTVIV, IdOficinaDesembolso, 2022-12TemperaturaAnomalia, 2023-03TemperaturaAnomalia, 2023-06TemperaturaAnomalia, 2023-09TemperaturaAnomalia
7	*W*	SalarioNormalizado, SegmentoCartera, MaxMontoInterno, Monto, Bal_TotalActivosNormalizado, Bal_PatrimonioNegocioNormalizado, LiquidezDisponibleNormalizado, NetoIngresosNegocioNormalizado, iAntiguedadBancarizado, ScoreOriginacionMicro, Score_Sobreendeudamiento, IdOrigenProspeccion, IdDestinoCredito, TasaEfectiva, EdadDesembolsoNormalizada, Profesion, Sexo, Dependientes, NivelInstruccion, Ocupacion, DirDARes, DirDTRes, DirDTVIV, DirNARes, DirNTVIV, IdOficinaDesembolso, 2022-12BloqueoFClimatico, 2023-01BloqueoFClimatico, 2023-02BloqueoFClimatico, 2023-03BloqueoFClimatico, 2023-04BloqueoFClimatico, 2023-05BloqueoFClimatico, 2023-06BloqueoFClimatico, 2023-07BloqueoFClimatico, 2023-08BloqueoFClimatico, 2023-09BloqueoFClimatico

Continúa en la página siguiente

Tabla 41 – continúa de la página anterior

Item	Escenario	Variables x Escenario
8	$U + T$	SalarioNormalizado, SegmentoCartera, MaxMontoInterno, Monto, Bal_TotalActivosNormalizado, Bal_PatrimonioNegocioNormalizado, LiquidezDisponibleNormalizado, NetoIngresosNegocioNormalizado, iAntiguedadBancarizado, ScoreOriginacionMicro, Score_Sobreendeudamiento, IdOrigenProspeccion, IdDestinoCredito, TasaEfectiva, EdadDesembolsoNormalizada, Profesion, Sexo, Dependientes, NivelInstruccion, Ocupacion, DirDARes, DirDTRes, DirDTVIV, DirNARes, DirNTVIV, IdOficinaDesembolso, 2022-12Convulsionsocial, 2023-01Convulsionsocial, 2023-02Convulsionsocial, 2023-03Convulsionsocial, 2023-04Convulsionsocial, 2023-05Convulsionsocial, 2023-06Convulsionsocial, 2023-07Convulsionsocial, 2023-08Convulsionsocial, 2023-09Convulsionsocial, 2022-12TemperaturaAnomalia, 2023-03TemperaturaAnomalia, 2023-06TemperaturaAnomalia, 2023-09TemperaturaAnomalia
9	$T + W$	SalarioNormalizado, SegmentoCartera, MaxMontoInterno, Monto, Bal_TotalActivosNormalizado, Bal_PatrimonioNegocioNormalizado, LiquidezDisponibleNormalizado, NetoIngresosNegocioNormalizado, iAntiguedadBancarizado, ScoreOriginacionMicro, Score_Sobreendeudamiento, IdOrigenProspeccion, IdDestinoCredito, TasaEfectiva, EdadDesembolsoNormalizada, Profesion, Sexo, Dependientes, NivelInstruccion, Ocupacion, DirDARes, DirDTRes, DirDTVIV, DirNARes, DirNTVIV, IdOficinaDesembolso, 2022-12TemperaturaAnomalia, 2023-03TemperaturaAnomalia, 2023-06TemperaturaAnomalia, 2023-09TemperaturaAnomalia, 2022-12BloqueoFClimatico, 2023-01BloqueoFClimatico, 2023-02BloqueoFClimatico, 2023-03BloqueoFClimatico, 2023-04BloqueoFClimatico, 2023-05BloqueoFClimatico, 2023-06BloqueoFClimatico, 2023-07BloqueoFClimatico, 2023-08BloqueoFClimatico, 2023-09BloqueoFClimatico

Continúa en la página siguiente

Tabla 41 – continúa de la página anterior

Item	Escenario	Variables x Escenario
10	$U + W$	SalarioNormalizado, SegmentoCartera, MaxMontoInterno, Monto, Bal_TotalActivosNormalizado, Bal_PatrimonioNegocioNormalizado, LiquidezDisponibleNormalizado, NetoIngresosNegocioNormalizado, iAntiguedadBancarizado, ScoreOriginacionMicro, Score_Sobreendeudamiento, IdOrigenProspeccion, IdDestinoCredito, TasaEfectiva, EdadDesembolsoNormalizada, Profesion, Sexo, Dependientes, NivelInstruccion, Ocupacion, DirDARes, DirDTRes, DirDTVIV, DirNARes, DirNTVIV, IdOficinaDesembolso, 2022-12Convulsionsocial, 2023-01Convulsionsocial, 2023-02Convulsionsocial, 2023-03Convulsionsocial, 2023-04Convulsionsocial, 2023-05Convulsionsocial, 2023-06Convulsionsocial, 2023-07Convulsionsocial, 2023-08Convulsionsocial, 2023-09Convulsionsocial, 2022-12BloqueoFClimatico, 2023-01BloqueoFClimatico, 2023-02BloqueoFClimatico, 2023-03BloqueoFClimatico, 2023-04BloqueoFClimatico, 2023-05BloqueoFClimatico, 2023-06BloqueoFClimatico, 2023-07BloqueoFClimatico, 2023-08BloqueoFClimatico, 2023-09BloqueoFClimatico
11	$U + T + W$	SalarioNormalizado, SegmentoCartera, MaxMontoInterno, Monto, Bal_TotalActivosNormalizado, Bal_PatrimonioNegocioNormalizado, LiquidezDisponibleNormalizado, NetoIngresosNegocioNormalizado, iAntiguedadBancarizado, ScoreOriginacionMicro, Score_Sobreendeudamiento, IdOrigenProspeccion, IdDestinoCredito, TasaEfectiva, EdadDesembolsoNormalizada, Profesion, Sexo, Dependientes, NivelInstruccion, Ocupacion, DirDARes, DirDTRes, DirDTVIV, DirNARes, DirNTVIV, IdOficinaDesembolso, 2022-12Convulsionsocial, 2023-01Convulsionsocial, 2023-02Convulsionsocial, 2023-03Convulsionsocial, 2023-04Convulsionsocial, 2023-05Convulsionsocial, 2023-06Convulsionsocial, 2023-07Convulsionsocial, 2023-08Convulsionsocial, 2023-09Convulsionsocial, 2022-12TemperaturaAnomalia, 2023-03TemperaturaAnomalia, 2023-06TemperaturaAnomalia, 2023-09TemperaturaAnomalia, 2022-12BloqueoFClimatico, 2023-01BloqueoFClimatico, 2023-02BloqueoFClimatico, 2023-03BloqueoFClimatico, 2023-04BloqueoFClimatico, 2023-05BloqueoFClimatico, 2023-06BloqueoFClimatico, 2023-07BloqueoFClimatico, 2023-08BloqueoFClimatico, 2023-09BloqueoFClimatico

Anexo B

Código fuente para la aplicación de Test de Estacionariedad de Dickey-Fuller y Test de Causalidad de Granger

En este anexo se presenta el código fuente completo utilizado para la determinación del Test de Estacionariedad de Dickey-Fuller y Test de Causalidad de Granger. El código está implementado en Python y documentado para su comprensión y posible reproducción.

Archivo: `T_Estacionariedad_Causalidad.py`

```
# -*- coding: utf-8 -*-
"""EcuacionesExtructurales_v10_Final_v1.ipynb

Automatically generated by Colab.

Original file is located at
    https://colab.research.google.com/drive/1z8R3fPmEJFaHQs9uNMPS4he6zuDuI_An

# **TEST DE ESTACIONARIEDAD:** DICK FULLER

# **Paso 1:** Conectarse con Repositorio de Datos
"""

import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.stattools import adfuller, coint

import pandas as pd

from google.colab import drive
drive.mount('/content/drive')

"""# **Paso 2:** Analizar ruptura estructural de actividades mas relevantes"""

```

```
#Funcion para cargar dataframe

def load_csv(file_path):
    """
    Load a CSV file into a Pandas DataFrame.

    Parameters:
    file_path (str): The path to the CSV file.

    Returns:
    DataFrame: A Pandas DataFrame containing the data from the CSV file.
    """
    try:
        df = pd.read_csv(file_path)
        return df
    except Exception as e:
        return str(e)


# Función para normalizar cada fila
def normalizar_fila(fila):
    return fila / fila.sum()

"""# **Dias Producto Especifico**"""

#Dias Producto Especifico
#Lectura de Datos

#dfM1 = load_csv('/content/drive/MyDrive/Doc/DiasMoraxActividad.csv')
dfM1 = load_csv('/content/drive/MyDrive/Doc/DiasMoraxActividad_2016_2023.csv')


dfM1 = dfM1[(dfM1['Fecha'] >= '2018-04-01') & (dfM1['Fecha'] <= '2020-02-28')]

dfM1.head()

pivot = dfM1.pivot_table(values='Avg_atrasomaximo',
↪  index='CodigoProductoEspecifico', columns='Fecha', aggfunc='mean')

pivot = pivot.fillna(0)

pivot.head()

"""# **Paso 3:** Aplica Normalización"""

#Normalizacion

df_normalizado = pivot.apply(normalizar_fila, axis=1)
df_normalizado = df_normalizado.fillna(0)
print(df_normalizado)

"""# **Paso 4:** Aplica prueba de DF"""

#Aplicando el dickey_fuller: para validar estacionaria

#La tabla Prueba de Dickey-Fuller Aumentada proporciona las hipótesis, una
↪  estadística de prueba,
#un valor p y una recomendación sobre si considerar la diferenciación para hacer
↪  que la serie sea estacionaria.
#La estadística de la prueba proporciona una forma de evaluar la hipótesis nula.

import pandas as pd
import statsmodels.api as sm

# Suponiendo que df es tu DataFrame
# df = pd.read_csv('tu_archivo.csv')  # Carga tus datos aquí

```

```
def aplicar_dickey_fuller(serie):
    try:
        resultado = sm.tsa.stattools.adfuller(serie)
        return resultado[1]  # Devuelve el p-valor
    except ValueError as e:
        print(f"Error en la fila: {serie.name}. Mensaje: {e}")
        return None  # O puedes retornar otro valor que indique el error


df_seleccionado = df_normalizado.iloc[:, 1:]

# Aplicar la prueba a cada fila
p_valores = df_normalizado.apply(aplicar_dickey_fuller, axis=1)

# Analizar los p-valores para determinar la estacionariedad

"""# **Paso 5:** Crea archivo con resultados"""

#Crea Vista

d_valores = p_valores.to_frame()

# Digamos que quieres ordenar por la columna en la posición N (por ejemplo, 0 para la primera columna)
numero_columna = 0  # Cambia esto por el número de columna que deseas
nombre_columna = d_valores.columns[numero_columna]

#df_ordenado = d_valores.sort_values(by='0', ascending=False)

# Mostrar el DataFrame ordenado
#print(df_ordenado)

#print (d_valores)
d_valores.head()

!pip install openpyxl

#d_valores.to_excel('test_archivo_20200316_20230930.xlsx')
#d_valores.to_excel('test_archivo_20200401_20221130.xlsx')
#d_valores.to_excel('test_archivo_20200401_20221031.xlsx')
d_valores.to_excel('test_archivo_20180401_20200228.xlsx')
#d_valores.to_excel('test_archivo_20221201_20230630.xlsx')

"""# **TEST DE CAUSALIDAD:** GRANGER
# **ANALISIS DE CORRELACION:**

# **Paso 1:** Carga Datos de Series Temporales (FE y Actividades economicas)
"""

#Funcion calcula linea de tendencia Mora

def CalculaTendiencia_DiasMora(pdf_Mora, pCodigoProductoEspecifico, pfecha_idx):
    """
    Load a CSV file into a Pandas DataFrame.

    Parameters:
    file_path (str): The path to the CSV file.

    Returns:
    DataFrame: A Pandas DataFrame containing the data from the CSV file.
    """
    try:
        dfM_temp = pdf_Mora[pdf_Mora['CodigoProductoEspecifico'] == pCodigoProductoEspecifico]
        fecha_idx_temp = pfecha_idx
        #dfM_temp_filtrado['Fecha'] = pd.to_datetime(dfM_temp['Fecha'])
```

```
153         dfM_temp_resultado_Mora =
            ↪  dfM_temp.groupby('Fecha')['Avg_atrasomaximo'].mean()
154         dfM_temp_resultado_Mora = dfM_temp_resultado_Mora.reset_index()
155
156         dfM_temp_resultado_Mora.set_index('Fecha', inplace=True)
157         dfM_temp_resultado_Mora = dfM_temp_resultado_Mora.reindex(fecha_idx_temp,
            ↪  fill_value=0)
158
159         resultadoMSuma_temp = dfM_temp_resultado_Mora.values.sum()
160
161         resultadoM_temp = dfM_temp_resultado_Mora.values / resultadoMSuma_temp
162
163         return resultadoM_temp
164     except Exception as e:
165         return str(e)
166
167 #Carga Dataset
168
169 # Data de actividades economicas
170 df_Mora = load_csv('/content/drive/MyDrive/Doc/DiasMoraxActividad.csv')
171
172 # Data de pandemia - Positivos
173 df_CovidPositivos = pd.read_csv('/content/drive/MyDrive/Doc/positivos_covid.csv',
    ↪  delimiter=';')
174 df_CovidPositivos['date'] = pd.to_datetime(df_CovidPositivos['FECHA_RESULTADO'],
    ↪  format='%Y%m%d')
175 df_CovidPositivos = df_CovidPositivos.sort_values(by='date')
176
177 # Data de pandemia - Fallecidos
178 df_CovidFallecidos =
    ↪  pd.read_csv('/content/drive/MyDrive/Doc/fallecidos_sinadef.csv', delimiter='|')
179
180
181 # Data de Bloqueo de vias - Convulsion Social
182 df_ConvulsionSocial = pd.read_csv('/content/drive/MyDrive/Doc/
183 ALERTAS_INTERRUMPIDO_RESTRINGIDO_2022-06_2023-11-01.csv', delimiter='|')
184 df_ConvulsionSocial = df_ConvulsionSocial[df_ConvulsionSocial['motivo'] ==
    ↪  'HUMANO']
185
186 # Data de Bloqueo de vias - Factores Climaticos
187 df_BloqueoClimatico = pd.read_csv('/content/drive/MyDrive/Doc/
188 ALERTAS_INTERRUMPIDO_RESTRINGIDO_2022-06_2023-11-01.csv', delimiter='|')
189 df_BloqueoClimatico = df_BloqueoClimatico[df_BloqueoClimatico['motivo'] ==
    ↪  'CLIMATOLOGICO']
190
191 # Data de Anomalia de temperatura
192 df_Temperatura = pd.read_csv('/content/drive/MyDrive/Doc/Temperatura4.csv',
    ↪  delimiter=';')
193 df_Temperatura['Fecha'] = pd.to_datetime(df_Temperatura['Fecha'])
194 df_Temperatura = df_Temperatura.fillna(0)
195 df_Temperatura = df_Temperatura.sort_values(by='Fecha')
196
197
198 # Definicion de Fecha de Analisis
199 fecha_min = pd.to_datetime('2020-01-01')
200 fecha_max = pd.to_datetime('2023-11-01')
201 fecha_idx = pd.date_range(fecha_min, fecha_max)
202
203 fecha_minMora = pd.to_datetime('2020-01-01')
204 fecha_maxMora = pd.to_datetime('2023-11-01')
205 fecha_idxMora = pd.date_range(fecha_minMora, fecha_maxMora)
206
207 #Filtra Fechas - Mora
208 df_Mora = df_Mora[(df_Mora['Fecha'] >= fecha_minMora.strftime('%Y-%m-%d')) &
    ↪  (df_Mora['Fecha'] <= fecha_maxMora.strftime('%Y-%m-%d'))]
209 df_Mora['Fecha'] = pd.to_datetime(df_Mora['Fecha'])
210
```

```
#Filtrar algunas actividades economicas relevantes
df_Mora_10133 = df_Mora[df_Mora['CodigoProductoEspecifico'] == 10133]
dfM_10364 = df_Mora[df_Mora['CodigoProductoEspecifico'] == 10364]
dfM_411 = df_Mora[df_Mora['CodigoProductoEspecifico'] == 411]
dfM_10453 = df_Mora[df_Mora['CodigoProductoEspecifico'] == 10453]
dfM_10384 = df_Mora[df_Mora['CodigoProductoEspecifico'] == 10384]
dfM_10436 = df_Mora[df_Mora['CodigoProductoEspecifico'] == 10436]
dfM_10449 = df_Mora[df_Mora['CodigoProductoEspecifico'] == 10449]


#Filtra Fechas
df_Mora_filtrado =  df_Mora #df_Mora_10133[(df_Mora_10133['Fecha'] >= fecha_min.strftime('%Y-%m-%d')) & (df_Mora_10133['Fecha'] <= fecha_max.strftime('%Y-%m-%d'))]
df_CovidFallecidos_filtrado = df_CovidFallecidos[(df_CovidFallecidos['FECHA'] >= fecha_min.strftime('%Y-%m-%d')) & (df_CovidFallecidos['FECHA'] <= fecha_max.strftime('%Y-%m-%d'))]
df_ConvulsionSocial_filtrado = df_ConvulsionSocial[(df_ConvulsionSocial['fecha_actualizacion'] >= fecha_min.strftime('%Y-%m-%d')) & (df_ConvulsionSocial['fecha_actualizacion'] <= fecha_max.strftime('%Y-%m-%d'))]
df_BloqueoClimatico_filtrado = df_BloqueoClimatico[(df_BloqueoClimatico['fecha_actualizacion'] >= fecha_min.strftime('%Y-%m-%d')) & (df_BloqueoClimatico['fecha_actualizacion'] <= fecha_max.strftime('%Y-%m-%d'))]
df_Temperatura_filtrado = df_Temperatura[(df_Temperatura['Fecha'] >= fecha_min.strftime('%Y-%m-%d')) & (df_Temperatura['Fecha'] <= fecha_max.strftime('%Y-%m-%d'))]

# Forzar la columna 'fecha' a timpid e dato datetime
df_CovidPositivos['date'] = pd.to_datetime(df_CovidPositivos['date'])
df_Mora_filtrado['Fecha'] = pd.to_datetime(df_Mora_filtrado['Fecha'])
df_CovidFallecidos_filtrado['FECHA'] = pd.to_datetime(df_CovidFallecidos_filtrado['FECHA'])
df_ConvulsionSocial_filtrado['fecha_actualizacion'] = pd.to_datetime(df_ConvulsionSocial_filtrado['fecha_actualizacion'])
df_BloqueoClimatico_filtrado['fecha_actualizacion'] = pd.to_datetime(df_BloqueoClimatico_filtrado['fecha_actualizacion'])
df_Temperatura_filtrado['Fecha'] = pd.to_datetime(df_Temperatura_filtrado['Fecha'])

# Agrupar por 'fecha' y sumar los valores en la columna 'valor'
resultado_CovidPositivos = df_CovidPositivos.groupby('date')['SEXO'].count()
resultado_Mora = df_Mora_filtrado.groupby('Fecha')['Avg_atrasomaximo'].mean()
resultado_Fallecidos = df_CovidFallecidos_filtrado.groupby('FECHA')['SEXO'].count()
resultado_ConvulsionSocial = df_ConvulsionSocial_filtrado.groupby('fecha_actualizacion')['afectacion'].count()
resultado_BloqueoClimatico = df_BloqueoClimatico_filtrado.groupby('fecha_actualizacion')['afectacion'].count()
resultado_Temperatura = df_Temperatura_filtrado.groupby('Fecha')['TempDiff'].mean()

resultado_CovidPositivos = resultado_CovidPositivos.reset_index()
resultado_Mora = resultado_Mora.reset_index()
resultado_Fallecidos = resultado_Fallecidos.reset_index()
resultado_ConvulsionSocial = resultado_ConvulsionSocial.reset_index()
resultado_BloqueoClimatico = resultado_BloqueoClimatico.reset_index()
resultado_Temperatura = resultado_Temperatura.reset_index()

resultado_CovidPositivos.set_index('date', inplace=True)
resultado_CovidPositivos = resultado_CovidPositivos.reindex(fecha_idx, fill_value=0)

resultado_Mora.set_index('Fecha', inplace=True)
resultado_Mora = resultado_Mora.reindex(fecha_idxMora, fill_value=0)

resultado_Fallecidos.set_index('FECHA', inplace=True)
resultado_Fallecidos = resultado_Fallecidos.reindex(fecha_idx, fill_value=0)

```

```
resultado_ConvulsionSocial.set_index('fecha_actualizacion', inplace=True)
resultado_ConvulsionSocial = resultado_ConvulsionSocial.reindex(fecha_idx, fill_value=0)

resultado_BloqueoClimatico.set_index('fecha_actualizacion', inplace=True)
resultado_BloqueoClimatico = resultado_BloqueoClimatico.reindex(fecha_idx, fill_value=0)

resultado_Temperatura.set_index('Fecha', inplace=True)
resultado_Temperatura = resultado_Temperatura.reindex(fecha_idx, fill_value=0)

#Calcula totales
resultadoSuma = resultado_CovidPositivos.values.sum()
resultadoMSuma = resultado_Mora.values.sum()
resultadoFSuma = resultado_Fallecidos.values.sum()
resultadoCS = resultado_ConvulsionSocial.values.sum()
#resultadoCS = resultadoCS * 2 #se amplia para mejorar la visivilidad en los graficos
resultadoBC = resultado_BloqueoClimatico.values.sum()
resultadoTemp = resultado_Temperatura.values.sum()
#resultadoTemp = resultadoTemp / 2 #se reduce para mejorar la visivilidad en los graficos

#Aplica normalizacion por suma
resultado = resultado_CovidPositivos.values / resultadoSuma
resultadoM = resultado_Mora.values / resultadoMSuma
resultadoF = resultado_Fallecidos.values / resultadoFSuma
resultadoCS = resultado_ConvulsionSocial.values / resultadoCS
resultadoBC = resultado_BloqueoClimatico.values / resultadoBC
resultadoTemp = resultado_Temperatura.values / resultadoTemp


#Actividades economicas no estacionarias
resultadoM_10364 = CalculaTendiencia_DiasMora(df_Mora, 10364, fecha_idxMora)
resultadoM_411 = CalculaTendiencia_DiasMora(df_Mora, 411, fecha_idxMora)
resultadoM_10384 = CalculaTendiencia_DiasMora(df_Mora, 10384, fecha_idxMora)
resultadoM_10436 = CalculaTendiencia_DiasMora(df_Mora, 10436, fecha_idxMora)
resultadoM_10449 = CalculaTendiencia_DiasMora(df_Mora, 10449, fecha_idxMora)

#Actividades economicas estacionarias
resultadoM_10366 = CalculaTendiencia_DiasMora(df_Mora, 10366, fecha_idxMora)
resultadoM_10324 = CalculaTendiencia_DiasMora(df_Mora, 10324, fecha_idxMora)
resultadoM_10354 = CalculaTendiencia_DiasMora(df_Mora, 10354, fecha_idxMora)
resultadoM_403 = CalculaTendiencia_DiasMora(df_Mora, 403, fecha_idxMora)
resultadoM_10453 = CalculaTendiencia_DiasMora(df_Mora, 10453, fecha_idxMora)
resultadoM_10449 = CalculaTendiencia_DiasMora(df_Mora, 10449, fecha_idxMora)
resultadoM_10319 = CalculaTendiencia_DiasMora(df_Mora, 10319, fecha_idxMora)


#Imprimer numero de series temporales por cada dataframe
print(len(resultado),len(resultadoM),len(resultadoF),len(resultadoCS),
len(resultadoM_10364), len(df_Mora), fecha_idx)

"""# **Paso 3:** Diagrama FE"""

n = len(resultadoM)
t = np.linspace(0, 10, n)
x = resultadoM
y = resultado
z = resultadoF
a = resultadoCS
bc = resultadoBC
tempe = resultadoTemp

a_10364 = resultadoM_10364
a_411 = resultadoM_411
a_10453 = resultadoM_10453
```

```
a_10384 = resultadoM_10384
a_10436 = resultadoM_10436

# Perform Augmented Dickey-Fuller test
adf_x = adfuller(x)
adf_y = adfuller(y)
adf_z = adfuller(z)
#adf_a = adfuller(w)


# Perform Cointegration test
coint_t, p_value, crit_value = coint(x, y)

# Plot the two time series
plt.figure(figsize=(14, 7))
plt.subplot(2, 1, 1)
#plt.plot(t, x, label='Mora Act.Eco')
plt.plot(t, y, label='COVID Positivos')
plt.plot(t, z, label='COVID Fallecidos')

plt.plot(t, a, label='Bloqueo Convulsion Social')
plt.plot(t, bc, label='Bloqueo Climatico')
plt.plot(t, tempe, label='Anomalia Temperatura')


#plt.plot(t, a_10364, label='a_10364')
#plt.plot(t, a_411, label='a_411')
#plt.plot(t, a_10384, label='a_10384')
#plt.plot(t, a_10436, label='a_10436')

#plt.plot(t, a_provision, label='a_provision')
#plt.plot(t, a_devengado, label='a_devengado')

plt.title('Time Series of X and Y and Z')
plt.legend()

# Show if the series are stationary or not
plt.subplot(2, 1, 2)
plt.title('ADF Test Results')
plt.bar(['X', 'Y', 'Z'], [adf_x[0], adf_y[0], adf_z[0]], color=['blue',
↪  'orange','red'])
plt.axhline(y=crit_value[1], linestyle='--', color='r', label='Critical Value 5%')
plt.legend()
plt.tight_layout()
plt.show()

# Show cointegration test results
adf_x, adf_y, coint_t, p_value, crit_value

"""# **Paso 4:** Diagrama FE + algunas actividades economicas
# Nota: para mejor visualizacion revisar los divisores de Anomalia de Temperatura y
↪  Bloqueo por Cambio Climatico
"""

n = len(resultadoM)
t = np.linspace(0, 10, n)
x = resultadoM
y = resultado
z = resultadoF
a = resultadoCS
bc = resultadoBC
tempe = resultadoTemp

a_10364 = resultadoM_10364
a_411 = resultadoM_411
a_10453 = resultadoM_10453
a_10384 = resultadoM_10384
```

```
a_10436 = resultadoM_10436

# Perform Augmented Dickey-Fuller test
adf_x = adfuller(x)
adf_y = adfuller(y)
adf_z = adfuller(z)
#adf_a = adfuller(w)


# Perform Cointegration test
coint_t, p_value, crit_value = coint(x, y)

# Plot the two time series
plt.figure(figsize=(14, 7))
plt.subplot(2, 1, 1)
#plt.plot(t, x, label='Mora Act.Eco')
plt.plot(t, y, label='COVID Positivos')
plt.plot(t, z, label='COVID Fallecidos')

plt.plot(t, a, label='Bloqueo Convulsion Social')
plt.plot(t, bc, label='Bloqueo Climatico')
plt.plot(t, tempe, label='Anomalia Temperatura')


plt.plot(t, a_10364, label='a_10364')
plt.plot(t, a_411, label='a_411')
plt.plot(t, a_10384, label='a_10384')
plt.plot(t, a_10436, label='a_10436')

#plt.plot(t, a_provision, label='a_provision')
#plt.plot(t, a_devengado, label='a_devengado')

plt.title('Time Series of X and Y and Z')
plt.legend()

# Show if the series are stationary or not
plt.subplot(2, 1, 2)
plt.title('ADF Test Results')
plt.bar(['X', 'Y', 'Z'], [adf_x[0], adf_y[0], adf_z[0]], color=['blue',
↪  'orange','red'])
plt.axhline(y=crit_value[1], linestyle='--', color='r', label='Critical Value 5%')
plt.legend()
plt.tight_layout()
plt.show()

# Show cointegration test results
adf_x, adf_y, coint_t, p_value, crit_value

"""# **Paso 5:** Crea Dataframe para analisis de Correlacion"""

#2. Importar las librerías necesarias

# Crear datos

#Factores Exteros
pCovidPositivos = resultado
pCovidFallecidos = resultadoF
pConvulsionSocial = resultadoCS
pBloqueoClimatico = resultadoCS
pTemperaturaAnomalia = resultadoTemp

pCovidTotal = resultado + resultadoF + resultadoCS + resultadoTemp +
↪  pBloqueoClimatico


## Convertir en una lista simple
pCovidTotal = [elemento for sublista in pCovidTotal for elemento in sublista]
```

```
pCovidPositivos = [elemento for sublista in pCovidPositivos for elemento in sublista]
pCovidFallecidos = [elemento for sublista in pCovidFallecidos for elemento in sublista]
pConvulsionSocial = [elemento for sublista in pConvulsionSocial for elemento in sublista]
pBloqueoClimatico = [elemento for sublista in pBloqueoClimatico for elemento in sublista]
pTemperaturaAnomalia = [elemento for sublista in pTemperaturaAnomalia for elemento in sublista]


#MORA TOTAL
pMoraTotal = resultadoM

## Convertir en una lista simple
pMoraTotal = [elemento for sublista in pMoraTotal for elemento in sublista]


#Variables Sensibles
pa_10364 = resultadoM_10364
pa_10384 = resultadoM_10436
pa_411 = resultadoM_411
pa_10436 = resultadoM_10384
pa_10449 = resultadoM_10449

## Convertir en una lista simple
pa_10364 = [elemento for sublista in pa_10364 for elemento in sublista]
pa_10384 = [elemento for sublista in pa_10384 for elemento in sublista]
pa_411 = [elemento for sublista in pa_411 for elemento in sublista]
pa_10436 = [elemento for sublista in pa_10436 for elemento in sublista]
pa_10449 = [elemento for sublista in pa_10449 for elemento in sublista]

#Variables No Sensibles
pa_10366 = resultadoM_10366
pa_10324 = resultadoM_10324
pa_10354 = resultadoM_10354
pa_403 = resultadoM_403
pa_10453 = resultadoM_10453
pa_10449 = resultadoM_10449
pa_10319 = resultadoM_10319

## Convertir en una lista simple
pa_10366 = [elemento for sublista in pa_10366 for elemento in sublista]
pa_10324 = [elemento for sublista in pa_10324 for elemento in sublista]
pa_10354 = [elemento for sublista in pa_10354 for elemento in sublista]
pa_403 = [elemento for sublista in pa_403 for elemento in sublista]
pa_10453 = [elemento for sublista in pa_10453 for elemento in sublista]
pa_10449 = [elemento for sublista in pa_10449 for elemento in sublista]
pa_10319 = [elemento for sublista in pa_10319 for elemento in sublista]


dates = pd.date_range(start='2020-01-01', periods=1401, freq='D')


# Asegurarse de que 'data' tiene un índice DatetimeIndex
data = pd.DataFrame({
    #'FE_Total': pCovidTotal,
    'FE_CovidPositivos': pCovidPositivos,
    'FE_CovidFallecidos': pCovidFallecidos,
    'FE_ConvulsionSocial': pConvulsionSocial,
    'FE_BloqueoClimatico': pBloqueoClimatico,
    'FE_TemperaturaAnomalia': pTemperaturaAnomalia,
    #'FI_MoraTotal': pMoraTotal,
    #'FIS_10364': pa_10364,
    #'FIS_10384': pa_10384,
    #'FIS_411': pa_411,
```

```
    #'FIS_10436': pa_10436,
    #'FIS_10449': pa_10449,
    #'FINS_10366': pa_10366,
    #'FINS_10324': pa_10324,
    #'FINS_10354': pa_10354,
    #'FINS_403': pa_403,
    #'FINS_10453': pa_10453,
    #'FINS_10449': pa_10449,
    #'FINS_10319': pa_10319,
    #'FDev_Total': pa_10319,
    #'FProv_Total': pa_10319,
    #'FDev_FS': pa_10319,
    #'FProv_FS': pa_10319
}, index=dates)

# Asegurarse de que 'data' tiene un índice DatetimeIndex
data.index = pd.to_datetime(data.index)

plt.figure(figsize=(10, 5))
data.plot()
plt.title('Simulated Time Series Data')
plt.show()

data.info()
df_Mora.info()

df_Mora_CodigoProductoEspecifico =
↪  df_Mora.drop_duplicates(subset='CodigoProductoEspecifico', keep='first')

df_Mora_CodigoProductoEspecifico100 = df_Mora_CodigoProductoEspecifico.copy()
df_Mora_CodigoProductoEspecifico80 = df_Mora_CodigoProductoEspecifico.copy()
df_Mora_CodigoProductoEspecifico99 = df_Mora_CodigoProductoEspecifico.copy()

codigos80 = [411,10366,10324,403,10453,10364,10354,10352,10365,10457,10344,
10384,10319,10396,412,10443,10463,10441,10449,10436,516,10320,10382,10478,
10414,134,10316,405,10408,10322,10464,10353,10340,10370,10386,10327,10337,
10416,10465,10474,10306,10400,10162,10376,10350,410,10401,10425,10367,10358,
10459,10371,10387,10378,10448,10309,10447,408,10388,10467,10326,10317,10380,
10343,10375,10356,10164,10452,10342,938,10334,10300,10341,10328,10351,10379,
10395,5220]

codigos99 = [411,10366,10324,403,10453,10364,10354,10352,10365,10457,10344,
10384,10319,10396,412,10443,10463,10441,10449,10436,516,10320,10382,10478,
10414,134,10316,405,10408,10322,10464,10353,10340,10370,10386,10327,10337,
10416,10465,10474,10306,10400,10162,10376,10350,410,10401,10425,10367,10358,
10459,10371,10387,10378,10448,10309,10447,408,10388,10467,10326,10317,10380,
10343,10375,10356,10164,10452,10342,938,10334,10300,10341,10328,10351,10379,
10395,5220,10115,10437,10170,10024,10377,10147,10472,10438,10445,10362,10460,
10374,10308,10318,10369,10399,10415,407,10458,10479,10477,10423,10181,9309,
10383,10361,10037,10475,513,10169,10346,10372,10473,10143,10599,10182,10336,
10468,10469,10397,512,138,10456,10461,10093,10454,10301,10241,10424,10311,
10339,10368,10393,514,10360,10026,10481,10471,10000,10303,10329,10141,5270,
10355,10484,10455,10435,10157,10349,10107,10120,10347,10432,10136,10149,10363,
10331,10433,10406,10389,10011,10394,10232,10526,10345,517,10496,10084,10450,
10315,10323,10403,10392,10185,10476,10390,10466,10207,8519,10505,10385,10357,
10407,5233,10411,10307,10426,10434,10150,10012,10597,10595,10314,10440,10175,
10332,10439,10604,10154,10062,10075,10193,7499,10422,10163,10421,10480,9999,
10335,10310,10158,10409,10572,10271,10113,10430,10381,10417,10517,10105,10140,
10278,10305,10609,10330,10470,10442,10236,10418,5232,10176,10130,10398,10131,
1511,10173,10263,5520,10259,10172,10222,10238,10338,401,1721,10168,10089,10189,
10040,10110,10495,10359,10013,10404,10427,10554,10165,10272,10085,10183,10289,
10542,10212,10312,5131,10284,10325,10348,10419,135,5234,10391,10498,5020,10030,
10159,5239,10156,5030,6023,10138,10155,10523,10009,10446,10246,122,10191,10413,
10483,10086,10260,5260,7421,10055,10152,10010,1712,10124,10015,10121,10431,7412,
10022,7491,10087,10053,10462,10091,10114,10214,10313,10304,10412,10405,1549,5211,
10023,10041,10145,10205,3720,7511,10103,10116,10134,10555,10594]
```

```
df_Mora_CodigoProductoEspecifico80 = df_Mora_CodigoProductoEspecifico80
[df_Mora_CodigoProductoEspecifico80['CodigoProductoEspecifico'].isin(codigos80)]
df_Mora_CodigoProductoEspecifico99 = df_Mora_CodigoProductoEspecifico99
[df_Mora_CodigoProductoEspecifico99['CodigoProductoEspecifico'].isin(codigos99)]

df_Mora_CodigoProductoEspecifico100.info()
df_Mora_CodigoProductoEspecifico99.info()
df_Mora_CodigoProductoEspecifico80.info()

# Lista para almacenar los DataFrames resultantes
df_Mora_CodigoProductoEspecifico_Detalle100 = []
df_Mora_CodigoProductoEspecifico_Detalle_Cab100 = []

df_Mora_CodigoProductoEspecifico_Detalle99 = []
df_Mora_CodigoProductoEspecifico_Detalle_Cab99 = []

df_Mora_CodigoProductoEspecifico_Detalle80 = []
df_Mora_CodigoProductoEspecifico_Detalle_Cab80 = []

# Iterar sobre cada valor único de CodigoProductoEspecifico
for codigo1 in df_Mora_CodigoProductoEspecifico100['CodigoProductoEspecifico']:
    # Aplicar la función y almacenar el resultado en la lista
    df_resultado = CalculaTendiencia_DiasMora(df_Mora, codigo1, fecha_idx)
    df_resultado = [elemento for sublista in df_resultado for elemento in sublista]
    df_Mora_CodigoProductoEspecifico_Detalle_Cab100.append(codigo1)
    df_Mora_CodigoProductoEspecifico_Detalle100.append(df_resultado)

# Iterar sobre cada valor único de CodigoProductoEspecifico
for codigo1 in df_Mora_CodigoProductoEspecifico99['CodigoProductoEspecifico']:
    # Aplicar la función y almacenar el resultado en la lista
    df_resultado = CalculaTendiencia_DiasMora(df_Mora, codigo1, fecha_idx)
    df_resultado = [elemento for sublista in df_resultado for elemento in sublista]
    df_Mora_CodigoProductoEspecifico_Detalle_Cab99.append(codigo1)
    df_Mora_CodigoProductoEspecifico_Detalle99.append(df_resultado)

# Iterar sobre cada valor único de CodigoProductoEspecifico
for codigo1 in df_Mora_CodigoProductoEspecifico80['CodigoProductoEspecifico']:
    # Aplicar la función y almacenar el resultado en la lista
    df_resultado = CalculaTendiencia_DiasMora(df_Mora, codigo1, fecha_idx)
    df_resultado = [elemento for sublista in df_resultado for elemento in sublista]
    df_Mora_CodigoProductoEspecifico_Detalle_Cab80.append(codigo1)
    df_Mora_CodigoProductoEspecifico_Detalle80.append(df_resultado)

def procesar_codigo_producto_especifico(df_codigo_producto, df_Mora, fecha_idx):
    """
    Procesa los códigos específicos de producto para un DataFrame dado.

    Args:
    - df_codigo_producto: DataFrame que contiene la columna
    ↪  'CodigoProductoEspecifico'.
    - df_Mora: DataFrame principal que se usa dentro de la función
    ↪  CalculaTendiencia_DiasMora.
    - fecha_idx: Índice de fecha que se pasa a la función
    ↪  CalculaTendiencia_DiasMora.

    Returns:
    - detalle_cab: Lista de códigos específicos procesados.
    - detalle: Lista de resultados procesados.
    """
    detalle_cab = []
    detalle = []

    # Iterar sobre cada valor único de CodigoProductoEspecifico
    for codigo1 in df_codigo_producto['CodigoProductoEspecifico']:
        # Aplicar la función y almacenar el resultado en la lista
        df_resultado = CalculaTendiencia_DiasMora(df_Mora, codigo1, fecha_idx)
```

```
646             df_resultado = [elemento for sublista in df_resultado for elemento in
                ↪  sublista]
647             detalle_cab.append(codigo1)
648             detalle.append(df_resultado)
649 
650         return detalle_cab, detalle
651 
652     # Procesar cada DataFrame individualmente
653     df_Mora_CodigoProductoEspecifico_Detalle_Cab100,
    ↪  df_Mora_CodigoProductoEspecifico_Detalle100 =
    ↪  procesar_codigo_producto_especifico(df_Mora_CodigoProductoEspecifico100,
    ↪  df_Mora, fecha_idx)
654     df_Mora_CodigoProductoEspecifico_Detalle_Cab99,
    ↪  df_Mora_CodigoProductoEspecifico_Detalle99 =
    ↪  procesar_codigo_producto_especifico(df_Mora_CodigoProductoEspecifico99,
    ↪  df_Mora, fecha_idx)
655     df_Mora_CodigoProductoEspecifico_Detalle_Cab80,
    ↪  df_Mora_CodigoProductoEspecifico_Detalle80 =
    ↪  procesar_codigo_producto_especifico(df_Mora_CodigoProductoEspecifico80,
    ↪  df_Mora, fecha_idx)
656 
657     df_Mora_CodigoProductoEspecifico100.info()
658     df_Mora_CodigoProductoEspecifico99.info()
659     df_Mora_CodigoProductoEspecifico80.info()
660 
661     # Paso 1: Importar las bibliotecas necesarias
662     import pandas as pd
663     import matplotlib.pyplot as plt
664     import seaborn as sns
665 
666     data.info()
667 
668     data100 = data.copy()
669     data99 = data.copy()
670     data80 = data.copy()
671 
672     import pandas as pd
673 
674     # Suponiendo que df_original es tu DataFrame original
675     # Suponiendo que lista_de_listas es tu lista de listas con datos de columnas
676     # Suponiendo que lista_nombres_columnas contiene los nombres de las columnas para
    ↪  df_original
677 
678     # Verificar que ambas listas tienen la misma longitud
679     if len(df_Mora_CodigoProductoEspecifico_Detalle100) !=
    ↪  len(df_Mora_CodigoProductoEspecifico_Detalle_Cab100):
680         print("Error: Las listas no tienen la misma longitud")
681     else:
682         # Iterar sobre los datos de las columnas y sus respectivos nombres
683         for datos_columna, nombre_columna in
        ↪  zip(df_Mora_CodigoProductoEspecifico_Detalle100,
        ↪  df_Mora_CodigoProductoEspecifico_Detalle_Cab100):
684             # Convertir la lista de datos en una Serie de pandas
685             serie = pd.Series(datos_columna)
686             # Asegurarse de que la serie tiene el mismo índice que df_original
687             serie.index = data.index
688             # Asignar la serie a df_original con el nombre de la columna
                ↪  correspondiente
689             data100[nombre_columna] = serie
690 
691     # Mostrar el resultado
692     print("Número de columnas después de la adición 100:", data100.shape[1])
693 
694     # Verificar que ambas listas tienen la misma longitud
695     if len(df_Mora_CodigoProductoEspecifico_Detalle99) !=
    ↪  len(df_Mora_CodigoProductoEspecifico_Detalle_Cab99):
696         print("Error: Las listas no tienen la misma longitud")
```

```
else:
    # Iterar sobre los datos de las columnas y sus respectivos nombres
    for datos_columna, nombre_columna in zip(df_Mora_CodigoProductoEspecifico_Detalle99, df_Mora_CodigoProductoEspecifico_Detalle_Cab99):
        # Convertir la lista de datos en una Serie de pandas
        serie = pd.Series(datos_columna)
        # Asegurarse de que la serie tiene el mismo índice que df_original
        serie.index = data.index
        # Asignar la serie a df_original con el nombre de la columna correspondiente
        data99[nombre_columna] = serie

# Mostrar el resultado
print("Número de columnas después de la adición 99:", data99.shape[1])

# Verificar que ambas listas tienen la misma longitud
if len(df_Mora_CodigoProductoEspecifico_Detalle80) != len(df_Mora_CodigoProductoEspecifico_Detalle_Cab80):
    print("Error: Las listas no tienen la misma longitud")
else:
    # Iterar sobre los datos de las columnas y sus respectivos nombres
    for datos_columna, nombre_columna in zip(df_Mora_CodigoProductoEspecifico_Detalle80, df_Mora_CodigoProductoEspecifico_Detalle_Cab80):
        # Convertir la lista de datos en una Serie de pandas
        serie = pd.Series(datos_columna)
        # Asegurarse de que la serie tiene el mismo índice que df_original
        serie.index = data.index
        # Asignar la serie a df_original con el nombre de la columna correspondiente
        data80[nombre_columna] = serie

# Mostrar el resultado
print("Número de columnas después de la adición 80:", data80.shape[1])

import pandas as pd

def agregar_columnas(data, detalle_list, cab_list):
    """
    Agrega columnas a un DataFrame a partir de listas de detalles y sus correspondientes nombres de columnas.

    Args:
    - data: DataFrame al que se agregarán las nuevas columnas.
    - detalle_list: Lista de listas con los datos para cada columna.
    - cab_list: Lista con los nombres de las columnas a agregar.

    Returns:
    - data: DataFrame con las nuevas columnas agregadas.
    """
    if len(detalle_list) != len(cab_list):
        print(f"Error: Las listas de detalles y cabeceras no tienen la misma longitud.")
        return data

    for datos_columna, nombre_columna in zip(detalle_list, cab_list):
        # Convertir la lista de datos en una Serie de pandas
        serie = pd.Series(datos_columna)
        # Asegurarse de que la serie tiene el mismo índice que el DataFrame original
        serie.index = data.index
        # Asignar la serie al DataFrame con el nombre de la columna correspondiente
        data[nombre_columna] = serie

    # Mostrar el número de columnas después de la adición
    print(f"Número de columnas después de la adición: {data.shape[1]}")
```

```
754
755         return data
756
757     # Aplicar la función a cada conjunto de datos
758     data100 = agregar_columnas(data100, df_Mora_CodigoProductoEspecifico_Detalle100,
        ↪  df_Mora_CodigoProductoEspecifico_Detalle_Cab100)
759     data99 = agregar_columnas(data99, df_Mora_CodigoProductoEspecifico_Detalle99,
        ↪  df_Mora_CodigoProductoEspecifico_Detalle_Cab99)
760     data80 = agregar_columnas(data80, df_Mora_CodigoProductoEspecifico_Detalle80,
        ↪  df_Mora_CodigoProductoEspecifico_Detalle_Cab80)
761
762     data100.info()
763     data99.info()
764     data80.info()
765
766     correlation_matrix = data.corr()
767     plt.figure(figsize=(20, 16))
768     sns.heatmap(correlation_matrix, annot=False, cmap='RdYlGn' )
769     plt.title('Matriz de correlación')
770     plt.show()
771
772     #plot_color_gradients('Diverging', #['RdYlGn'])
773
774     import pandas as pd
775     import numpy as np
776     from statsmodels.tsa.stattools import grangercausalitytests, InfeasibleTestError
777
778     def granger_matrix_p_lag_binary(data, main_vector, start_date=None, end_date=None,
        ↪  maxlag=8, frequency='M', exclude_vectors=None):
779         # Filtrar los datos para el rango de fechas especificado
780         if start_date:
781             data = data[data.index >= pd.to_datetime(start_date)]
782         if end_date:
783             data = data[data.index <= pd.to_datetime(end_date)]
784
785         # Excluir los vectores especificados
786         if exclude_vectors:
787             data = data.drop(columns=exclude_vectors)
788
789         # Agrupar los datos por la frecuencia especificada (por defecto, mensual)
790         grouped = data.groupby(pd.Grouper(freq=frequency))
791
792         # Preparar las columnas para los resultados, nombrándolas por mes y año
793         columns = [f'{date.year}-{date.month:02d}' for date in grouped.groups.keys()]
794         results_df = pd.DataFrame(index=data.columns, columns=columns, dtype='object')
795
796         # Iterar sobre cada grupo de tiempo
797         for name, group in grouped:
798             column_name = f'{name.year}-{name.month:02d}'
799
800             if np.all(group[main_vector] == 0):
801                 results_df[column_name] = 'NA'  # No aplicable si el vector principal
                    ↪  es cero
802             else:
803                 for target in data.columns:
804                     if target != main_vector:  # Excluir main_vector de la comparación
805                         if len(group) < maxlag or group[main_vector].nunique() == 1 or
                            ↪  group[target].nunique() == 1:
806                             results_df.at[target, column_name] = 'NA'  # No aplicable
                                ↪  si no hay suficientes datos o valores constantes
807                         else:
808                             try:
809                                 result = grangercausalitytests(group[[main_vector,
                                    ↪  target]], maxlag=maxlag, verbose=False)
810                                 min_p_value = float('inf')
811                                 # Encontrar el menor p-valor no cero
812                                 for lag in range(1, maxlag + 1):
```

```
                                p_value = result[lag][0]['ssr_chi2test'][1]
                                if p_value < min_p_value:
                                    min_p_value = p_value
                            # Determinar si el p-valor es menor o igual a 0.05 y
                            ↪  almacenar el resultado binario
                            results_df.at[target, column_name] = 1 if min_p_value
                            ↪  <= 0.05 else 0
                        except (ValueError, InfeasibleTestError):
                            results_df.at[target, column_name] = 'NA'

    return results_df

# Asumiendo que 'data' ya está definido y contiene los vectores apropiados
#FE_CovidPositivos
main_vector = 'FE_CovidPositivos'
exclude_vectors = ['FE_Total','FE_CovidFallecidos',
↪  'FE_IndiceSoberano','FE_ConvulsionSocial','FE_TemperaturaAnomalia']

 # ['FE_Total', 'FE_CovidPositivos', 'FE_CovidFallecidos',
 ↪  'FE_IndiceSoberano','FE_ConvulsionSocial']
start_date = '2020-01-01'
end_date = '2023-09-30'

result_matrix_Positivos = granger_matrix_p_lag_binary(data, main_vector,
↪  start_date, end_date, 3,'M',exclude_vectors=exclude_vectors)
#print(result_matrix)

# Asumiendo que 'data' ya está definido y contiene los vectores apropiados
#FE_CovidPositivos
main_vector = 'FE_CovidFallecidos'
exclude_vectors = ['FE_Total','FE_CovidPositivos',
↪  'FE_IndiceSoberano','FE_ConvulsionSocial','FE_TemperaturaAnomalia']

result_matrix_Fallecidos = granger_matrix_p_lag_binary(data, main_vector,
↪  start_date, end_date, 3,'M',exclude_vectors=exclude_vectors)
#print(result_matrix)

# Asumiendo que 'data' ya está definido y contiene los vectores apropiados
#FE_CovidPositivos
main_vector = 'FE_IndiceSoberano'
exclude_vectors = ['FE_Total','FE_CovidPositivos',
↪  'FE_CovidFallecidos','FE_ConvulsionSocial','FE_TemperaturaAnomalia']

result_matrix_IndiceSoberano = granger_matrix_p_lag_binary(data, main_vector,
↪  start_date, end_date, 3,'M',exclude_vectors=exclude_vectors)
#print(result_matrix)

# Asumiendo que 'data' ya está definido y contiene los vectores apropiados
#FE_CovidPositivos
main_vector = 'FE_ConvulsionSocial'
exclude_vectors = ['FE_Total','FE_CovidPositivos',
↪  'FE_CovidFallecidos','FE_IndiceSoberano','FE_TemperaturaAnomalia']

result_matrix_ConvulsionSocial = granger_matrix_p_lag_binary(data, main_vector,
↪  start_date, end_date, 3,'M',exclude_vectors=exclude_vectors)
#print(result_matrix)

# Asumiendo que 'data' ya está definido y contiene los vectores apropiados
#FE_CovidPositivos
main_vector = 'FE_TemperaturaAnomalia'
exclude_vectors = ['FE_Total','FE_CovidPositivos',
↪  'FE_CovidFallecidos','FE_IndiceSoberano','FE_ConvulsionSocial']

result_matrix_TemperaturaAnomalia = granger_matrix_p_lag_binary(data, main_vector,
↪  start_date, end_date, 3,'M',exclude_vectors=exclude_vectors)
#print(result_matrix)

```

```
# Asumiendo que 'data' ya está definido y contiene los vectores apropiados
#FE_CovidPositivos
main_vector = 'FE_Total'
exclude_vectors = ['FE_TemperaturaAnomalia','FE_CovidPositivos',
↪  'FE_CovidFallecidos','FE_IndiceSoberano','FE_ConvulsionSocial']

result_matrix_Total = granger_matrix_p_lag_binary(data, main_vector, start_date,
↪  end_date, 3,'M',exclude_vectors=exclude_vectors)
#print(result_matrix)

# Asumiendo que 'data' ya está definido y contiene los vectores apropiados
#FE_CovidPositivos
main_vector = 'FE_TemperaturaAnomalia'
exclude_vectors = ['FE_Total','FE_CovidPositivos',
↪  'FE_CovidFallecidos','FE_IndiceSoberano','FE_ConvulsionSocial']

result_matrix_TemperaturaAnomalia = granger_matrix_p_lag_binary(data, main_vector,
↪  start_date, end_date, 1,'M',exclude_vectors=exclude_vectors)
#print(result_matrix)


print(result_matrix_TemperaturaAnomalia)

result_matrix_Positivos.to_excel('Result_matrix_Positivos.xlsx')
result_matrix_Fallecidos.to_excel('result_matrix_Fallecidos.xlsx')
result_matrix_IndiceSoberano.to_excel('result_matrix_IndiceSoberano.xlsx')
result_matrix_ConvulsionSocial.to_excel('result_matrix_ConvulsionSocial.xlsx')
result_matrix_TemperaturaAnomalia.to_excel('result_matrix_TemperaturaAnomalia.xlsx')
result_matrix_Total.to_excel('result_matrix_Total.xlsx')

correlation_matrix = data.corr()

# Suponiendo que `definitions` es un diccionario que contiene las definiciones de
↪  las variables
definitions = {
    'Variable_1': 'Esta es la definición de la Variable_1.',
    'Variable_2': 'Esta es la definición de la Variable_2.',
    'Variable_3': 'Esta es la definición de la Variable_2.',
    'Variable_4': 'Esta es la definición de la Variable_2.',
    'Variable_5': 'Esta es la definición de la Variable_2.',
    'Variable_6': 'Esta es la definición de la Variable_2.',
    'Variable_7': 'Esta es la definición de la Variable_2.',
    'Variable_8': 'Esta es la definición de la Variable_2.',
    'Variable_9': 'Esta es la definición de la Variable_2.',
    'Variable_10': 'Esta es la definición de la Variable_2.',
    'Variable_11': 'Esta es la definición de la Variable_2.',
    '411.0' : 'FORMAL DEPENDENT',
'V10366' : 'WINERY - STORES',
'V10324' : 'SALE OF CLOTHES, CLOTHING ITEMS',
'V403' : 'INFORMAL DEPENDENT',
'V10453' : 'GENERAL CONTRACTOR',
'V10364' : 'FOOD SALE',
'V10354' : 'MARKETING AND BREEDING OF ANIMALS',
'V10352' : 'SALE OF VEGETABLES AND TUBERS',
'V10365' : 'BAZAAR',
'V10457' : 'RESTAURANT/SODA FOUNTAIN',
'V10344' : 'SALE OF FRUITS',
'V10384' : 'CONSTRUCTION WORKER',
'V10319' : 'SALE OF SODA, BEER, LIQUORS, CIGARS',
'V10396' : 'AGRICULTURE LABORERS',
'V412' : 'FINANCIAL INCOME',
'V10443' : 'DEALERS IN GENERAL',
'V10463' : 'PROPERTY RENTAL',
'V10441' : 'PRIVATE TRANSPORTATION - TAXI',
'V10449' : 'CLOTHING SERVICE',
'V516' : 'SELLERS',
'V10436' : 'MECHANICAL WORKSHOP / CAR REPAIR',
```

```
'V10320' : 'SALE OF PERFUMERY, COSMETICS',
'V10478' : 'HAIRDRESSER, BEAUTY TREATMENT, SPA',
'V10382' : 'DOMESTIC EMPLOYEE',
'V10414' : 'MANUFACTURING OF CLOTHING',
'V134' : 'PRIVATE EMPLOYEE',
'V405' : 'RETIRED',
'V10316' : 'SALE OF PREPARED FOOD',
'V10322' : 'SALE OF FOOTWEAR',
'V10408' : 'BAKERY',
'V10464' : 'RENT A CAR',
'V10353' : 'SALE OF GROCERIES',
'V10370' : 'HARDWARE AND PAINTS',
'V10340' : 'SALE OF POULTRY MEAT',
'V10327' : 'SELLING HOUSEHOLD ITEMS',
'V10386' : 'COOK / KITCHEN ASSISTANT / WAITERS',
'V10337' : 'SALE OF NATURAL PRODUCTS',
'V10416' : 'CARPENTRY',
'V10474' : 'MEDICAL/DENTAL AND OPTIC SERVICES',
'V10465' : 'RENT MACHINES, TOOLS AND EQUIPMENT',
'V10306' : 'BOOKSHOP',
'V10400' : 'CLOTHING IN GENERAL',
'V10350' : 'SALE OF FISH AND SEAFOOD',
'V410' : 'GENERIC',
'V10376' : 'TAXI DRIVER / MOTOTAXI DRIVER (OWN UNIT)',
'V10425' : 'METAL CARPENTRY',
'V10401' : 'SEWING / EMBROIDERY / TAILORING SERVICE',
'V10367' : 'PHARMACY AND PHARMACY',
'V10358' : 'SALE OF SPARE PARTS, AUTO PARTS AND VEHICLE ACCESSORIES',
'V10459' : 'MOTOTAXI TRANSPORTATION',
'V10387' : 'SEWING BOX',
'V10162' : 'CORN CROP',
'V10371' : 'SALE OF CANDY AND CONFITERIA',
'V10378' : 'WATCHMAN / GOALKEEPERS',
'V10309' : 'SALE OF CLEANING SUPPLIES',
'V10448' : 'SERVICES FOR AGRICULTURE',
'V10447' : 'BUFFET SERVICES / CATERING / DECORATIONS',
'V408' : 'HOUSEWORK',
'V10388' : 'ELECTRICIAN',
'V10467' : 'MACHINERY AND EQUIPMENT ASSEMBLY SERVICE',
'V10326' : 'SALE OF CELL PHONE CARDS, PHONE CARDS AND ACCESSORIES',
'V10317' : 'SALE OF GRAINS',
'V10380' : 'HOME NURSES',
'V938' : 'INDEPENDENT PROFESSIONAL',
'V10343' : 'SALE OF FLOWERS, PLANTS, GARDENING ARTICLE',
'V10375' : 'PUBLIC TRANSPORTATION AND FREIGHT DRIVER',
'V10356' : 'RECYCLING, SALE OF SCRAP, USED PRODUCTS.',
'V10452' : 'PRINTING SERVICES, GRAPHIC SERVICES, TYPING, COPIES',
'V10342' : 'SPICERY SALE',
'V10334' : 'SALE OF JEWELRY, JEWELRY',
'V10328' : 'SALE OF ELECTRONIC ITEMS',
'V10341' : 'SALE OF RED MEATS',
'V10300' : 'FORAGE SALE, ALFALFA',
'V10351' : 'SALE OF PLASTICS',
'V10379' : 'TUTOR',
'V10395' : 'MECHANIC / IRONER',
'V10437' : 'GAME',
'V10164' : 'APPLE CULTIVATION',
'V10377' : 'TAXI DRIVER (RENTALED UNIT)',
'V10472' : 'INITIAL, PRIMARY AND SECONDARY EDUCATION SERVICES',
'V10438' : 'REPAIR OF APPLIANCES',
'V10445' : 'ENTERTAINMENT / RECREATION SERVICE'
    # Agrega las definiciones de todas las variables en tu conjunto de datos
}


plt.figure(figsize=(20, 16))

```

```
996  import matplotlib.pyplot as plt
997  import seaborn as sns
998
999
1000
1001 # Renombrar las columnas del DataFrame
1002 data_renamed = data.rename(columns=definitions)
1003
1004 # Calcular la matriz de correlación
1005 correlation_matrix = data_renamed.corr()
1006
1007 plt.figure(figsize=(20, 16))
1008 sns.heatmap(correlation_matrix, annot=False, cmap='RdYlGn')
1009 plt.title('Matriz de correlación')
1010 plt.show()
1011
1012 # Visualización de una serie de tiempo como ejemplo
1013 plt.figure(figsize=(14, 7))
1014 plt.plot(data.index, data['FE_Total'], label='Casos COVID-19')
1015 plt.plot(data.index, data['FI_MoraTotal'], label='Días de mora - Actividad 1')
1016 plt.legend()
1017 plt.title('Series de tiempo de COVID-19 y días de mora para Actividad 1')
1018 plt.xlabel('Fecha')
1019 plt.ylabel('Cantidad')
1020 plt.show()
1021
1022 print(type(data))
1023
1024 import pandas as pd
1025 import statsmodels.api as sm
1026
1027 # Configurar pandas para mostrar todas las filas y columnas
1028 pd.set_option('display.max_rows', None)
1029 pd.set_option('display.max_columns', None)
1030 # Supongamos que `data` es tu DataFrame con las 54 columnas y el índice de fecha
1031
1032 # Lista para guardar los resultados
1033 results = []
1034
1035 # Iterar sobre cada columna de actividades económicas
1036 for col in data.columns:
1037     if col != 'FE_Total':
1038         activity_col = col
1039
1040         # Crear una columna de casos de COVID-19 con un rezago de 1 día
1041         combined_data = data[['FE_Total', activity_col]].copy()
1042         combined_data['lagged_covid'] = combined_data['FE_Total'].shift(1)
1043         combined_data = combined_data.dropna()  # Eliminar filas con valores nulos
             ↪  debido al rezago
1044
1045         # Análisis de Correlación
1046         correlation = combined_data['FE_Total'].corr(combined_data[activity_col])
1047
1048         # Análisis de Regresión
1049         X = combined_data[['lagged_covid']]
1050         y = combined_data[activity_col]
1051         X = sm.add_constant(X)
1052         model = sm.OLS(y, X).fit()
1053
1054         # Guardar resultados
1055         results.append({
1056             'actividad': activity_col,
1057             'correlacion': correlation,
1058             'coeficiente': model.params['lagged_covid'],
1059             'p_valor': model.pvalues['lagged_covid'],
1060             'r_cuadrado': model.rsquared
1061         })
```

```
1062
1063  # Convertir resultados a DataFrame para análisis posterior
1064  results_df = pd.DataFrame(results)
1065  print(results_df)
1066
1067  # Análisis de Cointegración para cada actividad económica
1068
1069  import pandas as pd
1070  from statsmodels.tsa.vector_ar.vecm import coint_johansen
1071
1072  # Lista para guardar los resultados de cointegración
1073  cointegration_results = []
1074
1075  excluir_columnas = ['FE_ConvulsionSocial', 'FE_Total','FE_CovidPositivos',
1076  'FE_CovidFallecidos','FE_IndiceSoberano','FE_TemperaturaAnomalia']
1077
1078  # Iterar sobre cada columna de actividades económicas
1079  for col in data.columns:
1080      if col not in excluir_columnas:
1081          activity_col = col
1082
1083          # Crear el DataFrame combinado
1084          combined_data = data[['FE_ConvulsionSocial', activity_col]].copy()
1085          combined_data = combined_data.dropna()  # Eliminar filas con valores nulos
1086
1087          # Prueba de cointegración de Johansen
1088          johansen_test = coint_johansen(combined_data, det_order=0, k_ar_diff=1)
1089
1090          # Guardar resultados
1091          cointegration_results.append({
1092              'actividad': activity_col,
1093              'trace_stat': johansen_test.lr1[0],
1094              'critical_value_5pct': johansen_test.cvt[0, 1]  # Valor crítico al 5%
1095          })
1096
1097  # Convertir resultados a DataFrame para análisis posterior
1098  cointegration_results_df = pd.DataFrame(cointegration_results)
1099  print(cointegration_results_df.head())
1100
1101  # Exportar los resultados a un archivo CSV
1102  cointegration_results_df.to_csv('resultados_cointegracion_CV.csv', index=False)
1103
1104  # Mostrar el DataFrame en el notebook (opcional)
1105  from IPython.display import display
1106
1107  display(cointegration_results_df)
```

Anexo C

Código fuente para el proceso de enriquecimiento de Dataset con datos de *FE*

En este anexo se presenta el código fuente completo utilizado para enriquecer el Dataset con datos de *FE COVID-19* Covid, convulsión social y anomalías por factores climáticos. El código está implementado en Python y documentado para su comprensión y posible reproducción.

Archivo: `Enriquec_Dataset_FE.py`

```python
# -*- coding: utf-8 -*-
"""Actualiza_FE_Dataset_v1.ipynb

Automatically generated by Colab.

Original file is located at
    https://colab.research.google.com/drive/1kRfARVDOozHZyQFiYLDYjO6QD-xZEbl5
"""

#Paso 1 Conectar a Drive
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.stattools import adfuller, coint

import pandas as pd

from google.colab import drive
drive.mount('/content/drive')

# Función para convertir los valores
def convert_to_datetime(val):
    if isinstance(val, str):
        return pd.to_datetime(val)
    elif isinstance(val, (int, float)):
        return pd.to_datetime(val, unit='s')
    else:
        return pd.NaT  # Not a Time para valores que no se pueden convertir

```

```
#Paso 4: Cargar y preparar los datos


# Asegúrate de usar la ruta correcta a tu archivo CSV
#file_path = '/content/drive/MyDrive/Doc/Data2/2023_09_v2.rpt'
file_path = '/content/drive/MyDrive/Doc/Data2/Data011.rpt'
data = pd.read_csv(file_path, delimiter = ';')

# Asegúrate de usar el formato de fecha correcto
# Suponiendo que la columna de fecha se llama 'date_column'
#data['FechaOtorgamiento'] = pd.to_datetime(data['FechaOtorgamiento'])
data['FechaOtorgamiento'] = data['FechaOtorgamiento'].apply(convert_to_datetime)

#Definir el Rango de Fechas
import datetime

fecha_inicio = datetime.datetime.strptime("2020-01-01", "%Y-%m-%d")  # Fecha de inicio
fecha_fin = datetime.datetime.strptime("2023-09-30", "%Y-%m-%d")  # Fecha de fin

#Filtrar Rango
#data = data[(data['FechaOtorgamiento'] >= fecha_inicio) & (data['FechaOtorgamiento'] <= fecha_fin)]

# Eliminar múltiples columnas
#data = data.drop(['DirDTVIV','DirNTVIV','CodigoFuenteIngreso','EstadoCredito','DiasMora','Sexo',
'RCC0','RCC1','RCC2','RCC3','RCC4','FechaOtorgamiento','CodigoProductoEspecifico'], axis=1)
#data = data.drop(['Salario','iInstituciones_sinQ_corte','Nacionalidad','DirNTRes','DiasMora',
'EstadoC','RCC0','RCC1','RCC2','RCC3','RCC4','DeudaTotal',        'MaxNroInst','FechaOtorgamiento',
'CodigoProductoEspecifico','FechaCancelacion'], axis=1)
data = data.drop(['Fechaventa','DUbigeo1','DUbigeo2','DUbigeo3','NUbigeo1','NUbigeo2','NUbigeo3',
'DeudaTotal','MaxNroInst','RCC0','RCC1','RCC2','RCC3','RCC4'], axis=1)




# Elimina filas que contengan algún NaN
#data.dropna(inplace=True)
#data = data.dropna()


# Aplica pd.to_numeric y verifica los datos en todas las columnas (ajusta según sea necesario)
for column in data.columns:
    data[column] = pd.to_numeric(data[column], errors='coerce')


# Mostrar las primeras filas para verificar que los datos se han cargado correctamente
#print(data.head())


data.info()


#print(data.isnull().sum())
#data.info()

print(data.info())

import pandas as pd

# Cargar la matriz de resultados desde el archivo Excel
```

```
result_matrixTotal = pd.read_excel('/content/drive/MyDrive/Doc/Data2/result_matrix_Total.xlsx', index_col=0)
result_matrix_TemperaturaAnomalia = pd.read_excel('/content/drive/MyDrive/Doc/Data2/result_matrix_TemperaturaAnomalia.xlsx', index_col=0)
result_matrix_IndiceSoberano = pd.read_excel('/content/drive/MyDrive/Doc/Data2/result_matrix_IndiceSoberano.xlsx', index_col=0)
result_matrix_Fallecidos = pd.read_excel('/content/drive/MyDrive/Doc/Data2/result_matrix_Fallecidos.xlsx', index_col=0)
result_matrix_ConvulsionSocial = pd.read_excel('/content/drive/MyDrive/Doc/Data2/result_matrix_ConvulsionSocial.xlsx', index_col=0)
result_matrix_Positivos  = pd.read_excel('/content/drive/MyDrive/Doc/Data2/Result_matrix_Positivos_1.xlsx', index_col=0)

# Cargar la matriz de resultados desde el archivo Excel
#result_matrixTotal_pvalue = pd.read_excel('/content/drive/MyDrive/Doc/Data2/result_matrix_Total_pvalue.xlsx', index_col=0)
#result_matrix_TemperaturaAnomalia_pvalue = pd.read_excel('/content/drive/MyDrive/Doc/Data2/result_matrix_TemperaturaAnomalia_pvalue.xlsx', index_col=0)
#result_matrix_IndiceSoberano_pvalue = pd.read_excel('/content/drive/MyDrive/Doc/Data2/result_matrix_IndiceSoberano_pvalue.xlsx', index_col=0)
#result_matrix_Fallecidos_pvalue = pd.read_excel('/content/drive/MyDrive/Doc/Data2/result_matrix_Fallecidos_pvalue.xlsx', index_col=0)
#result_matrix_ConvulsionSocial_pvalue = pd.read_excel('/content/drive/MyDrive/Doc/Data2/result_matrix_ConvulsionSocial_pvalue.xlsx', index_col=0)
#result_matrix_Positivos_pvalue  = pd.read_excel('/content/drive/MyDrive/Doc/Data2/Result_matrix_Positivos_pvalue.xlsx', index_col=0)

# Convertir la columna de fecha a tipo datetime
#data['FechaOtorgamiento'] = pd.to_datetime(dataset_adicional['FechaOtorgamiento'])
# Convertir la columna de fecha a tipo datetime
data['FechaOtorgamiento'] = pd.to_datetime(data['FechaOtorgamiento'])


# Extraer año y mes
data['year_month'] = data['FechaOtorgamiento'].dt.strftime('%Y-%m')

# Función para obtener el valor correspondiente de la matriz de resultados, aceptando result_matrix como parámetro
def obtener_valor(row, result_matrix):
    codigo = row['CodigoProductoEspecifico']
    year_month = row['year_month']
    if codigo in result_matrix.index and year_month in result_matrix.columns:
        return result_matrix.loc[codigo, year_month]
    else:
        return np.nan  #-1.000 #'NA'

# Aplicar la función al dataset adicional, pasando result_matrix como argumento adicional
#data['valor_FE_Total'] = data.apply(obtener_valor, axis=1, result_type='expand', args=(result_matrixTotal,))
#data['valor_FE_TemperaturaAnomalia'] = data.apply(obtener_valor, axis=1, result_type='expand', args=(result_matrix_TemperaturaAnomalia,))
#data['valor_FE_IndiceSoberano'] = data.apply(obtener_valor, axis=1, result_type='expand', args=(result_matrix_IndiceSoberano,))
#data['valor_FE_result_ConvulsionSocial'] = data.apply(obtener_valor, axis=1, result_type='expand', args=(result_matrix_ConvulsionSocial,))
#data['valor_FE_Covid_Fallecidos'] = data.apply(obtener_valor, axis=1, result_type='expand', args=(result_matrix_Fallecidos,))
#data['valor_FE_Covid_Positivos'] = data.apply(obtener_valor, axis=1, result_type='expand', args=(result_matrix_Positivos,))


#data['valor_FE_TemperaturaAnomalia'] = data['valor_FE_TemperaturaAnomalia'].fillna(-1.0000)
#data['valor_FE_IndiceSoberano'] = data['valor_FE_IndiceSoberano'].fillna(-1.0000)
```

```
#data['valor_FE_result_ConvulsionSocial'] = data['valor_FE_result_ConvulsionSocial'].fillna(-1.0000)
#data['valor_FE_Covid_Fallecidos'] = data['valor_FE_Covid_Fallecidos'].fillna(-1.0000)
#data['valor_FE_Covid_Positivos'] = data['valor_FE_Covid_Positivos'].fillna(-1.0000)
#data['valor_FE_Total'] = data['valor_FE_Total'].fillna(-1.0000)

# Aplicar la función al dataset adicional, pasando result_matrix como argumento adicional _pvalue
#data['valor_FE_Total_pvalue'] = data.apply(obtener_valor, axis=1, result_type='expand', args=(result_matrixTotal_pvalue,))
#data['valor_FE_TemperaturaAnomalia_pvalue'] = data.apply(obtener_valor, axis=1, result_type='expand', args=(result_matrix_TemperaturaAnomalia_pvalue,))
#data['valor_FE_IndiceSoberano_pvalue'] = data.apply(obtener_valor, axis=1, result_type='expand', args=(result_matrix_IndiceSoberano_pvalue,))
#data['valor_FE_result_ConvulsionSocial_pvalue'] = data.apply(obtener_valor, axis=1, result_type='expand', args=(result_matrix_ConvulsionSocial_pvalue,))
#data['valor_FE_Covid_Fallecidos_pvalue'] = data.apply(obtener_valor, axis=1, result_type='expand', args=(result_matrix_Fallecidos_pvalue,))
#data['valor_FE_Covid_Positivos_pvalue'] = data.apply(obtener_valor, axis=1, result_type='expand', args=(result_matrix_Positivos_pvalue,))


#data['valor_FE_TemperaturaAnomalia_pvalue'] = data['valor_FE_TemperaturaAnomalia_pvalue'].fillna(-1.000)
#data['valor_FE_IndiceSoberano_pvalue'] = data['valor_FE_IndiceSoberano_pvalue'].fillna(1.000)
#data['valor_FE_result_ConvulsionSocial_pvalue'] = data['valor_FE_result_ConvulsionSocial_pvalue'].fillna(-1.000)
#data['valor_FE_Covid_Fallecidos_pvalue'] = data['valor_FE_Covid_Fallecidos_pvalue'].fillna(-1.000)
#data['valor_FE_Covid_Positivos_pvalue'] = data['valor_FE_Covid_Positivos_pvalue'].fillna(-1.000)
#data['valor_FE_Total_pvalue'] = data['valor_FE_Total_pvalue'].fillna(-1.000)

#data['valor_FE_TemperaturaAnomalia_pvalue'] = data['valor_FE_TemperaturaAnomalia_pvalue'] * 100.000
#data['valor_FE_IndiceSoberano_pvalue'] = data['valor_FE_IndiceSoberano_pvalue'] * 100.000
#data['valor_FE_result_ConvulsionSocial_pvalue'] = data['valor_FE_result_ConvulsionSocial_pvalue'] * 100.000
#data['valor_FE_Covid_Fallecidos_pvalue'] = data['valor_FE_Covid_Fallecidos_pvalue'] * 100.000
#data['valor_FE_Covid_Positivos_pvalue'] = data['valor_FE_Covid_Positivos_pvalue'] * 100.000
#data['valor_FE_Total_pvalue'] = data['valor_FE_Total_pvalue'] * 100.000




# Mostrar el dataset adicional con la nueva columna
#print(data)

print(type(data))

result_matrixTotal.info()
result_matrix_TemperaturaAnomalia.info()
result_matrix_IndiceSoberano.info()
result_matrix_Fallecidos.info()
result_matrix_ConvulsionSocial.info()
result_matrix_Positivos.info()

import pandas as pd

def merge_dataframes_with_suffix(data, df_convulsionsocial, suffix):
    """
```

```
    Función para unir dos dataframes basados en la columna 'CodigoProductoEspecifico' y renombrar las columnas 'AÑO-MES' añadiendo un sufijo.

    Parameters:
    data (pd.DataFrame): Primer dataframe que contiene la columna 'CodigoProductoEspecifico'.
    df_convulsionsocial (pd.DataFrame): Segundo dataframe cuya columna índice coincide con 'CodigoProductoEspecifico'.
    suffix (str): Sufijo a añadir a las columnas 'AÑO-MES'.

    Returns:
    pd.DataFrame: Dataframe resultante de la unión.
    """

    # Restablecer el índice del segundo DataFrame para que la columna índice se convierta en una columna normal
    df_convulsionsocial.reset_index(inplace=True)

    # Renombrar la columna índice de df_convulsionsocial para que coincida con 'CodigoProductoEspecifico'
    df_convulsionsocial.rename(columns={df_convulsionsocial.columns[0]: 'CodigoProductoEspecifico'}, inplace=True)

    # Renombrar las columnas 'AÑO-MES' añadiendo el sufijo
    new_column_names = {col: col + suffix for col in df_convulsionsocial.columns if col != 'CodigoProductoEspecifico'}
    df_convulsionsocial.rename(columns=new_column_names, inplace=True)

    # Realizar la unión basada en 'CodigoProductoEspecifico'
    merged_data = pd.merge(data, df_convulsionsocial, on='CodigoProductoEspecifico', how='left')

    # Mostrar las primeras filas del dataframe resultante para verificar
    #print(merged_data.head())

    return merged_data

merged_data = data

merged_data = merge_dataframes_with_suffix(merged_data, result_matrix_Fallecidos, 'Fallecidos')

merged_data = merge_dataframes_with_suffix(merged_data, result_matrixTotal, 'Convulsionsocial')
merged_data = merge_dataframes_with_suffix(merged_data, result_matrix_TemperaturaAnomalia, 'TemperaturaAnomalia')

merged_data = merge_dataframes_with_suffix(merged_data, result_matrix_Fallecidos, 'Fallecidos')
merged_data = merge_dataframes_with_suffix(merged_data, result_matrix_ConvulsionSocial, 'Convulsionsocial')

merged_data = merge_dataframes_with_suffix(merged_data, result_matrix_Positivos, 'Positivos')

merged_data.to_csv('/content/drive/MyDrive/Doc/Data2/Data011_with_FE_COVID1.rpt', sep=';', index=False)

merged_data.to_csv('/content/drive/MyDrive/Doc/Data2/Data011_with_FE_COVUL_TEMP1.rpt', sep=';', index=False)

merged_data.info()

merged_data.head()

# Configurar pandas para mostrar todas las filas y columnas
```

```
pd.set_option('display.max_rows', None)  # Para mostrar todas las filas
pd.set_option('display.max_columns', None)  # Para mostrar todas las columnas
pd.set_option('display.width', None)  # Para ajustar el ancho de la consola a la salida del DataFrame
pd.set_option('display.max_colwidth', None)  # Para mostrar contenido de columna completo

codigo_producto_especifico = 10236
valor_2023_07Positivos = merged_data[merged_data['CodigoProductoEspecifico'] == codigo_producto_especifico]['2023-07Positivos'].values

# Mostrar el valor
print(f"El valor de '2023-07Positivos' para CodigoProductoEspecifico {codigo_producto_especifico} es: {valor_2023_07Positivos}")

print(result_matrix_Positivos.info())

numero_filas = result_matrix_Positivos.shape[0]

# Mostrar el número de filas
print(f"El DataFrame tiene {numero_filas} filas.")
```

Anexo D

Dataset enriquecido con datos derivados de *FE COVID-19*, convulsión social y anomalías por factores climáticos

En este anexo se presenta el detalle de los dos primeros registros del Dataset ya enriquecido con las variables derivadas de los *FE COVID 19*, convulsión social y anomalías por factores climáticos.

Tabla 42
Descripción de registros de Dataset enriquecido

Ítem	Columna	Registro 1	Registro 2
1	PlazoReal	32	30
2	Monto	417.05	1395.36
3	ApoyoGobierno	0	0
4	SegmentoCartera	3	3
5	MaxMontoInterno	417.05	1395.36
6	Saldo	417.05	1468.71
7	SalarioNormalizado	$2{,}60 \times 10^{-6}$	$1{,}66 \times 10^{-6}$
8	Bal_TotalActivosNormalizado	0	0
9	Bal_PatrimonioNegocioNormalizado	0	0
10	LiquidezDisponibleNormalizado	0	0
11	NetoIngresosNegocioNormalizado	0	0
12	iAntiguedadBancarizado	36	36
13	ScoreOriginacionMicro	618	901

Continúa en la siguiente página

Tabla 42 – Continúa

Ítem	Columna	Registro 1	Registro 2
14	Score_Sobreendeudamiento	400	836
15	IdOrigenProspeccion	1	1
16	IdDestinoCredito	960	960
17	CodigoProductoEspecifico	10387	411
18	FechaOtorgamiento	27/07/2023	22/11/2022
19	FechaCancelacion		
20	TasaEfectiva	0.8791	0.9632
21	EdadDesembolsoNormalizada	$3{,}19 \times 10^{-6}$	$3{,}44 \times 10^{-6}$
22	EstadoCivil		
23	Profesion	15	33
24	Sexo	2	1
25	Dependientes	0	0
26	NivelInstruccion	1	8
27	Ocupacion	4	12
28	DirDARes	2018	0
29	DirDTRes	2	2
30	DirDTVIV	11	10
31	DirNARes	2018	-1
32	DirNTRes	3	3
33	DirNTVIV	1	5
34	IdOficinaDesembolso	12	10
35	AtrasoMaximo_202001FM		
36	Class_202001BM	-1	-1
37	Class_202001FM	-1	-1
38	Class_202001BMDV	-1	-1
39	Class_202002BM	-1	-1
40	Class_202002FM	-1	-1
41	Class_202002BMDV	-1	-1
42	Class_202003BM	-1	-1
43	Class_202003FM	-1	-1
44	Class_202003BMDV	-1	-1
45	Class_202004BM	-1	-1
46	Class_202004FM	-1	-1
47	Class_202004BMDV	-1	-1
48	Class_202005BM	-1	-1

Continúa en la siguiente página

Tabla 42 – Continúa

Ítem	Columna	Registro 1	Registro 2
49	Class_202005FM	-1	-1
50	Class_202005BMDV	-1	-1
51	Class_202006BM	-1	-1
52	Class_202006FM	-1	-1
53	Class_202006BMDV	-1	-1
54	Class_202007BM	-1	-1
55	Class_202007FM	-1	-1
56	Class_202007BMDV	-1	-1
57	Class_202008BM	-1	-1
58	Class_202008FM	-1	-1
59	Class_202008BMDV	-1	-1
60	Class_202009BM	-1	-1
61	Class_202009FM	-1	-1
62	Class_202009BMDV	-1	-1
63	Class_202010BM	-1	-1
64	Class_202010FM	-1	-1
65	Class_202010BMDV	-1	-1
66	Class_202011BM	-1	-1
67	Class_202011FM	-1	-1
68	Class_202011BMDV	-1	-1
69	Class_202012BM	-1	-1
70	Class_202012FM	-1	-1
71	Class_202012BMDV	-1	-1
72	Class_202101BM	-1	-1
73	Class_202101FM	-1	-1
74	Class_202101BMDV	-1	-1
75	Class_202102BM	-1	-1
76	Class_202102FM	-1	-1
77	Class_202102BMDV	-1	-1
78	Class_202103BM	-1	-1
79	Class_202103FM	-1	-1
80	Class_202103BMDV	-1	-1
81	Class_202104BM	-1	-1
82	Class_202104FM	-1	-1
83	Class_202104BMDV	-1	-1

Continúa en la siguiente página

Tabla 42 – Continúa

Ítem	Columna	Registro 1	Registro 2
84	Class_202105BM	-1	-1
85	Class_202105FM	-1	-1
86	Class_202105BMDV	-1	-1
87	Class_202106BM	-1	-1
88	Class_202106FM	-1	-1
89	Class_202106BMDV	-1	-1
90	Class_202107BM	-1	-1
91	Class_202107FM	-1	-1
92	Class_202107BMDV	-1	-1
93	Class_202108BM	-1	-1
94	Class_202108FM	-1	-1
95	Class_202108BMDV	-1	-1
96	Class_202109BM	-1	-1
97	Class_202109FM	-1	-1
98	Class_202109BMDV	-1	-1
99	Class_202110BM	-1	-1
100	Class_202110FM	-1	-1
101	Class_202110BMDV	-1	-1
102	Class_202111BM	-1	-1
103	Class_202111FM	-1	-1
104	Class_202111BMDV	-1	-1
105	Class_202112BM	-1	-1
106	Class_202112FM	-1	-1
107	Class_202112BMDV	-1	-1
108	Class_202201BM	-1	-1
109	Class_202201FM	-1	-1
110	Class_202201BMDV	-1	-1
111	Class_202202BM	-1	-1
112	Class_202202FM	-1	-1
113	Class_202202BMDV	-1	-1
114	Class_202203BM	-1	-1
115	Class_202203FM	-1	-1
116	Class_202203BMDV	-1	-1
117	Class_202204BM	-1	-1
118	Class_202204FM	-1	-1

Continúa en la siguiente página

Tabla 42 – Continúa

Ítem	Columna	Registro 1	Registro 2
119	Class_202204BMDV	-1	-1
120	Class_202205BM	-1	-1
121	Class_202205FM	-1	-1
122	Class_202205BMDV	-1	-1
123	Class_202206BM	-1	-1
124	Class_202206FM	-1	-1
125	Class_202206BMDV	-1	-1
126	Class_202207BM	-1	-1
127	Class_202207FM	-1	-1
128	Class_202207BMDV	-1	-1
129	Class_202208BM	-1	-1
130	Class_202208FM	-1	-1
131	Class_202208BMDV	-1	-1
132	Class_202209BM	-1	-1
133	Class_202209FM	-1	-1
134	Class_202209BMDV	-1	-1
135	Class_202210BM	-1	-1
136	Class_202210FM	-1	-1
137	Class_202210BMDV	-1	-1
138	AtrasoMaximo_202211FM		
139	Class_202211BM	-1	-1
140	Class_202211FM	-1	-1
141	Class_202211BMDV	-1	-1
142	Class_202212BM	-1	-1
143	Class_202212FM	-1	-1
144	Class_202212BMDV	-1	-1
145	Class_202301BM	-1	-1
146	Class_202301FM	-1	-1
147	Class_202301BMDV	-1	-1
148	Class_202302BM	-1	-1
149	Class_202302FM	-1	-1
150	Class_202302BMDV	-1	-1
151	Class_202303BM	-1	-1
152	Class_202303FM	-1	-1
153	Class_202303BMDV	-1	-1

Continúa en la siguiente página

Tabla 42 – Continúa

Ítem	Columna	Registro 1	Registro 2
154	Class_202304BM	-1	-1
155	Class_202304FM	-1	-1
156	Class_202304BMDV	-1	-1
157	Class_202305BM	-1	1
158	Class_202305FM	-1	1
159	Class_202305BMDV	-1	1
160	Class_202306BM	-1	-1
161	Class_202306FM	-1	-1
162	Class_202306BMDV	-1	-1
163	Class_202307BM	1	-1
164	Class_202307FM	1	-1
165	Class_202307BMDV	1	-1
166	Class_202308BM	1	-1
167	Class_202308FM	1	-1
168	Class_202308BMDV	1	-1
169	Class_202309BM	-1	-1
170	Class_202309FM	-1	-1
171	Class_202309BMDV	-1	-1
172	year_month	2023-07	2022-11
173	2020-01Fallecidos	1	0
174	2020-02Fallecidos	0	0
175	2020-03Fallecidos	0	0
176	2020-04Fallecidos	1	1
177	2020-05Fallecidos	1	0
178	2020-06Fallecidos	0	1
179	2020-07Fallecidos	1	1
180	2020-08Fallecidos	1	1
181	2020-09Fallecidos	1	1
182	2020-10Fallecidos	1	0
183	2020-11Fallecidos	0	1
184	2020-12Fallecidos	1	0
185	2021-01Fallecidos	0	1
186	2021-02Fallecidos	0	0
187	2021-03Fallecidos	1	0
188	2021-04Fallecidos	1	0

Continúa en la siguiente página

Tabla 42 – Continúa

Ítem	Columna	Registro 1	Registro 2
189	2021-05Fallecidos	1	0
190	2021-06Fallecidos	1	1
191	2021-07Fallecidos	1	0
192	2021-08Fallecidos	0	1
193	2021-09Fallecidos	0	0
194	2021-10Fallecidos	1	0
195	2021-11Fallecidos	1	0
196	2021-12Fallecidos	1	0
197	2022-01Fallecidos	1	1
198	2022-02Fallecidos	1	0
199	2022-03Fallecidos	0	0
200	2022-04Fallecidos	0	0
201	2022-05Fallecidos	0	0
202	2022-06Fallecidos	0	0
203	2022-07Fallecidos	1	0
204	2022-08Fallecidos	1	0
205	2022-09Fallecidos	1	1
206	2022-10Fallecidos	0	0
207	2022-11Fallecidos	1	0
208	2022-12Fallecidos	1	0
209	2023-01Fallecidos	1	0
210	2023-02Fallecidos	0	1
211	2023-03Fallecidos	0	0
212	2023-04Fallecidos	1	0
213	2023-05Fallecidos	0	0
214	2023-06Fallecidos	1	0
215	2023-07Fallecidos	0	0
216	2023-08Fallecidos	1	0
217	2023-09Fallecidos	1	0
218	2020-01Positivos		
219	2020-02Positivos		
220	2020-03Positivos	1	1
221	2020-04Positivos	1	1
222	2020-05Positivos	1	0
223	2020-06Positivos	0	0

Continúa en la siguiente página

Tabla 42 – Continúa

Ítem	Columna	Registro 1	Registro 2
224	2020-07Positivos	1	1
225	2020-08Positivos	0	0
226	2020-09Positivos	1	1
227	2020-10Positivos	0	0
228	2020-11Positivos	0	0
229	2020-12Positivos	0	0
230	2021-01Positivos	0	0
231	2021-02Positivos	0	0
232	2021-03Positivos	1	1
233	2021-04Positivos	0	1
234	2021-05Positivos	1	1
235	2021-06Positivos	0	1
236	2021-07Positivos	1	0
237	2021-08Positivos	1	0
238	2021-09Positivos	0	0
239	2021-10Positivos	0	1
240	2021-11Positivos	0	0
241	2021-12Positivos	0	1
242	2022-01Positivos	0	0
243	2022-02Positivos	1	0
244	2022-03Positivos	0	0
245	2022-04Positivos	0	1
246	2022-05Positivos	1	1
247	2022-06Positivos	0	0
248	2022-07Positivos	0	1
249	2022-08Positivos	1	0
250	2022-09Positivos	0	0
251	2022-10Positivos	0	0
252	2022-11Positivos	0	1
253	2022-12Positivos	0	0
254	2023-01Positivos	0	0
255	2023-02Positivos	1	0
256	2023-03Positivos	1	1
257	2023-04Positivos	1	0
258	2023-05Positivos	0	0

Continúa en la siguiente página

Tabla 42 – Continúa

Ítem	Columna	Registro 1	Registro 2
259	2023-06Positivos	0	0
260	2023-07Positivos	1	0
261	2023-08Positivos	1	0
262	2023-09Positivos	0	0
263	2020-03TemperaturaAnomalia	1	1
264	2020-06TemperaturaAnomalia		
265	2020-09TemperaturaAnomalia		
266	2020-12TemperaturaAnomalia	1	0
267	2021-03TemperaturaAnomalia	1	0
268	2021-06TemperaturaAnomalia	0	1
269	2021-09TemperaturaAnomalia	0	0
270	2021-12TemperaturaAnomalia	1	1
271	2022-03TemperaturaAnomalia	0	0
272	2022-06TemperaturaAnomalia	0	0
273	2022-09TemperaturaAnomalia	0	0
274	2022-12TemperaturaAnomalia	0	1
275	2023-03TemperaturaAnomalia	1	0
276	2023-06TemperaturaAnomalia	0	1
277	2023-09TemperaturaAnomalia	0	1
278	2020-01Convulsionsocial		
279	2020-02Convulsionsocial		
280	2020-03Convulsionsocial		
281	2020-04Convulsionsocial		
282	2020-05Convulsionsocial		
283	2020-06Convulsionsocial		
284	2020-07Convulsionsocial		
285	2020-08Convulsionsocial		
286	2020-09Convulsionsocial		
287	2020-10Convulsionsocial		
288	2020-11Convulsionsocial		
289	2020-12Convulsionsocial		
290	2021-01Convulsionsocial		
291	2021-02Convulsionsocial		
292	2021-03Convulsionsocial		
293	2021-04Convulsionsocial		

Continúa en la siguiente página

Tabla 42 – Continúa

Ítem	Columna	Registro 1	Registro 2
294	2021-05Convulsionsocial		
295	2021-06Convulsionsocial		
296	2021-07Convulsionsocial		
297	2021-08Convulsionsocial		
298	2021-09Convulsionsocial		
299	2021-10Convulsionsocial		
300	2021-11Convulsionsocial		
301	2021-12Convulsionsocial		
302	2022-01Convulsionsocial		
303	2022-02Convulsionsocial		
304	2022-03Convulsionsocial		
305	2022-04Convulsionsocial		
306	2022-05Convulsionsocial		
307	2022-06Convulsionsocial	0	0
308	2022-07Convulsionsocial	0	1
309	2022-08Convulsionsocial	0	0
310	2022-09Convulsionsocial	0	0
311	2022-10Convulsionsocial	0	0
312	2022-11Convulsionsocial	0	1
313	2022-12Convulsionsocial	0	0
314	2023-01Convulsionsocial	1	1
315	2023-02Convulsionsocial	1	1
316	2023-03Convulsionsocial	1	0
317	2023-04Convulsionsocial	0	0
318	2023-05Convulsionsocial	0	0
319	2023-06Convulsionsocial	1	0
320	2023-07Convulsionsocial	0	0
321	2023-08Convulsionsocial	0	0
322	2023-09Convulsionsocial	0	0
323	2020-01BloqueoFClimatico		
324	2020-02BloqueoFClimatico		
325	2020-03BloqueoFClimatico		
326	2020-04BloqueoFClimatico		
327	2020-05BloqueoFClimatico		
328	2020-06BloqueoFClimatico		

Continúa en la siguiente página

Tabla 42 – Continúa

Ítem	Columna	Registro 1	Registro 2
329	2020-07BloqueoFClimatico		
330	2020-08BloqueoFClimatico		
331	2020-09BloqueoFClimatico		
332	2020-10BloqueoFClimatico		
333	2020-11BloqueoFClimatico		
334	2020-12BloqueoFClimatico		
335	2021-01BloqueoFClimatico		
336	2021-02BloqueoFClimatico		
337	2021-03BloqueoFClimatico		
338	2021-04BloqueoFClimatico		
339	2021-05BloqueoFClimatico		
340	2021-06BloqueoFClimatico		
341	2021-07BloqueoFClimatico		
342	2021-08BloqueoFClimatico		
343	2021-09BloqueoFClimatico		
344	2021-10BloqueoFClimatico		
345	2021-11BloqueoFClimatico		
346	2021-12BloqueoFClimatico		
347	2022-01BloqueoFClimatico		
348	2022-02BloqueoFClimatico		
349	2022-03BloqueoFClimatico		
350	2022-04BloqueoFClimatico		
351	2022-05BloqueoFClimatico		
352	2022-06BloqueoFClimatico	0	0
353	2022-07BloqueoFClimatico	0	0
354	2022-08BloqueoFClimatico	0	0
355	2022-09BloqueoFClimatico	0	1
356	2022-10BloqueoFClimatico	1	0
357	2022-11BloqueoFClimatico	0	0
358	2022-12BloqueoFClimatico	0	0
359	2023-01BloqueoFClimatico	0	0
360	2023-02BloqueoFClimatico	1	0
361	2023-03BloqueoFClimatico	0	0
362	2023-04BloqueoFClimatico	1	0
363	2023-05BloqueoFClimatico	0	0

Continúa en la siguiente página

Tabla 42 – Continúa

Ítem	Columna	Registro 1	Registro 2
364	2023-06BloqueoFClimatico	0	1
365	2023-07BloqueoFClimatico	0	0
366	2023-08BloqueoFClimatico	1	1
367	2023-09BloqueoFClimatico	1	1

Anexo E

Código fuente del proceso de entrenamiento y validación para *FE COVID-19*

En este anexo se presenta el código fuente completo utilizado para el entrenamiento y validación de modelos de *ML* aplicados al escenario de *COVID 19*, correspondiente al dataset final. El código está implementado en Python y documentado para su comprensión y posible reproducción.

Archivo: `Class_usando_ml_covid.py`

```
# -*- coding: utf-8 -*-
"""Clasificación usando ML_MOV_v1_FINAL.ipynb

Automatically generated by Colab.

Original file is located at
    https://colab.research.google.com/drive/1bLBGsJDwoXc1pNrbWXbUjw37yffyEhOC
"""

#Paso 1 Conectar a Drive
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.stattools import adfuller, coint

import pandas as pd

from google.colab import drive
drive.mount('/content/drive')

# Función para convertir los valores a datetime
def convert_to_datetime(val):
    if isinstance(val, str):
        return pd.to_datetime(val)
    elif isinstance(val, (int, float)):
        return pd.to_datetime(val, unit='s')
    else:
        return pd.NaT  # Not a Time para valores que no se pueden convertir

```

```
#Paso 2 Definiciones Actividades economicas por segmentos de importancia

#Filtrar 80%
Filtro80 = [411,10366,10324,403,10453,10364,10354,10352,10365,10457,10344,10384,
10319,10396,412,10443,10463,10441,10449,10436,516,10320,10382,10478,10414,134,
10316,405,10408,10322,10464,10353,10340,10370,10386,10327,10337,10416,10465,
10474,10306,10400,10162,10376,10350,410,10401,10425,10367,10358,10459,10371,
10387,10378,10448,10309,10447,408,10388,10467,10326,10317,10380,10343,10375,
10356,10164,10452,10342,938,10334,10300,10341,10328,10351,10379,10395,5220]


#Filtrar 99%
Filtro99 = [411,10366,10324,403,10453,10364,10354,10352,10365,10457,10344,
10384,10319,10396,412,10443,10463,10441,10449,10436,516,10320,10382,10478,
10414,134,10316,405,10408,10322,10464,10353,10340,10370,10386,10327,10337,
10416,10465,10474,10306,10400,10162,10376,10350,410,10401,10425,10367,10358,
10459,10371,10387,10378,10448,10309,10447,408,10388,10467,10326,10317,10380,
10343,10375,10356,10164,10452,10342,938,10334,10300,10341,10328,10351,10379,
10395,5220,10115,10437,10170,10024,10377,10147,10472,10438,10445,10362,10460,
10374,10308,10318,10369,10399,10415,407,10458,10479,10477,10423,10181,9309,
10383,10361,10037,10475,513,10169,10346,10372,10473,10143,10599,10182,10336,
10468,10469,10397,512,138,10456,10461,10093,10454,10301,10241,10424,10311,
10339,10368,10393,514,10360,10026,10481,10471,10000,10303,10329,10141,5270,
10355,10484,10455,10435,10157,10349,10107,10120,10347,10432,10136,10149,
10363,10331,10433,10406,10389,10011,10394,10232,10526,10345,517,10496,10084,
10450,10315,10323,10403,10392,10185,10476,10390,10466,10207,8519,10505,10385,
10357,10407,5233,10411,10307,10426,10434,10150,10012,10597,10595,10314,10440,
10175,10332,10439,10604,10154,10062,10075,10193,7499,10422,10163,10421,10480,
9999,10335,10310,10158,10409,10572,10271,10113,10430,10381,10417,10517,10105,
10140,10278,10305,10609,10330,10470,10442,10236,10418,5232,10176,10130,10398,
10131,1511,10173,10263,5520,10259,10172,10222,10238,10338,401,1721,10168,10089,
10189,10040,10110,10495,10359,10013,10404,10427,10554,10165,10272,10085,10183,
10289,10542,10212,10312,5131,10284,10325,10348,10419,135,5234,10391,10498,5020,
10030,10159,5239,10156,5030,6023,10138,10155,10523,10009,10446,10246,122,10191,
10413,10483,10086,10260,5260,7421,10055,10152,10010,1712,10124,10015,10121,10431,
7412,10022,7491,10087,10053,10462,10091,10114,10214,10313,10304,10412,10405,1549,
5211,10023,10041,10145,10205,3720,7511,10103,10116,10134,10555,10594]

#Paso 3: Cargar y preparar los datos

# Asegúrate de usar la ruta correcta a tu archivo CSV
#file_path = '/content/drive/MyDrive/Doc/Data2/2023_09_v2.rpt'
#file_path = '/content/drive/MyDrive/Doc/Data2/Data007.rpt'
#file_path = '/content/drive/MyDrive/Doc/Data2/Data008_with_FE_COVID1.rpt'
#file_path = '/content/drive/MyDrive/Doc/Data2/Data010_with_FE.rpt'
file_path = '/content/drive/MyDrive/Doc/Data2/Data011_with_FE.rpt'
data = pd.read_csv(file_path, delimiter = ';')

# Asegúrate de usar el formato de fecha correcto
# Suponiendo que la columna de fecha se llama 'date_column'
#data['FechaOtorgamiento'] = pd.to_datetime(data['FechaOtorgamiento'])
data['FechaOtorgamiento'] = data['FechaOtorgamiento'].apply(convert_to_datetime)

#Definir el Rango de Fechas
import datetime


#Declaración del estado de emergencia        16 de marzo de 2020- 28 de octubre de
↪  2022 (2 años, 7 meses y 12 días)
fecha_inicio = datetime.datetime.strptime("2020-04-01", "%Y-%m-%d")  # Fecha de
↪  inicio
fecha_fin = datetime.datetime.strptime("2021-10-31", "%Y-%m-%d")  # Fecha de fin

#Filtrar Rango Creditos que han sido otorgados antes de la fecha fin
#data = data[(data['FechaOtorgamiento'] >= fecha_inicio) &
↪  (data['FechaOtorgamiento'] <= fecha_fin)]
data = data[(data['FechaOtorgamiento'] <= fecha_fin)]
```

```
93
94  data.info()
95
96  #Paso 4: Elimina creditos cancelados antes del periodo de evaluación
97
98  data['FechaCancelacion'] = data['FechaCancelacion'].apply(convert_to_datetime)
99  data = data[(data['FechaCancelacion'] <= fecha_inicio) |
    ↪  (data['FechaCancelacion'].isna())]
100
101 #Paso 5: Elimina creditos con actividad economica incorrecta y creditos que en mes
    ↪  de inicio tengan ya un atraso mayor a 30 dias
102
103 data['CodigoProductoEspecifico'] = pd.to_numeric(data['CodigoProductoEspecifico'],
    ↪  errors='coerce')
104 data = data[data['CodigoProductoEspecifico'].notna()]
105 data = data[data['CodigoProductoEspecifico'].notna() &
    ↪  (data['CodigoProductoEspecifico'] != '')]
106 data = data[(data['AtrasoMaximo_202001FM'] < 30) |
    ↪  (data['AtrasoMaximo_202001FM'].isna())]
107
108 #data = data[(data['ApoyoGobierno'] == 0)]
109
110 #Paso 6: Procesar los datos
111 # Crear los subconjuntos
112 data100 = data.copy()
113 data99 = data.copy()
114 data80 = data.copy()
115
116 data99 = data99[data99['CodigoProductoEspecifico'].isin(Filtro99)]
117 data80 = data80[data80['CodigoProductoEspecifico'].isin(Filtro80)]
118
119 # Mostrar la información de los datos procesados
120 data100.info()
121 data99.info()
122 data80.info()
123
124 #Paso 7: Graficar Numero de creditos afectados con FE y Clasificacion Mensual
125
126 import pandas as pd
127 import matplotlib.pyplot as plt
128
129 def plot_class_fallecidos_positivos(df, tible):
130     """
131     Grafica la cantidad de valores 0 en las columnas 'Class_YYYYMMBM',
132     la cantidad de valores 1 en las columnas 'YYYY-MMFallecidos',
133     y la cantidad de valores 1 en las columnas 'YYYY-MMPositivos'.
134
135     Parameters:
136     df (pd.DataFrame): DataFrame que contiene las columnas a analizar.
137     """
138     # Filtrar columnas que cumplen con las condiciones
139     class_bm_columns = [col for col in df.columns if col.startswith('Class_') and
    ↪  col.endswith('BM')]
140     class_bm_columnsDV = [col for col in df.columns if col.startswith('Class_') and
    ↪  col.endswith('BMDV')]
141     fallecidos_columns = [col for col in df.columns if '-' in col and
    ↪  col.endswith('Fallecidos')]
142     positivos_columns = [col for col in df.columns if '-' in col and
    ↪  col.endswith('Positivos')]
143
144     # Inicializar diccionarios para almacenar los resultados
145     class_bm_counts = {}
146     class_bm_countsDV = {}
147     fallecidos_counts = {}
148     positivos_counts = {}
149
150     # Procesar columnas Class_YYYYMMBM
```

```
    for col in class_bm_columns:
        mes = col.split('_')[1][:6]  # Extraer AÑOMES
        count_zeros = (df[col] == 0).sum()  # Contar valores 0
        if mes in class_bm_counts:
            class_bm_counts[mes] += count_zeros
        else:
            class_bm_counts[mes] = count_zeros

    # Procesar columnas Class_YYYYMMBM
    for col in class_bm_columnsDV:
        mes = col.split('_')[1][:6]  # Extraer AÑOMES
        count_zeros = (df[col] == 0).sum()  # Contar valores 0
        #print(mes, count_zeros)
        if mes in class_bm_countsDV:
            class_bm_countsDV[mes] += count_zeros
        else:
            class_bm_countsDV[mes] = count_zeros

    # Procesar columnas YYYY-MMFallecidos
    for col in fallecidos_columns:
        parts = col.split('-')
        if len(parts) == 2:
            mes = parts[0] + parts[1][:2]  # Convertir AÑO-MES a AÑOMES
            count_ones = (df[col] == 1).sum()  # Contar valores 1
            if mes in fallecidos_counts:
                fallecidos_counts[mes] += count_ones
            else:
                fallecidos_counts[mes] = count_ones

    # Procesar columnas YYYY-MMPositivos
    for col in positivos_columns:
        parts = col.split('-')
        if len(parts) == 2:
            mes = parts[0] + parts[1][:2]  # Convertir AÑO-MES a AÑOMES
            count_ones = (df[col] == 1).sum()  # Contar valores 1
            if mes in positivos_counts:
                positivos_counts[mes] += count_ones
            else:
                positivos_counts[mes] = count_ones

    # Crear listas ordenadas de meses y conteos
    all_months = sorted(set(class_bm_counts.keys()).
        union(set(fallecidos_counts.keys())).
        union(set(positivos_counts.keys())).
        union(set(class_bm_countsDV.keys())))
    class_bm_counts_list = [class_bm_counts.get(mes, 0) for mes in all_months]
    class_bm_counts_listDV = [class_bm_countsDV.get(mes, 0) for mes in all_months]
    fallecidos_counts_list = [fallecidos_counts.get(mes, 0) for mes in all_months]
    positivos_counts_list = [positivos_counts.get(mes, 0) for mes in all_months]

    # Crear la gráfica
    plt.figure(figsize=(14, 7))
    plt.plot(all_months, class_bm_counts_list, label='Cantidad de 0 en
    ↪  Class_YYYYMMBM', marker='o')
    plt.plot(all_months, class_bm_counts_listDV, label='Cantidad de 0 en
    ↪  Class_DV_YYYYMMBM', marker='o')
    plt.plot(all_months, fallecidos_counts_list, label='Cantidad de 1 en
    ↪  YYYY-MMFallecidos', marker='o')
    plt.plot(all_months, positivos_counts_list, label='Cantidad de 1 en
    ↪  YYYY-MMPositivos', marker='o')
    plt.xlabel('AÑOMES')
    plt.ylabel('Cantidad')
    plt.title(tible + ' - Cantidad de 0 en Class_YYYYMMBM, en Class_DV_YYYYMMBM, 1
    ↪  en YYYY-MMFallecidos y YYYY-MMPositivos')
    plt.legend()
    plt.xticks(rotation=45)
    plt.grid(True)
```

```
    plt.show()

# Ejemplo de uso
# Supongamos que tienes un dataframe llamado df
# df = pd.read_csv('tu_archivo.csv')
df = data100.copy()
# Llamar a la función para crear la gráfica
plot_class_fallecidos_positivos(df, 'Data100')

df = data99.copy()
# Llamar a la función para crear la gráfica
plot_class_fallecidos_positivos(df, 'Data99')

df = data80.copy()
# Llamar a la función para crear la gráfica
plot_class_fallecidos_positivos(df, 'Data80')

#Funcion retraza la secuencia de fallecidos en 1 periodo
import pandas as pd

def shift_fallecidos_columns(df, prefix='Fallecidos'):
    """
    Función para reasignar valores en columnas con el formato 'AÑO-MESFallecidos'.
    Copia los valores de cada columna a la columna siguiente en orden cronológico.

    Parameters:
    df (pd.DataFrame): DataFrame que contiene las columnas con el formato
    ↪  'AÑO-MESFallecidos'.
    prefix (str): Sufijo que identifica las columnas a ser reasignadas. Por defecto
    ↪  es 'Fallecidos'.

    Returns:
    pd.DataFrame: DataFrame con los valores de las columnas reasignados.
    """
    # Crear una lista con los nombres de las columnas que contienen el prefijo
    fallecidos_columns = [col for col in df.columns if prefix in col]

    # Ordenar las columnas cronológicamente
    fallecidos_columns.sort()

    # Realizar la copia de valores
    for i in range(len(fallecidos_columns) - 1, 0, -1):
        df[fallecidos_columns[i]] = df[fallecidos_columns[i - 1]]

    return df

#Paso 8: Mueve la curva fallecidos a la derecha

data100 = shift_fallecidos_columns(data100)
data99 = shift_fallecidos_columns(data99)
data80 = shift_fallecidos_columns(data80)

#Paso 9: Graficar Numero de creditos afectados con FE y Clasificacion Mensual
#FE Fallecido movido a la derecha 1 periodo

import pandas as pd
import matplotlib.pyplot as plt

def plot_class_fallecidos_positivos(df, tible):
    """
    Grafica la cantidad de valores 0 en las columnas 'Class_YYYYMMBM',
    la cantidad de valores 1 en las columnas 'YYYY-MMFallecidos',
    y la cantidad de valores 1 en las columnas 'YYYY-MMPositivos'.

    Parameters:
    df (pd.DataFrame): DataFrame que contiene las columnas a analizar.
    """
```

```
    # Filtrar columnas que cumplen con las condiciones
    class_bm_columns = [col for col in df.columns if col.startswith('Class_') and
    ↪  col.endswith('BM')]
    class_bm_columnsDV = [col for col in df.columns if col.startswith('Class_') and
    ↪  col.endswith('BMDV')]
    fallecidos_columns = [col for col in df.columns if '-' in col and
    ↪  col.endswith('Fallecidos')]
    positivos_columns = [col for col in df.columns if '-' in col and
    ↪  col.endswith('Positivos')]

    # Inicializar diccionarios para almacenar los resultados
    class_bm_counts = {}
    class_bm_countsDV = {}
    fallecidos_counts = {}
    positivos_counts = {}

    # Procesar columnas Class_YYYYMMBM
    for col in class_bm_columns:
        mes = col.split('_')[1][:6]  # Extraer AÑOMES
        count_zeros = (df[col] == 0).sum()  # Contar valores 0
        if mes in class_bm_counts:
            class_bm_counts[mes] += count_zeros
        else:
            class_bm_counts[mes] = count_zeros

    # Procesar columnas Class_YYYYMMBM
    for col in class_bm_columnsDV:
        mes = col.split('_')[1][:6]  # Extraer AÑOMES
        count_zeros = (df[col] == 0).sum()  # Contar valores 0
        #print(mes, count_zeros)
        if mes in class_bm_countsDV:
            class_bm_countsDV[mes] += count_zeros
        else:
            class_bm_countsDV[mes] = count_zeros

    # Procesar columnas YYYY-MMFallecidos
    for col in fallecidos_columns:
        parts = col.split('-')
        if len(parts) == 2:
            mes = parts[0] + parts[1][:2]  # Convertir AÑO-MES a AÑOMES
            count_ones = (df[col] == 1).sum()  # Contar valores 1
            if mes in fallecidos_counts:
                fallecidos_counts[mes] += count_ones
            else:
                fallecidos_counts[mes] = count_ones

    # Procesar columnas YYYY-MMPositivos
    for col in positivos_columns:
        parts = col.split('-')
        if len(parts) == 2:
            mes = parts[0] + parts[1][:2]  # Convertir AÑO-MES a AÑOMES
            count_ones = (df[col] == 1).sum()  # Contar valores 1
            if mes in positivos_counts:
                positivos_counts[mes] += count_ones
            else:
                positivos_counts[mes] = count_ones

    # Crear listas ordenadas de meses y conteos
    all_months = sorted(set(class_bm_counts.keys()).
        union(set(fallecidos_counts.keys())).
        union(set(positivos_counts.keys())).
        union(set(class_bm_countsDV.keys())))
    class_bm_counts_list = [class_bm_counts.get(mes, 0) for mes in all_months]
    class_bm_counts_listDV = [class_bm_countsDV.get(mes, 0) for mes in all_months]
    fallecidos_counts_list = [fallecidos_counts.get(mes, 0) for mes in all_months]
    positivos_counts_list = [positivos_counts.get(mes, 0) for mes in all_months]

```

```
    # Crear la gráfica
    plt.figure(figsize=(14, 7))
    plt.plot(all_months, class_bm_counts_list, label='Cantidad de 0 en Class_YYYYMMBM', marker='o')
    plt.plot(all_months, class_bm_counts_listDV, label='Cantidad de 0 en Class_DV_YYYYMMBM', marker='o')
    plt.plot(all_months, fallecidos_counts_list, label='Cantidad de 1 en YYYY-MMFallecidos', marker='o')
    plt.plot(all_months, positivos_counts_list, label='Cantidad de 1 en YYYY-MMPositivos', marker='o')
    plt.xlabel('AÑOMES')
    plt.ylabel('Cantidad')
    plt.title(tible + ' - Cantidad de 0 en Class_YYYYMMBM, en Class_DV_YYYYMMBM, 1 en YYYY-MMFallecidos y YYYY-MMPositivos')
    plt.legend()
    plt.xticks(rotation=45)
    plt.grid(True)
    plt.show()

# Ejemplo de uso
# Supongamos que tienes un dataframe llamado df
# df = pd.read_csv('tu_archivo.csv')
df = data100.copy()
# Llamar a la función para crear la gráfica
plot_class_fallecidos_positivos(df, 'Data100')

df = data99.copy()
# Llamar a la función para crear la gráfica
plot_class_fallecidos_positivos(df, 'Data99')

df = data80.copy()
# Llamar a la función para crear la gráfica
plot_class_fallecidos_positivos(df, 'Data80')

#Paso 10: Agrega la variable Class que indica si en el periodo en evaluacion se presentaron impacto

import pandas as pd
import numpy as np

def filter_class_bm_columns(df, start_date, end_date):
    """
    Filtra las columnas del DataFrame que empiezan con "Class_" y terminan en "BM"
    dentro del rango de fechas especificado y determina el valor de ClassBM.

    Parameters:
    df (pd.DataFrame): DataFrame que contiene las columnas a filtrar.
    start_date (str): Fecha de inicio en formato "YYYY-MM".
    end_date (str): Fecha de fin en formato "YYYY-MM".

    Returns:
    pd.DataFrame: DataFrame con la nueva columna ClassBM.
    """
    # Convertir las fechas de inicio y fin a enteros para la comparación
    start_date_int = int(start_date.replace("-", ""))
    end_date_int = int(end_date.replace("-", ""))

    # Filtrar columnas que empiezan con "Class_" y terminan en "BM" dentro del rango de fechas
    class_bm_columns = [col for col in df.columns if col.startswith('Class_') and col.endswith('BM')
                        and start_date_int <= int(col[6:12]) <= end_date_int]

    # Función para determinar el valor de ClassBM
    def determine_class_bm(row):
        if (row[class_bm_columns] == 0).any():
            return 0
```

```
        elif (row[class_bm_columns] == 1).any():
            return 1
        else:
            return -1

    # Aplicar la función para crear la nueva columna ClassBM
    df['ClassBM'] = df.apply(determine_class_bm, axis=1)
    return df

# Ejemplo de uso
# Supongamos que tienes DataFrames llamados data100, data99, y data80
start_date = "2020-04"
end_date = "2022-01"

data100 = filter_class_bm_columns(data100, start_date, end_date)
data99 = filter_class_bm_columns(data99, start_date, end_date)
data80 = filter_class_bm_columns(data80, start_date, end_date)

# Mostrar las primeras filas del DataFrame para verificar
#print(data100.head())

#Paso 11: Grafica histograma de la variable agregada


# Asegúrate de que tienes las bibliotecas necesarias instaladas
!pip install pandas matplotlib

import pandas as pd
import matplotlib.pyplot as plt

import pandas as pd
import matplotlib.pyplot as plt

# Suponiendo que tienes un dataframe llamado df


# Función para crear histogramas apilados
def plot_stacked_histogram_Binary(df, date_column, binary_column, title):
    # Asegúrate de que la columna 'FechaOtorgamiento' esté en formato datetime
    df['FechaOtorgamiento'] = pd.to_datetime(df['FechaOtorgamiento'])

    # Agrupa por fecha y cuenta los valores de la columna binaria
    df_grouped = df.groupby([df[date_column].dt.date,
    ↪  binary_column]).size().unstack(fill_value=0)

    # Normaliza para que sumen 100%
    df_grouped = df_grouped.div(df_grouped.sum(axis=1), axis=0)

    # Traza el histograma apilado
    ax = df_grouped.plot(kind='bar', stacked=True, color=['blue', 'orange'],
    ↪  figsize=(14, 7))

    # Añadir etiquetas y título
    ax.set_xlabel('Fecha de Otorgamiento')
    ax.set_ylabel('Proporción')
    ax.set_title(title)
    ax.legend(title=binary_column, bbox_to_anchor=(1.05, 1), loc='upper left')

    plt.show()

import pandas as pd
import matplotlib.pyplot as plt

# Función para crear histogramas apilados
#Tres valores -1, 0, 1
def plot_stacked_histogram(df, date_column, class_column, title):
    """
```

```
    Función para crear histogramas apilados con una columna de clasificación que contiene valores -1, 0 y 1.

    Parameters:
    df (pd.DataFrame): DataFrame que contiene los datos.
    date_column (str): Nombre de la columna de fechas.
    class_column (str): Nombre de la columna de clasificación.
    title (str): Título del gráfico.
    """

    # Asegúrate de que la columna de fechas esté en formato datetime
    df[date_column] = pd.to_datetime(df[date_column])

    # Agrupa por fecha y cuenta los valores de la columna de clasificación
    df_grouped = df.groupby([df[date_column].dt.date, class_column]).size().unstack(fill_value=0)

    # Normaliza para que sumen 100%
    df_grouped = df_grouped.div(df_grouped.sum(axis=1), axis=0)

    # Traza el histograma apilado
    ax = df_grouped.plot(kind='bar', stacked=True, color=['blue', 'orange', 'green'], figsize=(14, 7))

    # Añadir etiquetas y título
    ax.set_xlabel('Fecha de Otorgamiento')
    ax.set_ylabel('Proporción')
    ax.set_title(title)
    ax.legend(title=class_column, bbox_to_anchor=(1.05, 1), loc='upper left')

    plt.show()

# Crear histogramas para valor_FE_Covid_Fallecidos
plot_stacked_histogram(data100, 'FechaOtorgamiento', 'ClassBM', 'Histograma 100% apilado de Class por Fecha de Otorgamiento')

# Crear histogramas para valor_FE_Covid_Fallecidos
#plot_stacked_histogram(data100, 'FechaOtorgamiento', 'valor_FE_Covid_Fallecidos', 'Histograma 100% apilado de valor_FE_Covid_Fallecidos por Fecha de Otorgamiento')

# Crear histogramas para valor_FE_Covid_Positivos
#plot_stacked_histogram(data100, 'FechaOtorgamiento', 'valor_FE_Covid_Positivos', 'Histograma 100% apilado de valor_FE_Covid_Positivos por Fecha de Otorgamiento')

# Crear histogramas para valor_FE_Covid_Fallecidos
plot_stacked_histogram(data99, 'FechaOtorgamiento', 'ClassBM', 'Histograma 99% apilado de Class por Fecha de Otorgamiento')

# Crear histogramas para valor_FE_Covid_Fallecidos
#plot_stacked_histogram(data99, 'FechaOtorgamiento', 'valor_FE_Covid_Fallecidos', 'Histograma 99% apilado de valor_FE_Covid_Fallecidos por Fecha de Otorgamiento')

# Crear histogramas para valor_FE_Covid_Positivos
#plot_stacked_histogram(data99, 'FechaOtorgamiento', 'valor_FE_Covid_Positivos', 'Histograma 99% apilado de valor_FE_Covid_Positivos por Fecha de Otorgamiento')

# Crear histogramas para valor_FE_Covid_Fallecidos
plot_stacked_histogram(data80, 'FechaOtorgamiento', 'ClassBM', 'Histograma 80% apilado de Class por Fecha de Otorgamiento')

# Crear histogramas para valor_FE_Covid_Fallecidos
#plot_stacked_histogram(data80, 'FechaOtorgamiento', 'valor_FE_Covid_Fallecidos', 'Histograma 80% apilado de valor_FE_Covid_Fallecidos por Fecha de Otorgamiento')
```

```
517
518 # Crear histogramas para valor_FE_Covid_Positivos
519 #plot_stacked_histogram(data80, 'FechaOtorgamiento', 'valor_FE_Covid_Positivos',
    ↪ 'Histograma 80% apilado de valor_FE_Covid_Positivos por Fecha de Otorgamiento')
520
521 #Paso 12: Define las columnas por escenario a validar
522
523 columns_to_keep_SF = [
524     #'IdOrigenCredito', 'PlazoReal', 'IdPeriodicidad', 'NumeroPeriodos',
    ↪ 'IdProducto',
525     'SalarioNormalizado',
526     'SegmentoCartera', 'MaxMontoInterno', 'Monto', 'Bal_TotalActivosNormalizado',
527     'Bal_PatrimonioNegocioNormalizado', 'LiquidezDisponibleNormalizado',
    ↪ 'NetoIngresosNegocioNormalizado',
528     'iAntiguedadBancarizado', 'ScoreOriginacionMicro',
529     'Score_Sobreendeudamiento', 'IdOrigenProspeccion', 'IdDestinoCredito',
530     'TasaEfectiva', 'EdadDesembolsoNormalizada',
531     'Profesion', 'Sexo', 'Dependientes', 'NivelInstruccion', 'Ocupacion',
532     'DirDARes', 'DirDTRes', 'DirDTVIV', 'DirNARes',
533     'DirNTVIV', 'IdOficinaDesembolso']
534
535 columns_to_keep_Fallecidos = [
536     #'IdOrigenCredito', 'PlazoReal', 'IdPeriodicidad', 'NumeroPeriodos',
    ↪ 'IdProducto',
537     'SalarioNormalizado',
538     'SegmentoCartera', 'MaxMontoInterno', 'Monto', 'Bal_TotalActivosNormalizado',
539     'Bal_PatrimonioNegocioNormalizado', 'LiquidezDisponibleNormalizado',
    ↪ 'NetoIngresosNegocioNormalizado',
540     'iAntiguedadBancarizado', 'ScoreOriginacionMicro',
541     'Score_Sobreendeudamiento', 'IdOrigenProspeccion', 'IdDestinoCredito',
542     'TasaEfectiva', 'EdadDesembolsoNormalizada',
543     'Profesion', 'Sexo', 'Dependientes', 'NivelInstruccion', 'Ocupacion',
544     'DirDARes', 'DirDTRes', 'DirDTVIV', 'DirNARes',
545     'DirNTVIV', 'IdOficinaDesembolso',
546     '2020-03Fallecidos', '2020-04Fallecidos', '2020-05Fallecidos',
547     '2020-06Fallecidos', '2020-07Fallecidos', '2020-08Fallecidos',
    ↪ '2020-09Fallecidos',
548     '2020-10Fallecidos', '2020-11Fallecidos', '2020-12Fallecidos',
    ↪ '2021-01Fallecidos',
549     '2021-02Fallecidos', '2021-03Fallecidos', '2021-04Fallecidos',
    ↪ '2021-05Fallecidos',
550     '2021-06Fallecidos', '2021-07Fallecidos', '2021-08Fallecidos',
    ↪ '2021-09Fallecidos',
551     '2021-10Fallecidos', '2021-11Fallecidos', '2021-12Fallecidos',
    ↪ '2022-01Fallecidos']
552     #,
553     #'2022-02Fallecidos', '2022-03Fallecidos', '2022-04Fallecidos',
    ↪ '2022-05Fallecidos',
554     #'2022-06Fallecidos', '2022-07Fallecidos', '2022-08Fallecidos',
    ↪ '2022-09Fallecidos']
555
556 columns_to_keep_Positivos = [
557     #'IdOrigenCredito', 'PlazoReal', 'IdPeriodicidad', 'NumeroPeriodos',
    ↪ 'IdProducto',
558     'SalarioNormalizado',
559     'SegmentoCartera', 'MaxMontoInterno', 'Monto', 'Bal_TotalActivosNormalizado',
560     'Bal_PatrimonioNegocioNormalizado', 'LiquidezDisponibleNormalizado',
    ↪ 'NetoIngresosNegocioNormalizado',
561     'iAntiguedadBancarizado', 'ScoreOriginacionMicro',
562     'Score_Sobreendeudamiento', 'IdOrigenProspeccion', 'IdDestinoCredito',
563     'TasaEfectiva', 'EdadDesembolsoNormalizada',
564     'Profesion', 'Sexo', 'Dependientes', 'NivelInstruccion', 'Ocupacion',
565     'DirDARes', 'DirDTRes', 'DirDTVIV', 'DirNARes',
566     'DirNTVIV', 'IdOficinaDesembolso',
567     '2020-03Positivos', '2020-04Positivos', '2020-05Positivos',
568     '2020-06Positivos', '2020-07Positivos', '2020-08Positivos', '2020-09Positivos',
569     '2020-10Positivos', '2020-11Positivos', '2020-12Positivos', '2021-01Positivos',
```

```
    '2021-02Positivos', '2021-03Positivos', '2021-04Positivos', '2021-05Positivos',
    '2021-06Positivos', '2021-07Positivos', '2021-08Positivos', '2021-09Positivos',
    '2021-10Positivos', '2021-11Positivos', '2021-12Positivos', '2022-01Positivos']
    #,
    #'2022-02Positivos', '2022-03Positivos', '2022-04Positivos',
    ↪  '2022-05Positivos',
    #'2022-06Positivos', '2022-07Positivos', '2022-08Positivos',
    ↪  '2022-09Positivos']

columns_to_keep_Positivos_Fallecidos = [
    #'IdOrigenCredito', 'PlazoReal', 'IdPeriodicidad', 'NumeroPeriodos',
    ↪  'IdProducto',
    'SalarioNormalizado',
    'SegmentoCartera', 'MaxMontoInterno', 'Monto',  'Bal_TotalActivosNormalizado',
    'Bal_PatrimonioNegocioNormalizado', 'LiquidezDisponibleNormalizado',
    ↪  'NetoIngresosNegocioNormalizado',
    'iAntiguedadBancarizado', 'ScoreOriginacionMicro',
    'Score_Sobreendeudamiento', 'IdOrigenProspeccion', 'IdDestinoCredito',
    'TasaEfectiva', 'EdadDesembolsoNormalizada',
    'Profesion', 'Sexo', 'Dependientes', 'NivelInstruccion', 'Ocupacion',
    'DirDARes', 'DirDTRes', 'DirDTVIV', 'DirNARes',
    'DirNTVIV', 'IdOficinaDesembolso',
    '2020-03Fallecidos', '2020-04Fallecidos', '2020-05Fallecidos',
    '2020-06Fallecidos', '2020-07Fallecidos', '2020-08Fallecidos',
    ↪  '2020-09Fallecidos',
    '2020-10Fallecidos', '2020-11Fallecidos', '2020-12Fallecidos',
    ↪  '2021-01Fallecidos',
    '2021-02Fallecidos', '2021-03Fallecidos', '2021-04Fallecidos',
    ↪  '2021-05Fallecidos',
    '2021-06Fallecidos', '2021-07Fallecidos', '2021-08Fallecidos',
    ↪  '2021-09Fallecidos',
    '2021-10Fallecidos', '2021-11Fallecidos', '2021-12Fallecidos',
    ↪  '2022-01Fallecidos',
    '2020-03Positivos', '2020-04Positivos', '2020-05Positivos',
    '2020-06Positivos', '2020-07Positivos', '2020-08Positivos', '2020-09Positivos',
    '2020-10Positivos', '2020-11Positivos', '2020-12Positivos', '2021-01Positivos',
    '2021-02Positivos', '2021-03Positivos', '2021-04Positivos', '2021-05Positivos',
    '2021-06Positivos', '2021-07Positivos', '2021-08Positivos', '2021-09Positivos',
    '2021-10Positivos', '2021-11Positivos', '2021-12Positivos', '2022-01Positivos']

    #,
    #'2022-02Positivos', '2022-03Positivos', '2022-04Positivos',
    ↪  '2022-05Positivos',
    #'2022-06Positivos', '2022-07Positivos', '2022-08Positivos',
    ↪  '2022-09Positivos']

# Cargar el archivo CSV en un DataFrame (reemplaza 'tu_archivo.csv' con el nombre
↪  de tu archivo)
#df = pd.read_csv('/path/to/tu_archivo.csv')

# Filtrar el DataFrame para conservar solo las columnas especificadas
#filtered_df = df[columns_to_keep]

# Mostrar las primeras filas del DataFrame filtrado para verificar
#print(filtered_df.head())

#Paso 13: Instalacion de componentes ML


!pip install keras

!pip install scikit-learn keras tensorflow

!pip install xgboost

!pip install xgboost

```

```
!pip install scikit-learn

!pip install imbalanced-learn

from keras.layers import Input, Dense, Conv1D, MaxPooling1D, Flatten

pd.set_option('display.max_rows', None)

#Paso 14: Define la funcion para realizar la evaluacion del dataset usando algoritmos ML
#Se aplican tecnicas de:
#FILTRADO DE VARIBALES CON CORRELACION > 90%, CON BALANCEO, CONJUNTO DE VALIDACION, KFOLD CROSS VALIDATION

import pandas as pd
import numpy as np
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis, QuadraticDiscriminantAnalysis
from sklearn.linear_model import Perceptron, RidgeClassifier
from sklearn.neural_network import MLPClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.neighbors import KNeighborsClassifier
from xgboost import XGBClassifier
from sklearn.metrics import roc_auc_score, accuracy_score
from sklearn.model_selection import train_test_split, StratifiedKFold
from keras.models import Sequential
from keras.layers import Dense, Conv1D, MaxPooling1D, Flatten
from keras.optimizers import Adam
from imblearn.over_sampling import SMOTE

def remove_collinear_features(df, threshold=0.9):
    corr_matrix = df.corr().abs()
    upper = corr_matrix.where(np.triu(np.ones(corr_matrix.shape), k=1).astype(bool))
    to_drop = [column for column in upper.columns if any(upper[column] > threshold)]
    return df.drop(columns=to_drop), to_drop

def create_ffn_model(input_dim):
    model = Sequential()
    model.add(Input(shape=(input_dim,)))
    model.add(Dense(64, activation='relu'))
    model.add(Dense(32, activation='relu'))
    model.add(Dense(1, activation='sigmoid'))
    model.compile(optimizer=Adam(), loss='binary_crossentropy', metrics=['AUC'])
    return model

def create_cnn_model(input_shape):
    model = Sequential()
    model.add(Input(shape=input_shape))
    model.add(Conv1D(32, kernel_size=3, activation='relu'))
    model.add(MaxPooling1D(pool_size=2))
    model.add(Conv1D(64, kernel_size=3, activation='relu'))
    model.add(MaxPooling1D(pool_size=2))
    model.add(Flatten())
    model.add(Dense(64, activation='relu'))
    model.add(Dense(1, activation='sigmoid'))
    model.compile(optimizer=Adam(), loss='binary_crossentropy', metrics=['AUC'])
    return model

def evaluate_model_cv(data, target_variable, include_variables, model_name, n_splits=5):
    X = data[include_variables]
    y = data[target_variable]

    # Eliminar variables colineales
    X, removed_features = remove_collinear_features(X)
```

```
    if removed_features:
        print(f'Removed collinear features for {model_name}: {removed_features}')

    skf = StratifiedKFold(n_splits=n_splits, shuffle=True, random_state=42)
    aucs_val, accs_val, aucs_test, accs_test = [], [], [], []

    for train_index, test_index in skf.split(X, y):
        X_temp_train, X_test = X.iloc[train_index], X.iloc[test_index]
        y_temp_train, y_test = y.iloc[train_index], y.iloc[test_index]
        X_train, X_val, y_train, y_val = train_test_split(X_temp_train,
        ↪  y_temp_train, test_size=0.25, random_state=42)

        # Aplicar SMOTE para balancear las clases
        smote = SMOTE(random_state=42)
        X_train, y_train = smote.fit_resample(X_train, y_train)

        if model_name == 'lda':
            model = LinearDiscriminantAnalysis()
        elif model_name == 'qda':
            model = QuadraticDiscriminantAnalysis()
        elif model_name == 'perceptron':
            model = Perceptron()
        elif model_name == 'mlp':
            model = MLPClassifier(hidden_layer_sizes=(100,), max_iter=500,
            ↪  random_state=42)
        elif model_name == 'ridge':
            model = RidgeClassifier()
        elif model_name == 'rf':
            model = RandomForestClassifier(n_estimators=100, random_state=42)
        elif model_name == 'xgb':
            model = XGBClassifier(eval_metric='logloss')
        elif model_name == 'knn':
            model = KNeighborsClassifier()
        elif model_name == 'deep_learning':
            model = create_ffn_model(X_train.shape[1])
            model.fit(X_train, y_train, epochs=10, batch_size=32,
            ↪  validation_data=(X_val, y_val), verbose=0)
            y_pred_prob_val = model.predict(X_val).ravel()
            y_pred_prob_test = model.predict(X_test).ravel()
            auc_val = roc_auc_score(y_val, y_pred_prob_val)
            auc_test = roc_auc_score(y_test, y_pred_prob_test)
            y_pred_val = (y_pred_prob_val > 0.5).astype(int)
            y_pred_test = (y_pred_prob_test > 0.5).astype(int)
            acc_val = accuracy_score(y_val, y_pred_val)
            acc_test = accuracy_score(y_test, y_pred_test)
            aucs_val.append(auc_val)
            accs_val.append(acc_val)
            aucs_test.append(auc_test)
            accs_test.append(acc_test)
            continue
        elif model_name == 'cnn':
            X_train_cnn = X_train.values.reshape(X_train.shape[0],
            ↪  X_train.shape[1], 1)
            X_val_cnn = X_val.values.reshape(X_val.shape[0], X_val.shape[1], 1)
            X_test_cnn = X_test.values.reshape(X_test.shape[0], X_test.shape[1], 1)
            model = create_cnn_model((X_train.shape[1], 1))
            model.fit(X_train_cnn, y_train, epochs=10, batch_size=32,
            ↪  validation_data=(X_val_cnn, y_val), verbose=0)
            y_pred_prob_val = model.predict(X_val_cnn).ravel()
            y_pred_prob_test = model.predict(X_test_cnn).ravel()
            auc_val = roc_auc_score(y_val, y_pred_prob_val)
            auc_test = roc_auc_score(y_test, y_pred_prob_test)
            y_pred_val = (y_pred_prob_val > 0.5).astype(int)
            y_pred_test = (y_pred_prob_test > 0.5).astype(int)
            acc_val = accuracy_score(y_val, y_pred_val)
            acc_test = accuracy_score(y_test, y_pred_test)
            aucs_val.append(auc_val)
```

```
748             accs_val.append(acc_val)
749             aucs_test.append(auc_test)
750             accs_test.append(acc_test)
751             continue
752         else:
753             raise ValueError(f"Modelo no soportado: {model_name}")
754 
755         model.fit(X_train, y_train)
756         if model_name == 'perceptron' or model_name == 'ridge':
757             y_pred_prob_val = model.decision_function(X_val)
758             y_pred_prob_test = model.decision_function(X_test)
759         else:
760             y_pred_prob_val = model.predict_proba(X_val)[:, 1]
761             y_pred_prob_test = model.predict_proba(X_test)[:, 1]
762 
763         auc_val = roc_auc_score(y_val, y_pred_prob_val)
764         auc_test = roc_auc_score(y_test, y_pred_prob_test)
765         y_pred_val = model.predict(X_val)
766         y_pred_test = model.predict(X_test)
767         acc_val = accuracy_score(y_val, y_pred_val)
768         acc_test = accuracy_score(y_test, y_pred_test)
769 
770         aucs_val.append(auc_val)
771         accs_val.append(acc_val)
772         aucs_test.append(auc_test)
773         accs_test.append(acc_test)
774 
775     return aucs_val, accs_val, aucs_test, accs_test
776 
777 def evaluate_all_models(database, target_variable, feature_sets, n_splits=5):
778     models = ['lda', 'qda', 'perceptron', 'mlp', 'ridge', 'rf', 'xgb',
    ↪  'deep_learning', 'knn', 'cnn']
779     results = []
780 
781     df = pd.DataFrame(database)
782     df = df[df[target_variable] != -1]
783 
784     for model_name in models:
785         for description, include_variables in feature_sets.items():
786             subset_df = df.dropna(subset=include_variables)
787             aucs_val, accs_val, aucs_test, accs_test = evaluate_model_cv(subset_df,
                ↪  target_variable, include_variables, model_name, n_splits)
788             for fold in range(n_splits):
789                 results.append({
790                     'Model': model_name,
791                     'Features': description,
792                     'Fold': fold + 1,
793                     'AUC_Val': aucs_val[fold],
794                     'ACC_Val': accs_val[fold],
795                     'AUC_Test': aucs_test[fold],
796                     'ACC_Test': accs_test[fold]
797                 })
798             print(f'{model_name} - AUC_Val {description}: {np.mean(aucs_val):.4f},
                ↪  ACC_Val {description}: {np.mean(accs_val):.4f}, AUC_Test
                ↪  {description}: {np.mean(aucs_test):.4f}, ACC_Test {description}:
                ↪  {np.mean(accs_test):.4f}')
799 
800     results_df = pd.DataFrame(results)
801     return results_df
802 
803 def evaluate_multiple_databases(databases, target_variable, feature_sets,
    ↪  n_splits=5):
804     all_results = []
805     for db_name, database in databases.items():
806         print(f"\nEvaluating database: {db_name}")
807         results_df = evaluate_all_models(database, target_variable, feature_sets,
            ↪  n_splits)
```

```
808         results_df['Database'] = db_name
809         all_results.append(results_df)
810         print(results_df)
811     return pd.concat(all_results, ignore_index=True)
812 
813 #Paso 15: Aplicacion del modelo de aprendizaje hasta con 10 FOLD, a los escenarios
    ↪  y datasets definidos
814 
815 
816 # Definir los conjuntos de características
817 feature_sets = {
818     'Sin FE': columns_to_keep_SF,
819     'valor_FE_Covid_Fallecidos': columns_to_keep_Fallecidos,
820     'valor_FE_Covid_Positivos': columns_to_keep_Positivos,
821     'valor_FE_Covid_Fallecidos valor_FE_Covid_Positivos':
    ↪  columns_to_keep_Positivos_Fallecidos
822 }
823 
824 # Ejecutar la evaluación con los diferentes conjuntos de características en
    ↪  múltiples bases de datos
825 databases = {
826     'data100': data100,
827     'data99': data99,
828     'data80': data80
829 }
830 
831 results_df = evaluate_multiple_databases(databases, 'ClassBM', feature_sets, 10)
832 print(results_df)
833 
834 results_df.to_csv('/content/drive/MyDrive/Doc/Data2/ML_COVID_MOV_20240909.csv',
    ↪  sep=';', index=False)
```

Anexo F

Código fuente del proceso de entrenamiento y validación para *FE* convulsión social y anomalías por factores climáticos

En este anexo se presenta el código fuente completo utilizado para el entrenamiento y validación de modelos de *ML* aplicados al escenario de , correspondiente al dataset final. El código está implementado en Python y documentado para su comprensión y posible reproducción.

Archivo: `Class_usando_ml_conv_cclimat.py`

```python
# -*- coding: utf-8 -*-
"""Clasificación usando ML_CONV_CCLIMAT_v1_FINAL99.ipynb

Automatically generated by Colab.

Original file is located at
    https://colab.research.google.com/drive/1kkgZ-PPIaOjSXZVBqUhZapLdrpaBuXqH
"""

#Paso 1 Conectar a Drive
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.stattools import adfuller, coint

import pandas as pd

from google.colab import drive
drive.mount('/content/drive')

# Función para convertir los valores a datetime
def convert_to_datetime(val):
    if isinstance(val, str):
        return pd.to_datetime(val)
    elif isinstance(val, (int, float)):
```

```
        return pd.to_datetime(val, unit='s')
    else:
        return pd.NaT  # Not a Time para valores que no se pueden convertir

#Paso 2 Definiciones Actividades economicas por segmentos de importancia

#Filtrar 80%
Filtro80 = [411,10366,10324,403,10453,10364,10354,10352,10365,10457,10344,
10384,10319,10396,412,10443,10463,10441,10449,10436,516,10320,10382,10478,
10414,134,10316,405,10408,10322,10464,10353,10340,10370,10386,10327,10337,
10416,10465,10474,10306,10400,10162,10376,10350,410,10401,10425,10367,10358,
10459,10371,10387,10378,10448,10309,10447,408,10388,10467,10326,10317,10380,
10343,10375,10356,10164,10452,10342,938,10334,10300,10341,10328,10351,10379,
10395,5220]


#Filtrar 99%
Filtro99 = [411,10366,10324,403,10453,10364,10354,10352,10365,10457,10344,
10384,10319,10396,412,10443,10463,10441,10449,10436,516,10320,10382,10478,
10414,134,10316,405,10408,10322,10464,10353,10340,10370,10386,10327,10337,
10416,10465,10474,10306,10400,10162,10376,10350,410,10401,10425,10367,10358,
10459,10371,10387,10378,10448,10309,10447,408,10388,10467,10326,10317,10380,
10343,10375,10356,10164,10452,10342,938,10334,10300,10341,10328,10351,10379,
10395,5220,10115,10437,10170,10024,10377,10147,10472,10438,10445,10362,10460,
10374,10308,10318,10369,10399,10415,407,10458,10479,10477,10423,10181,9309,
10383,10361,10037,10475,513,10169,10346,10372,10473,10143,10599,10182,10336,
10468,10469,10397,512,138,10456,10461,10093,10454,10301,10241,10424,10311,
10339,10368,10393,514,10360,10026,10481,10471,10000,10303,10329,10141,5270,
10355,10484,10455,10435,10157,10349,10107,10120,10347,10432,10136,10149,10363,
10331,10433,10406,10389,10011,10394,10232,10526,10345,517,10496,10084,10450,10315,
10323,10403,10392,10185,10476,10390,10466,10207,8519,10505,10385,10357,10407,5233,
10411,10307,10426,10434,10150,10012,10597,10595,10314,10440,10175,10332,10439,
10604,10154,10062,10075,10193,7499,10422,10163,10421,10480,9999,10335,10310,
10158,10409,10572,10271,10113,10430,10381,10417,10517,10105,10140,10278,10305,
10609,10330,10470,10442,10236,10418,5232,10176,10130,10398,10131,1511,10173,
10263,5520,10259,10172,10222,10238,10338,401,1721,10168,10089,10189,10040,10110,
10495,10359,10013,10404,10427,10554,10165,10272,10085,10183,10289,10542,10212,10312,
5131,10284,10325,10348,10419,135,5234,10391,10498,5020,10030,10159,5239,10156,5030,
6023,10138,10155,10523,10009,10446,10246,122,10191,10413,10483,10086,10260,5260,
7421,10055,10152,10010,1712,10124,10015,10121,10431,7412,10022,7491,10087,10053,
10462,10091,10114,10214,10313,10304,10412,10405,1549,5211,10023,10041,10145,10205,
3720,7511,10103,10116,10134,10555,10594]

#Paso 3: Cargar y preparar los datos

# Asegúrate de usar la ruta correcta a tu archivo CSV
#file_path = '/content/drive/MyDrive/Doc/Data2/2023_09_v2.rpt'
#file_path = '/content/drive/MyDrive/Doc/Data2/Data007.rpt'
#file_path = '/content/drive/MyDrive/Doc/Data2/Data008_with_FE_COVID1.rpt'
#file_path = '/content/drive/MyDrive/Doc/Data2/Data010_with_FE.rpt'
file_path = '/content/drive/MyDrive/Doc/Data2/Data011_with_FE.rpt'
data = pd.read_csv(file_path, delimiter = ';')

# Asegúrate de usar el formato de fecha correcto
# Suponiendo que la columna de fecha se llama 'date_column'
#data['FechaOtorgamiento'] = pd.to_datetime(data['FechaOtorgamiento'])
data['FechaOtorgamiento'] = data['FechaOtorgamiento'].apply(convert_to_datetime)

#Definir el Rango de Fechas
import datetime


#Declaración del estado de emergencia        16 de marzo de 2020- 28 de octubre de
↪  2022 (2 años, 7 meses y 12 días)
fecha_inicio = datetime.datetime.strptime("2022-12-01", "%Y-%m-%d")  # Fecha de
↪  inicio
fecha_fin = datetime.datetime.strptime("2023-06-30", "%Y-%m-%d")  # Fecha de fin
```

```
90
91  #Filtrar Rango Creditos que han sido otorgados antes de la fecha fin
92  #data = data[(data['FechaOtorgamiento'] >= fecha_inicio) &
    ↪  (data['FechaOtorgamiento'] <= fecha_fin)]
93  data = data[(data['FechaOtorgamiento'] <= fecha_fin)]
94
95  data.info()
96
97  #Paso 4: Elimina creditos cancelados antes del periodo de evaluación
98
99  data['FechaCancelacion'] = data['FechaCancelacion'].apply(convert_to_datetime)
100 data = data[(data['FechaCancelacion'] <= fecha_inicio) |
    ↪  (data['FechaCancelacion'].isna())]
101
102 #Paso 5: Elimina creditos con actividad economica incorrecta y creditos que en mes
    ↪  de inicio tengan ya un atraso mayor a 30 dias
103
104 data['CodigoProductoEspecifico'] = pd.to_numeric(data['CodigoProductoEspecifico'],
    ↪  errors='coerce')
105 data = data[data['CodigoProductoEspecifico'].notna()]
106 data = data[data['CodigoProductoEspecifico'].notna() &
    ↪  (data['CodigoProductoEspecifico'] != '')]
107 data = data[(data['AtrasoMaximo_202211FM'] < 30) |
    ↪  (data['AtrasoMaximo_202211FM'].isna())]
108
109 #data = data[(data['ApoyoGobierno'] == 0)]
110
111 #Paso 6: Procesar los datos
112 # Crear los subconjuntos
113 data100 = data.copy()
114 data99 = data.copy()
115 data80 = data.copy()
116
117 data99 = data99[data99['CodigoProductoEspecifico'].isin(Filtro99)]
118 data80 = data80[data80['CodigoProductoEspecifico'].isin(Filtro80)]
119
120 # Mostrar la información de los datos procesados
121 data100.info()
122 data99.info()
123 data80.info()
124
125 #Paso 7: Graficar Numero de creditos afectados con FE y Clasificacion Mensual
126
127 import pandas as pd
128 import matplotlib.pyplot as plt
129
130 def plot_class_FE(df, tible):
131     """
132     Grafica la cantidad de valores 0 en las columnas 'Class_YYYYMMBM',
133     la cantidad de valores 1 en las columnas 'YYYY-MMFallecidos',
134     y la cantidad de valores 1 en las columnas 'YYYY-MMPositivos'.
135
136     Parameters:
137     df (pd.DataFrame): DataFrame que contiene las columnas a analizar.
138     """
139     # Filtrar columnas que cumplen con las condiciones
140     class_bm_columns = [col for col in df.columns if col.startswith('Class_') and
    ↪  col.endswith('BM')]
141     class_bm_columnsDV = [col for col in df.columns if col.startswith('Class_') and
    ↪  col.endswith('BMDV')]
142     convulsionsocial_columns = [col for col in df.columns if '-' in col and
    ↪  col.endswith('Convulsionsocial')]
143     temperaturaanomalia_columns = [col for col in df.columns if '-' in col and
    ↪  col.endswith('TemperaturaAnomalia')]
144     bloqueosfclimatico_columns = [col for col in df.columns if '-' in col and
    ↪  col.endswith('BloqueoFClimatico')]
145
```

```
    # Inicializar diccionarios para almacenar los resultados
    class_bm_counts = {}
    class_bm_countsDV = {}
    convulsionsocial_counts = {}
    temperaturaanomalia_counts = {}
    bloqueosfclimatico_counts = {}

    # Procesar columnas Class_YYYYMMBM
    for col in class_bm_columns:
        mes = col.split('_')[1][:6]  # Extraer AÑOMES
        count_zeros = (df[col] == 0).sum()  # Contar valores 0
        if mes in class_bm_counts:
            class_bm_counts[mes] += count_zeros
        else:
            class_bm_counts[mes] = count_zeros

    # Procesar columnas Class_YYYYMMBM
    for col in class_bm_columnsDV:
        mes = col.split('_')[1][:6]  # Extraer AÑOMES
        count_zeros = (df[col] == 0).sum()  # Contar valores 0
        #print(mes, count_zeros)
        if mes in class_bm_countsDV:
            class_bm_countsDV[mes] += count_zeros
        else:
            class_bm_countsDV[mes] = count_zeros

    # Procesar columnas YYYY-MM Convulsionsocial
    for col in convulsionsocial_columns:
        parts = col.split('-')
        if len(parts) == 2:
            mes = parts[0] + parts[1][:2]  # Convertir AÑO-MES a AÑOMES
            count_ones = (df[col] == 1).sum()  # Contar valores 1
            if mes in convulsionsocial_counts:
                convulsionsocial_counts[mes] += count_ones
            else:
                convulsionsocial_counts[mes] = count_ones

    # Procesar columnas YYYY-MM TemperaturaAnomalia
    for col in temperaturaanomalia_columns:
        parts = col.split('-')
        if len(parts) == 2:
            mes = parts[0] + parts[1][:2]  # Convertir AÑO-MES a AÑOMES
            count_ones = (df[col] == 1).sum()  # Contar valores 1
            if mes in temperaturaanomalia_counts:
                temperaturaanomalia_counts[mes] += count_ones
            else:
                temperaturaanomalia_counts[mes] = count_ones

    # Procesar columnas YYYY-MM BloqueoFClimatico
    for col in bloqueosfclimatico_columns:
        parts = col.split('-')
        if len(parts) == 2:
            mes = parts[0] + parts[1][:2]  # Convertir AÑO-MES a AÑOMES
            count_ones = (df[col] == 1).sum()  # Contar valores 1
            if mes in bloqueosfclimatico_counts:
                bloqueosfclimatico_counts[mes] += count_ones
            else:
                bloqueosfclimatico_counts[mes] = count_ones

    # Crear listas ordenadas de meses y conteos
    all_months = sorted(set(class_bm_counts.keys()).
        union(set(convulsionsocial_counts.keys())).
        union(set(bloqueosfclimatico_counts.keys())).
        union(set(temperaturaanomalia_counts.keys())).
        union(set(class_bm_countsDV.keys())))
    class_bm_counts_list = [class_bm_counts.get(mes, 0) for mes in all_months]
    class_bm_counts_listDV = [class_bm_countsDV.get(mes, 0) for mes in all_months]
```

```
    convulsionsocial_counts_list = [convulsionsocial_counts.get(mes, 0) for mes in all_months]
    temperaturaanomalia_counts_list = [temperaturaanomalia_counts.get(mes, 0) for mes in all_months]
    bloqueosfclimatico_counts_list = [bloqueosfclimatico_counts.get(mes, 0) for mes in all_months]

    # Crear la gráfica
    plt.figure(figsize=(14, 7))
    plt.plot(all_months, class_bm_counts_list, label='Cantidad de 0 en Class_YYYYMMBM', marker='o')
    plt.plot(all_months, class_bm_counts_listDV, label='Cantidad de 0 en Class_DV_YYYYMMBM', marker='o')
    plt.plot(all_months, convulsionsocial_counts_list, label='Cantidad de 1 en YYYY-MM Convulsion Social', marker='o')
    plt.plot(all_months, temperaturaanomalia_counts_list, label='Cantidad de 1 en YYYY-MM Anomalia Temp', marker='o')
    plt.plot(all_months, bloqueosfclimatico_counts_list, label='Cantidad de 1 en YYYY-MM BLQ F Climaticos', marker='o')
    plt.xlabel('AÑOMES')
    plt.ylabel('Cantidad')
    plt.title(tible + ' - Cantidad de 0 en Class_YYYYMMBM, en Class_DV_YYYYMMBM, 1 en YYYY-MMFallecidos y YYYY-MMPositivos')
    plt.legend()
    plt.xticks(rotation=45)
    plt.grid(True)
    plt.show()

# Ejemplo de uso
# Supongamos que tienes un dataframe llamado df
# df = pd.read_csv('tu_archivo.csv')
df = data100.copy()
# Llamar a la función para crear la gráfica
plot_class_FE(df, 'Data100')

df = data99.copy()
# Llamar a la función para crear la gráfica
plot_class_FE(df, 'Data99')

df = data80.copy()
# Llamar a la función para crear la gráfica
plot_class_FE(df, 'Data80')

#Funcion retraza la secuencia de fallecidos en 1 periodo
import pandas as pd

def shift_fallecidos_columns(df, prefix='Fallecidos'):
    """
    Función para reasignar valores en columnas con el formato 'AÑO-MESFallecidos'.
    Copia los valores de cada columna a la columna siguiente en orden cronológico.

    Parameters:
    df (pd.DataFrame): DataFrame que contiene las columnas con el formato 'AÑO-MESFallecidos'.
    prefix (str): Sufijo que identifica las columnas a ser reasignadas. Por defecto es 'Fallecidos'.

    Returns:
    pd.DataFrame: DataFrame con los valores de las columnas reasignados.
    """
    # Crear una lista con los nombres de las columnas que contienen el prefijo
    fallecidos_columns = [col for col in df.columns if prefix in col]

    # Ordenar las columnas cronológicamente
    fallecidos_columns.sort()

    # Realizar la copia de valores
```

```
    for i in range(len(fallecidos_columns) - 1, 0, -1):
        df[fallecidos_columns[i]] = df[fallecidos_columns[i - 1]]

    return df

#Paso 8: Mueve la curva fallecidos a la derecha

#data100 = shift_fallecidos_columns(data100)
#data99 = shift_fallecidos_columns(data99)
#data80 = shift_fallecidos_columns(data80)

#Paso 9: Graficar Numero de creditos afectados con FE y Clasificacion Mensual
#FE Fallecido movido a la derecha 1 periodo

import pandas as pd
import matplotlib.pyplot as plt

def plot_class_FE(df, tible):
    """
    Grafica la cantidad de valores 0 en las columnas 'Class_YYYYMMBM',
    la cantidad de valores 1 en las columnas 'YYYY-MMFallecidos',
    y la cantidad de valores 1 en las columnas 'YYYY-MMPositivos'.

    Parameters:
    df (pd.DataFrame): DataFrame que contiene las columnas a analizar.
    """
    # Filtrar columnas que cumplen con las condiciones
    class_bm_columns = [col for col in df.columns if col.startswith('Class_') and
    ↪  col.endswith('BM')]
    class_bm_columnsDV = [col for col in df.columns if col.startswith('Class_') and
    ↪  col.endswith('BMDV')]
    convulsionsocial_columns = [col for col in df.columns if '-' in col and
    ↪  col.endswith('Convulsionsocial')]
    temperaturaanomalia_columns = [col for col in df.columns if '-' in col and
    ↪  col.endswith('TemperaturaAnomalia')]
    bloqueosfclimatico_columns = [col for col in df.columns if '-' in col and
    ↪  col.endswith('BloqueoFClimatico')]

    # Inicializar diccionarios para almacenar los resultados
    class_bm_counts = {}
    class_bm_countsDV = {}
    convulsionsocial_counts = {}
    temperaturaanomalia_counts = {}
    bloqueosfclimatico_counts = {}

    # Procesar columnas Class_YYYYMMBM
    for col in class_bm_columns:
        mes = col.split('_')[1][:6]  # Extraer AÑOMES
        count_zeros = (df[col] == 0).sum()  # Contar valores 0
        if mes in class_bm_counts:
            class_bm_counts[mes] += count_zeros
        else:
            class_bm_counts[mes] = count_zeros

    # Procesar columnas Class_YYYYMMBM
    for col in class_bm_columnsDV:
        mes = col.split('_')[1][:6]  # Extraer AÑOMES
        count_zeros = (df[col] == 0).sum()  # Contar valores 0
        #print(mes, count_zeros)
        if mes in class_bm_countsDV:
            class_bm_countsDV[mes] += count_zeros
        else:
            class_bm_countsDV[mes] = count_zeros

    # Procesar columnas YYYY-MM Convulsionsocial
    for col in convulsionsocial_columns:
        parts = col.split('-')
```

```
331             if len(parts) == 2:
332                 mes = parts[0] + parts[1][:2]  # Convertir AÑO-MES a AÑOMES
333                 count_ones = (df[col] == 1).sum()  # Contar valores 1
334                 if mes in convulsionsocial_counts:
335                     convulsionsocial_counts[mes] += count_ones
336                 else:
337                     convulsionsocial_counts[mes] = count_ones
338 
339         # Procesar columnas YYYY-MM TemperaturaAnomalia
340         for col in temperaturaanomalia_columns:
341             parts = col.split('-')
342             if len(parts) == 2:
343                 mes = parts[0] + parts[1][:2]  # Convertir AÑO-MES a AÑOMES
344                 count_ones = (df[col] == 1).sum()  # Contar valores 1
345                 if mes in temperaturaanomalia_counts:
346                     temperaturaanomalia_counts[mes] += count_ones
347                 else:
348                     temperaturaanomalia_counts[mes] = count_ones
349 
350         # Procesar columnas YYYY-MM BloqueoFClimatico
351         for col in bloqueosfclimatico_columns:
352             parts = col.split('-')
353             if len(parts) == 2:
354                 mes = parts[0] + parts[1][:2]  # Convertir AÑO-MES a AÑOMES
355                 count_ones = (df[col] == 1).sum()  # Contar valores 1
356                 if mes in bloqueosfclimatico_counts:
357                     bloqueosfclimatico_counts[mes] += count_ones
358                 else:
359                     bloqueosfclimatico_counts[mes] = count_ones
360 
361         # Crear listas ordenadas de meses y conteos
362         all_months = sorted(set(class_bm_counts.keys()).
363             union(set(convulsionsocial_counts.keys())).
364             union(set(bloqueosfclimatico_counts.keys())).
365             union(set(temperaturaanomalia_counts.keys())).
366             union(set(class_bm_countsDV.keys())))
367         class_bm_counts_list = [class_bm_counts.get(mes, 0) for mes in all_months]
368         class_bm_counts_listDV = [class_bm_countsDV.get(mes, 0) for mes in all_months]
369         convulsionsocial_counts_list = [convulsionsocial_counts.get(mes, 0) for mes in
            ↪  all_months]
370         temperaturaanomalia_counts_list = [temperaturaanomalia_counts.get(mes, 0) for
            ↪  mes in all_months]
371         bloqueosfclimatico_counts_list = [bloqueosfclimatico_counts.get(mes, 0) for mes
            ↪  in all_months]
372 
373         # Crear la gráfica
374         plt.figure(figsize=(14, 7))
375         plt.plot(all_months, class_bm_counts_list, label='Cantidad de 0 en
            ↪  Class_YYYYMMBM', marker='o')
376         plt.plot(all_months, class_bm_counts_listDV, label='Cantidad de 0 en
            ↪  Class_DV_YYYYMMBM', marker='o')
377         plt.plot(all_months, convulsionsocial_counts_list, label='Cantidad de 1 en
            ↪  YYYY-MM Convulsion Social', marker='o')
378         plt.plot(all_months, temperaturaanomalia_counts_list, label='Cantidad de 1 en
            ↪  YYYY-MM Anomalia Temp', marker='o')
379         plt.plot(all_months, bloqueosfclimatico_counts_list, label='Cantidad de 1 en
            ↪  YYYY-MM BLQ F Climaticos', marker='o')
380         plt.xlabel('AÑOMES')
381         plt.ylabel('Cantidad')
382         plt.title(tible + ' - Cantidad de 0 en Class_YYYYMMBM, en Class_DV_YYYYMMBM, 1
            ↪  en YYYY-MMFallecidos y YYYY-MMPositivos')
383         plt.legend()
384         plt.xticks(rotation=45)
385         plt.grid(True)
386         plt.show()
387 
388 # Ejemplo de uso
```

```python
# Supongamos que tienes un dataframe llamado df
# df = pd.read_csv('tu_archivo.csv')
df = data100.copy()
# Llamar a la función para crear la gráfica
plot_class_FE(df, 'Data100')

df = data99.copy()
# Llamar a la función para crear la gráfica
plot_class_FE(df, 'Data99')

df = data80.copy()
# Llamar a la función para crear la gráfica
plot_class_FE(df, 'Data80')

#Paso 10: Agrega la variable Class que indica si en el periodo en evaluacion se presentaron impacto

import pandas as pd
import numpy as np

def filter_class_bm_columns(df, start_date, end_date):
    """
    Filtra las columnas del DataFrame que empiezan con "Class_" y terminan en "BM"
    dentro del rango de fechas especificado y determina el valor de ClassBM.

    Parameters:
    df (pd.DataFrame): DataFrame que contiene las columnas a filtrar.
    start_date (str): Fecha de inicio en formato "YYYY-MM".
    end_date (str): Fecha de fin en formato "YYYY-MM".

    Returns:
    pd.DataFrame: DataFrame con la nueva columna ClassBM.
    """
    # Convertir las fechas de inicio y fin a enteros para la comparación
    start_date_int = int(start_date.replace("-", ""))
    end_date_int = int(end_date.replace("-", ""))

    # Filtrar columnas que empiezan con "Class_" y terminan en "BM" dentro del rango de fechas
    class_bm_columns = [col for col in df.columns if col.startswith('Class_') and col.endswith('BM')
                        and start_date_int <= int(col[6:12]) <= end_date_int]

    # Función para determinar el valor de ClassBM
    def determine_class_bm(row):
        if (row[class_bm_columns] == 0).any():
            return 0
        elif (row[class_bm_columns] == 1).any():
            return 1
        else:
            return -1

    # Aplicar la función para crear la nueva columna ClassBM
    df['ClassBM'] = df.apply(determine_class_bm, axis=1)
    return df

# Ejemplo de uso
# Supongamos que tienes DataFrames llamados data100, data99, y data80
start_date = "2022-12"
end_date = "2023-09"

data100 = filter_class_bm_columns(data100, start_date, end_date)
data99 = filter_class_bm_columns(data99, start_date, end_date)
data80 = filter_class_bm_columns(data80, start_date, end_date)

# Mostrar las primeras filas del DataFrame para verificar
#print(data100.head())
```

```
453
454 #Paso 11: Grafica histograma de la variable agregada
455
456 # Asegúrate de que tienes las bibliotecas necesarias instaladas
457 !pip install pandas matplotlib
458
459 import pandas as pd
460 import matplotlib.pyplot as plt
461
462 import pandas as pd
463 import matplotlib.pyplot as plt
464
465 # Suponiendo que tienes un dataframe llamado df
466
467
468 # Función para crear histogramas apilados
469 def plot_stacked_histogram_Binary(df, date_column, binary_column, title):
470     # Asegúrate de que la columna 'FechaOtorgamiento' esté en formato datetime
471     df['FechaOtorgamiento'] = pd.to_datetime(df['FechaOtorgamiento'])
472
473     # Agrupa por fecha y cuenta los valores de la columna binaria
474     df_grouped = df.groupby([df[date_column].dt.date,
    ↪  binary_column]).size().unstack(fill_value=0)
475
476     # Normaliza para que sumen 100%
477     df_grouped = df_grouped.div(df_grouped.sum(axis=1), axis=0)
478
479     # Traza el histograma apilado
480     ax = df_grouped.plot(kind='bar', stacked=True, color=['blue', 'orange'],
    ↪  figsize=(14, 7))
481
482     # Añadir etiquetas y título
483     ax.set_xlabel('Fecha de Otorgamiento')
484     ax.set_ylabel('Proporción')
485     ax.set_title(title)
486     ax.legend(title=binary_column, bbox_to_anchor=(1.05, 1), loc='upper left')
487
488     plt.show()
489
490 import pandas as pd
491 import matplotlib.pyplot as plt
492
493 # Función para crear histogramas apilados
494 #Tres valores -1, 0, 1
495 def plot_stacked_histogram(df, date_column, class_column, title):
496     """
497     Función para crear histogramas apilados con una columna de clasificación que
    ↪  contiene valores -1, 0 y 1.
498
499     Parameters:
500     df (pd.DataFrame): DataFrame que contiene los datos.
501     date_column (str): Nombre de la columna de fechas.
502     class_column (str): Nombre de la columna de clasificación.
503     title (str): Título del gráfico.
504     """
505
506     # Asegúrate de que la columna de fechas esté en formato datetime
507     df[date_column] = pd.to_datetime(df[date_column])
508
509     # Agrupa por fecha y cuenta los valores de la columna de clasificación
510     df_grouped = df.groupby([df[date_column].dt.date,
    ↪  class_column]).size().unstack(fill_value=0)
511
512     # Normaliza para que sumen 100%
513     df_grouped = df_grouped.div(df_grouped.sum(axis=1), axis=0)
514
515     # Traza el histograma apilado
```

```
516         ax = df_grouped.plot(kind='bar', stacked=True, color=['blue', 'orange',
            ↪  'green'], figsize=(14, 7))
517
518         # Añadir etiquetas y título
519         ax.set_xlabel('Fecha de Otorgamiento')
520         ax.set_ylabel('Proporción')
521         ax.set_title(title)
522         ax.legend(title=class_column, bbox_to_anchor=(1.05, 1), loc='upper left')
523
524         plt.show()
525
526     # Crear histogramas para valor_FE_Covid_Fallecidos
527     plot_stacked_histogram(data100, 'FechaOtorgamiento', 'ClassBM', 'Histograma 100%
        ↪  apilado de Class por Fecha de Otorgamiento')
528
529     # Crear histogramas para valor_FE_Covid_Fallecidos
530     #plot_stacked_histogram(data100, 'FechaOtorgamiento', 'valor_FE_Covid_Fallecidos',
        ↪  'Histograma 100% apilado de valor_FE_Covid_Fallecidos por Fecha de
        ↪  Otorgamiento')
531
532     # Crear histogramas para valor_FE_Covid_Positivos
533     #plot_stacked_histogram(data100, 'FechaOtorgamiento', 'valor_FE_Covid_Positivos',
        ↪  'Histograma 100% apilado de valor_FE_Covid_Positivos por Fecha de
        ↪  Otorgamiento')
534
535     # Crear histogramas para valor_FE_Covid_Fallecidos
536     plot_stacked_histogram(data99, 'FechaOtorgamiento', 'ClassBM', 'Histograma 99%
        ↪  apilado de Class por Fecha de Otorgamiento')
537
538     # Crear histogramas para valor_FE_Covid_Fallecidos
539     #plot_stacked_histogram(data99, 'FechaOtorgamiento', 'valor_FE_Covid_Fallecidos',
        ↪  'Histograma 99% apilado de valor_FE_Covid_Fallecidos por Fecha de
        ↪  Otorgamiento')
540
541     # Crear histogramas para valor_FE_Covid_Positivos
542     #plot_stacked_histogram(data99, 'FechaOtorgamiento', 'valor_FE_Covid_Positivos',
        ↪  'Histograma 99% apilado de valor_FE_Covid_Positivos por Fecha de Otorgamiento')
543
544     # Crear histogramas para valor_FE_Covid_Fallecidos
545     plot_stacked_histogram(data80, 'FechaOtorgamiento', 'ClassBM', 'Histograma 80%
        ↪  apilado de Class por Fecha de Otorgamiento')
546
547     # Crear histogramas para valor_FE_Covid_Fallecidos
548     #plot_stacked_histogram(data80, 'FechaOtorgamiento', 'valor_FE_Covid_Fallecidos',
        ↪  'Histograma 80% apilado de valor_FE_Covid_Fallecidos por Fecha de
        ↪  Otorgamiento')
549
550     # Crear histogramas para valor_FE_Covid_Positivos
551     #plot_stacked_histogram(data80, 'FechaOtorgamiento', 'valor_FE_Covid_Positivos',
        ↪  'Histograma 80% apilado de valor_FE_Covid_Positivos por Fecha de Otorgamiento')
552
553     #Paso 12: Define las columnas por escenario a validar
554
555     columns_to_keep_SF = [
556         #'IdOrigenCredito', 'PlazoReal', 'IdPeriodicidad', 'NumeroPeriodos',
            ↪  'IdProducto',
557         'SalarioNormalizado',
558         'SegmentoCartera', 'MaxMontoInterno', 'Monto',  'Bal_TotalActivosNormalizado',
559         'Bal_PatrimonioNegocioNormalizado', 'LiquidezDisponibleNormalizado',
            ↪  'NetoIngresosNegocioNormalizado',
560         'iAntiguedadBancarizado', 'ScoreOriginacionMicro',
561         'Score_Sobreendeudamiento', 'IdOrigenProspeccion', 'IdDestinoCredito',
562         'TasaEfectiva', 'EdadDesembolsoNormalizada',
563         'Profesion', 'Sexo', 'Dependientes', 'NivelInstruccion', 'Ocupacion',
564         'DirDARes', 'DirDTRes', 'DirDTVIV', 'DirNARes',
565         'DirNTVIV', 'IdOficinaDesembolso']
566
```

```
columns_to_keep_Convulsionsocial = [
    #'IdOrigenCredito', 'PlazoReal', 'IdPeriodicidad', 'NumeroPeriodos',
    ↪  'IdProducto',
    'SalarioNormalizado',
    'SegmentoCartera', 'MaxMontoInterno', 'Monto',  'Bal_TotalActivosNormalizado',
    'Bal_PatrimonioNegocioNormalizado', 'LiquidezDisponibleNormalizado',
    ↪  'NetoIngresosNegocioNormalizado',
    'iAntiguedadBancarizado', 'ScoreOriginacionMicro',
    'Score_Sobreendeudamiento', 'IdOrigenProspeccion', 'IdDestinoCredito',
    'TasaEfectiva', 'EdadDesembolsoNormalizada',
    'Profesion', 'Sexo', 'Dependientes', 'NivelInstruccion', 'Ocupacion',
    'DirDARes', 'DirDTRes', 'DirDTVIV', 'DirNARes',
    'DirNTVIV', 'IdOficinaDesembolso',
    '2022-12Convulsionsocial', '2023-01Convulsionsocial',
    '2023-02Convulsionsocial', '2023-03Convulsionsocial',
    ↪  '2023-04Convulsionsocial', '2023-05Convulsionsocial',
    '2023-06Convulsionsocial', '2023-07Convulsionsocial',
    ↪  '2023-08Convulsionsocial', '2023-09Convulsionsocial']

columns_to_keep_BloqueoFClimatico = [
    #'IdOrigenCredito', 'PlazoReal', 'IdPeriodicidad', 'NumeroPeriodos',
    ↪  'IdProducto',
    'SalarioNormalizado',
    'SegmentoCartera', 'MaxMontoInterno', 'Monto',  'Bal_TotalActivosNormalizado',
    'Bal_PatrimonioNegocioNormalizado', 'LiquidezDisponibleNormalizado',
    ↪  'NetoIngresosNegocioNormalizado',
    'iAntiguedadBancarizado', 'ScoreOriginacionMicro',
    'Score_Sobreendeudamiento', 'IdOrigenProspeccion', 'IdDestinoCredito',
    'TasaEfectiva', 'EdadDesembolsoNormalizada',
    'Profesion', 'Sexo', 'Dependientes', 'NivelInstruccion', 'Ocupacion',
    'DirDARes', 'DirDTRes', 'DirDTVIV', 'DirNARes',
    'DirNTVIV', 'IdOficinaDesembolso',
    '2022-12BloqueoFClimatico', '2023-01BloqueoFClimatico',
    '2023-02BloqueoFClimatico', '2023-03BloqueoFClimatico',
    ↪  '2023-04BloqueoFClimatico', '2023-05BloqueoFClimatico',
    '2023-06BloqueoFClimatico', '2023-07BloqueoFClimatico',
    ↪  '2023-08BloqueoFClimatico', '2023-09BloqueoFClimatico']

columns_to_keep_TemperaturaAnomalia = [
    #'IdOrigenCredito', 'PlazoReal', 'IdPeriodicidad', 'NumeroPeriodos',
    ↪  'IdProducto',
    'SalarioNormalizado',
    'SegmentoCartera', 'MaxMontoInterno', 'Monto',  'Bal_TotalActivosNormalizado',
    'Bal_PatrimonioNegocioNormalizado', 'LiquidezDisponibleNormalizado',
    ↪  'NetoIngresosNegocioNormalizado',
    'iAntiguedadBancarizado', 'ScoreOriginacionMicro',
    'Score_Sobreendeudamiento', 'IdOrigenProspeccion', 'IdDestinoCredito',
    'TasaEfectiva', 'EdadDesembolsoNormalizada',
    'Profesion', 'Sexo', 'Dependientes', 'NivelInstruccion', 'Ocupacion',
    'DirDARes', 'DirDTRes', 'DirDTVIV', 'DirNARes',
    'DirNTVIV', 'IdOficinaDesembolso',
    '2022-12TemperaturaAnomalia',
    '2023-03TemperaturaAnomalia',
    '2023-06TemperaturaAnomalia', '2023-09TemperaturaAnomalia']


columns_to_keep_TemperaturaAnomalia_BloqueFClimatico = [
    #'IdOrigenCredito', 'PlazoReal', 'IdPeriodicidad', 'NumeroPeriodos',
    ↪  'IdProducto',
    'SalarioNormalizado',
    'SegmentoCartera', 'MaxMontoInterno', 'Monto',  'Bal_TotalActivosNormalizado',
    'Bal_PatrimonioNegocioNormalizado', 'LiquidezDisponibleNormalizado',
    ↪  'NetoIngresosNegocioNormalizado',
    'iAntiguedadBancarizado', 'ScoreOriginacionMicro',
    'Score_Sobreendeudamiento', 'IdOrigenProspeccion', 'IdDestinoCredito',
    'TasaEfectiva', 'EdadDesembolsoNormalizada',
    'Profesion', 'Sexo', 'Dependientes', 'NivelInstruccion', 'Ocupacion',
```

```
    'DirDARes', 'DirDTRes', 'DirDTVIV', 'DirNARes',
    'DirNTVIV', 'IdOficinaDesembolso',
    '2022-12TemperaturaAnomalia',
    '2023-03TemperaturaAnomalia',
    '2023-06TemperaturaAnomalia', '2023-09TemperaturaAnomalia',
    '2022-12BloqueoFClimatico', '2023-01BloqueoFClimatico',
    '2023-02BloqueoFClimatico', '2023-03BloqueoFClimatico',
    ↪ '2023-04BloqueoFClimatico', '2023-05BloqueoFClimatico',
    '2023-06BloqueoFClimatico', '2023-07BloqueoFClimatico',
    ↪ '2023-08BloqueoFClimatico', '2023-09BloqueoFClimatico']


columns_to_keep_Convulsionsocial_BloqueFClimatico = [
    #'IdOrigenCredito', 'PlazoReal', 'IdPeriodicidad', 'NumeroPeriodos',
    ↪ 'IdProducto',
    'SalarioNormalizado',
    'SegmentoCartera', 'MaxMontoInterno', 'Monto', 'Bal_TotalActivosNormalizado',
    'Bal_PatrimonioNegocioNormalizado', 'LiquidezDisponibleNormalizado',
    ↪ 'NetoIngresosNegocioNormalizado',
    'iAntiguedadBancarizado', 'ScoreOriginacionMicro',
    'Score_Sobreendeudamiento', 'IdOrigenProspeccion', 'IdDestinoCredito',
    'TasaEfectiva', 'EdadDesembolsoNormalizada',
    'Profesion', 'Sexo', 'Dependientes', 'NivelInstruccion', 'Ocupacion',
    'DirDARes', 'DirDTRes', 'DirDTVIV', 'DirNARes',
    'DirNTVIV', 'IdOficinaDesembolso',
    '2022-12Convulsionsocial', '2023-01Convulsionsocial',
    '2023-02Convulsionsocial', '2023-03Convulsionsocial',
    ↪ '2023-04Convulsionsocial', '2023-05Convulsionsocial',
    '2023-06Convulsionsocial', '2023-07Convulsionsocial',
    ↪ '2023-08Convulsionsocial', '2023-09Convulsionsocial',
    '2022-12BloqueoFClimatico', '2023-01BloqueoFClimatico',
    '2023-02BloqueoFClimatico', '2023-03BloqueoFClimatico',
    ↪ '2023-04BloqueoFClimatico', '2023-05BloqueoFClimatico',
    '2023-06BloqueoFClimatico', '2023-07BloqueoFClimatico',
    ↪ '2023-08BloqueoFClimatico', '2023-09BloqueoFClimatico']

columns_to_keep_Convulsionsocial_TemperaturaAnomalia = [
    #'IdOrigenCredito', 'PlazoReal', 'IdPeriodicidad', 'NumeroPeriodos',
    ↪ 'IdProducto',
    'SalarioNormalizado',
    'SegmentoCartera', 'MaxMontoInterno', 'Monto', 'Bal_TotalActivosNormalizado',
    'Bal_PatrimonioNegocioNormalizado', 'LiquidezDisponibleNormalizado',
    ↪ 'NetoIngresosNegocioNormalizado',
    'iAntiguedadBancarizado', 'ScoreOriginacionMicro',
    'Score_Sobreendeudamiento', 'IdOrigenProspeccion', 'IdDestinoCredito',
    'TasaEfectiva', 'EdadDesembolsoNormalizada',
    'Profesion', 'Sexo', 'Dependientes', 'NivelInstruccion', 'Ocupacion',
    'DirDARes', 'DirDTRes', 'DirDTVIV', 'DirNARes',
    'DirNTVIV', 'IdOficinaDesembolso',
    '2022-12Convulsionsocial', '2023-01Convulsionsocial',
    '2023-02Convulsionsocial', '2023-03Convulsionsocial',
    ↪ '2023-04Convulsionsocial', '2023-05Convulsionsocial',
    '2023-06Convulsionsocial', '2023-07Convulsionsocial',
    ↪ '2023-08Convulsionsocial', '2023-09Convulsionsocial',
    '2022-12TemperaturaAnomalia',
    '2023-03TemperaturaAnomalia',
    '2023-06TemperaturaAnomalia', '2023-09TemperaturaAnomalia']


columns_to_keep_TemperaturaAnomalia_Convulsionsocial_BloqueFClimatico = [
    #'IdOrigenCredito', 'PlazoReal', 'IdPeriodicidad', 'NumeroPeriodos',
    ↪ 'IdProducto',
    'SalarioNormalizado',
    'SegmentoCartera', 'MaxMontoInterno', 'Monto', 'Bal_TotalActivosNormalizado',
    'Bal_PatrimonioNegocioNormalizado', 'LiquidezDisponibleNormalizado',
    ↪ 'NetoIngresosNegocioNormalizado',
    'iAntiguedadBancarizado', 'ScoreOriginacionMicro',
```

```
    'Score_Sobreendeudamiento', 'IdOrigenProspeccion', 'IdDestinoCredito',
    'TasaEfectiva', 'EdadDesembolsoNormalizada',
    'Profesion', 'Sexo', 'Dependientes', 'NivelInstruccion', 'Ocupacion',
    'DirDARes', 'DirDTRes', 'DirDTVIV', 'DirNARes',
    'DirNTVIV', 'IdOficinaDesembolso',
    '2022-12Convulsionsocial', '2023-01Convulsionsocial',
    '2023-02Convulsionsocial', '2023-03Convulsionsocial',
    '2023-04Convulsionsocial', '2023-05Convulsionsocial',
    '2023-06Convulsionsocial', '2023-07Convulsionsocial',
    '2023-08Convulsionsocial', '2023-09Convulsionsocial',
    '2022-12TemperaturaAnomalia',
    '2023-03TemperaturaAnomalia',
    '2023-06TemperaturaAnomalia', '2023-09TemperaturaAnomalia',
    '2022-12BloqueoFClimatico', '2023-01BloqueoFClimatico',
    '2023-02BloqueoFClimatico', '2023-03BloqueoFClimatico',
    '2023-04BloqueoFClimatico', '2023-05BloqueoFClimatico',
    '2023-06BloqueoFClimatico', '2023-07BloqueoFClimatico',
    '2023-08BloqueoFClimatico', '2023-09BloqueoFClimatico']

# Cargar el archivo CSV en un DataFrame (reemplaza 'tu_archivo.csv' con el nombre de tu archivo)
#df = pd.read_csv('/path/to/tu_archivo.csv')

# Filtrar el DataFrame para conservar solo las columnas especificadas
#filtered_df = df[columns_to_keep]

# Mostrar las primeras filas del DataFrame filtrado para verificar
#print(filtered_df.head())

#Paso 13: Instalacion de componentes ML

!pip install keras

pip install scikit-learn keras tensorflow

!pip install xgboost

#Se instala 2 veces para corregir el error de dependencias
!pip install xgboost

!pip install scikit-learn

!pip install imbalanced-learn

from keras.layers import Input, Dense, Conv1D, MaxPooling1D, Flatten

pd.set_option('display.max_rows', None)

#Paso 14: Define la funcion para realizar la evaluacion del dataset usando algoritmos ML
#Se aplican tecnicas de:
#FILTRADO DE VARIBALES CON CORRELACION > 90%, CON BALANCEO, CONJUNTO DE VALIDACION, KFOLD CROSS VALIDATION

import pandas as pd
import numpy as np
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis, QuadraticDiscriminantAnalysis
from sklearn.linear_model import Perceptron, RidgeClassifier
from sklearn.neural_network import MLPClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.neighbors import KNeighborsClassifier
from xgboost import XGBClassifier
from sklearn.metrics import roc_auc_score, accuracy_score
from sklearn.model_selection import train_test_split, StratifiedKFold
from keras.models import Sequential
from keras.layers import Dense, Conv1D, MaxPooling1D, Flatten
```

```
734 from keras.optimizers import Adam
735 from imblearn.over_sampling import SMOTE
736
737 def remove_collinear_features(df, threshold=0.9):
738     corr_matrix = df.corr().abs()
739     upper = corr_matrix.where(np.triu(np.ones(corr_matrix.shape),
        ↪  k=1).astype(bool))
740     to_drop = [column for column in upper.columns if any(upper[column] >
        ↪  threshold)]
741     return df.drop(columns=to_drop), to_drop
742
743 def create_ffn_model(input_dim):
744     model = Sequential()
745     model.add(Input(shape=(input_dim,)))
746     model.add(Dense(64, activation='relu'))
747     model.add(Dense(32, activation='relu'))
748     model.add(Dense(1, activation='sigmoid'))
749     model.compile(optimizer=Adam(), loss='binary_crossentropy', metrics=['AUC'])
750     return model
751
752 def create_cnn_model(input_shape):
753     model = Sequential()
754     model.add(Input(shape=input_shape))
755     model.add(Conv1D(32, kernel_size=3, activation='relu'))
756     model.add(MaxPooling1D(pool_size=2))
757     model.add(Conv1D(64, kernel_size=3, activation='relu'))
758     model.add(MaxPooling1D(pool_size=2))
759     model.add(Flatten())
760     model.add(Dense(64, activation='relu'))
761     model.add(Dense(1, activation='sigmoid'))
762     model.compile(optimizer=Adam(), loss='binary_crossentropy', metrics=['AUC'])
763     return model
764
765 def evaluate_model_cv(data, target_variable, include_variables, model_name,
    ↪  n_splits=5):
766     X = data[include_variables]
767     y = data[target_variable]
768
769     # Eliminar variables colineales
770     X, removed_features = remove_collinear_features(X)
771     if removed_features:
772         print(f'Removed collinear features for {model_name}: {removed_features}')
773
774     skf = StratifiedKFold(n_splits=n_splits, shuffle=True, random_state=42)
775     aucs_val, accs_val, aucs_test, accs_test = [], [], [], []
776
777     for train_index, test_index in skf.split(X, y):
778         X_temp_train, X_test = X.iloc[train_index], X.iloc[test_index]
779         y_temp_train, y_test = y.iloc[train_index], y.iloc[test_index]
780         X_train, X_val, y_train, y_val = train_test_split(X_temp_train,
            ↪  y_temp_train, test_size=0.25, random_state=42)
781
782         # Aplicar SMOTE para balancear las clases
783         smote = SMOTE(random_state=42)
784         X_train, y_train = smote.fit_resample(X_train, y_train)
785
786         if model_name == 'lda':
787             model = LinearDiscriminantAnalysis()
788         elif model_name == 'qda':
789             model = QuadraticDiscriminantAnalysis()
790         elif model_name == 'perceptron':
791             model = Perceptron()
792         elif model_name == 'mlp':
793             model = MLPClassifier(hidden_layer_sizes=(100,), max_iter=500,
                ↪  random_state=42)
794         elif model_name == 'ridge':
795             model = RidgeClassifier()
```

```
796             elif model_name == 'rf':
797                 model = RandomForestClassifier(n_estimators=100, random_state=42)
798             elif model_name == 'xgb':
799                 model = XGBClassifier(eval_metric='logloss')
800             elif model_name == 'knn':
801                 model = KNeighborsClassifier()
802             elif model_name == 'deep_learning':
803                 model = create_ffn_model(X_train.shape[1])
804                 model.fit(X_train, y_train, epochs=10, batch_size=32,
                    ↪  validation_data=(X_val, y_val), verbose=0)
805                 y_pred_prob_val = model.predict(X_val).ravel()
806                 y_pred_prob_test = model.predict(X_test).ravel()
807                 auc_val = roc_auc_score(y_val, y_pred_prob_val)
808                 auc_test = roc_auc_score(y_test, y_pred_prob_test)
809                 y_pred_val = (y_pred_prob_val > 0.5).astype(int)
810                 y_pred_test = (y_pred_prob_test > 0.5).astype(int)
811                 acc_val = accuracy_score(y_val, y_pred_val)
812                 acc_test = accuracy_score(y_test, y_pred_test)
813                 aucs_val.append(auc_val)
814                 accs_val.append(acc_val)
815                 aucs_test.append(auc_test)
816                 accs_test.append(acc_test)
817                 continue
818             elif model_name == 'cnn':
819                 X_train_cnn = X_train.values.reshape(X_train.shape[0],
                    ↪  X_train.shape[1], 1)
820                 X_val_cnn = X_val.values.reshape(X_val.shape[0], X_val.shape[1], 1)
821                 X_test_cnn = X_test.values.reshape(X_test.shape[0], X_test.shape[1], 1)
822                 model = create_cnn_model((X_train.shape[1], 1))
823                 model.fit(X_train_cnn, y_train, epochs=10, batch_size=32,
                    ↪  validation_data=(X_val_cnn, y_val), verbose=0)
824                 y_pred_prob_val = model.predict(X_val_cnn).ravel()
825                 y_pred_prob_test = model.predict(X_test_cnn).ravel()
826                 auc_val = roc_auc_score(y_val, y_pred_prob_val)
827                 auc_test = roc_auc_score(y_test, y_pred_prob_test)
828                 y_pred_val = (y_pred_prob_val > 0.5).astype(int)
829                 y_pred_test = (y_pred_prob_test > 0.5).astype(int)
830                 acc_val = accuracy_score(y_val, y_pred_val)
831                 acc_test = accuracy_score(y_test, y_pred_test)
832                 aucs_val.append(auc_val)
833                 accs_val.append(acc_val)
834                 aucs_test.append(auc_test)
835                 accs_test.append(acc_test)
836                 continue
837             else:
838                 raise ValueError(f"Modelo no soportado: {model_name}")
839 
840             model.fit(X_train, y_train)
841             if model_name == 'perceptron' or model_name == 'ridge':
842                 y_pred_prob_val = model.decision_function(X_val)
843                 y_pred_prob_test = model.decision_function(X_test)
844             else:
845                 y_pred_prob_val = model.predict_proba(X_val)[:, 1]
846                 y_pred_prob_test = model.predict_proba(X_test)[:, 1]
847 
848             auc_val = roc_auc_score(y_val, y_pred_prob_val)
849             auc_test = roc_auc_score(y_test, y_pred_prob_test)
850             y_pred_val = model.predict(X_val)
851             y_pred_test = model.predict(X_test)
852             acc_val = accuracy_score(y_val, y_pred_val)
853             acc_test = accuracy_score(y_test, y_pred_test)
854 
855             aucs_val.append(auc_val)
856             accs_val.append(acc_val)
857             aucs_test.append(auc_test)
858             accs_test.append(acc_test)
859 
```

```
860     return aucs_val, accs_val, aucs_test, accs_test
861
862 def evaluate_all_models(database, target_variable, feature_sets, n_splits=5):
863     models = ['lda', 'qda', 'perceptron', 'mlp', 'ridge', 'rf', 'xgb',
        ↪  'deep_learning', 'knn', 'cnn']
864     results = []
865
866     df = pd.DataFrame(database)
867     df = df[df[target_variable] != -1]
868
869     for model_name in models:
870         for description, include_variables in feature_sets.items():
871             subset_df = df.dropna(subset=include_variables)
872             aucs_val, accs_val, aucs_test, accs_test = evaluate_model_cv(subset_df,
                ↪  target_variable, include_variables, model_name, n_splits)
873             for fold in range(n_splits):
874                 results.append({
875                     'Model': model_name,
876                     'Features': description,
877                     'Fold': fold + 1,
878                     'AUC_Val': aucs_val[fold],
879                     'ACC_Val': accs_val[fold],
880                     'AUC_Test': aucs_test[fold],
881                     'ACC_Test': accs_test[fold]
882                 })
883             print(f'{model_name} - AUC_Val {description}: {np.mean(aucs_val):.4f},
                ↪  ACC_Val {description}: {np.mean(accs_val):.4f}, AUC_Test
                ↪  {description}: {np.mean(aucs_test):.4f}, ACC_Test {description}:
                ↪  {np.mean(accs_test):.4f}')
884
885     results_df = pd.DataFrame(results)
886     return results_df
887
888 def evaluate_multiple_databases(databases, target_variable, feature_sets,
    ↪  n_splits=5):
889     all_results = []
890     for db_name, database in databases.items():
891         print(f"\nEvaluating database: {db_name}")
892         results_df = evaluate_all_models(database, target_variable, feature_sets,
            ↪  n_splits)
893         results_df['Database'] = db_name
894         all_results.append(results_df)
895         print(results_df)
896     return pd.concat(all_results, ignore_index=True)
897
898 #Paso 15: Aplicacion del modelo de aprendizaje hasta con 10 FOLD, a los escenarios
    ↪  y datasets definidos
899
900 # Definir los conjuntos de características
901 feature_sets = {
902     'Sin FE': columns_to_keep_SF,
903     'valor_FE_Convulsionsocial': columns_to_keep_Convulsionsocial,
904     'valor_FE_TemperaturaAnomalia': columns_to_keep_TemperaturaAnomalia,
905     'valor_FE_BloqueoFClimatico': columns_to_keep_BloqueoFClimatico,
906     'valor_FE_TemperaturaAnomalia_FE_BloqueFClimatico':
        ↪  columns_to_keep_TemperaturaAnomalia_BloqueFClimatico,
907     'valor_FE_Convulsionsocial_FE_BloqueFClimatico':
        ↪  columns_to_keep_Convulsionsocial_BloqueFClimatico,
908     'valor_FE_Convulsionsocial_FE_TemperaturaAnomalia':
        ↪  columns_to_keep_Convulsionsocial_TemperaturaAnomalia,
909     'valor_FE_Convulsionsocial_FE_TemperaturaAnomalia_FE_BloqueFClimatico':
        ↪  columns_to_keep_TemperaturaAnomalia_Convulsionsocial_BloqueFClimatico
910 }
911
912
913 # Ejecutar la evaluación con los diferentes conjuntos de características en
    ↪  múltiples bases de datos
```

```
databases = {
    'data100': data100,
    'data99': data99,
    'data80': data80
}

results_df = evaluate_multiple_databases(databases, 'ClassBM', feature_sets, 10)
print(results_df)

results_df.to_csv('/content/drive/MyDrive/Doc/Data2/ML_CONV_CCLIMAT_20240909_99.csv',
↪  sep=';', index=False)
```

Anexo G

Código fuente del proceso validación de significancia estadística

En este anexo se presenta el código fuente completo utilizado para la determinación de la significancia estadística usando la t de Student. El código está implementado en Python y documentado para su comprensión y posible reproducción.

Archivo: `Calcula_significancia_estad.py`

```
# -*- coding: utf-8 -*-
"""Calcula Nivel de Significancia de Resultados_Final_v2.ipynb

Automatically generated by Colab.

Original file is located at
    https://colab.research.google.com/drive/1Gbvm7HeVBOV6VFig3pHcXiaOk5GG_aBx
"""

#Paso 1 Conectar a Drive
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.stattools import adfuller, coint

import pandas as pd

from google.colab import drive
drive.mount('/content/drive')

import pandas as pd
import numpy as np
from scipy import stats

#file_path = '/content/drive/MyDrive/Doc/Data2/Resultados de ML_Folds.xlsx'
#file_path = '/content/drive/MyDrive/Doc/Data2/Resultados de ML_Folds_FINAL_V5.xlsx'
#file_path = '/content/drive/MyDrive/Doc/Data2/Resultados de ML_Folds_FINAL_V6.xlsx'
file_path = '/content/drive/MyDrive/Doc/Data2/Resultados de ML_Folds_FINAL_V7.xlsx'


# Cargar los datos desde un archivo CSV
data = pd.read_excel(file_path)
```

```
32
33  data.head()
34
35  data.info()
36
37  print(data['Tipo Eval'].unique())
38
39  # Configurar pandas para mostrar todas las filas y columnas
40  pd.set_option('display.max_rows', None)
41  pd.set_option('display.max_columns', None)
42
43  #T student con una sola cola a la izquierda
44
45  import pandas as pd
46  import numpy as np
47  from scipy import stats
48
49  # Configurar pandas para mostrar todas las filas y columnas
50  pd.set_option('display.max_rows', None)
51  pd.set_option('display.max_columns', None)
52
53  # Cargar los datos desde un archivo CSV
54  # data = pd.read_csv('/mnt/data/data.csv')
55
56  # Convertir las columnas de porcentaje a float
57  #data[['ACC_Test']] = data[['ACC_Test']].replace('%', '', regex=True).astype(float)
    ↪  / 100.0
58  #data[['AUC_Test']] = data[['AUC_Test']].replace('%', '', regex=True).astype(float)
    ↪  / 100.0
59  #data[['ACC_Val']] = data[['ACC_Val']].replace('%', '', regex=True).astype(float) /
    ↪  100.0
60  #data[['AUC_Val']] = data[['AUC_Val']].replace('%', '', regex=True).astype(float) /
    ↪  100.0
61
62  # Filtrar los escenarios SFE y otros
63  sfe_data = data[data['Escenario'] == 'Sin FE']
64  other_data = data[data['Escenario'] != 'Sin FE']
65
66  # Obtener la lista de tipos de evaluación y modelos únicos
67  eval_types = data['Tipo Eval'].unique()
68  models = data['Modelo'].unique()
69
70  # Inicializar un DataFrame para los resultados
71  results = pd.DataFrame(columns=['Tipo Eval', 'Modelo', 'Escenario', 'Métrica',
    ↪  'Mean_SFE', 'Mean_Other', 'Max_SFE', 'Max_Other', 't_stat', 'p_value',
    ↪  'Significativo'])
72
73  # Función para realizar la prueba t de Student con una cola a la izquierda
    ↪  (muestras relacionadas) y guardar los resultados
74  def t_test_and_store(sfe_group, other_group, eval_type, model, scenario, metric):
75      # Realizar la prueba t para muestras relacionadas (pareadas)
76      t_stat, p_value = stats.ttest_rel(sfe_group[metric], other_group[metric])
77
78      # Ajustar el valor p para una prueba de una sola cola a la izquierda
79      p_value_one_tailed = p_value / 2 if t_stat < 0 else 1.0
80
81      # Determinar si es significativo
82      significant = p_value_one_tailed < 0.05
83
84      # Calcular medias y máximos
85      mean_sfe = sfe_group[metric].mean()
86      mean_other = other_group[metric].mean()
87      max_sfe = sfe_group[metric].max()
88      max_other = other_group[metric].max()
89
90      # Guardar los resultados
```

```
91      results.loc[len(results)] = [eval_type, model, scenario, metric, mean_sfe,
        ↪  mean_other, max_sfe, max_other, t_stat, p_value_one_tailed, significant]
92
93  # Lista de métricas a evaluar
94  metrics = ['ACC_Test', 'AUC_Test', 'ACC_Val', 'AUC_Val']
95
96  # Iterar sobre cada tipo de evaluación y modelo
97  for eval_type in eval_types:
98      for model in models:
99          # Filtrar los datos por tipo de evaluación y modelo
100         sfe_group = sfe_data[(sfe_data['Tipo Eval'] == eval_type) &
            ↪  (sfe_data['Modelo'] == model)]
101         for scenario in other_data['Escenario'].unique():
102             other_group = other_data[(other_data['Tipo Eval'] == eval_type) &
                ↪  (other_data['Modelo'] == model) & (other_data['Escenario'] ==
                ↪  scenario)]
103             if not sfe_group.empty and not other_group.empty:
104                 # Realizar la prueba t para cada métrica y guardar los resultados
105                 for metric in metrics:
106                     t_test_and_store(sfe_group, other_group, eval_type, model,
                        ↪  scenario, metric)
107
108 # Guardar los resultados en un archivo CSV
109 results.to_csv('t_test_metrics_results_with_max_one_tailed_V07.csv', index=False)
110 #results_df.to_csv('/content/drive/MyDrive/Doc/Data2/t_test_metrics_results_with_max_one_tailed_V07.
    ↪  sep=';', index=False)
111 # Mostrar los resultados
112 print(results)
113
114 results.info()
115
116 print(results['Tipo Eval'].unique())
```

Anexo H

Artículo: Machine Learning for Credit Risk Prediction: A Systematic Literature Review

Este artículo presenta una revisión sistemática de la literatura sobre el uso de *ML* en la predicción de riesgo crediticio. Se identifican 52 estudios relevantes publicados entre 2019 y 2023, analizando modelos, métricas, datasets, variables y limitaciones comunes. Destaca que los modelos de la categoría *Boosted* (p.ej., *XGBoost*) son los más investigados y que métricas como *AUC*, *ACC*, *Recall*, *F1* y *Precision* son las más empleadas. El trabajo enfatiza los retos en la selección de variables relevantes, manejo de datos desbalanceados y explicabilidad de modelos, ofreciendo una guía para desarrolladores y entidades financieras.

Referencia: Data 2023, 8, 169. https://doi.org/10.3390/data8110169

data

MDPI

Article

Machine Learning for Credit Risk Prediction: A Systematic Literature Review

Jomark Pablo Noriega [1,2,*,†], Luis Antonio Rivera [1,3,†] and José Alfredo Herrera [1,4,†]

1 Departamento Académico de Ciencia de la Computacion, Universidad Nacional Mayor de San Marcos, Decana de América, Lima 15081, Peru; lriverae@unmsm.edu.pe or rivera@uenf.br (L.A.R.); jherreraqu@unmsm.edu.pe (J.A.H.)
2 Financiera QAPAQ, Lima 150120, Peru
3 Centro de Ciências Exatas e Tecnologia, Universidade Estadual do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes 28013-602, Brazil
4 Programme in Biotechnology, Engineering and Chemical Technology, Universidad Pablo de Olavide, 41013 Sevilla, Spain
* Correspondence: jomark.noriega@unmsm.edu.pe or jnoriega@qapaq.pe; Tel.: +51-987-521-006
† These authors contributed equally to this work.

Abstract: In this systematic review of the literature on using Machine Learning (ML) for credit risk prediction, we raise the need for financial institutions to use Artificial Intelligence (AI) and ML to assess credit risk, analyzing large volumes of information. We posed research questions about algorithms, metrics, results, datasets, variables, and related limitations in predicting credit risk. In addition, we searched renowned databases responding to them and identified 52 relevant studies within the credit industry of microfinance. Challenges and approaches in credit risk prediction using ML models were identified; we had difficulties with the implemented models such as the black box model, the need for explanatory artificial intelligence, the importance of selecting relevant features, addressing multicollinearity, and the problem of the imbalance in the input data. By answering the inquiries, we identified that the Boosted Category is the most researched family of ML models; the most commonly used metrics for evaluation are Area Under Curve (AUC), Accuracy (ACC), Recall, precision measure F1 (F1), and Precision. Research mainly uses public datasets to compare models, and private ones to generate new knowledge when applied to the real world. The most significant limitation identified is the representativeness of reality, and the variables primarily used in the microcredit industry are data related to the Demographic, Operation, and Payment behavior. This study aims to guide developers of credit risk management tools and software towards the existing ability of ML methods, metrics, and techniques used to forecast it, thereby minimizing possible losses due to default and guiding risk appetite.

Keywords: loan; credit risk; prediction; machine learning; systematic literature review

check for updates

Citation: Noriega, J.P.; Rivera, L.A.; Herrera, J.A. Machine Learning for Credit Risk Prediction: A Systematic Literature Review. *Data* **2023**, *8*, 169. https://doi.org/10.3390/data8110169

Academic Editors: Henry Han, Qiannong (Chan) Gu, Diane Li, Tie Wei and Jeffrey Yi-Lin Forrest

Received: 5 August 2023
Revised: 13 October 2023
Accepted: 3 November 2023
Published: 7 November 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The digitalization of processes and AI are already part of our daily lives and have been developing in all areas with which we interact, especially during the period of the COVID-19 pandemic [1,2]. This trend continued with the promotion of online loans and Internet sales [3], and in turn, with the increase in demand for credit and crowdfunding through the Internet in the short term [4–6], and considering the pandemic as an external factor, it has considerably changed the economy, increasing uncertainty for financial institutions, and consequently the need to generate new models to manage it [4]. The World Bank (WB) emphasizes that the banking sector is crucial because it improves the well-being of a country's population and is essential for the growth of the economy [7].

In this competitive environment, financial institutions seek to differentiate themselves, generate shareholder value, improve the customer experience, and promote financial

inclusion. In this sense, they face the challenge of adopting data-driven innovation (DDI) to manage credit, customer and operational risk appetite, and seeking efficiency [8] while being supported by information technology—cloud services, Internet of Things (IoT), BIG DATA, AI and mobile telephony—known as the fourth industrial revolution [9]. This is followed by its sustainable evolution towards virtuous interaction between humans, technology, and the environment, called the fifth industrial revolution [10].

There are still many use cases for DDI and ML in financial institutions to solve [11,12]. There are gaps in the evaluation and precariousness in the results when processing large volumes of information [13] using ML algorithms for real-time applications [8]. On the other hand, using machine learning techniques, through association, collaborative filtering, along with recommended and personalized content, systems can identify individual preferences that could enrich the risk assessment [8]. Likewise, BIG DATA tools and technology could help improve forecasts in changing markets, considering their analysis potential [11,14].

Utilizing machine learning in business intelligence to reduce the uncertainty of payment behavior, also called credit risk, is a necessity in the microfinance industry since it allows for the analysis of large volumes of information generated, especially in the context of post-pandemic COVID-19 and technological development [4,15]. The challenges are determining which configurations of attributes and algorithms are best suited for the tasks, together with identifying limitations in the applications. Consequently, we propose the research topics shown in Table 1.

Table 1. Research Topics.

Motivation	Research Topics
We wish to know what models the industry and academics use to predict credit risk.	The algorithms, methods, and models used to predict credit risk.
We wish to know what metrics to use in the industry for academics to evaluate the performance of algorithms, methods, or models to predict credit risk.	The metrics to evaluate the performance of algorithms, methods, or models.
We wish to know the metrics accuracy, precision, F1 measure, and AUC of algorithms, methods, or models to predict credit risk.	The models' accuracy, precision, F1 measure, and AUC.
We wish to know what datasets to use in the industry and for academics to predict credit risk.	The datasets are used in the prediction of credit risk.
We wish to know what variables or features to use in the industry and for academics to predict credit risk.	The variables or features used in the prediction of credit risk.
We wish to know the main problems or limitations to predicting credit risk.	The main problems or limitations of predicting credit risk.

2. Materials and Methods

Based on the research topics, we pose the questions in Table 2 as a first step. Using the most relevant words related to our research topic, we build our search string and apply it to the recognized databases: IEEE Xplore, Scopus, and Web Of Science (WOS). We consider the studies belonging to articles from journals and conferences published from 2019 to May 2023 and are related to the computer science area as shown in Figure 1.

We identified 275 studies. There were 77 eliminations for being duplicates and 131 eliminations from applying the eligibility criteria (without full access to the document, without ranking in Scimagojr, review article) and 15 eliminations for not having relevance after the complete analysis of the document. We obtained as a result 52 revealing articles; we show in Table 3 the result of the application of the inclusion and exclusion criteria.

Table 2. Research Questions.

Id.	Metrics
RQ1	What are the algorithms, methods, and models used to predict credit risk?
RQ2	Which are the metrics to evaluate the performance of algorithms, methods, or models?
RQ3	What are these models' accuracy, precision, F1 measure, and AUC?
RQ4	What datasets are used in the prediction of credit risk?
RQ5	What variables or features are used to predict credit risk?
RQ6	What are the main problems or limitations of predicting credit risk?

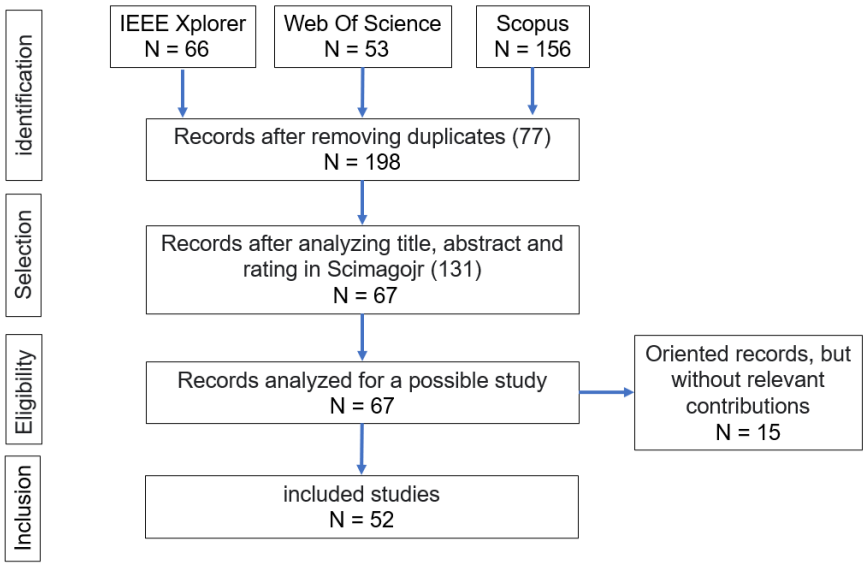


Figure 1. PRISMA Research Strategy.

Research string:

TITLE-ABS-KEY ((“credit” OR “loan”) AND (“machine learning”) AND (“model” OR “algorithm” OR “method”) AND (“financial”) AND (“credit risk”)) AND (LIMIT-TO (PUBYEAR, 2023) OR LIMIT-TO (PUBYEAR, 2022) OR LIMIT-TO (PUBYEAR, 2021) OR LIMIT-TO (PUBYEAR, 2020) OR LIMIT-TO (PUBYEAR, 2019)) AND (LIMIT-TO (DOCTYPE, “cp”) OR LIMIT-TO (DOCTYPE, “ar”)) AND (LIMIT-TO (SUBJAREA, “COMP”)).

Table 3. Application of inclusion and exclusion criteria.

Inclusion Criteria	Exclusion Criteria	#	%
Article of conference		2	0.73%
Article of journal		50	18.18%
	Article duplicated	77	28.00%
	Not related	15	5.45%
	Review article	1	0.36%
	Without access to the full document	57	20.73%
	Without rank in Scimagojr	73	26.55%
	Total	275	100.00%

Research Strategy

To determine the importance of studies, we assess that they include relevant conclusions or results; attributes or features of datasets; descriptions of the proposed model or algorithm; the metrics with which the models were evaluated; preprocessing techniques; identified problems or limitations; and future studies. We also consider when the article

includes more than one set of data or experiments; we select one of them to present the metrics considering the most relevant in terms of ACC, AUC, or another metric in the event the indicated ones are not used. According to our criteria and in the cases which it applies, we decided on the German Dataset, which is the dataset of the University of California Irvine (UCI), considering that it is imbalanced [16], of frequent use, and it will allow us to compare results between investigations that include it.

The main limitation when carrying out a systematic review is bias, and considering how the decisions of researchers influence the application of the method framed in experiences and previous knowledge. For example, the choice of topic, choice of electronic resources, proposal of research questions, methodology for data collection, and their evaluation. We have tried to meticulously follow the PRISMA method, taking into account the application of the most appropriate criteria and procedures in the construction of this document.

3. Current Research

The demand for online credit generates considerable information, with which, when analyzed using BIG DATA [17], designs new products, machine learning models, and credit risk assessment methods [18]. Consequently, in scenarios of increased demand, credit risks also escalate considerably, in a non-linear manner, considering the level of risk, rate, and terms of credit [19]. In the same manner, there is an expectation of an increase in fraud in the following years [20]. Another problem to consider is the consistency of the information recorded at the different stages of the process, such as sales data [21], cultural variables, environmental data [22], macroeconomics [23], innovation capacity management and development, exchange rate evolution, Gross Domestic Product (GDP) growth trends [23,24], economic activity, and experience [25].

The relevant problems are addressed by various research papers, using different ML approaches for the respective interpretations and decision-making [26]. However, there are also difficulties with the implemented models [27], which generally follow the black box model; that is, to predict, for example, the good and bad payers [28–30]. These models have presented problems, especially in difficult times, such as "the financial crisis of 2008" [1,31], since financial institutions focus on loans that generate the most income, being, therefore, of higher risk due to payment defaults [19,22]. Automatic evaluation models, based on credit data, could confuse good paying customers for bad ones [20], and apply penalties on possible benefits [32]. The low explainability of advanced non-linear ML models is slowing down their implementation [33]. The challenge is in the development of Explainable artificial intelligence (XAI), whose objective is to provide predictive models with inherent transparency, especially in complex cases [34,35]. In that sense, one could use Shapley Additive exPlanations (SHAP) or the Generalized Shapley Choquet integral (GSCI) to expose the parameter dependency [36,37], and a model interpretable payment risk forecast from a penalized logistic tree to enhance the performance of Logistic Regression (LR), and LR + Slack Based Measure (SBM) [38,39]. The MLIA model and variable logistic regression coefficient more intuitively reveal the contribution of a variable to the predicted outcome [40]. The "non-payment" problem is important because it could generate significant losses for financial entities [41,42].

Here, the challenge of machine learning is to consider the existing multicollinearity in the input data [43,44] where there is a high correlation of variables and some that is not useful for classification [2]. This imbalance in the data actually used, with probable overfit [20], could generate biases in machine learning [45,46], causing chaotic reputation management, and malicious or criminal deception [47]. In other words, the challenge is to determine the effective, relevant features [48], for example, for the training of neural networks (NNs) (which is performed end-to-end by interactions). These are additional constraints with desirable features known a priori in order to reduce the number of interactions and prioritize smoothness; they are factors of explanatory interest, in the domain, control, and generative modeling of features [49]. Other authors recommend the use of genetic

algorithms (GAs) to guide training, with data sequences that have the best result [50]; the use of hive intelligence (HI) is also highlighted for this purpose [51–53]; Boost Category models (BCat) such as Adaboost (ADAB), XGBoost (XGB), LightGBM (LGMB), and Bagging (BAG) [30,40,54–56]; multi-classification (MC) and information fusion models (MIFCA) [44]; there is also noise removal using fuzzy logic (FL) and contribution to the identification of main attributes [28,37,57]. It is worth noting the possibility to evaluate the interaction of borrowers within their supply chain to enrich predictive models [17,58]. The use of images, interviews, text and social information, and interaction with mobile technology would give the credit risk assessment a multidimensional and multi-focus characteristic [15,59,60]; this indicates the inclusion of integrated accounting information with statistical variables of profitability, liquidity, asset quality, management indices, capital adequacy, efficiency, SCORECARD scorecard and the maximization of the internal rate of return (IRR), risks of the industry, and GDP [1,23].

Some authors maintain that the most relevant characteristics are gender, educational level, mortgage credit, microfinance, debt balance, and days past due [61]. Other authors maintain that the most relevant variables are the days of default, especially those greater than 90 days, to determine non-payment behavior and consider that discriminatory variables such as gender, age, and marital status should not be considered [55,61]. Consequently, the challenge of validating the quality of the features arises [39,62]. For this part, [17] maintains that the number of influential variables for risk assessment has increased, and the linear and non-linear time series relationships have increased their complexity.

There are also methods to address the imbalance problem: Random Undersampling (RUS), Near Miss (NMISS), Cluster Centroid (CC), Random Oversampling (ROS), Adaptive Synthetic (ADASYN), K-Nearest Neighbor (KNN) SMOTE (KN-SMOTE), Synthetic Minority Oversampling Technique (SMOTE), Borderline-Smot (B-SMOT), and Smotetomek (SMOTE-T) [63]. Other authors propose the use of CS-classifiers [64], and KFold [29,55]. However, other authors propose considering imbalance and missing data as characteristics to take into account in the [54,56], evaluation. As an initial part of the model evaluations in the experiments, the importance of optimizing the hyperparameters highlight using, for example, GA [15], K-Fold CV [38], random search (RS) [20], grid search (GS) [57,63], and other methods [56].

In Table 4, we summarize the challenges in credit risk prediction and the ML methods and techniques proposed by the authors to address them. The main challenge is high uncertainty. The authors propose various classification models, including Boosted category models, neural networks, deep learning enhanced with fuzzy logic, genetic algorithms, and hive intelligence. For the low explainability of the results, the authors propose applying Explainable artificial intelligence, Shapley Additive exPlanations, Generalized Shapley Choquet integral, and MLIA. To address the complexity of ML models, the authors propose optimizing the hyperparameters of the models supported with KFold CV, genetic algorithms, and grid search. For the multivariate origin of the data, the authors propose applying BIG DATA to take advantage of its prominent volume characteristics and its unstructured nature. To address the natural characteristics of unbalanced data, the authors propose the application of SMOTE, RUS, ROS, KFold, and ADASYN.

Table 4. Challenge of Credit Risk Prediction.

It.	Challenges	ML Methods & Techniques	#	Author
1	Uncertainty	NN, DL, BCAT, MC, MIFCA, FL, GA, HI	28	[2,15,17,19,20,28,30,36–58]
2	Explainability	XAI, SHAP, GSCI, MLIA	18	[1,19,20,22,27–40]
3	Complexity	GA, K-Fold CV, RS, GS	9	[15,17,20,23,24,38,56,57,63]
4	Multivariate Data	BIG DATA	9	[1,15,17,23–25,59–61]
5	Unbalanced Data	SMOTE, RUS, ROS, KFold, CS, ADASYN	7	[20,29,54–56,63,64]

4. Results

4.1. Answer to RQ1: What Are the Algorithms, Methods, and Models Used to Predict Credit Risk?

In their research to forecast credit risk, the authors use ML models: 72.76% non-assembled (N-Ass) and 27.24% assembled (Ass), which are shown in Table 5.

Table 5. Family of algorithms, methods, and models.

It.	Family	#			%		
		Ass	N-Ass	Total	Ass	N-Ass	Total
1	Boosted Category	36	46	82	11.96%	15.28%	27.24%
2	Collective Intelligence	7	-	7	2.33%	0.00%	2.33%
3	Fuzzy Logic	10	-	10	3.32%	0.00%	3.32%
4	NN/DL	8	28	36	2.66%	9.30%	11.96%
5	Other Model	3	10	13	1.00%	3.32%	4.32%
6	Traditional	18	135	153	5.98%	44.85%	50.83%
	Total	82	219	301	27.24%	72.76%	100.00%

For better presentation we group them into the following families: Boosted Category, the models related to the Boosted algorithm; Collective Intelligence, models related to collective or swarm intelligence including Ant Colony Optimization (ACO), Bat Algorithm (BAT), and Particle Swarm Optimization (PSO); Fuzzy Logic, models related to Fuzzy Logic including Clustering-Based Fuzzy Classification (CBFC); NN/DL, models related to neural networks or Deep Learning (DL) including Back Propagation Neural Network (BPNN), Artificial Neural Network (ANN), Multilayer Perceptron (MLP), Wide and Deep Neural Networks (Wide&Deep), Gated Recurrent Unit (GRU), Geometric Deep Learning (GDL), Graph Neural Networks (GNNs), Deep Genetic Hierarchical Network of Learners (DG), and Convolutional Neural Networks (CNNs); Other Models, for the models not cataloged; and Traditional, for the models cataloged but not related to previous models including Decision Tree (DT), Decision Tree C4.5 (C4.5), Classification And Regression Tree (CART), K-Means (KM), Linear Discriminant Analysis (LDA), Non-Linear Logistic Regression (NLR), Naive Bayes (NB), Random Forest (RF), Random Tree (RT), Support Vector Machine (SVM), and Sequential Minimal Optimization (SMO). Of the total models used, 50.83% correspond to the Traditional model family, 27.24% to the Boosted Category, and 11.96% to NN/DL.

Analyzing the groups separately corresponding to N-Ass as shown in Figure 2, the Category Boosted family was used in 21% of the studies, NN/DL has a use rate of 12.8%, Traditional has a use rate of 61.6%, and Other Models have a use rate of 4.6%.

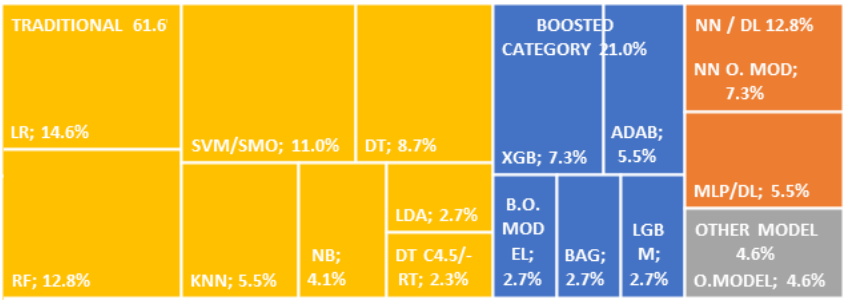


Figure 2. Non-Assembled Models.

In Figure 3, the Category Boosted family was used 43.9% of the time, NN/DL has a use rate of 9.8%, Traditional has a use rate of 22%, Other Models have a use rate of 3.7%, Fuzzy Logic has a use rate of 12.2%, and Collective Intelligence is used in 8.5% of the total. These results will demonstrate that N-Ass models are mostly used in credit risk prediction. However, the authors use these as a baseline to compare them with Ass models generated by nesting N-Ass models and improving them using fuzzy logic and hive intelligence.

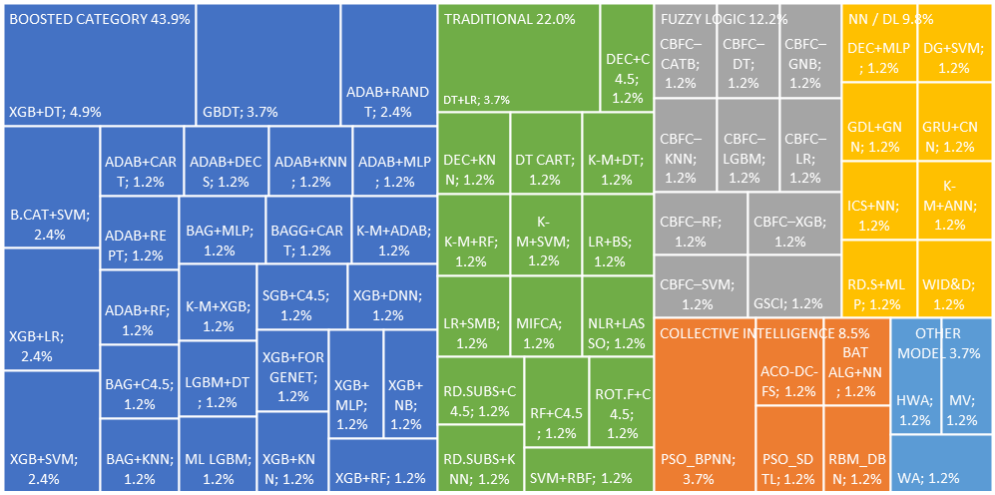


Figure 3. Assembled Models.

4.2. Answer to RQ2: Which Are the Metrics to Evaluate the Performance of Algorithms, Methods, or Models?

We collected the metrics used by the researchers in their articles, taking into account that for the cases in which they evaluate more than one dataset and more than one model is applied, a pair is searched, looking forthe best ACC and AUC values, or another value in the cases where these are not used, as shown in Figure 4.

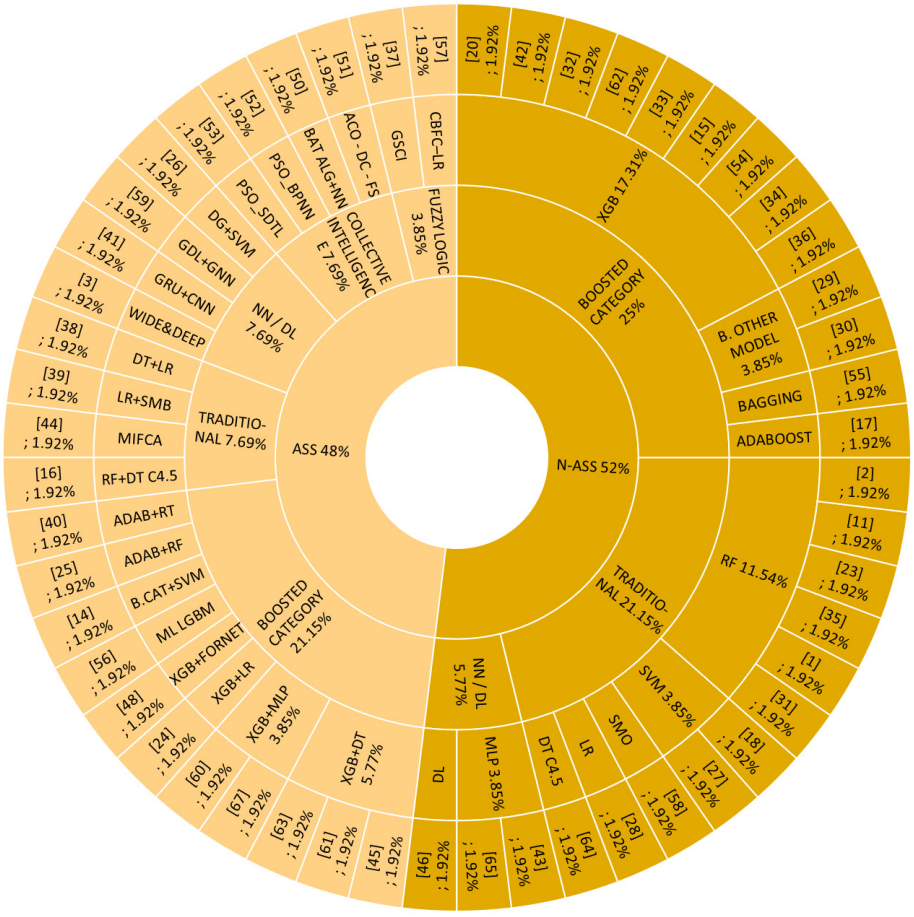


Figure 4. Best Models with family and author.

From this simplification, the authors propose 48% assembled models and 52% non-assembled models.

Of the total assembled models, the Boosted Category has 21%; Collective Intelligence has 8%; NN/DL has 8%, Traditional has 8%, and Fuzzy Logic has 4%. Of the non-assembled models, the Boosted Category makes up 25%; Traditional is 21%; and NN/DL is 6%.

In these models, the metrics used by the authors are dividied into AUC with 16%, ACC with 14%, F1 measure with 11%, others with 11%, Precision with 10%, Recall with 9%, True Positive Rate (TPR) with 7%, True Negative Rate (TNR) with 6%, Geometric Mean (GMEAN) with 4%, Kolmogorov–Smirnov (KS) with 3%, Brier Score (BS) with 3%, GINNI Score (GINNI) with 2%, and Root Mean Squared Error (RMSE), KAPPA coefficient (KAPPA), and Mean Absolute Error (MAE), whose group participation is 2%; the details are shown in Table 6.

These results could demonstrate that the most-used metrics for credit risk prediction are AUC and ACC since they allow for the comparison of different models. However, AUC is prioritized because the distribution of the classes does not influence it and has better behavior when using unbalanced data.

Table 6. Metrics.

It.	Metrics	#	%	Author
1	AUC	34	16.11%	[1,2,11,16,17,20,24,25,27,28,30,32–43,46,50,54–58,61–64]
2	ACC	30	14.22%	[3,11,15,18,25–32,35,36,36,37,39,44,46,48,50–53,55,57,58,60–63,65]
3	F1	24	11.37%	[11,17,18,24,27,29,30,32,36,37,39,41,42,44,48,51,56,58,60–65]
4	Precis.	22	10.43%	[3,11,17,18,24,27–30,32,36,39,41,42,44,46,58,60–62,65]
5	Recall	19	9.00%	[11,24,27–30,36,39,41,42,44,46,58,60–63,65,66]
6	TPR	14	6.64%	[16,18,24,25,31,32,35,37,45,50,51,55,57,65]
7	TNR	13	6.16%	[16,25,31,32,35,37,39,45,50,51,55,57,65]
8	GMEAN	9	4.27%	[20,31,35,39,45,48,50,63,64]
9	KS	7	3.32%	[24,36,38,39,43,55,56]
10	BS	6	2.84%	[33,37,38,55,57,62]
11	GINNI	5	2.37%	[2,24,38,43,46]
12	RMSE	2	0.95%	[53,67]
13	KAPPA	1	0.47%	[51]
14	MAE	1	0.47%	[67]
15	Other	24	11.37%	[2,14,23,25,28,33,35,36,38,41,43,45,46,51,53,54,57,59–62,64,65,67]
	Total	211	100.00%	

4.3. Answer to RQ3: What Are These Models' Accuracy, Precision, F1 Measure, and AUC?

We listed the values of the five metrics most used by the authors in their research: AUC, ACC, F1 measure, Precision, and Recall, as shown in Table 6. Furthermore, we have taken the values in each case according to the tuple defined in the question. To compare experiments, the characteristics of the dataset and metric to be evaluated should correspond. Below, we show only the metrics' values the authors used and evaluated in their research in Table 7. If the metrics have an empty value, the authors do not consider them in their experiment design.

Table 7. Metrixs author.

It.	Dataset	Author	Metrics' Values				
			ACC	Precision	F1	Recall	AUC
1	UCI Taiwan	[29]	85.00	70.00	50.00	62.00	-
2	UCI German	[63]	83.50	82.10	84.40	86.80	91.00
3	UCI German	[25]	82.80	-	-	-	91.20
4	UCI German	[50]	81.18	-	-	-	85.38
5	UCI German	[51]	76.60	-	84.74	-	-

Table 7. *Cont.*

It.	Dataset	Author	Metrics' Values				
			ACC	Precision	F1	Recall	AUC
6	UCI German	[28]	75.80	54.20	-	82.00	85.90
7	UCI German	[57]	74.90	-	-	-	75.80
8	UCI German	[16]	-	-	-	-	79.40
9	Lending Club	[32]	92.60	97.90	92.20	-	97.00
10	Lending Club	[65]	84.40	88.99	91.42	93.98	-
11	Lending Club	[30]	76.10	75.98	75.95	76.35	76.80
12	Lending Club	[48]	88.77	-	94.14	-	-
13	Lending Club	[31]	74.90	-	-	-	-
14	Lending Club	[35]	64.00	-	-	-	71.70
15	Lending Club	[36]	63.60	85.30	73.50	64.50	67.40
16	Lending Club	[46]	-	18.25	-	46.88	63.63
17	Lending Club	[56]	-	-	2.72	-	75.86
18	K Prosper	[3]	78.50	54.70	-	-	-
19	K Prosper	[17]	-	79.00	71.00	65.00	80.00
20	K Give Me	[61]	88.30	78.50	77.60	76.70	93.30
21	RenRenDai	[37]	93.35	-	73.12	-	82.64
22	BR	[55]	96.68	-	-	-	89.63
23	AVG Used	[11]	92.80	31.60	33.40	35.50	82.80
24	AVG Used	[64]	-	-	91.89	-	96.19
25	UCI Austr...	[26]	97.39	-	-	-	-
26	Tsinghua	[52]	91.23	-	-	-	-
27	Tsinghua	[62]	77.20	75.90	77.54	79.38	85.01
28	Private Data	[18]	98.34	100.00	96.00	-	-
29	Private Data	[53]	98.00	-	-	-	-
30	Private Data	[60]	97.80	98.90	98.70	98.90	-
31	Private Data	[15]	90.10	-	-	-	-
32	Private Data	[27]	84.29	82.63	84.68	86.83	84.29
33	Private Data	[44]	84.15	82.15	83.40	84.68	-
34	Private Data	[58]	83.00	83.50	83.00	83.00	83.30
35	Private Data	[39]	77.49	79.87	85.59	92.18	79.00
36	Private Data	[24]	-	87.15	84.56	83.91	83.59
37	Private Data	[54]	-	-	-	-	46.10
38	Private Data	[1]	-	-	-	-	75.40
39	Private Data	[38]	-	-	-	-	85.68
40	Private Data	[33]	-	-	-	-	93.39
41	Private Data	[34]	-	-	-	-	93.00
42	Private Data	[42]	-	42.81	52.00	67.01	78.00
43	Private Data	[40]	-	-	-	-	71.32
44	Private Data	[2]	-	-	-	-	91.40
45	Private Data	[41]	-	88.00	88.00	88.00	93.00
46	Private Data	[43]	-	-	-	-	77.56
47	Private Data	[20]	-	-	-	-	95.50

Considering that there are experiments in which the same dataset is used, we can compare their prediction capabilities. This occurs in the case of the UCI German Dataset, in which the XGB + DT model has an ACC of 84 [63], against the LR models [28], and Random Forest (RF) + C4.5 [16]. Other less-used datasets on which we can compare the metrics are the Tsinghua—Pekin U RESSET database and the Kaggle Prosper dataset, as shown in Table 7. It is essential to indicate that the Lending Club dataset is possible to compare; however, it is necessary to consider the different date ranges for the investigations that the authors used [35,36].

4.4. Answer to RQ4: What Datasets Are Used in the Prediction of Credit Risk?

The datasets used by the authors can be divided into 53.85% public and 46.15% private, seeking mainly in the first case to validate new models. To be able to compare them with experiments, we show the data in Figure 5—private datasets used to extract knowledge by revalidating existing models in real scenarios. In the public group, the most used is the UCI German dataset with a usage of 15.38%, possibly due to its characteristics [16]; the second most used is the Lending Club platform, which is a P2P loan platform used in 15.38% of the studies. Some authors in their experiments validate both public and private data, and Ass and N-Ass models, and determine their behavior in different scenarios.

The financial industry needs machine learning tools to support credit risk prediction, as many experiments with private datasets demonstrate. However, since these use sensitive data, their access is limited, which could slow down community efforts to improve prediction using actual cases.

4.5. Answer to RQ5: What Variables or Features Are Used to Predict Credit Risk?

The authors propose many variables, using different methods to identify the variables with the best predictive capacity. GA, FL, hive intelligence, statistical methods, and functions are used to determine the correlation; in Table 8 we display the details.

To simplify the analysis, we have grouped the proposed variables into Demographic, which has a 54.09% share; Operation with 29.18%; Payment behavior with 7.62%; External factors with 6.69%; Unstructured data with 1.30%, and Transaction with 1.12%, as shown in Table 9.

The preference in the use of demographic variables for the prediction of credit risk would have to be explained because these can represent the behavior, preferences, and socioeconomic profile of the client or the segments to which they belong; their inclusion contributes positively to the models, but it is not sufficient. It must be accompanied by Payment behavior variables, characteristics of the operation, and environmental variables that could influence the results.

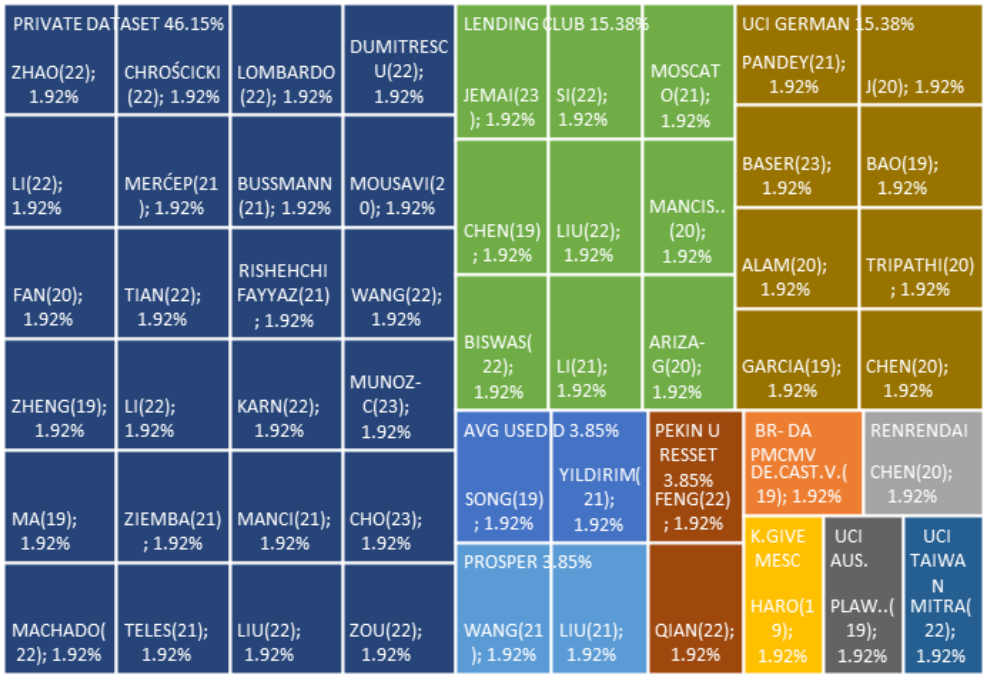


Figure 5. Dataset used.

Table 8. Features.

It.	Features Group	Feature	#	%
1	Demographic	External Debt Value/historical	27	5.02%
2	Demographic	Domestic Debt Value/historical	27	5.02%
3	Operation	Loan value	24	4.46%
4	Demographic	Average/Total revenue	20	3.72%
5	Demographic	Residence/Registered Assets	19	3.53%
6	Demographic	Economic Activity/Experience	18	3.35%
7	Demographic	Family Income	18	3.35%
8	Payment behavior	Days in arrears/Range Days in arrears	17	3.16%
9	Operation	Historical use of debt	16	2.97%
10	Operation	Destination of the Credit/Purpose	16	2.97%
11	Operation	Interest Rate	16	2.97%
12	External factors	Debt Profitability	16	2.97%
13	Demographic	Total Debt/Income/DTI	15	2.79%
14	Demographic	Gender/Sex	14	2.60%
15	Demographic	Risk Segment/Buro Rating/Score	14	2.60%
16	Demographic	Age/Date of Birth	13	2.42%
17	Operation	Checking/Savings Account	13	2.42%
18	Operation	Credit Line Limit	13	2.42%
19	Demographic	Civil Status	12	2.23%
20	Demographic	Mortgage Debt	12	2.23%
21	Operation	Monthly Fees	12	2.23%
22	Payment behavior	Collection status	11	2.04%
23	Payment behavior	Unpaid Installment Number	11	2.04%
24	Demographic	Financial maturity	9	1.67%
25	Demographic	Residence type	9	1.67%
26	Demographic	Fee value	9	1.67%
27	External factors	Inventory turnover	9	1.67%
28	Demographic	Labor Old	7	1.30%
29	Demographic	Education Level	7	1.30%
30	Others	Others	114	21.21%
		Total	538	100.00%

Table 9. Features Group.

It.	Features Group	#	%
1	Demographic	291	54.09%
2	Operation	157	29.18%
3	Payment behavior	41	7.62%
4	External factors	36	6.69%
5	Unstructured data	7	1.30%
6	Transaction	6	1.12%
	Total	538	100.00%

4.6. Answer to RQ6: What Are the Main Problems or Limitations of Predicting Credit Risk?

In the reviewed articles, the authors state limitations or problems that they have faced during the experiments, although in each case there are nuances. We have grouped them into the following: Representativeness of reality, seen in 32% of the studies, refers to the fact that many of the existing variables do not reflect the true nature of the information. Unbalanced data seen in 28%, refers to the fact that, according to some authors, the usage of highly unbalanced data significantly reduces the performance of the models. Inconsistency in information recording is noted in 17%, where reference is made to the fact that the existing records have been entered with errors, bias, and noise that generates the need to apply cleaning techniques, with the risk of losing certain information. Lack of ability to explain the proposed results in 13% refers to the explainability limitation that most robust

ML models have. Availability of information and centralized processing in 6% refers to the need to process information centrally, which can generate additional losses, noise, and delays. Adaptability in processing structured and unstructured information in 4%, refers to the need to process structured and unstructured data within the operation process. We show the results in Table 10.

Table 10. Limitations.

It.	Limits Identified	#	%
1	Representativeness of reality	39	31.71%
2	Unbalanced data	35	28.46%
3	Inconsistency in information recording	21	17.07%
4	Lack of ability to explain the proposed results	16	13.01%
5	Availability of information and centralized processing	7	5.69%
6	Adaptability in processing struct. and unstruct. information	5	4.07%
	Total	123	100.00%

For credit risk prediction models, the difficulties center on the consistency of available information. Furthermore, the capacity of the models to process it is a limitation, since in this industry, its fundamental nature corresponds to unbalanced data.

5. Additional Findings

During the development of this RSL, we identified preprocessing techniques that the authors refer to during their experiments.

The techniques most used by the authors for estimating the hyperparameters are KFold in 58.33% of the studies and Grid Search in 22.22%. We display the details in Table 11.

The most-used algorithms to mitigate the problem of imbalance in the datasets are SMOTE at 29.55%, followed by KFold with 20.45%, RUS at 11.76%, and ROS similarly at 11.36%. We show the detail in Table 12. Also used are ADASYN, some variants of SMOTE like B-SMOTE, SMOTE-T, adapted classification algorithms like KN-SMOTE, Under-Bagging, and techniques to identify missing values such as NMISS, CC, CS-Classifiers, and RESAMPLE.

Credit risk prediction can be schematized as a classification problem with unbalanced samples [20]; in that sense, the authors' preprocessing techniques can be considered the baseline for new research.

Table 11. Techniques for determination of hyperparameters.

It.	Method	#	%	Author
1	KFold CV	21	58.33%	[1,2,15,18,26,28,29,31,33,35,38,40,41,43,45,46,50,55,59,62,67]
2	Grid Search Method	8	22.22%	[11,15,16,44,48,56,57,63]
3	LightGBM Bayesian Optimization	2	5.56%	[20,56]
4	Genetic Algorithm	2	5.56%	[26,52]
5	Random Search	1	2.78%	[20]
6	Ant Colony Optimization	1	2.78%	[53]
7	Other	1	2.78%	[65]
	Total	36	100.00%	

Table 12. Dataset balancing techniques.

It.	Method	#	%	Author
1	SMOTE	13	29.55%	[2,11,15,20,25,30,32,35,36,44,45,63,64]
2	KFold	9	20.45%	[14,16,24–26,29,50,55,57]
3	ROS	5	11.36%	[2,20,35,45,63]
4	RUS	5	11.36%	[2,20,45,63,64]
5	ADASYN	2	4.55%	[35,63]

Table 12. *Cont.*

It.	Method	#	%	Author
6	SMOTEBoost	2	4.55%	[20,64]
7	B-SMOT	1	2.27%	[63]
8	CC	1	2.27%	[63]
9	CS-Classifiers	1	2.27%	[64]
10	KN-SMOTE	1	2.27%	[63]
11	NMISS	1	2.27%	[63]
12	RESAMPLE	1	2.27%	[48]
13	SMOTE-T	1	2.27%	[63]
14	Under-Bagging	1	2.27%	[64]
	Total	44	100.00%	

6. Discussion

The most widely used ML models to assess credit risk correspond to the Traditional family, possibly due to their easy implementation. On the other hand, those with the best prediction results correspond to the Boosted Category, both in the Ass and in the N-Ass groups. This trend is evidenced in Table 13, where this family obtains 24 evaluations out of 52 and is constantly growing through recent years. Another trend identified is that better results are acquired with the N-Ass models [67], for example, in the experiments which obtained AUC 91.00 with the AdaB + RF model [63] and 91.20 with XGB + DT [25], respectively. As shown in Table 7 and Figure 4, these results could be explained by gradient-based optimization features, parallelism, high-volume throughput, and missing values. The most-used metrics correspond to AUC, ACC, F1 measure, Recall, and Precision; the authors propose more specialized metrics according to the situation being evaluated. AUC and ACC in 16.11% and 14.22% of the studies, are mainly due to their ability to measure the capacity of different types of ML models. In the first case, it does not vary before the transformation by normalization, which allows analyzing unbalanced data with a high cost of classification; in the second case, it works better with balanced data, and that of easy explanation.

Table 13. Model trends.

It.	Family	2019	2020	2021	2022	2023	Total
1	Boosted Category	4	4	5	10	1	24
2	Traditional	4	1	5	4	1	15
3	NN/DL	1	1	2	2	1	7
4	Collective Intelligence		2		2		4
5	Fuzzy Logic		1			1	2
	Total	9	9	12	18	4	52

Likewise, of the experiments carried out, the most significant databases used are 53.85% public datasets, while 46.15% correspond to private ones; the first serves to evaluate the predictive capacity of the new models proposed by the authors, comparing the results with previous experiences and the second to generate new knowledge through its application in the real world. In the experiments, the authors identify as the main problem in the datasets for the design of ML models the misrepresentation of reality due to possible bias, inconsistency, or error when recording the information. The second problem corresponds to the imbalance of the data, which can impair the excellent performance of the models. To face this problem, the SMOTE algorithm is mainly used, and for the optimization of the hyperparameters, the Kfold CV and Grid Search Method techniques are utilized. However, some authors propose hive intelligence [53] and genetic algorithms [26,52] to guide optimization. Finally, the most-used variables that best represent the intention and ability to pay, which in turn originate credit risk, correspond to the Demographic, Transaction, and

Payment behavior features. These encompass the main characteristics expected to predict it; see Table 8. However, the corresponding external features and unstructured data must be considered, bearing in mind the former, the influence in the hyperconnected world, the growing development of DDI, and the processing capacity of BIG DATA.

7. Conclusions and Future Research

In this systematic review article of the literature on credit risk prediction using ML, we reach the following conclusions:

- The Boosted Category is a family of ML models being investigated. They are the most used in Ass and N-Ass situations, highlighting the XGB model, with the tendency being its use in Ass. This category in the experiments obtained better results than other models due to its ability to process categorical variables—numerical, with noise, missing, and unbalanced data—and applying regularization could avoid overfitting. However, since they are complex models, they are challenging to interpret and are not very tolerant of applying atypical values.
- The five most-used metrics are AUC, ACC, Recall, F1 measure, and Precision, although, in practice, the problem must considered when choosing the most appropriate metrics. However, AUC stands out for its ability to not be influenced by the distribution of classes and preferable behavior when processing unbalanced data.
- Public datasets are more utilized; of this group, the commonly used are UCI German Dataset and Landing Club Dataset. Their main use is to validate the behavior against other models under the same conditions. Private datasets generate knowledge from the application to a specific situation.
- For the evaluation of credit risks through ML, demographic variables are mainly used, which represent behavior, preferences, socioeconomic profile, and operations that represent the characteristics of the financial product acquired. However, this information is insufficient and is complemented by external variables and those related to unstructured data such as images, video, or others generated from hyperconnectivity, which is supported by DDI and BIG DATA development and processing.
- The main problems are the representativeness of reality, the imbalance of data for the training, and the inconsistency in recording information. All cases arise due to biases, errors, or problems in recording information.
- The most widely used method to solve the imbalance problem is SMOTE to optimize the performance of ML models, while the methods to determine the hyperparameters are KFold-CV and Grid Search to guide their optimization.

The credit risk prediction contribution corresponds to the stage where the credit originates. In this sense, we propose to extend the application of ML to precise credit datasets from specialized companies, including these models in other processes, such as credit collection and customer retention, considering the regulatory impositions governments are implementing to mitigate possible losses in the industry. Credit risk prediction can be enhanced with BIG DATA analysis, especially on unstructured data such as images, text, writing, sentiment, and hive intelligence, to assess adaptability to changing scenarios [17]. Finally, and in the same sense, including variables that represent the state of the environment could contribute to reducing uncertainty in this sector in the face of unexpected external events.

Author Contributions: Conceptualization, J.P.N. and J.A.H.; methodology, J.P.N.; validation, J.P.N., L.A.R. and J.A.H.; formal analysis, J.P.N. and J.A.H.; investigation, J.P.N.; resources, J.P.N.; writing—original draft preparation, J.P.N.; writing—review and editing, J.P.N., L.A.R. and J.A.H.; visualization, J.P.N.; supervision, J.A.H.; project administration, J.P.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lombardo, G.; Pellegrino, M.; Adosoglou, G.; Cagnoni, S.; Pardalos, P.M.; Poggi, A. Machine Learning for Bankruptcy Prediction in the American Stock Market: Dataset and Benchmarks. *Future Internet* **2022**, *14*, 244. [CrossRef]
2. Ziemba, P.; Becker, J.; Becker, A.; Radomska-Zalas, A.; Pawluk, M.; Wierzba, D. Credit decision support based on real set of cash loans using integrated machine learning algorithms. *Electronics* **2021**, *10*, 2099. [CrossRef]
3. Liu, C.; Ming, Y.; Xiao, Y.; Zheng, W.; Hsu, C.H. Finding the next interesting loan for investors on a peer-to-peer lending platform. *IEEE Access* **2021**, *9*, 111293–111304. [CrossRef]
4. Chen, C.; Lin, K.; Rudin, C.; Shaposhnik, Y.; Wang, S.; Wang, T. A holistic approach to interpretability in financial lending: Models, visualizations, and summary-explanations. *Decis. Support Syst.* **2022**, *152*, 113647. [CrossRef]
5. Shih, D.H.; Wu, T.W.; Shih, P.Y.; Lu, N.A.; Shih, M.H. A Framework of Global Credit-Scoring Modeling Using Outlier Detection and Machine Learning in a P2P Lending Platform. *Mathematics* **2022**, *10*, 2282. [CrossRef]
6. Zhang, Z.; Jia, X.; Chen, S.; Li, M.; Wang, F. Dynamic Prediction of Internet Financial Market Based on Deep Learning. *Comput. Intell. Neurosci.* **2022**, *2022*, 1465394. [CrossRef]
7. BM Panorama General. Available online: https://www.bancomundial.org/es/topic/financialsector/overview (accessed on 22 December 2021).
8. Hani, U.; Wickramasinghe, A.; Kattiyapornpong, U.; Sajib, S. The future of data-driven relationship innovation in the microfinance industry. In *Annals of Operations Research*; Springer: Dordrecht, The Netherlands, 2022; pp. 1–27.
9. Zhang, C.; Zhong, H.; Hu, A. A Method for Financial System Analysis of Listed Companies Based on Random Forest and Time Series. *Mob. Inf. Syst.* **2022**, *2022*, 6159459. [CrossRef]
10. Majerník, M.; Daneshjo, N.; Malega, P.; Drábik, P.; Barilová, B. Sustainable development of the intelligent industry from Industry 4.0 to Industry 5.0. *Adv. Sci. Technol. Res. J.* **2022**, *16*. [CrossRef]
11. Yıldırım, M.; Okay, F.Y.; Øzdemir, S. Big data analytics for default prediction using graph theory. *Expert Syst. Appl.* **2021**, *176*, 114840. [CrossRef]
12. Bi, W.; Liang, Y. Risk Assessment of Operator's Big Data Internet of Things Credit Financial Management Based on Machine Learning. *Mob. Inf. Syst.* **2022**, *2022*, 5346995. [CrossRef]
13. Hariri, R.H.; Fredericks, E.M.; Bowers, K.M. Uncertainty in big data analytics: Survey, opportunities, and challenges. *J. Big Data* **2019**, *6*, 44. [CrossRef]
14. Chen, Z.; Chen, W.; Shi, Y. Ensemble learning with label proportions for bankruptcy prediction. *Expert Syst. Appl.* **2020**, *146*, 113155. [CrossRef]
15. Fan, S.; Shen, Y.; Peng, S. Improved ML-based technique for credit card scoring in internet financial risk control. *Complexity* **2020**, *2020*, 8706285. [CrossRef]
16. García, V.; Marques, A.I.; Sánchez, J.S. Exploring the synergetic effects of sample types on the performance of ensembles for credit risk and corporate bankruptcy prediction. *Inf. Fusion* **2019**, *47*, 88–101. [CrossRef]
17. Wang, M.; Yang, H. Research on personal credit risk assessment model based on instance-based transfer learning. In Proceedings of the Intelligence Science III: 4th IFIP TC 12 International Conference, ICIS 2020, Durgapur, India, 24–27 February 2021; Revised Selected Papers 4; Springer: Dordrecht, The Netherlands, 2021; pp. 159–169.
18. Teles, G.; Rodrigues, J.J.; Rabêlo, R.A.; Kozlov, S.A. Comparative study of support vector machines and random forests machine learning algorithms on credit operation. *Softw. Pract. Exp.* **2021**, *51*, 2492–2500. [CrossRef]
19. Orlova, E.V. Decision-making techniques for credit resource management using machine learning and optimization. *Information* **2020**, *11*, 144. [CrossRef]
20. Zou, Y.; Gao, C.; Gao, H. Business failure prediction based on a cost-sensitive extreme gradient boosting machine. *IEEE Access* **2022**, *10*, 42623–42639. [CrossRef]
21. Fritz-Morgenthal, S.; Hein, B.; Papenbrock, J. Financial risk management and explainable, trustworthy, responsible AI. *Front. Artif. Intell.* **2022**, *5*, 779799. [CrossRef]
22. Sun, M.; Li, Y. Credit Risk Simulation of Enterprise Financial Management Based on Machine Learning Algorithm. *Mob. Inf. Syst.* **2022**, *2022*, 9007140. [CrossRef]
23. Mousavi, M.M.; Lin, J. The application of PROMETHEE multi-criteria decision aid in financial decision making: Case of distress prediction models evaluation. *Expert Syst. Appl.* **2020**, *159*, 113438. [CrossRef]
24. Zhao, L.; Yang, S.; Wang, S.; Shen, J. Research on PPP Enterprise Credit Dynamic Prediction Model. *Appl. Sci.* **2022**, *12*, 10362. [CrossRef]
25. Pandey, M.K.; Mittal, M.; Subbiah, K. Optimal balancing & efficient feature ranking approach to minimize credit risk. *Int. J. Inf. Manag. Data Insights* **2021**, *1*, 100037.
26. Pławiak, P.; Abdar, M.; Acharya, U.R. Application of new deep genetic cascade ensemble of SVM classifiers to predict the Australian credit scoring. *Appl. Soft Comput.* **2019**, *84*, 105740. [CrossRef]

27. Cho, S.H.; Shin, K.s. Feature-Weighted Counterfactual-Based Explanation for Bankruptcy Prediction. *Expert Syst. Appl.* **2023**, *216*, 119390. [CrossRef]
28. Bao, W.; Lianju, N.; Yue, K. Integration of unsupervised and supervised machine learning algorithms for credit risk assessment. *Expert Syst. Appl.* **2019**, *128*, 301–315. [CrossRef]
29. Mitra, R.; Goswami, A.; Tiwari, M.K. Financial supply chain analysis with borrower identification in smart lending platform. *Expert Syst. Appl.* **2022**, *208*, 118026. [CrossRef]
30. Jemai, J.; Zarrad, A. Feature Selection Engineering for Credit Risk Assessment in Retail Banking. *Information* **2023**, *14*, 200. [CrossRef]
31. Chen, S.F.; Chakraborty, G.; Li, L.H. Feature selection on credit risk prediction for peer-to-peer lending. In Proceedings of the New Frontiers in Artificial Intelligence: JSAI-isAI 2018 Workshops, JURISIN, AI-Biz, SKL, LENLS, IDAA, Yokohama, Japan, 12–14 November 2018; Revised Selected Papers; Springer: Dordrecht, The Netherlands, 2019; pp. 5–18.
32. Si, Z.; Niu, H.; Wang, W. Credit Risk Assessment by a Comparison Application of Two Boosting Algorithms. In *Fuzzy Systems and Data Mining VIII*; IOS Press: Amsterdam, The Netherlands, 2022; pp. 34–40.
33. Merćep, A.; Mrčela, L.; Birov, M.; Kostanjčar, Z. Deep neural networks for behavioral credit rating. *Entropy* **2020**, *23*, 27. [CrossRef]
34. Bussmann, N.; Giudici, P.; Marinelli, D.; Papenbrock, J. Explainable machine learning in credit risk management. *Comput. Econ.* **2021**, *57*, 203–216. [CrossRef]
35. Moscato, V.; Picariello, A.; Sperlí, G. A benchmark of machine learning approaches for credit score prediction. *Expert Syst. Appl.* **2021**, *165*, 113986. [CrossRef]
36. Ariza-Garzón, M.J.; Arroyo, J.; Caparrini, A.; Segovia-Vargas, M.J. Explainability of a machine learning granting scoring model in peer-to-peer lending. *IEEE Access* **2020**, *8*, 64873–64890. [CrossRef]
37. Chen, X.; Li, S.; Xu, X.; Meng, F.; Cao, W. A novel GSCI-based ensemble approach for credit scoring. *IEEE Access* **2020**, *8*, 222449–222465. [CrossRef]
38. Dumitrescu, E.; Hué, S.; Hurlin, C.; Tokpavi, S. Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. *Eur. J. Oper. Res.* **2022**, *297*, 1178–1192. [CrossRef]
39. Li, D.; Li, L. Research on Efficiency in Credit Risk Prediction Using Logistic-SBM Model. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 5986295. [CrossRef]
40. Ma, X.; Lv, S. Financial credit risk prediction in internet finance driven by machine learning. *Neural Comput. Appl.* **2019**, *31*, 8359–8367. [CrossRef]
41. Karn, A.L.; Sachin, V.; Sengan, S.; Gandhi, I.; Ravi, L.; Sharma, D.K.; Subramaniyaswamy, V. Designing a Deep Learning-Based Financial Decision Support System for Fintech to Support Corporate Customer's Credit Extension. *Malays. J. Comput. Sci.* **2022**, *2022*, 116–131. [CrossRef]
42. Zheng, B. Financial default payment predictions using a hybrid of simulated annealing heuristics and extreme gradient boosting machines. *Int. J. Internet Technol. Secur. Trans.* **2019**, *9*, 404–425. [CrossRef]
43. Mancisidor, R.A.; Kampffmeyer, M.; Aas, K.; Jenssen, R. Learning latent representations of bank customers with the variational autoencoder. *Expert Syst. Appl.* **2021**, *164*, 114020. [CrossRef]
44. Wang, T.; Liu, R.; Qi, G. Multi-classification assessment of bank personal credit risk based on multi-source information fusion. *Expert Syst. Appl.* **2022**, *191*, 116236. [CrossRef]
45. Liu, W.; Fan, H.; Xia, M.; Pang, C. Predicting and interpreting financial distress using a weighted boosted tree-based tree. *Eng. Appl. Artif. Intell.* **2022**, *116*, 105466. [CrossRef]
46. Andrade Mancisidor, R.; Kampffmeyer, M.; Aas, K.; Jenssen, R. Deep generative models for reject inference in credit scoring. *Knowl.-Based Syst.* **2020**, *196*, 105758. [CrossRef]
47. Wu, Z. Using machine learning approach to evaluate the excessive financialization risks of trading enterprises. *Comput. Econ.* **2021**, *59*, 1607–1625. [CrossRef]
48. Liu, J.; Zhang, S.; Fan, H. A two-stage hybrid credit risk prediction model based on XGBoost and graph-based deep neural network. *Expert Syst. Appl.* **2022**, *195*, 116624. [CrossRef]
49. Shu, R. *Deep Representations with Learned Constraints*; Stanford University: Stanford, CA, USA, 2022.
50. Tripathi, D.; Edla, D.R.; Kuppili, V.; Bablani, A. Evolutionary extreme learning machine with novel activation function for credit scoring. *Eng. Appl. Artif. Intell.* **2020**, *96*, 103980. [CrossRef]
51. Uj, A.; Nmb, E.; Ks, C.; Skl, D. Financial crisis prediction model using ant colony optimization-ScienceDirect. *Int. J. Inf. Manag.* **2020**, *50*, 538–556.
52. Feng, Y. Bank Green Credit Risk Assessment and Management by Mobile Computing and Machine Learning Neural Network under the Efficient Wireless Communication. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 3444317. [CrossRef]
53. Tian, J.; Li, L. Digital universal financial credit risk analysis using particle swarm optimization algorithm with structure decision tree learning-based evaluation model. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 4060256. [CrossRef]
54. Chrościcki, D.; Chlebus, M. The Advantage of Case-Tailored Information Metrics for the Development of Predictive Models, Calculated Profit in Credit Scoring. *Entropy* **2022**, *24*, 1218. [CrossRef]
55. de Castro Vieira, J.R.; Barboza, F.; Sobreiro, V.A.; Kimura, H. Machine learning models for credit analysis improvements: Predicting low-income families' default. *Appl. Soft Comput.* **2019**, *83*, 105640. [CrossRef]

56. Li, Z.; Zhang, J.; Yao, X.; Kou, G. How to identify early defaults in online lending: A cost-sensitive multi-layer learning framework. *Knowl.-Based Syst.* **2021**, *221*, 106963. [CrossRef]
57. Koç, O.; Başer, F.; Kestel, S.A. Credit Risk Evaluation Using Clustering Based Fuzzy Classification Method. *Expert Syst. Appl.* **2023**, *223*, 119882.
58. Rishehchi Fayyaz, M.; Rasouli, M.R.; Amiri, B. A data-driven and network-aware approach for credit risk prediction in supply chain finance. *Ind. Manag. Data Syst.* **2021**, *121*, 785–808. [CrossRef]
59. Muñoz-Cancino, R.; Bravo, C.; Ríos, S.A.; Graña, M. On the combination of graph data for assessing thin-file borrowers' creditworthiness. *Expert Syst. Appl.* **2023**, *213*, 118809. [CrossRef]
60. Li, Y.; Stasinakis, C.; Yeo, W.M. A hybrid XGBoost-MLP model for credit risk assessment on digital supply chain finance. *Forecasting* **2022**, *4*, 184–207. [CrossRef]
61. Haro, B.; Ortiz, C.; Armas, J. Predictive Model for the Evaluation of Credit Risk in Banking Entities Based on Machine Learning. In *Brazilian Technology Symposium*; Springer: Dordrecht, The Netherlands, 2018; pp. 605–612.
62. Qian, H.; Wang, B.; Yuan, M.; Gao, S.; Song, Y. Financial distress prediction using a corrected feature selection measure and gradient boosted decision tree. *Expert Syst. Appl.* **2022**, *190*, 116202. [CrossRef]
63. Alam, T.M.; Shaukat, K.; Hameed, I.A.; Luo, S.; Sarwar, M.U.; Shabbir, S.; Li, J.; Khushi, M. An investigation of credit card default prediction in the imbalanced datasets. *IEEE Access* **2020**, *8*, 201173–201198. [CrossRef]
64. Song, Y.; Peng, Y. A MCDM-based evaluation approach for imbalanced classification methods in financial risk prediction. *IEEE Access* **2019**, *7*, 84897–84906. [CrossRef]
65. Biswas, N.; Mondal, A.S.; Kusumastuti, A.; Saha, S.; Mondal, K.C. Automated credit assessment framework using ETL process and machine learning. *Innov. Syst. Softw. Eng.* **2022**, 1–14. [CrossRef]
66. Wang, Y. Research on supply chain financial risk assessment based on blockchain and fuzzy neural networks. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 5565980. [CrossRef]
67. Machado, M.R.; Karray, S. Assessing credit risk of commercial customers using hybrid machine learning algorithms. *Expert Syst. Appl.* **2022**, *200*, 116889. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Anexo I

Artículo: From Crisis to Algorithm: Credit Delinquency Prediction in Peru Under Critical External Factors Using Machine Learning

Este estudio desarrolla una metodología basada en escenarios para integrar factores externos críticos —*COVID-19*, anomalías climáticas y convulsión social— en modelos de Machine Learning para predecir morosidad crediticia en Perú. Utilizando 8.2 millones de registros y el marco *CRISP-DM*, se incorporan análisis de estacionariedad (*Dickey–Fuller*), causalidad (*Granger*), y técnicas de explicabilidad (*SHAP*, *LIME*). Se demuestra que la inclusión de factores externos mejora significativamente el desempeño y la adaptabilidad de modelos como *CNN* y *XGBoost*, proporcionando una base metodológica replicable para la gestión de riesgo en entornos volátiles.

Referencia: Data 2025, 10, 63. https://doi.org/10.3390/data10050063

data

MDPI

Article

From Crisis to Algorithm: Credit Delinquency Prediction in Peru Under Critical External Factors Using Machine Learning

Jomark Noriega [1,2,*,†], Luis Rivera [1,3], Jorge Castañeda [4,†] and José Herrera [1,5]

[1] Facultad de Ingeniería de Sistemas e Informática, Escuela de Posgrado, Universidad Nacional Mayor de San Marcos, Campus Ciudad Universitaria, Calle Germán Amézaga 375, Lima 15801, Peru; rivera@uenf.br (L.R.); jherreraqu@unmsm.edu.pe (J.H.)
[2] Financiera QAPAQ, Lima 150120, Peru
[3] Center of Sciences and Technology (CCT), Laboratory of Mathematical Sciences (LCMAT), State University of North Fluminense Darcy Ribeiro (UENF), Campos dos Goytacazes, Rio de Janeiro 28015-602, Brazil
[4] Facultad de Ingeniería, Escuela de Ingeniería Informática, Universidad San Ignacio de Loyola, Campus 2 La Molina, Av. Fontana 550, La Molina 15026, Peru; jcastanedaa@usil.edu.pe
[5] Facultad de Ciencias Informáticas, Departamento de Lenguajes y Programación, Escuela de Postgrado, Universidad Pablo de Olavide, Campus Carretera de Utrera km 1, Andalucía, 41013 Sevilla, Spain
* Correspondence: jomark.noriega@unmsm.edu.pe or jnoriega@pucp.pe; Tel.: +51-987-521-006
† These authors contributed equally to this work.

Abstract: Robust credit risk prediction in emerging economies increasingly demands the integration of external factors (EFs) beyond borrowers' control. This study introduces a scenario-based methodology to incorporate EF—namely COVID-19 severity (mortality and confirmed cases), climate anomalies (temperature deviations, weather-induced road blockages), and social unrest—into machine learning (ML) models for credit delinquency prediction. The approach is grounded in a CRISP-DM framework, combining stationarity testing (Dickey–Fuller), causality analysis (Granger), and post hoc explainability (SHAP, LIME), along with performance evaluation via AUC, ACC, KS, and F1 metrics. The empirical analysis uses nearly 8.2 million records compiled from multiple sources, including 367,000 credit operations granted to individuals and microbusiness owners by a regulated Peruvian financial institution (FMOD) between January 2020 and September 2023. These data also include time series of delinquency by economic activity, external factor indicators (e.g., mortality, climate disruptions, and protest events), and their dynamic interactions assessed through Granger causality to evaluate both the intensity and propagation of external shocks. The results confirm that EF inclusion significantly enhances model performance and robustness. Time-lagged mortality (COVID MOV) emerges as the most powerful single predictor of delinquency, while compound crises (climate and unrest) further intensify default risk—particularly in portfolios without public support. Among the evaluated models, *CNN* and *XGB* consistently demonstrate superior adaptability, defined as their ability to maintain strong predictive performance across diverse stress scenarios—including pandemic, climate, and unrest contexts—and to dynamically adjust to varying input distributions and portfolio conditions. Post hoc analyses reveal that EF effects dynamically interact with borrower income, indebtedness, and behavioral traits. This study provides a scalable, explainable framework for integrating systemic shocks into credit risk modeling. The findings contribute to more informed, adaptive, and transparent lending decisions in volatile economic contexts, relevant to financial institutions, regulators, and risk practitioners in emerging markets.

Keywords: loan; credit risk; prediction; machine learning; external factors

check for updates

Academic Editor: Irene Finocchi

Received: 19 February 2025
Revised: 21 April 2025
Accepted: 25 April 2025
Published: 28 April 2025

Citation: Noriega, J.; Rivera, L.; Castañeda, J.; Herrera, J. From Crisis to Algorithm: Credit Delinquency Prediction in Peru Under Critical External Factors Using Machine Learning. *Data* **2025**, *10*, 63. https://doi.org/10.3390/data10050063

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Currently, artificial intelligence (AI), machine learning (ML), and process digitization are fundamental elements of daily life. This trend began in the 1990s with the global expansion of the Internet [1] and notably intensified during the COVID-19 pandemic [2]. Simultaneously, the increase in the use of online credit has generated large volumes of data [3], leading the financial industry to the imperative need to develop advanced technological tools for risk prediction and effective uncertainty management [4] that include exogenous variables in the process [5]. Presently, scoring models and machine learning techniques are common for risk prediction, applying innovations based on exhaustive data analysis, data-driven innovation (DDI) [6], and the big data paradigm—characterized by the use of massive, high-dimensional, and heterogeneous datasets for predictive modeling and decision-making [7].

In the context of rising demand, credit risk levels show a significant increase [8,9]. Despite the growing integration of DDI and ML in the financial sector, substantial challenges persist, leaving numerous applications unexplored. The application of ML techniques during unpredictable events, such as the 2008 housing crisis [2,10], the COVID-19 pandemic [11], and the impacts of climate change [12], has exposed critical limitations in risk assessment methodologies and revealed considerable uncertainties in their outputs [2,13–15]. These challenges underscore the need to examine the influence of external factors (EF) on the accuracy and robustness of ML models in credit risk evaluation [6,16], particularly concerning the inclusion of sequential or time-series data, which are commonly encountered in credit risk scenarios [17]. Moreover, these issues emphasize the need to ensure explainability in ML-driven predictions [18], both to facilitate informed decision-making by stakeholders and to maintain transparency for all parties involved [19].

The main objective of this article is to address the integration of variables sensitive to external risk factors into machine learning models for credit risk assessment, a critical yet insufficiently explored topic in the academic context. Peru, as a case study, represents a particularly relevant scenario: it was the country with the highest COVID-19 death rate per 100,000 inhabitants during the pandemic [20]; it faced the periodic and devastating effects of the El Niño phenomenon [12]; and it experienced intense social unrest following the attempted coup in December 2022 [21]. These events have directly impacted the microcredit market by targeting small businesses, a sector highly vulnerable to risk volatility [12].

While the influence of external factors—such as macroeconomic variables and pandemic-related disruptions—on credit risk has been the subject of substantial academic research, few studies have explored the simultaneous integration of multiple external shocks (pandemics, climate anomalies, and social unrest) into ML-based credit risk models, especially in emerging economies such as Peru. This study addresses that gap, offering valuable insights into the applicability and extrapolation of multifactor risk models in vulnerable financial systems [12]. To this end, EF-related variables are integrated into credit risk models through causality and stationarity tests [22]. These variables are derived from public time-series data and combined with proprietary credit data provided by a regulated Peruvian financial institution—hereafter referred to as FMOD—used under a confidentiality agreement. This strategy not only enhances the predictive power of ML models but also establishes a replicable methodology for integrating diverse EF indicators into future risk prediction studies.

The FMOD dataset plays a pivotal role in this study. It incorporates a unique variable that identifies loans supported by government intervention during periods influenced by EF, such as climate change, social unrest, and the COVID-19 pandemic. This variable allows for the detection and detailed analysis of the potential biases introduced by mitigation programs designed to address economic disruptions. By explicitly accounting for these interventions,

the FMOD dataset provides a nuanced understanding of how state-led programs shape credit risk dynamics and borrower behavior under varying external conditions.

Moreover, the FMOD dataset includes monthly updated records of delinquency days, enabling granular analysis across distinct periods. This temporal precision facilitates an in-depth exploration of the interplay between EF and default patterns, thus enhancing the methodological rigor of credit risk assessment. By integrating time-series data with real-world financial indicators, the FMOD dataset offers a robust and replicable framework for evaluating the influence of EF and government programs on financial stability and credit behavior in emerging markets.

Finally, this study evaluates the impact of EF on the predictive accuracy and explainability of ML models. Particular attention is paid to identifying the most influential factors when incorporating external variables into predictive frameworks. By addressing these challenges, this research not only fills a critical gap in the existing literature but also advances the understanding of the effects of EF on credit risk in emerging economies. This contributes actionable insights for policymakers and financial institutions aiming to enhance resilience in volatile economic environments.

2. Related Work

Credit risk, an inherent component of financial activities, arises when borrowers fail to meet their credit obligations, typically due to inability rather than unwillingness to pay [23]. This form of credit deterioration is strongly associated with demographic and macroeconomic variables—such as household income and employment—which are, in turn, sensitive to external systemic shocks [9,23]. Enhancing the explainability of these complex relationships is essential to support data-driven decision-making for policymakers, financial institutions, and affected communities.

The recent literature has increasingly focused on the role of exogenous compound shocks—namely, pandemics, climate change, and social unrest—as emerging sources of credit instability. These phenomena often co-occur and interact in nonlinear ways, producing cascading effects on borrower behavior, portfolio quality, and systemic financial resilience.

- COVID-19 and Credit Risk: The COVID-19 pandemic exposed critical limitations in traditional credit risk models. Government interventions—such as moratoriums, liquidity injections, and public credit guarantees—generated structural breaks in borrower behavior, which challenged the assumptions of pre-pandemic models and led to inaccurate estimations of default risk [11]. These disruptions particularly affected financially vulnerable groups as noted in recent empirical analyses [24].
  Several studies have highlighted the necessity of using adaptive, data-driven approaches—such as deep learning and regularized regression models—to handle the nonlinearity and volatility introduced by systemic shocks [25–27]. In the Peruvian context, unemployment shocks and loan restructuring policies were found to be key drivers of delinquency in consumer and microfinance credit portfolios during the pandemic period [28,29]. Evidence also suggests that regularized machine learning models, including Lasso and Ridge regression, offer robust predictive performance under uncertainty, particularly in large-scale government programs such as “Reactiva Perú” [27].
- Climate-Related Risks: Climate change poses both physical risks (e.g., floods and droughts) and transition risks (e.g., carbon pricing and regulatory shifts), each with distinct implications for credit stability. Physical hazards can disrupt supply chains, damage infrastructure, and reduce household and firm-level repayment capacity [30],

while transition risks expose high-emission sectors to market penalties and regulatory uncertainty [31,32].

Empirical evidence shows that extreme climate events elevate loan impairment rates and increase provisioning requirements, particularly for microfinance institutions operating in vulnerable regions [12]. These dynamics underscore the urgency of integrating climate stress testing, early-warning systems, and network-aware modeling into credit risk frameworks [33,34].

This study [35] analyzes the interaction between macroeconomic shocks and climate-driven economic pressures within the Peruvian context, using a vector autoregression (VAR) model with national time-series data. Their findings suggest that expansionary monetary policies during periods of climate-related stress may inadvertently elevate the credit risk of microfinance portfolios. This reinforces the need to develop predictive frameworks that account for both environmental and policy-related drivers of credit instability in emerging economies.

- Social Unrest and Political Instability: Social unrest—often triggered by inequality, political crises, or economic shocks—can significantly amplify credit risk by disrupting income flows, consumption patterns, and investor confidence. Empirical evidence from Chile shows that the 2019 protests led to a measurable increase in household default probability, with partial mitigation through pandemic-era relief programs [36]. Similarly, civil unrest has been shown to induce volatility in financial markets, exemplified by reverse herding behaviors among investors [37], and to influence regulatory responses such as foreclosure bans, which may inadvertently elevate long-term credit risk [38]. In the Middle East and North Africa (MENA) region, waves of political instability have prompted structural reforms such as bank privatization, producing mixed outcomes in terms of credit risk and financial system resilience [39].

  While regional and global analyses are increasingly available, country-level studies for Peru remain scarce. However, recent academic research suggests that social and health crises in Peru prompted a behavioral shift among microentrepreneurs, particularly those in lower-poverty districts, who increasingly substituted personal loans with government-backed business credit during times of elevated uncertainty. This substitution pattern reflects a form of borrower adaptation and highlights the indirect impact of public policy interventions—such as state guarantee programs—on the structure and resilience of credit portfolios in emerging markets [29].
- Compound External Shocks: An emerging body of research highlights that the simultaneous occurrence of exogenous stressors—such as pandemics, climate anomalies, and social unrest—can generate nonlinear, systemic risks that conventional credit risk models are ill-equipped to anticipate [40]. These compound shocks interact across temporal and spatial dimensions, producing cascading effects on borrower behavior, institutional solvency, and portfolio stability. Moreover, societal vulnerabilities such as institutional fragility, inequality, and economic informality intensify the transmission channels and feedback loops of these risks, particularly in emerging economies.

  To address this complexity, researchers have proposed integrated modeling frameworks that incorporate early-warning signals, tipping points, and network-aware propagation mechanisms spanning environmental, economic, and political domains [40]. These multidimensional approaches are considered essential for the development of next-generation credit risk systems capable of operating under deep systemic uncertainty.

  In the Peruvian context, recent analyses emphasize the need for credit scoring systems to account for borrower heterogeneity under compound volatility conditions [41]. Empirical studies support the use of hybrid models that integrate Explainable Artifi-

cial Intelligence (XAI) techniques—such as SHapley Additive Explanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME)—into traditional risk evaluation processes, thereby enhancing model transparency and predictive performance in highly unstable environments [23]. These insights reinforce the urgency of designing robust, adaptive credit assessment tools tailored to complex and evolving risk landscapes.

Despite the growing body of work on machine learning applications for credit scoring in Peru, most studies focus on conventional financial variables without incorporating external systemic stressors. Prior research has evaluated multiple machine learning algorithms for microcredit assessment in rural areas but typically omits exogenous factors such as pandemic shocks or climate anomalies [42]. This reveals a gap in studies that simultaneously integrate multiple external variables into ML-based credit risk models, particularly within the Peruvian financial system.

The present study addresses this limitation by incorporating time-series-based exogenous indicators—including COVID-19 deaths, temperature anomalies, and climate-induced road blockades—into a wide range of ML models. We assess their influence on both predictive accuracy and model interpretability.

3. Materials and Methods

3.1. Research Questions and Data Description

Based on this assessment, this article poses the research questions outlined in Table 1, centering on the interplay between external factors and credit risk assessment. Particular emphasis is placed on the importance of explainability in ML-driven models, as understanding how EF impacts credit risk is critical for decision-making and actionable insights.

Table 1. Research questions.

Id	Question
RQ1	How did the COVID-19 pandemic, through factors such as positive cases and mortality rates, influence credit defaults in different economic activities?
RQ2	To what extent do climate change indicators, such as temperature anomalies and road blockages due to weather, impact credit delinquency patterns?
RQ3	What is the relationship between credit delinquency and social unrest, considering disruptions to economic activities and societal stability?
RQ4	How do the combined effects of external factors (COVID-19, climate change, and social unrest) contribute to variations in credit delinquency, and what are the most influential factors?

Note: This table outlines the research questions (RQs) focusing on how external factors such as COVID-19, climate change, and social unrest influence credit default and delinquency patterns. It explores their combined effects and identifies the most influential factors.

To explore these relationships, a study focusing on Peru from January 2020 to September 2023 was conducted. The analysis was based on a comprehensive dataset comprising nearly 8.2 million records, as summarized in Table 2. This dataset integrates diverse EF, such as COVID-19 positive cases and deaths, road blockages, and temperature anomalies, along with financial delinquency (see Table 3) and credit activity data (see Table 4), ensuring a multifaceted approach to understanding credit risk dynamics under external influences.

Table 2. Datasets.

It.	Dataset and EF	Source	Rows 1000×	Use to Evaluate
1	COVID-19 Positive Cases	https://www.datosabiertos.gob.pe/	4523	Time series of positive COVID cases
2	COVID-19 Deaths	https://www.datosabiertos.gob.pe/	1064	Time series of deaths from COVID
3	Road Blockages	https://www.gob.pe/mtc	271	Time series of roadblocks due to social unrest and climate factors
4	Temperature Anomaly	https://www.senamhi.gob.pe/	22	Time-series temperature anomaly climatic factor
5	Delinquency Activity	Model Finance	1945	Time series of days of non-payment of economic activities related to credits
6	Credits with EF	Model Finance	367	Labeled Observations w/ External Factors
	Total Rows		8192	

Note: This table summarizes the datasets used in the study, including their sources, the number of rows (in thousands), and their role in modeling credit risk under the influence of EF such as the COVID-19 pandemic, climate anomalies, and social unrest. Public datasets are accessible through the listed official URLs and are also included in the public repository. Proprietary financial data, used under a confidentiality agreement, have been anonymized and aggregated to comply with data protection regulations. The repository additionally contains the metadata, modeling programs, and experimental results used in this study: https://zenodo.org/records/14890903.

Figure 1 illustrates the trends of various external factors, including COVID-19 cases (positive and dead), social unrest, road blockages due to social unrest and weather, and temperature anomalies. Peaks in certain factors, such as COVID-19 cases and social unrest, were visible during specific periods, indicating their significant occurrence and potential impact on other events or systems.

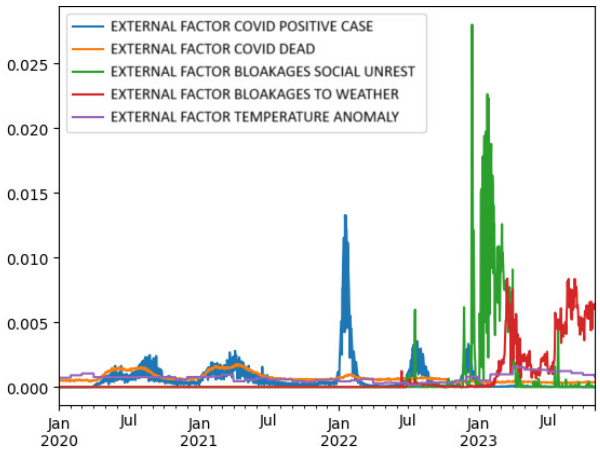


Figure 1. Time series visualization of external factors from January 2020 to September 2023. The plot includes normalized series (min–max scaled) for external factors such as COVID-19 confirmed cases, COVID-19 deaths, protest-related road blockages, weather-induced road blockages, and temperature anomalies. The y-axis ranges from 0 to 1, representing relative intensity over time for each factor. This normalization facilitates the comparison of dynamic patterns across factors with different units and magnitudes.

The main features of the credit dataset provided by FMOD are summarized in Table 5. This summary complements the structural descriptions in Tables 3 and 4.

Table 3. Description of variables in the time-series dataset on credit delinquency by economic activity.

Variable	Type	Description
Date	Date	Date of credit data registration (YYYY-MM-DD).
EconomicActivityCode	int	Identifier for the borrower's economic activity.
IncomeSourceCode	int	Code representing the borrower's income source (e.g., self-employed, salaried).
LoanPurposeCode	int	Identifier for the purpose or destination of the loan.
AvgOverdueDays	int	Average number of overdue days among loans in the grouped record.
MaxOverdueDays	int	Maximum number of overdue days observed in the group.
LoanCount	int	Number of loans aggregated in the same activity group and date.

Table 4. Summary of variables used in the modeling experiments.

Variable	Type	Description
PortfolioSegment	int	Portfolio segment category (see Appendix Table A23)
InterestRate	dec	Effective interest rate applied to the credit
LoanAmount	dec	Approved loan amount
LoanTermDays	int	Actual loan term in days
MaxInternalAmount	dec	Maximum historical internal debt of the client
GovernmentBackedID	bit	Indicates if the loan was government-backed (1 = yes)
GenderID	int	Code representing borrower gender (see Appendix Table A24)
NormalizedSalary	dec	Client's salary normalized by total portfolio
NormalizedTotalAssets	dec	Business asset value normalized by total
NormalizedBusinessEquity	dec	Business equity normalized by total
NormalizedAvailableLiquidity	dec	Business liquidity normalized by total
NormalizedNetBusinessIncome	dec	Net business income normalized by total
BankingHistoryMonths	int	Number of months with banking history
MicroOriginationScore	int	Microcredit origination score
OverIndebtednessScore	int	Over-indebtedness score
ProspectionSourceID	int	Credit acquisition channel code
CreditDestinationID	int	Intended use or destination of the credit
NormalizedDisbursementAge	int	Client's age at disbursement (normalized)
OccupationCode	int	Code representing the borrower's occupation (see Appendix Table A25)
Dependents	int	Number of reported dependents
EducationLevelCode	int	Education level code (see Appendix Table A26)
EmploymentCode	int	Employment status code (see Appendix Table A27)
ResidenceStartYear	int	Year when current residence started
ResidenceOwnershipCode	int	Type of residence ownership (see Appendix Table A28)
ResidenceTypeCode	int	Type of residence (see Appendix Table A29)
DisbursementOfficeID	int	Office code where the loan was disbursed
DelinquencyDays_MMYYYY	int	Number of days past due per month from January 2020 to September 2023
ExternalFactorImpact_MMYYYY	bit	Indicates impact of external factor per month from January 2020 to September 2023

Table 5. Descriptive statistics of the credit dataset used for modeling (FMOD).

Indicator	Value
Effective Interest Rate (Average)	69.98%
Average Loan Amount	S/. 5554
Average Loan Term (in days)	364
Number of Credit Records	367,346
Number of Government-Backed Loans	21,769
Number of Non-Government-Backed Loans	345,577
Number of Loans to Male Borrowers	141,857
Number of Loans to Female Borrowers	225,489
Loan Balance with Less Than 30 Days Past Due (Jan 2020)	S/. 299,040,327.46
Loan Balance with 30–120 Days Past Due (Jan 2020)	S/. 27,787,259.90
Delinquency Rate Over 30 Days (Jan 2020)	8.50%

Note. According to the credit classification policy of FMOD, credits with more than 120 days of delinquency are classified as "charged-off" and fully provisioned (100%). As such, they are excluded from the calculation of the delinquency rate.

Table 3 presents the structure of the time-series dataset used to analyze credit delinquency patterns across economic activities. Each row corresponds to an aggregated daily observation, capturing both the average and maximum number of overdue days among loans, grouped by income source, loan purpose, and economic sector. These variables are essential for modeling how different segments of the economy responded to external shocks over time.

By combining these features, the study constructed temporal indicators of financial stress, which were then aligned with a time series of external factors (e.g., COVID-19 indicators and climate anomalies). This temporal alignment enabled the application of causality analyses—such as Granger causality tests—and enhanced the interpretability of the model's behavior under compound external disruptions.

This approach not only evaluates the predictive accuracy of machine learning models but also examines their capacity to provide transparent explanations for the relationships uncovered, ensuring their practical relevance for decision-making in volatile and complex environments.

Additionally, Table 4 summarizes the set of features used in the individual-level credit risk modeling. These include socio-demographic variables, financial indicators, and normalized business metrics. Each row corresponds to a credit evaluated during the study period, enriched with external factor indicators to support scenario-based experimentation.

Among the set of features, two representative variables were constructed with monthly frequency for this study: 'DelinquencyDays_MMYYYY', which captures the number of overdue days per month, and 'ExternalFactorImpact_MMYYYY', which flags the presence of external stressors—such as COVID-19, climate anomalies, or social unrest—during the same period. This temporal alignment enabled a dynamic analysis of credit behavior under external disruptions, strengthening both the predictive modeling and the interpretability of the results.

3.2. Cross-Industry Standard Process for Data Mining

To evaluate the influence of EF on credit risk, we adopt the Cross-Industry Standard Process for Data Mining (CRISP-DM), a widely used framework that structures data mining projects into six iterative stages: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment [43]. This methodological foundation provides a robust and flexible structure for guiding machine learning workflows across domains, and has also been successfully applied in prior studies analyzing external factors during crisis scenarios [44,45].

In the context of this study, the application of CRISP-DM allowed us to address two specific challenges associated with modeling credit risk under compound external shocks. First, during the Data Understanding and Preparation phases, the incorporation of exogenous time-series variables—such as pandemic indicators, climate anomalies, and social unrest—required additional preprocessing steps, including stationarity testing, temporal alignment, and Granger causality analysis. Second, during the Modeling and Evaluation phases, we conducted comparative experiments across multiple scenarios (with and without EF), enabling us to assess the marginal contribution and explanatory power of these external variables under volatile conditions.

As illustrated in Figure 2, the CRISP-DM framework provided a structured and iterative foundation for organizing each phase of the experiment. The process began with Business and Data Understanding, which guided the initial dataset exploration and external factor characterization. During Data Preparation, we performed normalization, standardization, and the integration of exogenous time-series variables. This integration was based on a monthly alignment strategy.

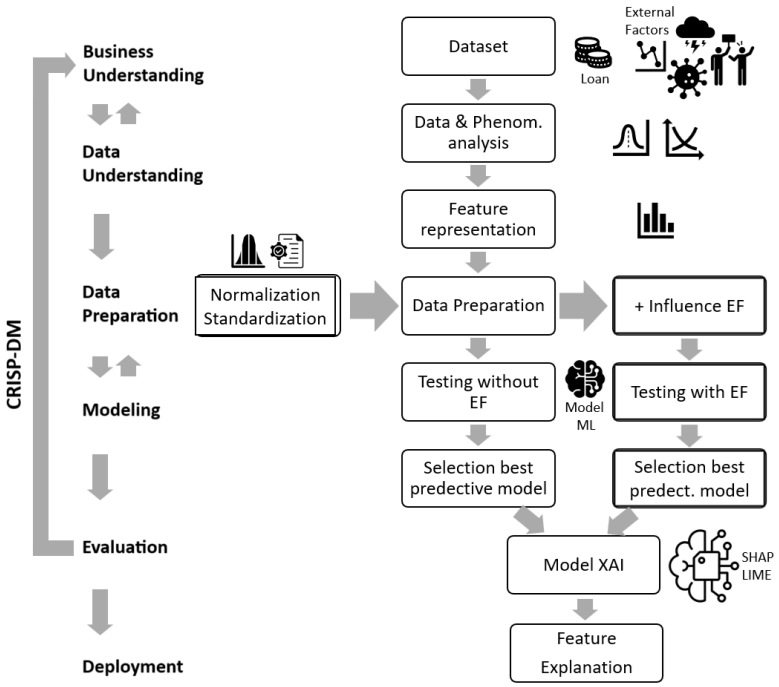


Figure 2. CRISP-DM framework applied to EF-driven credit risk modeling. This diagram illustrates how the CRISP-DM phases structure the experimental workflow, from data understanding and preparation—including the monthly alignment and integration of exogenous time-series variables—to scenario-based modeling (with and without external factors) and the use of explainable AI techniques (SHAP and LIME) during evaluation. The variables DelinquencyDays_MMYYYY and ExternalFactorImpact_MMYYYY are constructed to support a time-aware analysis of credit behavior and its relationship to validated external shocks across economic sectors.

Importantly, the 'ExternalFactorImpact_MMYYYY' variable was not assigned uniformly across the dataset. Instead, it was derived by evaluating the declared economic activity of each credit and determining whether that activity exhibited a statistically significant relationship with a given external factor. This was performed using Granger causality tests between the monthly time series of average delinquency for each economic activity and the corresponding external factor series (e.g., COVID-19 deaths and road blockages). Only when a causal relationship was validated did the external factor impact flag become active for loans in that activity during the affected periods.

Each loan record was then linked to the corresponding external factor conditions (e.g., COVID-19, road blockages, and temperature anomalies) based on the month of disbursement or performance observation. This was operationalized through two representative monthly variables: 'DelinquencyDays_MMYYYY', capturing overdue status, and 'ExternalFactorImpact_MMYYYY', flagging the presence of external stressors. This structure enabled the inclusion of both defaulting and non-defaulting loans, allowing the model to learn from the evolution of payment behavior under varying external conditions.

In the Modeling and Evaluation stages, we compared predictive performance across parallel scenarios—with and without EF—to assess their marginal contribution. Finally, during the Evaluation phase, Explainable Artificial Intelligence (XAI) techniques, specifically SHAP and LIME, were incorporated to enhance transparency and facilitate the interpretation of model decisions concerning external shocks.

3.2.1. Workflow Description

1. Dataset Definition and Feature Selection: The process begins by determining the relevant datasets and conducting a comprehensive analysis of the phenomena to

be evaluated (EF in this study). Feature selection was guided by the established methodologies [46], identifying attributes with high relevance to credit risk prediction, including both internal (e.g., loan characteristics) and external (e.g., EF) variables. The selected features were refined using feature importance techniques, such as Recursive Feature Elimination (RFE) and Mutual Information (MI), ensuring that only the most significant predictors were retained.

2. Data Preparation: Preprocessing involved standardizing and normalizing the dataset (using sum normalization) to homogenize the data scales. Additionally, missing data were handled using multiple imputation techniques to maintain dataset integrity [46].
3. Integration of External Factors: To evaluate the impact of EF on credit risk prediction, additional variables corresponding to EF (e.g., COVID-19, temperature anomalies, and social unrest) were incorporated into the dataset. These variables were derived from time-series analyses and causal relationships, validated through statistical methods such as the Dickey–Fuller test and the Granger causality test [47].

The Dickey–Fuller test was employed to verify the stationarity of the time-series data, ensuring that relationships between variables were not spurious and that the models yielded reliable results [47]. The stationarity of time series is critical, particularly when external factors such as COVID-19 cases, mortality rates, and roadblocks caused by weather events influence the temporal dynamics. The test is described mathematically as

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \delta_1 \Delta Y_{t-1} + \cdots + \delta_p \Delta Y_{t-p} + \epsilon_t \quad (1)$$

The terms represent the following:

- ΔY_t: first difference in time series ($Y_t - Y_{t-1}$);
- α: a constant;
- βt: a trend term;
- Y_{t-1}: lagged value of the time series;
- δ_i: coefficients of lagged differences;
- ϵ_t: the error term;
- p: the number of lags.

Following the confirmation of stationarity, the Granger causality test was applied to evaluate the causal influence of these external factors on economic activities. The test, widely used in econometric studies, is defined as

$$Y_t = \beta_0 + \sum_{i=1}^{k} \beta_i Y_{t-i} + \sum_{j=1}^{k} \gamma_j X_{t-j} + \eta_t \quad (2)$$

The terms represent the following:

- X_{t-j}: lagged values of series X_t;
- β_0, β_i, and γ_j: the coefficients;
- η_t: the error term.

Lag variables were calculated to capture delayed effects, with lag periods tailored to each EF (e.g., quarterly for temperature anomalies, and monthly for social unrest and COVID-19). These statistical methods ensured rigorous validation of the causal relationship between EF and economic activities that influence credit defaults.

4. ML Model Training and Hyperparameter Tuning: Advanced ML models were trained on datasets both with and without EF to assess their predictive contributions. Hyperparameter tuning was conducted using grid search and Bayesian optimization [48] to enhance the model performance. To address the class imbalance in the dataset,

techniques such as the Synthetic Minority Oversampling Technique (SMOTE) [49] and class-weighted loss functions were employed [50,51].

5. Evaluation and Explainability (XAI): Models were evaluated using metrics such as Accuracy (ACC), Area Under the Curve (AUC), Kolmogorov–Smirnov (KS) and F1-score (F1) across multiple folds (10-fold cross-validation). To enhance interpretability, post hoc explainability techniques such as SHAP and LIME were applied [52]. These methods highlight the most influential features in predicting credit delinquency both before and after incorporating EF [19].
6. Comparison of Scenarios: The workflow was executed in parallel for scenarios excluding EF and those incorporating EF. This allowed for a direct comparison of the model performance and the added value of EF inclusion.

3.2.2. Key Enhancements in the Current Study

- The integration of explainability methods (SHAP and LIME) provides actionable insights for decision-makers in financial institutions.
- Comprehensive feature selection processes were implemented, ensuring that only the most predictive attributes were retained for model development.
- Rigorous hyperparameter optimization to maximize model performance and robustness.

By employing this adapted CRISP-DM workflow, this study contributes a systematic and replicable methodology for evaluating EF in credit risk contexts, particularly during crises.

Figure 3 depicts the impact of COVID-19 on economic activities by comparing delinquency rates with the economic impacts over time. The graph highlights fluctuations in economic impact trends, particularly during pandemic peaks, and distinguishes patterns between financial and personal impacts (F + P) and movement-related impacts (CR Impact F + Mov).

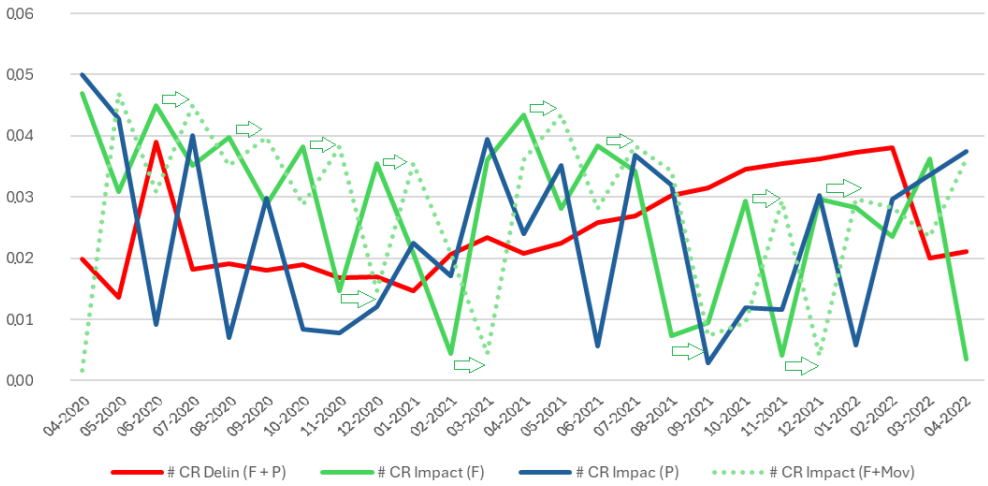


Figure 3. Impact on economic activities by external factor COVID-19. This figure presents normalized credit risk impact values across economic activity types under different pandemic-related external factor scenarios. The vertical axis represents a normalized metric (ranging from 0 to 1) that quantifies the relative intensity of impact observed for each external factor. External factor notations: F = COVID-19 Deaths, P = COVID-19 Positive Cases, F + P = Combined Impact of COVID-19 Deaths and Positive Cases, F + Mov = Combined Impact including Mobility Restrictions.

Figure 4 illustrates the impact of social unrest and climatic factors on economic activity from December 2022 to September 2023. It compares delinquency rates with the economic impact of social unrest, weather, and temperature anomalies, and shows significant spikes during periods of intensified external disruptions.

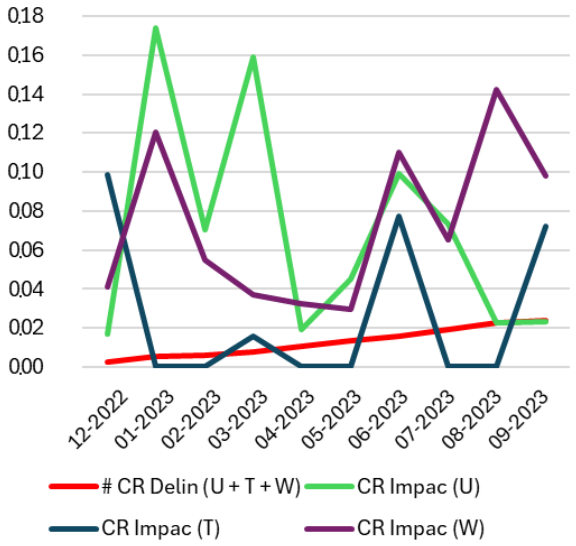


Figure 4. Impact on economic activities by external factors: social unrest and climatic events. This figure presents normalized credit risk impact values across economic activity types under different external factor scenarios. The vertical axis represents a normalized impact metric (ranging from 0 to 1) that reflects the relative intensity of the effect across sectors. External factor notations: U = social unrest, T = temperature anomaly, W = weather-induced road blockages, U + T + W = combined effect of social unrest, temperature anomaly, and road blockages.

3.2.3. Research Strategy

The evaluation period for this study is defined based on the behavior of the monthly classification variables, starting from the onset of the impact of EF until stabilization as shown in Figures 3 and 4. Stabilization is defined as the point at which the slope of the curve representing the proportion of bad payers (red line) began to level off.

Evaluation Periods

- COVID-19: The evaluation period spans March 2020 to January 2022. Loans disbursed until October 2021 are included to account for a 60-day grace period before delinquencies appeared, ensuring unbiased results (Figure 3).
  During the analysis of the COVID-19 external factors, it was identified that the trends in the number of positive cases and deaths often showed opposing behaviors. To address this, an additional experiment was conducted, in which the death curve was shifted from one period to the right, referred to as the COVID MOV. This adjustment considered the estimated incubation period of the disease and the time to potential fatality, typically ranging from 2 to 3 weeks. By aligning the death data with this temporal delay, this analysis aimed to capture the causal relationship between disease progression and its impact on credit delinquency rates. This methodological adjustment is illustrated in Figure 3 (green dotted line).
- Social Unrest and Climate Factors: Data from December 2022 to September 2023 were analyzed, with loans disbursed up to June 2023 considered under the same grace period logic (Figure 4).

In addition, the dataset includes a variable that identifies whether a credit was backed by a government guarantee. Based on this variable, two experimental configurations were defined: one including all government-backed credits (WGB_CR), and another excluding them (WOGB_CR), to evaluate potential biases introduced by public credit support programs.

To assess the influence of external stressors on credit risk, multiple experimental scenarios were constructed using different combinations of observed external factors (see Table 6). These include a baseline scenario SFE, as well as scenarios incorporating COVID-19-related metrics such as fatalities (F), confirmed positive cases (P), and a combined F + P scenario. Additionally, a variation labeled COVID MOV was considered, shifting fatality data forward by one month to account for incubation and reporting delays.

Table 6. Acronyms of evaluated scenarios.

Item	Scenario	Acronym
1	Without EF	SFE
2	EF COVID Dead	F
3	EF COVID Positive	P
4	EF COVID Dead and Positive	F + P
5	EF Social Unrest	U
6	EF Temperature Anomaly	T
7	EF Road blockage weather (RBW)	W
8	EF S.Unrest and Temp.Anomaly	U + T
9	EF Temp. Anomaly and RBW	T + W
10	EF S. Unrest and RBW	U + W
11	EF S. Unrest, Temp. Anomaly and RBW	U + T + W
12	With govt-backed Credit	WGB_CR
13	Without govt-backed Credit	WOGB_CR

Note: This table lists the acronyms for the evaluated scenarios, such as social unrest (U), temperature anomaly (T), and road blockages (RBW). Combined scenarios such as U + T and U + T + W are also defined.

Other scenarios focused on sociopolitical and environmental disruptions, including social unrest indicators (U), temperature anomalies (T), and road blockages caused by extreme weather (W). Combined scenarios such as T + W (climatic anomalies and road disruptions) and U + T + W (integrating social unrest, climate variability, and climate-induced blockages) were also included to evaluate compound effects.

Data Preprocessing

- Data normalization and standardization were applied to homogenize the scales [53].
- Features with correlations above 90% were removed to prevent redundancy.
- SMOTE was used to balance the dataset [54].

Machine Learning Models

We implemented the following models to evaluate the binary classification tasks, prioritizing interpretability and predictive performance:

- Traditional Models: Linear Discriminant Analysis (*LDA*), Quadratic Discriminant Analysis (*QDA*), and k-Nearest Neighbors (*KNN*) [54].
- Advanced Models: Perceptron (*PERCEP*), Multi-layer Perceptron (*MLP*), Random Forest (*RF*) [3], Ridge Regression (*Ridge*) [55], and Extreme Gradient Boosting (*XGB*) [56].
- Explainable Models:
  - Explainable Neural Network (*XNN*), which directly integrates interpretability into the model architecture [17].
  - Explainable Boosting Machines (*EBM*), a method that combines high accuracy with feature interpretability through additive modeling [57].

- Feedforward Neural Networks: Variants of Feedforward Neural Networks (*FFNN*) were explored:
  - Standard Feedforward Neural Network (*FFN*): A classical architecture used as a baseline for comparison.
  - Feedforward Neural Network with Sparse Regularization (*FFNSP*): Incorporates sparsity constraints to enhance generalization and reduce overfitting.
  - Feedforward Neural Network with Local Penalty Regularization (*FFNLP*): Local penalty terms are applied to balance feature contributions and improve interpretability.
- Deep Learning: Convolutional Neural Networks (*CNN*) were included for their ability to capture complex relationships in both spatial and temporal data [25]. Although *CNN* are traditionally applied to image or time-series data, in this study, each tabular instance was reshaped into a 2D tensor of shape $(features, 1)$, allowing the use of 1D convolutional layers to capture localized patterns across the feature vector. This approach has been successfully applied in previous studies involving structured tabular data [58].

Each model was trained and evaluated using a stratified 10-fold cross-validation protocol, ensuring balanced class representation across folds [26]. In each fold, 90% of the data was used for model development and 10% was held out for testing. Within the development portion, 75% was used for training and 25% for internal validation, enabling unbiased hyperparameter tuning. We employed a combination of grid search and Bayesian optimization to identify optimal model configurations.

All performance metrics—including ACC, AUC, F1, and KS—were computed separately on the validation and test partitions for each fold. The final results reported in the manuscript correspond to the average values obtained across the 10 test folds.

Explainability with XAI

To address concerns about interpretability, SHAP and LIME were applied to identify the most influential features in predicting credit delinquency both before and after incorporating EF. Models such as XNN and EBM provided native interpretability, complementing these post hoc methods.

Statistical Evaluation

The impact of EF on model performance was assessed using Student's t-test (Equation (3)), applied to metrics such as ACC, AUC, KS and F1 from the cross-validation folds, with a significance level of $\alpha = 0.05$:

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \tag{3}$$

The terms represent the following:

- $\bar{X}$: mean of metrics from ten folds;
- μ: hypothetical mean under the null scenario;
- s: standard deviation of metrics;
- n: number of folds ($n = 10$).

Economic Activity Relevance

The dataset, comprising 367,000 loans (see Table 2), was analyzed based on its concentration across economic activities. This analysis is particularly relevant, as the impact of EF on these activities amplifies exposure to credit risk. The distribution of loans is as follows:

- 80% Concentration: 80% of the loans are concentrated in 78 economic activities.
- 99% Concentration: 99% of the loans are concentrated in 326 economic activities.
- 100% Concentration: 100% of the loans are distributed across 660 economic activities.

Figure 5 illustrates these groupings, highlighting the increased dispersion and noise as the number of activities increases. To ensure a comprehensive analysis, the results were further examined by including and excluding loans affected by government-backed programs, thereby isolating the potential biases introduced by such interventions.

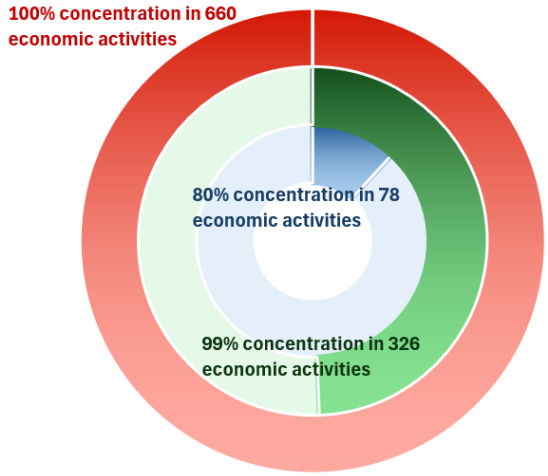


Figure 5. Relevance of Economic Activity.

Research Limitations and Biases

While this study provides valuable insights into the influence of EF on credit delinquency predictions, several limitations and potential biases must be considered:

- Context-Specific Findings: The analysis is based on data from the Peruvian context, which may limit the generalizability to other regions with different economic, social, and environmental dynamics. Future research should incorporate external validation by using datasets from diverse geographical areas to evaluate the transferability of the proposed methodology.
- Training Data Bias: Dataset imbalances and the under-reporting of delinquency during government interventions may introduce biases. Although SMOTE has been used to address class imbalances, residual biases related to data collection practices or policy impacts may persist.
- Model Transparency: While explainability techniques such as SHAP, LIME, and interpretable models (e.g., *XNN* and *EBM*) have been employed, the use of black-box models like *CNN* and *XGB* remains a challenge. This lack of inherent transparency could hinder their applicability to financial institutions, where explainability is crucial for regulatory compliance and stakeholder trust.
- Temporal Assumptions: The assumption that loans disbursed near the evaluation cut-off reflect the grace period impacts may not fully account for the delayed delinquency effects, particularly for loans with extended repayment terms.
- Data Quality and Noise: External factor measurements and institutional datasets may contain noise or inaccuracies that could affect model performance, despite preprocessing steps such as normalization and standardization. Future work should explore advanced noise-reduction techniques.

- Researcher Influence: Prior domain knowledge and experience can introduce subtle biases in model selection, EF, or preprocessing decisions. To mitigate this, the study adhered to best practices such as k-fold cross-validation, rigorous statistical testing, and methodological transparency.

Despite these limitations, the methodological rigor applied in this study, including robust preprocessing, comprehensive statistical validation, and the integration of explainable models, provides a solid foundation for future research. The findings highlight the need for the further exploration of EF impacts in broader contexts, paving the way for more robust, interpretable, and generalizable credit risk assessment frameworks.

4. Results

4.1. Stationarity Analysis and Dataset Integration

In this analysis, "economic activities" refer to the internal sector classification codes provided by the financial institution FMOD. These codes represent a customized segmentation structure used to group clients based on their type of business or productive sector, and are derived from operational risk models. Although not fully equivalent to official international standards, these codes are conceptually aligned with the International Standard Industrial Classification of All Economic Activities (ISIC Rev. 4), maintained by the United Nations Statistics Division. (https://unstats.un.org/unsd/classifications/Econ/isic, accessed on 5 April 2025).

Specifically, the evaluation window from April 2018 to February 2020 (as shown in Table 7) contains 718 unique activity codes, each linked to credit operations that remained active during that period. Other periods in the table include subsets of these activities, depending on whether data on overdue loans were available within each respective evaluation window.

Table 7. Economic activity delinquency stationarity.

Class.	**Without EF**		**With EF**	
Eval. Start	April 2018	January 2019	April 2020	December 2022
Eval. End	February 2020	December 2019	September 2023	June 2023
(#) Non-Stat.	647	625	465	404
(#) Stat.	49	40	137	29
(#) Indeterm.	22	21	15	41
(%) Non-Stat.	90%	91%	75%	85%
(%) Stat.	7%	6%	22%	6%
(%) Indeterm.	3%	3%	2%	9%
(#) Total	718	686	617	474
(%) Total	100%	100%	100%	100%

Note: This table compares the stationarity of credit delinquency series by economic activity, under scenarios with and without EF, across different evaluation periods. The first column (April 2018–February 2020) includes the full set of 718 distinct economic activities associated with credit operations that remained active during that period. The other columns reflect subsets of those activities, based on the data availability of each corresponding evaluation window. The symbol (#) indicates the number of activities in each category.

To assess the influence of EF on economic activity delinquency, we first applied the Dickey–Fuller stationarity test (see Equation (1)) to the normalized daily time series covering the period from January 2016 to September 2023 (see Table 2). These time series represent the number of delinquency days associated with the economic activities under study, examined both in periods with EF and without EF [47].

Key Stationarity Findings

Table 7 shows that, in periods without EF, between 90% and 91% of the activities exhibit non-stationary delinquency patterns, while 6–7% exhibit stationarity, and around 3% are indeterminate. In contrast, periods with EF present a lower proportion of non-stationary activities (75–85%) and a higher proportion of stationary (6–22%) or indeterminate patterns (2–9%). These differences suggest that exogenous shocks, such as pandemics, social unrest, or climatic disruptions, can induce more persistent or complex delinquency behaviors, consistent with similar observations in other emerging economies.

To assess the influence of EF on economic activity delinquency, the Dickey–Fuller stationarity test was applied (see Equation (1)) to the normalized time series of daily frequency. These series represent the days of delinquency associated with the analyzed economic activities. The data used for this analysis cover the period from January 2016 to September 2023 (see Table 2). The analysis was performed for both periods influenced by EF and those without EF [47].

The results indicate that during periods without EF influence, between 90% and 91% of economic activities exhibited non-stationary delinquency, while 7% and 6% showed stationarity. Additionally, 3% of the activities fell into the "indeterminate" category, where stationarity could not be conclusively determined. In periods with EF influence, the proportion of non-stationary activities decreased to 75% and 85%, whereas stationary activities increased to 22% and 6%. The "indeterminate" category also showed slight variations, with 2% and 9% of the activities falling into this group.

The "indeterminate" classification, though a minority, highlights cases where delinquency patterns are less predictable or fall outside the thresholds for stationarity detection. This suggests that EF may not only influence stationarity directly but may also introduce complexity in delinquency patterns, making it harder to classify definitively.

This behavior may be explained by the risk management practices of financial institutions during stable periods. In the absence of external shocks, institutions are more capable of identifying early signs of payment delays and can adjust their credit exposure accordingly—often by suspending or restricting lending to high-risk segments. These interventions disrupt emerging trends, resulting in more irregular default behavior that tends to deviate from stationary assumptions due to abrupt shifts in mean or variance.

Conversely, during systemic external shocks—such as pandemics or climate disruptions—multiple economic sectors experience distress concurrently, reducing the institution's ability to react. This leads to prolonged and homogeneous delinquency patterns across broader borrower groups, increasing the likelihood of stationarity in the associated time series.

These findings are summarized in Table 7, which provides a detailed breakdown of the stationarity classifications across the different periods of evaluation. The results underscore the dynamic impact of external factors on the stationarity of delinquency days in economic activities, reinforcing the need to account for these factors in economic analyses and credit risk prediction models.

4.2. Definition of Evaluated Scenarios and Causality Testing

To evaluate the seasonal impact of EF on the default behavior of economic activities, the Granger causality test (see Equation (2)) was applied to the time series representing the days of default for each economic activity, in conjunction with the time series of each EF. For simplicity, a binary value of 1 was assigned when the p-value was less than 0.05, indicating statistically significant causality, and 0 otherwise. This binary classification was performed every month to effectively capture temporal variations.

The results of this causality analysis were subsequently integrated into the credit dataset and associated with the corresponding economic activity and period under analysis.

Additionally, a global binary classification variable was calculated and appended to the dataset to identify the payment quality of credit holders (good or bad payers) based on monthly delinquency data. This enriched dataset provided the foundation for applying Algorithm 1 to the established datasets (see Figure 5) to compute the ACC, AUC, KS, and F1 metrics for the evaluated machine learning models. These models were assessed across the scenarios listed in Table 6. Based on the values obtained, the differences between the scenarios that include EF and SFE scenarios were calculated as shown in Table 8.

Table 8. Acronyms of the results for the evaluated scenarios.

Item	Results	Acronym
1	Result F minus SFE	F - SFE
2	Result P minus SFE	P - SFE
3	Result F + P minus SFE	F + P - SFE
4	Result U minus SFE	U - SFE
5	Result T minus SFE	T - SFE
6	Result W minus SFE	W - SFE
7	Result U + T minus SFE	U + T - SFE
8	Result T + W minus SFE	T + W - SFE
9	Result U + W minus SFE	U + W - SFE
10	Result U + T + W minus SFE	U + T + W - SFE

Note: This table provides acronyms for the results of the evaluated scenarios, comparing the differences between scenarios like "Result F minus SFE" and "Result U + T + W minus SFE". These comparisons are systematically abbreviated for clarity.

Algorithm 1 Dataset evaluation using ML algorithms.

EVALUATE(**data**, **target**, **features**, **model_name**)
1. Remove collinear variables (correlation $> 90\%$)
2. Perform Stratified k-Fold Cross-Validation ($k = 10$)
3. In each fold:
a. Use 90% for training, 10% for testing
b. Split training set: 75% train, 25% validation
c. Apply SMOTE for class balancing on train set
d. Train selected model: LDA, QDA, Perceptron, MLP, Ridge, RF, XGB, KNN, DL, CNN, XNN, FFN (Standard, Sparse, Local Penalty), EBM
e. Compute metrics: *AUC_Val*, *ACC_Val*, *AUC_Test*, *ACC_Test*, *KS_Val*, *KS_Test*, *F1_Val*, *F1_Test*
return Mean metrics across 10 folds

EVALUATE_ALL_MODELS(**data**, **target**, **features**)
1. Iterate over models and feature sets
2. Evaluate on datasets: **1. D100**, **2. D99**, **3.D80**
return DataFrame with results

EVALUATE_XAI_BEST_MODELS(**best_Model**)
1. Evaluate SHAP
2. Evaluate LIME
return DataFrame with SHAP and LIME results

FEATURES_EXPLANATION(**best_Model**)

The performance metrics (ACC, AUC, KS and F1) derived from the application of machine learning models for the binary classification of good and bad payers, evaluated across datasets stratified by relevance to established scenarios, are presented in Tables A1–A16 in Appendix A. These tables provide a detailed comparative analysis of the scenarios with and without the inclusion of EF related to COVID-19, social unrest, and climate change.

Additionally, Tables A17–A22 in Appendix A identify the machine learning models that achieved statistically significant improvements in their classification metrics when the EFs were incorporated. These tables delineate the conditions under which specific models demonstrate enhanced predictive capabilities, providing critical insights into the sensitivity and adaptability of various machine learning approaches to external factors.

The comprehensive dataset from these experiments forms a robust foundation for addressing the research questions that underpin this study. Given the extensive volume of information presented, the subsequent discussion focuses on the most significant results, with direct references to the relevant tables to highlight critical findings and their implications for predictive modeling under diverse scenarios.

RQ1. *How did the COVID-19 pandemic, through factors such as positive cases and mortality rates, influence credit defaults in different economic activities?*

The first research question is focused on understanding the impact of the COVID-19 pandemic on credit delinquency, particularly through the inclusion of variables such as confirmed infections and mortality. A comparative analysis was conducted between baseline scenarios SFE and those influenced by pandemic conditions, including a time-shifted mortality scenario (COVID MOV). Across all evaluation metrics—AUC, ACC, KS, and F1—the results consistently indicate a marked shift in credit risk behavior during pandemic periods.

Among the pandemic-related variables, mortality rates—especially when temporally aligned to reflect delayed economic effects—proved to be the most influential predictor of default. This effect was most evident in the 3. D80 dataset, which encompasses high-contact economic sectors that are particularly vulnerable to pandemic restrictions and health crises. In this scenario, the inclusion of lag-adjusted mortality data significantly enhanced model performance, suggesting that delayed health outcomes capture the real-world propagation of credit risk better than contemporaneous case counts alone.

The presence of government-backed credit mechanisms also played a decisive role. In datasets where such guarantees were included (`WGB_CR`), the predictive impact of the pandemic was partially cushioned, resulting in lower volatility and more stable classification outcomes. Conversely, in non-guaranteed portfolios (`WOGB_CR`), default risks increased substantially, reflecting the lack of institutional buffers.

In terms of model performance, *CNN* and *XGB* emerged as the most effective across all partitions, consistently achieving top-tier predictive accuracy and discrimination. While models such as *EBM* demonstrated improvements in specific scenarios—particularly under external stressors—others like *XNN* exhibited limited effectiveness, especially in pandemic-related contexts, which constrained their applicability despite their inherent interpretability features.

These findings are further reinforced by the interpretability analyses shown in Figure 6, where SHAP and LIME techniques reveal how the relative importance of predictive features shifts across scenarios. Variables related to debt burden and origination characteristics gained or lost prominence depending on whether pandemic indicators were present, reflecting the models' dynamic response to exogenous health shocks.

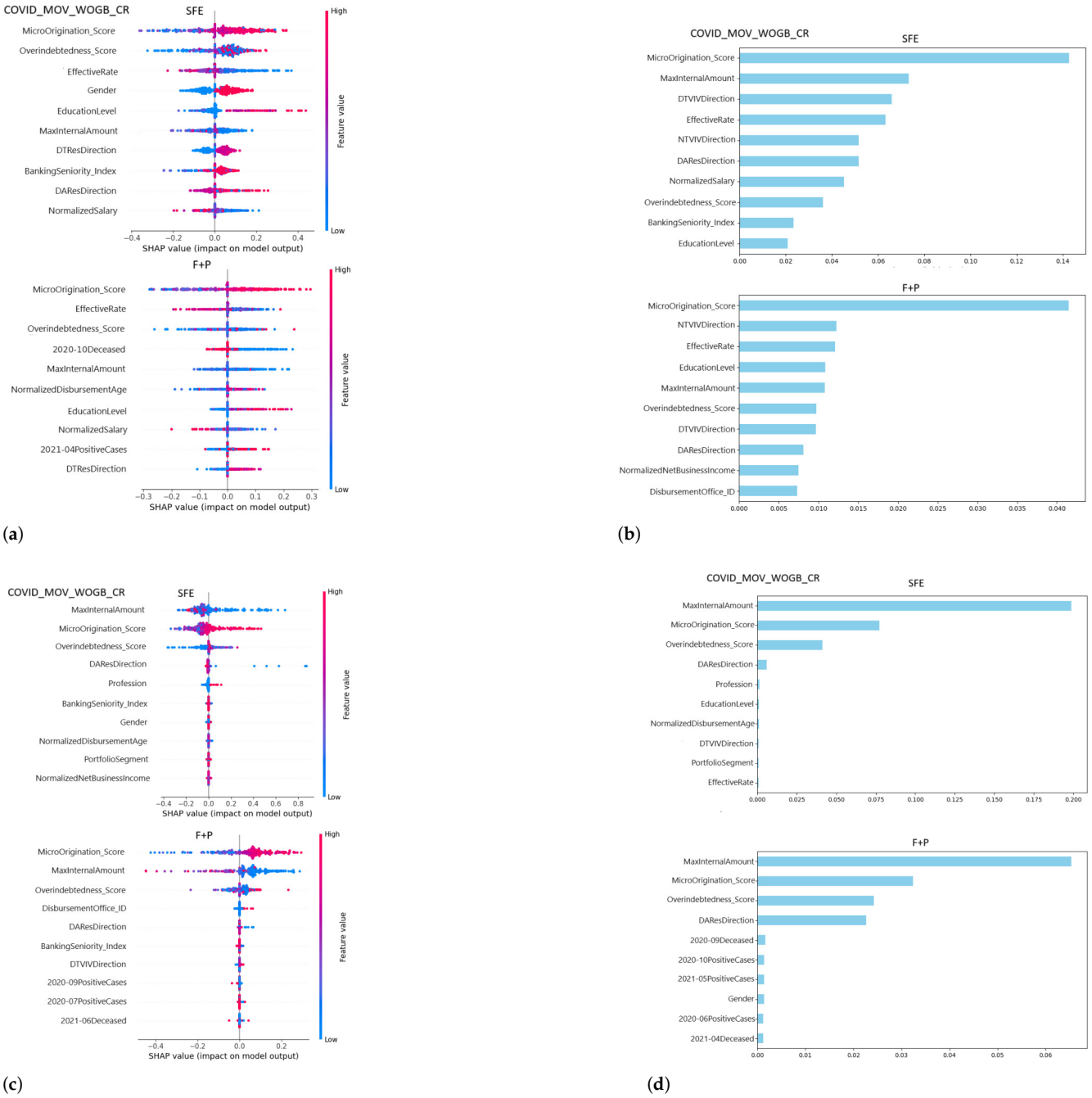


Figure 6. Top 10 features between without external factors (SFE) and COVID-19 stress scenarios (F + P). (**a**) XGB SHAP analysis, (**b**) XGB LIME analysis, (**c**) CNN SHAP analysis, and (**d**) CNN LIME analysis. Legend: F = COVID-19 confirmed cases; P = COVID-19 deceased; F + P = Combined scenaries; Mov = Mobility restrictions; WGB_CR = With Government-Backed Credit; WOGB_CR = Without Government-Backed Credit.

In summary, the COVID-19 pandemic had a substantial and measurable effect on credit delinquency patterns. Mortality—particularly when time-shifted—proved to be the most impactful factor, and its inclusion substantially improved model robustness. Government interventions, while partially effective, could not fully suppress the delinquency risks in the most exposed segments. These insights underscore the value of integrating lag-sensitive external variables into credit risk modeling, especially for sectors with heightened structural vulnerability.

Figure 6 provides interpretability insights into the predictive behavior of the *XGB* and *CNN* models under both baseline (SFE) and COVID-affected (F + P) scenarios, using SHAP and LIME explainability techniques.

Figure 6a,b illustrate the SHAP and LIME results for the *XGB* model, respectively. Under the SFE scenario, features such as `MicroOrigination_Score`, `Overindebtedness_Score`, and `MaxInternalAmount` exhibit the greatest impact on prediction outputs. In the F + P setting, variables directly linked to pandemic conditions—such as `2021-04PositiveCases` and `2020-10Deceased`—emerge alongside financial indicators, suggesting an integration of external health shocks into the model's risk prioritization.

Figure 6c,d display the corresponding SHAP and LIME analyses for the *CNN* model. Similar to *XGB*, the *CNN* model assigns greater relevance to features such as `MaxInternalAmount` and `MicroOrigination_Score` under the baseline (SFE) scenario. In the F + P scenario, feature importance shifts toward pandemic-related variables and demographic indicators (e.g., `2020-09PositiveCases`, `2020-09Deceased`, `DAResDirection`, `Gender`), reflecting the model's responsiveness to contextual changes introduced by COVID-19.

Together, these visualizations underscore the dynamic interaction between model architecture and scenario context, revealing how external stressors—particularly COVID-19—reshape the prioritization of financial and behavioral risk indicators within advanced machine learning models.

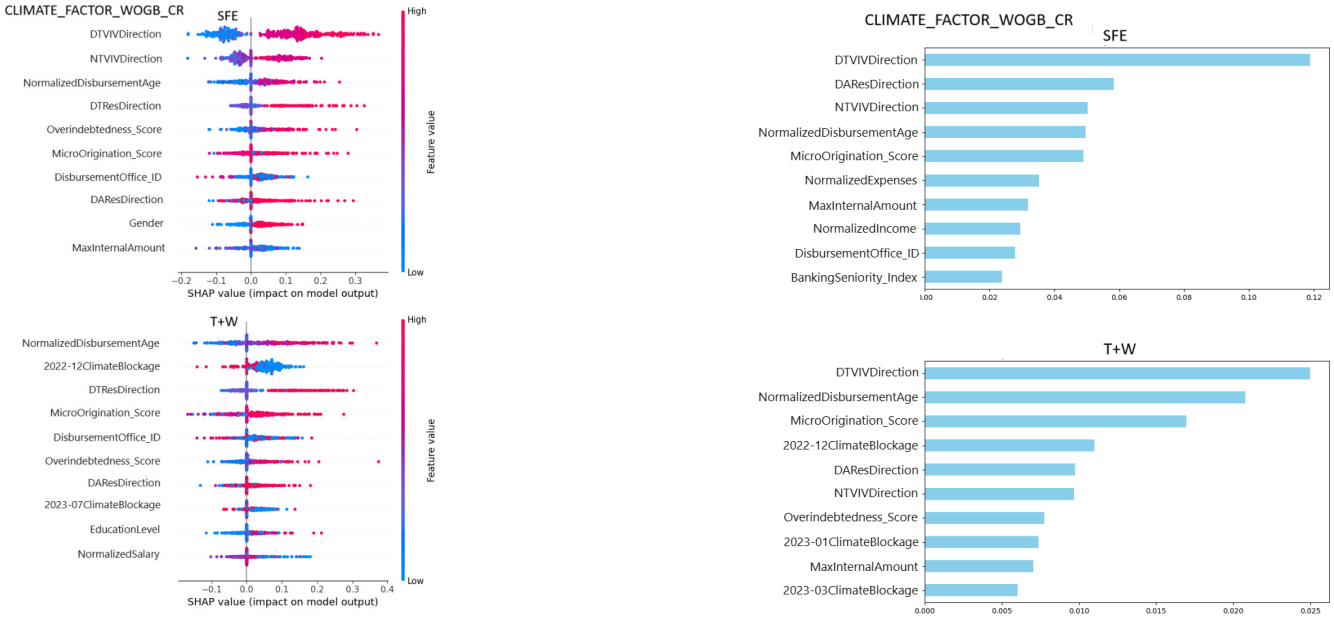
RQ2. *To what extent do climate change indicators, such as temperature anomalies and road blockages due to weather, impact credit delinquency patterns?*

The analysis of climate-related external factors—specifically temperature anomalies (T) and weather-induced road blockages (W)—reveals a measurable and statistically significant influence on credit delinquency behavior. While the magnitude of this effect is generally more moderate compared to pandemic-related variables, it remains consistent across scenarios and data partitions.

Tables A9–A12 (`WGB_CR`) and Tables A13–A16 (`WOGB_CR`) show notable improvements in classification metrics such as ACC, AUC, KS, and F1 upon including climate variables. These gains are particularly prominent in the absence of government guarantees, where borrowers are more directly affected by logistical disruptions. In portfolios without policy support (`WOGB_CR`), the effect of weather-related blockages is accentuated, often contributing to delays in income flows and operational closures—factors that increase credit risk exposure.

From a comparative perspective, the predictive gains attributed to climate anomalies may not be as large as those observed under pandemic mortality scenarios, yet they are highly context dependent. For instance, improvements in AUC of up to 3% and accuracy gains of over 30% have been recorded in neural network models, especially when the economic activities under analysis are concentrated in geographically exposed sectors. This supports the findings of [30], who demonstrated that road infrastructure disruptions can propagate liquidity shocks and elevate default risk, particularly in informal or underserved lending markets.

Figure 7 illustrates how the *XGB* model integrates climate variables into its decision structure, ranking them alongside key behavioral and demographic attributes. This interpretability result reaffirms that exogenous environmental stressors, while external to borrower control, are internalized by advanced algorithms as significant contributors to credit delinquency prediction.



(**a**) (**b**)

Figure 7. Top 10 features between scenarios without external factors (SFE) and with climate external factors (T + W): (**a**) XGB SHAP analysis. (**b**) XGB LIME analysis. Legend: SFE = Scenario without external factors; T = Temperature Anomaly; W = Climate Blockage; T + W = Combined scenarios; WGB_CR = With Government-Backed Credit; WOGB_CR = Without Government-Backed Credit.

In sum, the inclusion of climate indicators in credit risk models enhances their ability to reflect real-world disruptions and borrower vulnerability. The observed performance improvements are not only statistically valid but also economically meaningful, underscoring the importance of incorporating climate-related risks into financial predictive systems, particularly in emerging markets subject to environmental volatility.

Figure 7 provides interpretability insights into the behavior of the *XGB* model under climate-related stress conditions. Figure 7a presents the SHAP analysis comparing the baseline scenario SFE and the scenario that includes both temperature anomalies and weather-induced road blockages (T + W). The visualization highlights the ten most influential features in each setting, with color gradients indicating the direction and intensity of their contribution to the model's predictions. High feature values are represented in red and low values in blue, allow a clear interpretation of how variable magnitudes relate to increased or decreased credit risk.

Figure 7b displays the corresponding LIME analysis for the same dataset (`CLIMATE_FACTOR_WOGB_CR`). In both scenarios, `DTResDirection` emerges as the most critical driver of classification outcomes. Notably, under the T + W scenario, `Gender` gains relevance, suggesting that climate-related externalities may interact differently with sociodemographic characteristics. This contrast between the SFE and T + W contexts reinforces the need to account for latent vulnerabilities when modeling credit risk under environmental stress.

Together, these interpretability results confirm that the inclusion of climate factors not only enhances model accuracy but also reconfigures the internal prioritization of risk features. This dynamic adaptation reflects the capacity of machine learning models like *XGB* to absorb and reflect external shocks through learned decision boundaries.

RQ3. *What is the relationship between credit delinquency and social unrest, considering disruptions to economic activities and societal stability?*

Social unrest—manifested through protests, strikes, and other large-scale disruptions—can severely affect economic activity by limiting mobility, reducing consumer demand, and disrupting informal and formal markets. These effects become particularly salient in borrower segments highly dependent on local or face-to-face interactions.

Our analysis confirms that the inclusion of social unrest indicators (U) in the datasets yields statistically significant improvements in credit delinquency prediction. These improvements are evident in both `WGB_CR` and `WOGB_CR` portfolios, although the magnitude of the effect is more pronounced in the latter. This distinction highlights the protective role of policy guarantees, which tend to buffer the financial impact of systemic disruptions.

In terms of model responsiveness, *CNN*, and *XGB* exhibit the highest sensitivity to unrest scenarios, achieving notable gains in metrics such as AUC, ACC, KS, and F1. Classical models, including *LDA* and *Ridge*, also show moderate but statistically consistent improvements. Notably, in the 3. D80 dataset—focused on high-contact economic activities—the accuracy and discrimination capacity of models improve substantially under social unrest scenarios, suggesting that these sectors are disproportionately vulnerable to institutional instability.

Interpretability analyses provide further insight. As shown in Figure 8, SHAP values for *CNN* and *XGB* highlight a shift in feature relevance when social unrest variables are introduced. Features such as `Overindebtedness_Score` and `MaxInternalAmount` become more prominent, reflecting the increased importance of financial stress indicators in volatile sociopolitical contexts.

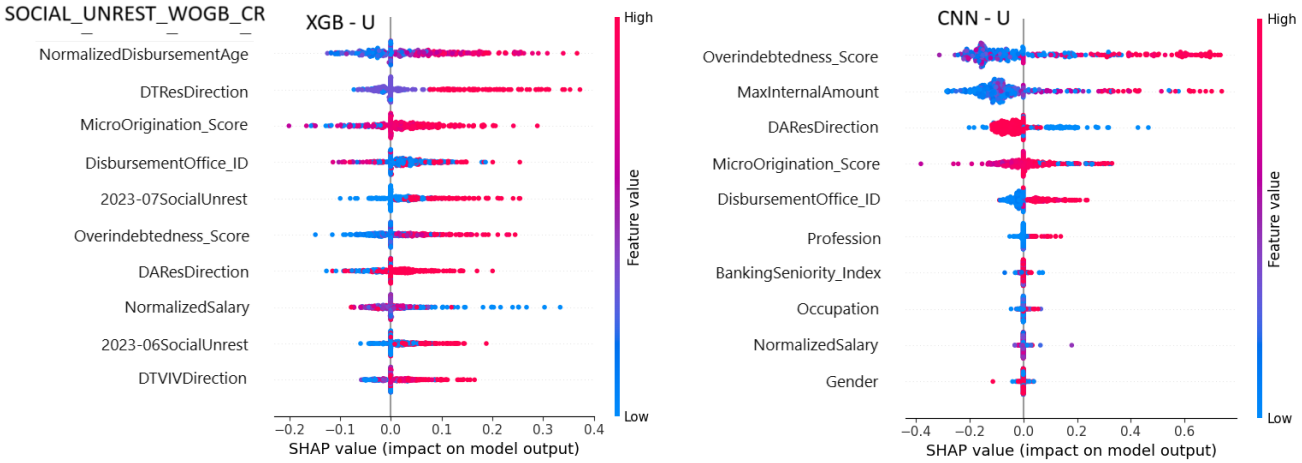


Figure 8. SHAP XGB and CNN: Top 10 Features in scenarios with social unrest (U). Legend: U = social unrest, represented by events of civil protest and political instability registered in 2023; WGB_CR = with government-backed credit; WOGB_CR = without government-backed credit.

In summary, the presence of social unrest introduces measurable and context-sensitive volatility in credit delinquency behavior. Its integration into predictive models not only improves classification performance but also enhances the capacity to identify borrower segments at heightened risk. These findings support the existing literature on the financial impacts of sociopolitical instability and underscore the importance of incorporating such variables into credit risk assessments, particularly in emerging economies where unrest events are recurrent and unevenly absorbed by the population.

Figure 8 presents the SHAP feature importance analysis for the *XGB* and *CNN* models under the "U scenarios" in the `SOCIAL_UNREST_WOGB_CR dataset`. The top 10 influential features are shown for each model, with `Overindebtedness_Score` and `MaxInternalAmount` being particularly significant in both, highlighting their impact on the model predictions.

RQ4. *How do the combined effects of external factors (COVID-19, climate change, and social unrest) contribute to variations in credit delinquency, and what are the most influential factors?*

The combination of external factors—pandemic severity, climatic anomalies, and sociopolitical instability—amplifies the complexity and unpredictability of credit behavior. When analyzed jointly, these variables generate compounded stress conditions that exert a stronger influence on delinquency patterns than any of them in isolation. This effect is particularly evident in the absence of government-backed guarantees (`WOGB_CR`), where borrowers are fully exposed to systemic shocks.

Empirical results from multifactor scenarios, such as U + T + W and Climate+Unrest, show a higher density of statistically significant gains across all evaluation metrics. Notably, mortality indicators (especially in lag-adjusted form, i.e., COVID MOV) remain the dominant predictor of default, reflecting the profound economic dislocation associated with fatal health outcomes. Climatic variables, although more moderate in their individual contribution, reinforce vulnerability in certain borrower segments, particularly when the infrastructure is disrupted by weather-related events. Social unrest introduces further pressure by destabilizing labor income and limiting business continuity.

Model behavior under combined factor scenarios reveals distinct patterns. *XGB* consistently ranks as the most resilient and adaptive model, achieving significance across all multifactor experiments. *CNN* also demonstrates strong performance, although with some sensitivity to the structure of the input data and the intensity of external stressors. While *XNN* was included as an interpretable neural model, its performance was inconsistent, limiting its contribution beyond methodological comparison. Classical models such as *Ridge* and *LDA* continue to offer baseline robustness, particularly in datasets with partial shielding (e.g., `WGB_CR`), although their gains remain smaller and more scenario dependent.

Figure 9 supports these findings by showing the SHAP analyses for *CNN*, *XGB*, and *XNN* in scenarios combining all external factors. Variables such as `MaxInternalAmount`, `Overindebtedness_Score`, and `DAResDirection` consistently emerge as top predictors, reflecting how borrower financial behavior interacts with macro-level disruptions. The shifting prominence of these features across models underscores the necessity of interpretability tools to contextualize model predictions in complex environments.

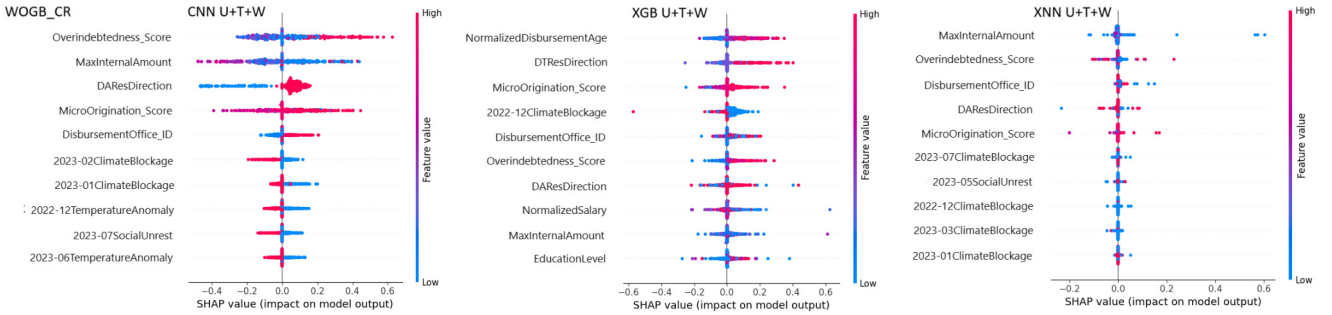


Figure 9. SHAP CNN, XGB, and XNN: Top 10 features in scenarios combining external factors. Legend: U = social unrest (e.g., protests and road blockages); T = temperature anomalies; W = climate blockages; U + T + W = combined scenario including all three external factors; WOGB_CR = credit risk portfolio excluding government-backed loans.

In conclusion, the combined influence of external shocks intensifies credit risk dynamics and reveals structural fragilities in both the borrower base and the credit portfolios. The models' ability to detect and adapt to these stressors is contingent not only on the algorithmic design but also on how exogenous variables are incorporated, preprocessed, and aligned temporally. These findings reinforce the importance of building integrated, context-aware credit risk models—particularly in economies where multiple external threats converge and the margin for institutional response is limited.

5. Discussion and Research Implications

This study examined how credit delinquency prediction is affected by the incorporation of EF, including pandemic severity, climate anomalies, and social unrest. The findings demonstrate that these variables introduce significant and context-dependent shifts in credit behavior, particularly in scenarios without government-backed guarantees. At the same time, the process revealed methodological challenges that are increasingly relevant in the application of ML in dynamic financial environments.

The first challenge relates to the quality and availability of data on external disruptions. Publicly available records on COVID-19 mortality, weather events, and protest activity often lack consistency and timeliness, complicating their integration into high-frequency credit datasets. Similar issues have been noted in prior empirical efforts conducted in Peru, where the construction of reliable time series for economic modeling has required substantial preprocessing [28,29]. In this study, the alignment of mortality data through temporal shifts proved essential, as it revealed the delayed but pronounced influence of pandemic fatalities on repayment capacity.

Another source of complexity was the divergence between credit portfolios with and without public guarantees. Government-backed loans mitigated part of the risk during systemic shocks but also altered the underlying statistical distribution of defaults. This distinction made model generalization more challenging but also more necessary for understanding the buffering effects of financial policy. Prior studies in Latin America suggest that the performance of ML classifiers can vary significantly depending on whether policy instruments like emergency credit programs are present [27].

Overfitting risks were also non-trivial, particularly when combining multiple exogenous factors in deep learning architectures. Although regularization and cross-validation strategies were implemented, the complexity of models such as *CNN* and *XNN* occasionally reduced transparency. While post hoc interpretability tools like SHAP and LIME help uncover shifts in feature relevance across scenarios, they do not entirely resolve concerns regarding model opacity—especially in contexts where regulatory clarity and explainability are essential [30,59].

Despite these challenges, several insights emerge. Mortality-related variables—especially when time-lagged—exert the strongest influence on credit behavior, particularly in high-contact economic sectors. Climate-related disruptions, though more moderate in effect, consistently improved predictive accuracy, especially in vulnerable portfolios. Social unrest also introduces nonlinear shocks, further reinforcing the need for adaptive models that account for real-world volatility.

The integration of multiple external stressors revealed a compounding effect on delinquency, a finding that conventional econometric approaches may fail to capture. Tree-based and neural models, particularly *XGB* and *CNN*, showed the greatest adaptability in this regard. At the same time, classical models like Ridge maintained stable performance under certain constrained scenarios, suggesting a role for hybrid modeling strategies.

In the Peruvian context, these findings align with recent efforts to modernize credit risk assessment through advanced analytics. Several local studies highlight the value of incorporating contextual variables and borrower segmentation into scoring systems, particularly in the wake of the pandemic and its economic aftermath [29,35]. Internationally, there is growing consensus around the importance of multidimensional modeling frameworks that include environmental and social dimensions alongside traditional financial indicators.

Overall, this study contributes to the ongoing development of risk prediction systems capable of responding to complex and evolving realities. It underscores the importance of aligning exogenous variables not only temporally but also conceptually with the credit cycle. Future implementations of ML in finance—particularly in emerging economies—should prioritize data interoperability, scenario partitioning, and explainable design to ensure both accuracy and transparency in risk-based decision-making.

6. Conclusions and Future Research

This study has demonstrated that credit delinquency prediction improves significantly when EF—such as COVID-19 severity, climate anomalies, and social unrest—are systematically integrated into machine learning workflows. Using a multi-scenario design and high-resolution time series, the models captured both the direct and delayed effects of these external shocks across various economic activities and portfolio compositions.

By incorporating explanatory tools such as SHAP and LIME, and validating the results with formal stationarity and causality tests, the study provided both predictive and interpretive value. Notably, the performance of key metrics—AUC, ACC, KS, and F1—was significantly enhanced in models exposed to time-shifted mortality and climate-related variables. This confirmed that delayed systemic shocks (e.g., COVID-related deaths) exert stronger predictive influence than contemporaneous infection rates, and that road infrastructure disruptions propagate liquidity risks, particularly in non-guaranteed portfolios.

The comparative evaluation revealed that *CNN* and *XGB* consistently outperformed other models across multiple relevance partitions and scenarios. While *EBM* proved especially useful in contexts requiring greater model transparency, *XNN* served as an exploratory reference for intrinsically interpretable architectures, although its predictive effectiveness was limited in high-volatility scenarios. Moreover, the role of government-backed guarantees emerged as central: their presence moderated the effects of external crises, whereas their absence magnified the volatility and default risks observed in the models.

Overall, these findings offer a scenario-driven framework for understanding how systemic shocks affect credit behavior. This framework is replicable and adaptable to emerging economies facing concurrent crises, combining domain-sensitive data transformations (e.g., lag structures) with advanced modeling and interpretability techniques. Importantly, the integration of KS and F1 metrics deepens the evaluation of not just model accuracy but also the discriminatory power and balance between false positives and false negatives—factors critical to financial risk management and regulatory compliance.

Looking forward, several research directions are both relevant and necessary. First, the use of static lag structures could be extended through adaptive or dynamic lag modeling, particularly in the case of gradual-onset crises such as climate change or prolonged social instability. Second, cross-regional replication would test the generalizability of these findings in other vulnerable markets, where the interaction between policy buffers and external shocks may differ. Third, in-depth analysis of credit policies—such as emergency moratoria or targeted guarantees—would enable the finer-grained modeling of institutional effectiveness under stress. Finally, explainability remains a methodological frontier: while SHAP and LIME provide insights, future work could explore intrinsically interpretable models or causal frameworks that support regulatory transparency and stakeholder trust.

In conclusion, this research confirms the value of integrating multidimensional exogenous variables into credit risk modeling, offering a robust and explainable approach to capturing volatility in increasingly complex economic environments. As global financial systems face overlapping crises, such tools will become indispensable for promoting resilience, fairness, and informed decision-making in credit allocation.

Author Contributions: Conceptualization, J.N. and J.H.; methodology, J.N.; validation, J.N., L.R. and J.H.; formal analysis, J.N., J.C. and J.H.; investigation, J.N.; resources, J.N.; writing—original draft preparation, J.N.; writing—review and editing, J.N., L.R., J.C. and J.H.; visualization, J.N. and J.C.; supervision, J.H.; project administration, J.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets presented in this study are partly publicly available and partly proprietary. Public datasets (COVID-19 cases, deaths, road blockages, and temperature anomalies) are accessible through the official URLs listed in Table 2. Proprietary financial datasets (credit delinquency activity and credits with external factors) were used under confidentiality agreements and have been anonymized and aggregated. Metadata, modeling programs, and experimental results have been deposited and are openly available at Zenodo: https://zenodo.org/records/14890903. Further inquiries can be directed to the corresponding author(s).

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Results Tables

Table A1. Average AUC and ACC across 10 folds of ML models with COVID EF WGB_CR, by evaluated scenario and relevance groups.

	AUC (%)							ACC (%)						
	Scenarios				Results			Scenarios				Results		
	SFE	**P**	**F + P**	**F**	**P - SFE**	**F + P - SFE**	**F - SFE**	**SFE**	**P**	**F + P**	**F**	**P - SFE**	**F + P - SFE**	**F - SFE**
WGB_CR														
1. D100														
cnn	72.82	73.63	74.09	74.09	0.82	1.28	1.27	67.49	70.07	70.18	70.14	2.58	2.69	2.65
ebm	81.13	81.23	81.42	81.26	0.10	0.29	0.13	74.84	74.94	75.05	74.93	0.09	0.20	0.08
ffn	72.30	72.56	72.39	73.38	0.26	0.10	1.08	66.74	65.24	66.15	66.14	−1.50	−0.59	−0.60
ffnlp	72.11	72.35	72.41	72.72	0.25	0.30	0.61	66.28	67.55	65.64	65.19	1.27	−0.64	−1.09
ffnsp	72.31	72.31	71.42	72.77	0.00	−0.89	0.47	66.40	66.45	64.67	65.59	0.05	−1.73	−0.81
knn	67.21	66.92	66.93	66.93	−0.29	−0.28	−0.28	63.23	63.00	62.97	62.98	−0.22	−0.26	−0.24
lda	67.23	68.01	68.54	68.25	0.78	1.31	1.02	63.09	63.74	64.09	63.77	0.65	1.00	0.67
mlp	69.30	70.83	71.17	70.47	1.53	1.87	1.17	60.82	63.26	63.58	62.87	2.44	2.76	2.06
percep	57.96	58.44	59.88	59.46	0.48	1.92	1.50	44.58	57.99	56.37	49.31	13.41	11.78	4.73
qda	67.44	66.98	65.81	67.17	−0.46	−1.63	−0.27	52.50	54.01	54.13	53.72	1.52	1.64	1.22
rf	81.74	81.51	81.21	81.46	−0.23	−0.53	−0.29	75.51	75.44	75.22	75.48	−0.07	−0.29	−0.03
ridge	66.94	67.55	68.05	67.81	0.62	1.12	0.87	62.93	63.39	63.70	63.43	0.46	0.77	0.50
xgb	80.76	81.00	81.10	81.07	0.23	0.33	0.30	74.77	74.98	75.03	74.98	0.21	0.26	0.22
xnn	50.14	50.38	50.10	50.37	0.24	−0.05	0.22	53.42	57.99	62.76	54.22	4.57	9.34	0.80
2. D99														
cnn	72.75	73.60	74.27	73.57	0.86	1.52	0.83	67.04	70.03	70.42	69.97	2.98	3.37	2.93
ebm	81.04	81.24	81.37	81.24	0.19	0.33	0.20	74.77	74.93	75.04	74.94	0.16	0.28	0.18
ffn	72.69	72.38	71.50	71.55	−0.31	−1.20	−1.14	66.07	65.94	65.86	65.75	−0.14	−0.21	−0.32
ffnlp	71.71	72.81	72.16	72.45	1.09	0.44	0.74	65.27	66.83	66.28	65.94	1.56	1.02	0.67
ffnsp	72.58	73.11	72.58	72.99	0.53	−0.00	0.41	63.93	65.34	66.81	66.17	1.41	2.87	2.24
knn	67.10	66.88	66.88	66.88	−0.22	−0.21	−0.21	63.24	63.02	63.02	63.04	−0.23	−0.22	−0.21
lda	67.17	67.92	68.46	68.19	0.75	1.29	1.02	63.07	63.68	64.04	63.78	0.60	0.96	0.70
mlp	69.79	68.93	69.62	68.70	−0.86	−0.16	−1.08	63.83	64.12	59.76	64.96	0.29	−4.07	1.13
percep	55.09	57.67	54.19	56.99	2.58	−0.90	1.90	55.48	58.56	53.93	50.60	3.07	−1.55	−4.88
qda	67.32	66.90	65.74	67.08	−0.42	−1.58	−0.23	52.41	53.98	54.02	53.63	1.57	1.62	1.22
rf	81.68	81.33	81.02	81.30	−0.35	−0.66	−0.38	75.63	75.32	75.09	75.38	−0.31	−0.54	−0.24
ridge	66.85	67.47	67.97	67.74	0.62	1.12	0.89	62.94	63.39	63.68	63.49	0.45	0.73	0.55
xgb	80.72	80.96	80.97	80.95	0.24	0.25	0.24	74.81	74.97	74.89	74.86	0.16	0.08	0.04
xnn	50.47	50.25	50.48	50.59	−0.23	0.00	0.11	57.50	53.55	59.22	54.37	−3.95	1.71	−3.14
3. D80														
cnn	72.68	73.18	73.79	73.50	0.50	1.11	0.82	66.38	70.28	70.65	70.44	3.90	4.26	4.06
ebm	80.73	81.03	81.16	81.05	0.29	0.43	0.31	74.92	75.16	75.29	75.19	0.24	0.37	0.27
ffn	71.76	72.16	72.38	72.83	0.41	0.62	1.07	63.24	66.92	66.19	67.18	3.68	2.95	3.94
ffnlp	71.14	72.00	71.66	71.97	0.86	0.51	0.83	65.73	66.90	66.45	65.44	1.16	0.72	−0.29
ffnsp	70.87	72.03	72.29	72.13	1.15	1.42	1.25	64.07	65.62	66.69	65.65	1.54	2.61	1.58
knn	66.57	66.57	66.57	66.57	−0.01	0.00	0.00	62.98	62.95	62.95	62.96	−0.02	−0.03	−0.01
lda	67.33	67.90	68.29	68.07	0.57	0.96	0.74	63.45	63.88	64.20	64.02	0.43	0.75	0.57
mlp	67.96	68.85	68.61	69.92	0.89	0.66	1.96	58.09	63.16	62.43	60.73	5.07	4.34	2.64
percep	57.44	55.39	59.27	56.66	−2.05	1.83	−0.78	52.88	57.22	50.16	56.66	4.35	−2.71	3.78
qda	67.69	66.51	64.11	66.30	−1.18	−3.58	−1.39	52.76	54.28	52.68	53.73	1.52	−0.08	0.97
rf	81.18	81.03	80.68	80.96	−0.15	−0.51	−0.22	75.47	75.57	75.26	75.52	0.11	−0.20	0.05
ridge	66.67	67.23	67.59	67.40	0.56	0.92	0.73	63.06	63.35	63.73	63.55	0.29	0.67	0.49
xgb	80.20	80.75	80.76	80.69	0.54	0.56	0.49	74.74	75.13	75.15	75.14	0.39	0.41	0.39
xnn	50.35	50.07	49.89	50.03	−0.28	−0.47	−0.33	63.26	63.80	60.07	53.46	0.54	−3.19	−9.80

Note: This table summarizes the average AUC and ACC metrics obtained across 10 folds for machine learning models applied in binary classification scenarios (good and bad payers). The analysis is focused on datasets enriched with EF related to COVID-19, specifically scenarios including credits with governmental benefits (WGB_CR). Key findings include the following: 1. Models such as *EBM* and *RF* demonstrate robust performance, achieving the highest AUC values (above 81%), indicating strong discriminative capability for credit classification under EF-influenced scenarios. 2. The *CNN* model shows marked improvement in ACC, with increases of up to 4.26 percentage points when incorporating EF, highlighting its sensitivity to the enriched dataset. 3. Conversely, models such as *KNN* and *XNN* show limited or negative responsiveness to EF enrichment, with minimal improvements or declines in metrics. 4. The inclusion of governmental benefits as part of the COVID-19 EF notably enhances the classification performance for most models, emphasizing the critical role of such interventions in predictive modeling.

Table A2. Average AUC and ACC across 10 folds of ML models with COVID EF WOGB_CR, by evaluated scenario and relevance groups.

	AUC (%)							ACC (%)						
	Scenarios				Results			Scenarios				Results		
	SFE	P	F + P	F	P - SFE	F + P - SFE	F - SFE	SFE	P	F + P	F	P - SFE	F + P - SFE	F - SFE
WOGB_CR														
1. D100														
cnn	72.92	74.10	74.17	74.28	1.18	1.25	1.35	67.53	72.93	72.67	73.11	5.41	5.15	5.58
ebm	81.38	81.29	81.45	81.31	−0.09	0.07	−0.07	76.54	76.41	76.51	76.48	−0.13	−0.03	−0.06
ffn	72.71	72.60	73.34	72.89	−0.11	0.63	0.18	67.46	67.29	67.49	68.33	−0.17	0.03	0.88
ffnlp	72.16	72.32	71.68	73.16	0.15	−0.49	1.00	66.86	66.56	63.51	66.75	−0.30	−3.35	−0.11
ffnsp	72.48	73.44	73.75	73.36	0.96	1.27	0.88	65.18	68.26	67.02	69.48	3.09	1.84	4.30
knn	66.21	65.98	65.98	65.99	−0.23	−0.23	−0.22	63.19	62.84	62.82	62.84	−0.35	−0.36	−0.35
lda	66.08	66.68	67.16	66.89	0.60	1.08	0.81	63.18	63.64	63.93	63.65	0.46	0.75	0.47
mlp	70.92	70.07	71.03	68.89	−0.85	0.11	−2.03	60.20	67.96	63.45	62.29	7.76	3.25	2.09
percep	55.94	54.88	57.04	55.72	−1.06	1.10	−0.22	54.77	59.56	50.92	57.58	4.78	−3.85	2.81
qda	66.82	65.97	64.75	66.18	−0.84	−2.07	−0.64	50.72	52.66	52.59	52.06	1.94	1.86	1.33
rf	79.53	79.09	78.75	78.99	−0.44	−0.78	−0.53	75.34	75.34	75.17	75.21	0.01	−0.16	−0.13
ridge	65.95	66.44	66.89	66.66	0.50	0.94	0.71	63.14	63.44	63.77	63.54	0.30	0.63	0.40
xgb	80.43	80.60	80.64	80.68	0.17	0.21	0.25	76.00	76.16	76.13	76.24	0.16	0.13	0.24
xnn	50.14	50.71	50.15	50.19	0.56	0.00	0.05	62.48	66.64	62.37	54.50	4.17	−0.10	−7.98
2. D99														
cnn	73.04	73.95	74.16	74.28	0.91	1.12	1.24	68.52	72.56	72.93	72.91	4.03	4.40	4.38
ebm	81.23	81.31	81.44	81.33	0.08	0.22	0.10	76.31	76.39	76.49	76.48	0.07	0.18	0.17
ffn	72.50	72.95	72.44	73.37	0.46	−0.05	0.87	64.38	65.93	64.91	68.10	1.54	0.53	3.72
ffnlp	72.72	72.92	73.36	72.79	0.20	0.64	0.07	68.01	65.29	69.55	70.48	−2.72	1.54	2.47
ffnsp	73.19	73.12	72.32	73.13	−0.07	−0.87	−0.06	67.88	61.75	68.83	65.88	−6.13	0.95	−2.00
knn	65.94	66.00	66.00	66.00	0.06	0.06	0.06	62.92	62.58	62.58	62.59	−0.35	−0.34	−0.33
lda	66.00	66.67	67.15	66.90	0.67	1.15	0.90	63.15	63.63	64.00	63.67	0.48	0.85	0.52
mlp	68.79	70.06	70.21	70.58	1.27	1.42	1.79	70.57	64.20	64.24	68.64	−6.37	−6.33	−1.93
percep	55.28	56.66	55.64	56.59	1.37	0.36	1.31	54.45	56.49	49.80	55.10	2.04	−4.65	0.65
qda	66.72	65.95	64.69	66.12	−0.77	−2.03	−0.60	50.57	52.65	52.77	52.16	2.08	2.20	1.60
rf	79.41	78.96	78.60	78.91	−0.46	−0.81	−0.51	75.31	75.19	74.90	75.24	−0.13	−0.41	−0.07
ridge	65.85	66.42	66.87	66.65	0.57	1.02	0.80	63.13	63.45	63.82	63.58	0.32	0.69	0.45
xgb	80.35	80.66	80.65	80.61	0.32	0.30	0.26	76.13	76.28	76.21	76.25	0.15	0.08	0.12
xnn	50.08	50.74	50.00	49.90	0.67	−0.08	−0.18	62.22	65.34	70.29	66.02	3.12	8.07	3.80
3. D80														
cnn	73.29	73.32	73.58	73.05	0.03	0.29	−0.24	68.07	72.86	73.28	72.64	4.78	5.20	4.56
ebm	80.75	81.05	81.13	81.02	0.30	0.37	0.27	76.37	76.72	76.74	76.69	0.35	0.37	0.32
ffn	71.19	73.05	72.68	73.08	1.86	1.49	1.88	63.49	66.53	66.51	68.70	3.04	3.02	5.22
ffnlp	72.25	72.57	71.92	73.00	0.32	−0.33	0.74	64.39	64.96	64.64	62.94	0.57	0.25	−1.45
ffnsp	71.72	73.07	72.27	72.45	1.35	0.55	0.72	68.29	64.39	64.76	65.33	−3.90	−3.53	−2.96
knn	65.37	65.36	65.36	65.37	−0.01	−0.01	−0.00	62.54	62.52	62.51	62.53	−0.02	−0.03	−0.00
lda	65.85	66.35	66.70	66.52	0.50	0.85	0.67	63.27	63.64	63.93	63.75	0.37	0.65	0.48
mlp	69.10	71.07	69.19	70.08	1.96	0.09	0.97	62.97	59.38	52.32	60.10	−3.59	−10.65	−2.87
percep	55.71	56.80	57.86	55.39	1.09	2.16	−0.31	51.81	55.32	52.05	50.98	3.51	0.24	−0.83
qda	66.65	65.29	63.02	65.30	−1.36	−3.63	−1.35	50.91	53.12	50.75	52.15	2.21	−0.16	1.25
rf	78.71	78.53	78.12	78.56	−0.18	−0.59	−0.15	75.29	75.54	75.39	75.65	0.25	0.10	0.36
ridge	65.48	65.97	66.26	66.13	0.48	0.78	0.65	63.14	63.37	63.65	63.58	0.23	0.51	0.43
xgb	79.57	80.16	80.16	80.20	0.59	0.59	0.63	75.92	76.45	76.44	76.45	0.54	0.52	0.54
xnn	50.76	50.57	49.89	49.89	−0.19	−0.87	−0.86	56.71	55.79	71.19	66.89	−0.92	14.48	10.18

Note: This table reports the average AUC and ACC across ten folds for binary credit classification (good vs. bad payers) in datasets enriched with COVID-19-related EF, but excluding credits with government-backed benefits (WOGB_CR). Notable observations include the following: 1. Models such as *EBM* and *XGB* maintain robust predictive capabilities, achieving AUC values above 80%, with minimal performance fluctuation in the absence of government-benefit data with EF. 2. Both *CNN* and *MLP* exhibit marked improvements in ACC, with the *CNN* model reaching gains of up to 5.58 percentage points in certain scenarios, thereby underscoring the impact of EF enrichment on classification accuracy. 3. The exclusion of government-benefit information adversely affects some models (e.g., *RF*, *QDA*, and *XNN*), which register declines in both AUC and ACC. This variability suggests that these models are more sensitive to the absence of such interventions. 4. Overall, these findings reveal that omitting data on governmental support introduces notable variability and potential biases in the models' predictive performance—while some models capitalize on reduced bias, others suffer from diminished input granularity. Collectively, these results illuminate the nuanced effects of excluding government-benefit data on credit risk classification and emphasize the integral role that such interventions play in mitigating bias and enhancing model accuracy.

Table A3. Average KS and F1 across 10 folds of ML models with COVID EF WGB_CR, by evaluated scenario and relevance groups.

	KS (%)							F1 (%)						
	Scenarios				Results			Scenarios				Results		
	SFE	**P**	**F + P**	**F**	**P - SFE**	**F + P - SFE**	**F - SFE**	**SFE**	**P**	**F + P**	**F**	**P - SFE**	**F + P - SFE**	**F - SFE**
1. D100														
cnn	36.00	36.34	37.65	37.16	0.34	1.65	1.16	74.04	78.43	78.30	78.28	4.39	4.26	4.24
ebm	46.69	48.66	49.02	48.59	1.97	2.33	1.90	80.49	81.36	81.48	81.42	0.86	0.99	0.92
ffn	34.10	34.29	34.81	34.48	0.19	0.71	0.38	70.15	73.61	70.72	72.97	3.46	0.57	2.82
ffnlp	35.44	35.13	34.95	34.57	−0.30	−0.49	−0.87	69.62	71.86	71.92	69.80	2.24	2.30	0.18
ffnsp	7.50	5.34	6.95	11.19	−2.16	−0.55	3.68	47.14	69.09	32.39	55.49	21.95	−14.75	8.35
knn	26.04	25.53	25.49	25.50	−0.51	−0.55	−0.54	69.55	69.31	69.27	69.28	−0.24	−0.28	−0.26
lda	24.32	25.24	26.28	25.80	0.93	1.97	1.48	69.92	70.25	70.56	70.28	0.33	0.64	0.36
mlp	30.67	28.94	30.37	30.44	−1.73	−0.30	−0.23	65.29	74.99	57.90	68.99	9.70	−7.39	3.70
percep	13.14	13.64	15.84	15.29	0.50	2.70	2.15	29.26	62.04	59.05	40.43	32.78	29.79	11.17
qda	26.02	25.55	23.25	25.70	−0.47	−2.77	−0.32	49.55	53.09	54.16	52.10	3.54	4.61	2.55
rf	49.27	48.88	48.19	48.76	−0.39	−1.07	−0.50	81.70	82.11	82.13	82.15	0.41	0.43	0.45
ridge	24.33	25.22	26.28	25.79	0.88	1.95	1.46	30.27	30.94	31.73	31.39	0.68	1.46	1.12
xgb	47.88	48.21	48.31	48.12	0.33	0.43	0.24	81.03	81.45	81.48	81.45	0.43	0.45	0.42
xnn	0.68	0.01	0.25	0.04	−0.67	−0.43	−0.64	73.61	79.33	72.10	71.51	5.72	−1.50	−2.09
2. D99														
cnn	35.96	36.57	36.01	36.50	0.61	0.04	0.53	71.65	78.71	74.94	78.32	7.05	3.29	6.67
ebm	46.33	48.39	48.65	48.27	2.06	2.32	1.94	80.67	81.33	81.47	81.27	0.66	0.79	0.60
ffn	34.10	35.21	34.90	34.31	1.11	0.80	0.22	71.22	72.95	70.99	73.04	1.73	−0.22	1.82
ffnlp	35.64	33.85	36.03	34.20	−1.78	0.39	−1.44	71.84	70.94	72.02	69.87	−0.90	0.18	−1.97
ffnsp	3.65	3.42	2.39	5.27	−0.23	−1.26	1.62	30.78	37.40	38.08	36.47	6.62	7.30	5.69
knn	25.85	25.61	25.62	25.65	−0.24	−0.23	−0.20	69.70	69.32	69.33	69.35	−0.38	−0.37	−0.35
lda	24.24	25.17	26.14	25.73	0.93	1.90	1.49	70.00	70.30	70.58	70.37	0.30	0.58	0.37
mlp	28.44	29.26	30.16	31.09	0.81	1.71	2.64	66.98	68.22	69.25	71.16	1.24	2.27	4.18
percep	9.56	12.65	9.00	12.14	3.09	−0.56	2.58	55.60	64.38	53.50	47.50	8.78	−2.10	−8.10
qda	25.86	25.52	23.10	25.68	−0.34	−2.77	−0.18	49.60	53.06	54.03	52.02	3.46	4.43	2.42
rf	48.94	48.29	48.09	48.41	−0.65	−0.85	−0.52	81.71	82.06	81.99	82.13	0.35	0.28	0.42
ridge	24.22	25.17	26.14	25.72	0.96	1.92	1.50	30.26	31.13	31.95	31.59	0.87	1.70	1.33
xgb	47.78	48.04	48.01	48.07	0.25	0.22	0.29	81.11	81.47	81.44	81.39	0.36	0.33	0.29
xnn	0.47	0.26	0.46	0.21	−0.21	−0.00	−0.26	56.15	64.27	64.65	56.02	8.12	8.51	−0.13
3. D80														
cnn	34.99	35.94	35.88	36.67	0.95	0.89	1.67	73.58	78.90	77.79	78.71	5.31	4.20	5.13
ebm	45.97	48.05	48.31	48.00	2.08	2.35	2.03	81.06	81.94	82.06	81.98	0.89	1.00	0.92
ffn	31.00	32.98	33.99	32.50	1.98	2.98	1.50	73.84	65.83	69.25	72.99	−8.01	−4.59	−0.85
ffnlp	32.14	32.53	34.70	33.98	0.39	2.55	1.83	67.56	72.39	75.08	73.50	4.83	7.52	5.94
ffnsp	4.74	9.54	4.96	4.68	4.80	0.22	−0.06	36.56	34.09	37.52	28.57	−2.47	0.96	−7.99
knn	25.09	25.07	25.08	25.07	−0.02	−0.01	−0.02	69.81	69.78	69.77	69.79	−0.03	−0.04	−0.02
lda	24.18	24.99	25.52	25.47	0.82	1.35	1.29	70.49	70.76	71.06	70.93	0.27	0.57	0.44
mlp	29.32	28.12	30.99	30.30	−1.21	1.67	0.97	66.46	75.92	73.22	65.00	9.46	6.76	−1.46
percep	12.46	9.77	15.04	11.38	−2.69	2.59	−1.08	49.76	59.25	44.51	59.75	9.49	−5.25	9.99
qda	26.25	24.52	20.38	24.33	−1.73	−5.87	−1.92	51.24	55.03	53.58	54.36	3.80	2.34	3.12
rf	48.16	48.21	47.44	47.94	0.06	−0.72	−0.22	81.87	82.66	82.60	82.60	0.79	0.73	0.73
ridge	24.17	24.98	25.53	25.45	0.81	1.36	1.28	31.33	32.18	32.88	32.36	0.85	1.55	1.03
xgb	47.03	47.83	47.66	47.80	0.80	0.63	0.77	81.40	82.02	82.05	82.02	0.62	0.65	0.62
xnn	1.73	1.53	0.21	0.06	−0.20	−1.51	−1.67	39.49	71.30	72.95	72.25	31.81	33.45	32.75

Note: This table summarizes the average KS and F1 metrics obtained across 10 folds for machine learning models applied in binary classification scenarios (good and bad payers). The analysis is focused on datasets enriched with EF related to COVID-19, specifically scenarios including credits with governmental benefits (WGB_CR). Key findings include the following: 1. Models such as *EBM* and *RF* demonstrate consistently strong KS and F1 scores, indicating robust performance under EF-influenced scenarios. 2. The *CNN* model show high F1 improvements, with increases of over 4 percentage points, highlighting its sensitivity to enriched datasets. 3. Conversely, models such as *KNN* and *XNN* show limited or negative responsiveness to EF enrichment. 4. The inclusion of governmental benefits as part of the COVID-19 EF notably improves the classification metrics, underscoring the importance of contextual variables in predictive modeling.

Table A4. Average KS and F1 across 10 folds of ML models with COVID EF WOGB_CR, by evaluated scenario and relevance groups.

	KS (%)							F1 (%)						
	Scenarios				Results			Scenarios				Results		
	SFE	P	F + P	F	P - SFE	F + P - SFE	F - SFE	SFE	P	F + P	F	P - SFE	F + P - SFE	F - SFE
WOGB_CR														
1. D100														
cnn	35.39	36.92	38.55	37.23	1.52	3.15	1.84	74.22	82.19	82.25	81.72	7.97	8.03	7.50
ebm	47.01	48.82	49.21	48.70	1.82	2.21	1.69	83.33	83.96	83.98	83.97	0.63	0.65	0.64
ffn	35.63	35.56	35.64	35.15	−0.06	0.02	−0.48	73.80	73.89	70.71	66.74	0.09	−3.09	−7.06
ffnlp	35.19	34.35	36.49	36.33	−0.84	1.30	1.14	74.45	72.76	73.73	73.33	−1.69	−0.72	−1.12
ffnsp	0.00	5.00	1.23	2.58	5.00	1.23	2.58	41.31	70.67	74.25	30.48	29.36	32.94	−10.84
knn	24.89	24.34	24.33	24.34	−0.55	−0.56	−0.56	71.07	70.69	70.67	70.69	−0.38	−0.40	−0.38
lda	22.95	23.93	24.61	24.39	0.98	1.65	1.43	71.66	71.87	72.18	71.93	0.21	0.52	0.27
mlp	29.53	30.31	27.95	33.21	0.78	−1.58	3.67	76.57	76.42	57.49	65.01	−0.15	−19.07	−11.55
percep	10.32	9.32	12.21	10.34	−1.00	1.89	0.02	55.33	65.60	48.97	64.21	10.27	−6.36	8.88
qda	25.55	23.93	21.31	23.99	−1.62	−4.24	−1.56	51.06	54.99	55.68	53.72	3.92	4.62	2.66
rf	45.86	45.21	44.57	44.64	−0.65	−1.29	−1.22	82.69	83.42	83.33	83.52	0.73	0.64	0.84
ridge	22.96	23.93	24.60	24.33	0.97	1.64	1.37	32.93	33.47	34.15	33.83	0.54	1.22	0.90
xgb	47.32	47.67	47.70	47.64	0.34	0.38	0.31	83.32	83.69	83.69	83.73	0.37	0.37	0.41
xnn	1.52	0.59	0.49	0.38	−0.94	−1.04	−1.14	40.03	59.14	71.82	50.73	19.11	31.79	10.70
2. D99														
cnn	35.88	37.51	35.57	37.84	1.63	−0.31	1.96	74.77	82.07	81.91	81.99	7.30	7.14	7.22
ebm	46.58	48.93	49.37	48.82	2.35	2.79	2.24	83.35	83.98	83.99	83.97	0.63	0.64	0.62
ffn	36.14	34.37	36.19	35.45	−1.77	0.05	−0.69	70.67	76.86	74.86	69.89	6.19	4.18	−0.78
ffnlp	35.57	34.92	35.60	34.61	−0.65	0.04	−0.96	75.97	71.13	68.27	72.44	−4.84	−7.70	−3.53
ffnsp	0.40	11.95	2.33	4.76	11.54	1.93	4.35	41.37	49.08	54.60	43.60	7.71	13.23	2.22
knn	24.20	23.94	23.95	23.97	−0.27	−0.25	−0.23	70.93	70.47	70.47	70.48	−0.46	−0.46	−0.45
lda	22.99	23.78	24.50	24.21	0.79	1.52	1.22	71.70	71.89	72.24	72.02	0.19	0.55	0.32
mlp	29.08	29.66	30.34	30.19	0.58	1.26	1.11	70.32	53.49	67.30	60.32	−16.83	−3.02	−10.00
percep	10.41	11.63	10.31	11.61	1.22	−0.10	1.20	58.58	62.03	47.10	58.86	3.45	−11.48	0.28
qda	25.43	23.97	21.23	23.91	−1.46	−4.20	−1.52	50.97	55.01	56.03	53.98	4.04	5.06	3.01
rf	45.52	45.09	44.30	44.72	−0.43	−1.22	−0.80	82.81	83.42	83.38	83.40	0.60	0.57	0.59
ridge	22.98	23.72	24.47	24.22	0.74	1.49	1.23	32.91	33.41	34.19	33.87	0.51	1.28	0.96
xgb	47.13	47.66	47.73	47.79	0.53	0.61	0.67	83.48	83.82	83.77	83.79	0.34	0.28	0.31
xnn	1.02	0.00	0.00	0.03	−1.02	−1.01	−0.98	74.61	74.30	82.55	66.02	−0.32	7.94	−8.59
3. D80														
cnn	35.73	36.31	37.13	36.87	0.58	1.40	1.15	76.63	81.99	82.32	82.06	5.36	5.70	5.43
ebm	45.55	47.97	48.30	47.92	2.43	2.75	2.38	83.62	84.43	84.45	84.42	0.81	0.83	0.81
ffn	33.41	35.00	34.05	34.86	1.59	0.64	1.45	77.49	75.95	76.77	75.96	−1.53	−0.71	−1.53
ffnlp	33.38	34.87	35.53	35.39	1.48	2.14	2.00	69.61	75.81	76.76	72.69	6.20	7.15	3.08
ffnsp	0.01	4.62	0.21	3.16	4.60	0.20	3.15	58.32	52.39	66.66	32.04	−5.94	8.33	−26.28
knn	23.42	23.39	23.35	23.40	−0.03	−0.06	−0.02	70.84	70.82	70.81	70.83	−0.02	−0.03	−0.01
lda	22.59	23.27	23.75	23.42	0.68	1.16	0.83	72.00	72.19	72.43	72.37	0.19	0.42	0.36
mlp	30.42	32.02	29.35	31.94	1.60	−1.07	1.52	64.56	81.48	59.48	77.58	16.92	−5.09	13.01
percep	10.39	11.01	13.18	10.16	0.63	2.79	−0.23	50.31	58.93	54.18	50.57	8.61	3.87	0.26
qda	24.95	22.50	18.76	22.76	−2.45	−6.18	−2.19	52.42	56.95	54.06	55.54	4.53	1.64	3.12
rf	44.14	44.19	43.50	43.92	0.05	−0.64	−0.22	83.05	83.90	83.91	83.91	0.85	0.86	0.87
ridge	22.57	23.26	23.73	23.41	0.68	1.16	0.83	33.39	34.27	34.83	34.40	0.88	1.44	1.01
xgb	45.79	46.94	46.78	47.09	1.15	0.99	1.30	83.57	84.22	84.22	84.23	0.66	0.65	0.66
xnn	0.74	0.47	0.97	0.51	−0.27	0.23	−0.23	66.67	59.76	70.13	67.23	−6.91	3.46	0.56

Note: This table reports the average KS and F1 across ten folds for binary credit classification (good vs. bad payers) in datasets enriched with COVID-19-related EF but excluding credits with government-backed benefits (WOGB_CR). Notable observations include the following: 1. Models such as *EBM* and *XGB* maintain robust predictive capabilities, achieving KS values above 45%, with minimal performance fluctuation in the absence of government-benefit data with EF. 2. Both *CNN* and *MLP* exhibit marked improvements in F1, with the *CNN* model reaching gains of up to 8.03 percentage points in certain scenarios, thereby underscoring the impact of EF enrichment on classification accuracy. 3. The exclusion of government-benefit information adversely affects some models (e.g., *RF*, *QDA*, *XNN*), which register declines in both KS and F1. This variability suggests that these models are more sensitive to the absence of such interventions. 4. Overall, these findings reveal that omitting data on governmental support introduces notable variability and potential biases in the models' predictive performance—while some models capitalize on reduced bias, others suffer from diminished input granularity. Collectively, these results illuminate the nuanced effects of excluding government-benefit data on credit risk classification and emphasize the integral role that such interventions play in mitigating bias and enhancing model accuracy.

Table A5. Average AUC and ACC across 10 folds of ML models with COVID MOV EF WGB_CR, by evaluated scenario and relevance groups.

	AUC (%)							ACC (%)						
	Scenarios				Results			Scenarios				Results		
	SFE	P	F + P	F	P - SFE	F + P - SFE	F - SFE	SFE	P	F + P	F	P - SFE	F + P - SFE	F - SFE
WGB_CR														
1. D100														
cnn	72.63	73.34	74.25	74.04	0.71	1.61	1.41	66.85	69.80	70.36	69.98	2.95	3.51	3.13
ebm	81.13	81.23	81.40	81.27	0.10	0.28	0.14	74.84	74.94	74.93	74.95	0.09	0.09	0.10
ffn	71.73	72.79	73.11	73.02	1.05	1.37	1.28	64.01	67.56	65.13	65.75	3.55	1.12	1.73
ffnlp	71.58	71.96	72.83	72.17	0.38	1.25	0.58	64.96	65.44	65.43	66.57	0.48	0.47	1.61
ffnsp	72.35	72.17	73.07	72.46	−0.18	0.72	0.10	64.93	66.32	65.79	66.45	1.39	0.86	1.52
knn	67.21	66.92	66.85	67.03	−0.29	−0.37	−0.18	63.23	63.00	62.92	63.03	−0.22	−0.30	−0.20
lda	67.23	68.01	68.56	68.27	0.78	1.33	1.04	63.09	63.74	64.12	63.85	0.65	1.03	0.76
mlp	69.30	70.83	71.62	70.01	1.53	2.31	0.71	60.82	63.26	62.93	60.51	2.44	2.12	−0.31
percep	57.96	58.44	58.16	56.56	0.48	0.20	−1.40	44.58	57.99	47.08	54.03	13.41	2.50	9.45
qda	67.44	66.98	65.86	67.20	−0.46	−1.58	−0.24	52.50	54.01	54.08	53.82	1.52	1.58	1.32
rf	81.74	81.51	81.20	81.46	−0.23	−0.54	−0.28	75.51	75.44	75.29	75.47	−0.07	−0.22	−0.04
ridge	66.94	67.55	68.07	67.83	0.62	1.14	0.89	62.93	63.39	63.78	63.55	0.46	0.85	0.62
xgb	80.76	81.00	81.05	81.09	0.23	0.28	0.33	74.77	74.98	75.00	75.05	0.21	0.23	0.28
xnn	50.21	50.42	50.69	49.86	0.21	0.47	−0.35	53.52	64.62	58.11	62.32	11.09	4.59	8.79
2. D99														
cnn	72.80	73.40	74.06	73.79	0.60	1.25	0.98	68.60	70.07	69.79	70.22	1.48	1.19	1.63
ebm	81.04	81.24	81.40	81.27	0.19	0.36	0.23	74.77	74.93	75.01	74.95	0.16	0.24	0.19
ffn	72.82	72.50	72.90	72.94	−0.33	0.08	0.11	65.07	67.39	65.81	67.26	2.31	0.74	2.19
ffnlp	72.93	72.33	73.50	71.57	−0.60	0.57	−1.35	64.79	66.33	66.36	66.76	1.54	1.56	1.96
ffnsp	71.88	72.73	72.13	72.97	0.85	0.25	1.09	66.23	66.61	65.97	67.17	0.38	−0.26	0.94
knn	67.10	66.88	67.10	66.91	−0.22	0.00	−0.19	63.24	63.02	63.17	63.06	−0.23	−0.08	−0.18
lda	67.17	67.92	68.53	68.22	0.75	1.36	1.05	63.07	63.68	64.12	63.79	0.60	1.05	0.72
mlp	69.79	68.93	69.57	69.17	−0.86	−0.22	−0.62	63.83	64.12	60.77	58.27	0.29	−3.06	−5.56
percep	55.09	57.67	53.79	59.21	2.58	−1.30	4.12	55.48	58.56	55.13	48.25	3.07	−0.36	−7.23
qda	67.32	66.90	65.80	67.15	−0.42	−1.52	−0.16	52.41	53.98	53.89	53.67	1.57	1.49	1.26
rf	81.68	81.33	81.18	81.41	−0.35	−0.50	−0.27	75.63	75.32	75.32	75.40	−0.31	−0.31	−0.23
ridge	66.85	67.47	68.04	67.78	0.62	1.19	0.93	62.94	63.39	63.77	63.49	0.45	0.82	0.55
xgb	80.72	80.96	81.10	81.11	0.24	0.38	0.40	74.81	74.97	75.19	75.00	0.16	0.38	0.19
xnn	49.91	50.11	50.00	49.88	0.20	0.09	−0.02	56.25	62.86	62.66	56.14	6.62	6.42	−0.11
3. D80														
cnn	71.98	73.18	73.19	73.24	1.21	1.21	1.26	66.51	70.27	70.06	70.11	3.75	3.55	3.60
ebm	80.73	81.03	81.17	81.06	0.29	0.44	0.32	74.92	75.16	75.27	75.20	0.24	0.35	0.28
ffn	71.97	71.95	72.62	71.74	−0.02	0.65	−0.23	65.38	64.76	65.96	67.04	−0.62	0.59	1.67
ffnlp	70.31	72.08	72.51	70.73	1.77	2.20	0.42	66.23	67.11	66.03	65.58	0.89	−0.20	−0.65
ffnsp	70.27	71.89	72.16	72.29	1.62	1.89	2.02	66.20	65.43	64.91	65.77	−0.77	−1.29	−0.43
knn	66.57	66.57	66.57	66.57	−0.01	−0.01	−0.00	62.98	62.95	62.95	62.97	−0.02	−0.03	−0.01
lda	67.33	67.90	68.29	68.07	0.57	0.96	0.74	63.45	63.88	64.20	64.10	0.43	0.75	0.65
mlp	67.96	68.85	68.47	67.87	0.89	0.52	−0.08	58.09	63.16	61.09	58.34	5.07	2.99	0.25
percep	57.44	55.39	56.68	57.06	−2.05	−0.76	−0.38	52.88	57.22	59.57	57.33	4.35	6.69	4.45
qda	67.69	66.51	64.09	66.23	−1.18	−3.60	−1.46	52.76	54.28	52.61	54.33	1.52	−0.15	1.57
rf	81.18	81.03	80.64	80.99	−0.15	−0.55	−0.20	75.47	75.57	75.25	75.52	0.11	−0.21	0.05
ridge	66.67	67.23	67.59	67.40	0.56	0.92	0.73	63.06	63.35	63.74	63.54	0.29	0.68	0.48
xgb	80.20	80.75	80.77	80.80	0.54	0.57	0.60	74.74	75.13	75.27	75.18	0.39	0.52	0.44
xnn	50.41	50.39	50.04	50.27	−0.02	−0.36	−0.14	53.24	65.69	60.33	63.61	12.45	7.09	10.37

Note: This table reports the average AUC and ACC across ten folds for binary credit classification (good vs. bad payers) in datasets enriched with COVID-19-related EF under the COVID MOV scenario, where the time series for fatalities was shifted by one period. Notable observations include the following: 1. Models such as *EBM*, *RF*, and *XGB* demonstrate strong predictive performance, achieving AUC values exceeding 80% across all scenarios, with minimal impact from the addition of EF data. 2. The *CNN* model exhibit consistent improvements in both AUC and ACC, with accuracy gains of up to 3.75 percentage points, highlighting its robustness to temporal adjustments in EF data. 3. Models such as *PERCEP* and *XNN* show significant sensitivity to the temporal shift, with marked fluctuations in both AUC and ACC. These results suggest potential challenges in handling shifted time-series data, possibly because of their reliance on specific temporal patterns. 4. Among the classical algorithms, *LDA* and *RIDGE* show modest but consistent improvements across scenarios, underscoring their adaptability to enriched datasets with temporal modifications. 5. Overall, the COVID MOV experiment underscores the critical influence of temporal changes in external factors. This experiment demonstrates that variables derived from the same EF but influenced by their temporal dynamics must be carefully considered in the experimental design to optimize the predictive performance of ML models. These findings highlight the nuanced relationship between temporal adjustments in EF data and model performance, emphasizing the importance of robust experimental frameworks for mitigating variability and maximizing predictive accuracy.

Table A6. Average AUC and ACC across 10 folds of ML models with COVID MOV EF WOGB_CR, by evaluated scenario and relevance groups.

	AUC (%)							ACC (%)						
	Scenarios				Results			Scenarios				Results		
	SFE	P	F + P	F	P - SFE	F + P - SFE	F - SFE	SFE	P	F + P	F	P - SFE	F + P - SFE	F - SFE
WOGB_CR														
1. D100														
cnn	73.31	74.18	74.61	74.10	0.87	1.30	0.80	69.26	72.83	73.05	72.78	3.57	3.80	3.52
ebm	81.38	81.29	81.49	81.40	−0.09	0.11	0.02	76.54	76.41	76.66	76.52	−0.13	0.12	−0.02
ffn	71.66	73.08	73.45	73.19	1.42	1.79	1.53	65.08	69.34	66.12	62.94	4.26	1.03	−2.14
ffnlp	73.32	73.03	73.66	71.79	−0.29	0.34	−1.53	68.13	70.61	66.03	63.72	2.48	−2.09	−4.41
ffnsp	73.51	72.10	73.20	72.72	−1.41	−0.30	−0.79	68.47	66.44	64.67	67.81	−2.03	−3.80	−0.66
knn	66.21	65.98	65.89	65.90	−0.23	−0.32	−0.31	63.19	62.84	62.65	62.70	−0.35	−0.54	−0.48
lda	66.08	66.68	67.18	66.95	0.60	1.10	0.87	63.18	63.64	64.00	63.70	0.46	0.82	0.52
mlp	70.92	70.07	70.02	69.70	−0.85	−0.90	−1.22	60.20	67.96	61.74	61.31	7.76	1.54	1.10
percep	55.94	54.88	55.19	57.55	−1.06	−0.75	1.61	54.77	59.56	50.09	56.79	4.78	−4.68	2.02
qda	66.82	65.97	64.74	66.34	−0.84	−2.08	−0.48	50.72	52.66	52.80	52.24	1.94	2.08	1.51
rf	79.53	79.09	78.49	79.01	−0.44	−1.03	−0.52	75.34	75.34	74.93	75.34	0.01	−0.40	0.00
ridge	65.95	66.44	66.90	66.71	0.50	0.95	0.77	63.14	63.44	63.79	63.56	0.30	0.65	0.42
xgb	80.43	80.60	80.65	80.70	0.17	0.22	0.26	76.00	76.16	76.18	76.34	0.16	0.18	0.34
xnn	49.96	50.00	49.85	49.92	0.04	−0.11	−0.05	58.19	66.17	65.91	57.93	7.98	7.72	−0.27
2. D99														
cnn	73.06	74.03	74.62	74.77	0.97	1.56	1.71	68.66	72.76	73.12	73.24	4.10	4.47	4.58
ebm	81.23	81.31	81.45	81.34	0.08	0.22	0.11	76.31	76.39	76.51	76.45	0.07	0.20	0.14
ffn	71.40	73.31	73.20	73.08	1.91	1.80	1.68	65.45	67.42	64.39	65.51	1.98	−1.05	0.06
ffnlp	72.79	72.79	73.32	72.59	0.01	0.53	−0.20	67.15	67.42	66.71	66.48	0.27	−0.44	−0.66
ffnsp	71.94	71.95	73.53	71.60	0.01	1.59	−0.33	60.05	64.21	66.13	65.70	4.17	6.08	5.65
knn	65.94	66.00	65.81	65.69	0.06	−0.13	−0.25	62.92	62.58	62.67	62.59	−0.35	−0.25	−0.33
lda	66.00	66.67	67.11	66.88	0.67	1.11	0.88	63.15	63.63	64.00	63.68	0.48	0.85	0.53
mlp	68.79	70.06	69.61	69.62	1.27	0.82	0.83	70.57	64.20	58.61	54.19	−6.37	−11.96	−16.38
percep	55.28	56.66	56.65	56.40	1.37	1.37	1.12	54.45	56.49	57.69	60.59	2.04	3.24	6.14
qda	66.72	65.95	64.75	66.15	−0.77	−1.97	−0.57	50.57	52.65	52.60	51.89	2.08	2.03	1.32
rf	79.41	78.96	78.61	78.90	−0.46	−0.80	−0.51	75.31	75.19	74.94	75.23	−0.13	−0.37	−0.08
ridge	65.85	66.42	66.82	66.64	0.57	0.97	0.79	63.13	63.45	63.82	63.49	0.32	0.69	0.35
xgb	80.35	80.66	80.61	80.56	0.32	0.26	0.22	76.13	76.28	76.15	76.20	0.15	0.02	0.07
xnn	49.89	49.89	50.00	50.22	0.00	0.12	0.33	53.92	70.09	66.22	66.57	16.17	12.30	12.65
3. D80														
cnn	72.58	73.36	74.04	73.73	0.78	1.46	1.15	66.42	72.66	73.58	73.19	6.25	7.16	6.78
ebm	80.75	81.05	81.11	81.06	0.30	0.36	0.31	76.37	76.72	76.73	76.73	0.35	0.36	0.36
ffn	72.47	72.15	73.07	71.92	−0.32	0.59	−0.55	64.08	67.19	65.91	68.57	3.11	1.83	4.49
ffnlp	70.26	72.56	73.17	71.30	2.31	2.92	1.04	61.75	67.37	66.64	63.25	5.62	4.89	1.50
ffnsp	71.99	72.70	72.66	72.22	0.72	0.67	0.23	69.54	67.24	68.99	67.02	−2.30	−0.56	−2.52
knn	65.37	65.36	65.36	65.37	−0.01	−0.01	−0.00	62.54	62.52	62.51	62.54	−0.02	−0.03	0.01
lda	65.85	66.35	66.70	66.51	0.50	0.85	0.66	63.27	63.64	63.94	63.75	0.37	0.66	0.47
mlp	69.10	71.07	70.63	71.45	1.96	1.53	2.34	62.97	59.38	60.65	65.13	−3.59	−2.33	2.16
percep	55.71	56.80	57.65	59.30	1.09	1.94	3.59	51.81	55.32	50.69	58.55	3.51	−1.12	6.74
qda	66.65	65.29	63.23	65.24	−1.36	−3.42	−1.41	50.91	53.12	51.39	52.64	2.21	0.49	1.73
rf	78.71	78.53	78.08	78.54	−0.18	−0.63	−0.17	75.29	75.54	75.36	75.57	0.25	0.08	0.28
ridge	65.48	65.97	66.26	66.12	0.48	0.78	0.63	63.14	63.37	63.63	63.57	0.23	0.49	0.43
xgb	79.57	80.16	80.21	80.22	0.59	0.64	0.65	75.92	76.45	76.51	76.39	0.54	0.59	0.47
xnn	50.38	50.23	50.56	50.42	−0.15	0.18	0.04	67.98	63.92	69.19	59.87	−4.06	1.20	−8.11

Note: This table reports the average AUC and ACC across ten folds for binary credit classification (good vs. bad payers) in datasets enriched with COVID-19-related EF under the COVID MOV scenario, where the time series for fatalities was shifted by one period and government-backed credits were excluded. Notable observations include the following: 1. Models such as *EBM*, *RF*, and *XGB* maintain consistent predictive performance, achieving AUC values above 80% across scenarios, with minimal impact from the exclusion of government-backed credits. 2. The *CNN* model demonstrates substantial improvements in ACC, with gains of up to 4.58 percentage points, further emphasizing its robustness in adapting to temporal adjustments and reduced data granularity. 3. Neural network-based models such as *MLP* and *XNN* exhibit contrasting behaviors; while *MLP* shows significant declines in both AUC and ACC in certain scenarios, *XNN* achieves notable gains in accuracy (up to 16.17 percentage points) under specific conditions. 4. Classical algorithms such as *LDA* and *RIDGE* consistently show moderate improvements, reflecting their adaptability to datasets that exclude external financial support. 5. Models such as *FFNLP* and *FFNSP* are sensitive to the removal of government-backed credits, with performance fluctuations in both AUC and ACC, underscoring the importance of these credits in predictive accuracy for some architectures. 6. Overall, the results reveal that excluding government-backed credits introduces variability in predictive performance, with tree-based models (*RF*, *XGB*) and the *CNN* model demonstrating greater resilience, whereas some neural network models face challenges adapting to the reduced dataset complexity. These findings highlight the nuanced effects of excluding government-backed credits on model performance and demonstrate the introduced bias. They also highlighted the need for robust EF and scenario-specific tuning to optimize the predictive accuracy.

Table A7. Average KS and F1 across 10 folds of ML models with COVID MOV EF WGB_CR, by evaluated scenario and relevance groups.

	KS (%)							F1 (%)						
	Scenarios				Results			Scenarios				Results		
	SFE	**P**	**F + P**	**F**	**P - SFE**	**F + P - SFE**	**F - SFE**	**SFE**	**P**	**F + P**	**F**	**P - SFE**	**F + P - SFE**	**F - SFE**
WGB_CR														
1. D100														
cnn	35.93	36.40	37.93	37.61	0.48	2.01	1.68	73.50	78.41	78.74	78.25	4.91	5.24	4.75
ebm	46.69	48.66	48.90	48.67	1.97	2.21	1.98	80.49	81.36	81.44	81.31	0.86	0.94	0.82
ffn	34.69	34.08	35.42	34.02	−0.61	0.73	−0.67	71.62	71.20	71.96	71.10	−0.42	0.34	−0.53
ffnlp	35.10	33.58	35.28	35.56	−1.52	0.17	0.46	70.62	70.73	69.83	70.58	0.11	−0.79	−0.03
ffnsp	4.95	7.36	5.47	4.80	2.41	0.52	−0.15	44.43	35.67	52.84	43.98	−8.76	8.41	−0.45
knn	26.04	25.53	25.38	25.68	−0.51	−0.65	−0.36	69.55	69.31	69.23	69.29	−0.24	−0.31	−0.26
lda	24.32	25.24	26.33	25.91	0.93	2.01	1.59	69.92	70.25	70.64	70.39	0.33	0.72	0.47
mlp	30.67	31.76	30.35	30.77	1.09	−0.32	0.09	65.29	69.17	64.42	68.99	3.88	−0.87	3.70
percep	13.14	13.64	13.12	12.61	0.50	−0.02	−0.53	29.26	62.04	36.15	53.45	32.78	6.89	24.19
qda	26.02	25.55	23.43	25.72	−0.47	−2.59	−0.30	49.55	53.09	54.08	52.18	3.54	4.53	2.63
rf	49.27	48.88	48.08	48.57	−0.39	−1.18	−0.70	81.70	82.11	82.13	82.10	0.41	0.43	0.40
ridge	24.33	25.22	26.32	25.89	0.88	1.99	1.56	30.27	30.94	31.76	31.65	0.68	1.49	1.38
xgb	47.88	48.21	48.11	48.15	0.33	0.23	0.27	81.03	81.45	81.47	81.48	0.43	0.44	0.46
xnn	0.03	0.64	0.02	0.04	0.61	−0.01	0.01	63.60	73.12	47.59	63.54	9.52	−16.01	−0.06
2. D99														
cnn	35.86	36.35	37.81	37.22	0.49	1.96	1.37	72.23	78.07	78.27	78.14	5.84	6.04	5.91
ebm	46.33	48.39	48.82	48.14	2.06	2.49	1.81	80.67	81.33	81.50	81.39	0.66	0.82	0.72
ffn	33.49	34.50	35.56	35.31	1.01	2.07	1.82	66.82	72.08	70.01	72.45	5.26	3.19	5.63
ffnlp	34.71	32.82	35.25	34.09	−1.89	0.55	−0.62	70.70	73.14	71.99	69.72	2.44	1.29	−0.98
ffnsp	2.49	7.29	2.46	3.07	4.81	−0.03	0.59	30.01	35.25	36.73	38.74	5.25	6.72	8.73
knn	25.85	25.61	25.90	25.66	−0.24	0.06	−0.19	69.70	69.32	69.46	69.38	−0.38	−0.24	−0.32
lda	24.24	25.17	26.26	25.78	0.93	2.01	1.54	70.00	70.30	70.67	70.38	0.30	0.67	0.38
mlp	28.44	29.26	32.14	29.49	0.81	3.70	1.04	66.98	68.22	71.81	63.57	1.24	4.83	−3.42
percep	9.56	12.65	9.52	14.94	3.09	−0.04	5.38	55.60	64.38	56.35	38.09	8.78	0.75	−17.51
qda	25.86	25.52	23.23	25.59	−0.34	−2.63	−0.27	49.60	53.06	53.72	52.02	3.46	4.12	2.42
rf	48.94	48.29	48.05	48.64	−0.65	−0.89	−0.30	81.71	82.06	82.13	82.15	0.35	0.42	0.45
ridge	24.22	25.17	26.27	25.78	0.96	2.05	1.56	30.26	31.13	31.76	31.65	0.87	1.50	1.40
xgb	47.78	48.04	48.24	48.43	0.25	0.45	0.65	81.11	81.47	81.65	81.49	0.36	0.54	0.38
xnn	0.23	0.10	0.85	0.97	−0.13	0.61	0.74	79.32	47.91	76.80	75.84	−31.41	−2.52	−3.48
3. D80														
cnn	34.94	36.15	36.87	36.25	1.21	1.93	1.31	69.39	78.62	79.17	79.28	9.23	9.77	9.88
ebm	45.97	48.05	48.30	48.01	2.08	2.34	2.05	81.06	81.94	82.06	81.93	0.89	1.00	0.88
ffn	33.87	30.69	33.29	33.45	−3.18	−0.59	−0.42	70.98	70.98	62.95	68.35	0.00	−8.03	−2.63
ffnlp	32.01	33.20	33.74	33.50	1.19	1.72	1.49	72.95	71.81	73.45	69.45	−1.14	0.50	−3.51
ffnsp	10.70	10.02	7.70	4.07	−0.68	−3.00	−6.63	51.26	64.39	48.97	55.36	13.13	−2.29	4.10
knn	25.09	25.07	25.07	25.08	−0.02	−0.02	−0.01	69.81	69.78	69.77	69.80	−0.03	−0.04	−0.01
lda	24.18	24.99	25.57	25.43	0.82	1.40	1.26	70.49	70.76	71.09	70.90	0.27	0.60	0.41
mlp	29.32	28.12	30.89	31.28	−1.21	1.57	1.95	66.46	75.92	69.76	62.11	9.46	3.30	−4.35
percep	12.46	9.77	10.97	11.87	−2.69	−1.49	−0.58	49.76	59.25	64.82	61.34	9.49	15.06	11.58
qda	26.25	24.52	20.48	24.00	−1.73	−5.78	−2.25	51.24	55.03	53.41	55.55	3.80	2.18	4.31
rf	48.16	48.21	47.35	47.91	0.06	−0.81	−0.25	81.87	82.66	82.62	82.64	0.79	0.75	0.76
ridge	24.17	24.98	25.55	25.42	0.81	1.38	1.25	31.33	32.18	32.89	32.31	0.85	1.56	0.98
xgb	47.03	47.83	47.72	47.71	0.80	0.68	0.68	81.40	82.02	82.14	82.06	0.62	0.74	0.66
xnn	1.79	0.00	1.82	0.45	−1.79	0.03	−1.34	73.57	80.31	78.82	72.13	6.75	5.26	−1.43

Note: This table reports the average KS and F1 across ten folds for binary credit classification (good vs. bad payers) in datasets enriched with COVID-19-related EF under the COVID MOV scenario, where the time series for fatalities was shifted by one period. Notable observations include the following: 1. Models such as *EBM*, *RF*, and *XGB* demonstrate strong predictive performance, achieving KS values exceeding 45% and F1 scores above 80% across most scenarios, with minimal impact from the addition of EF data. 2. The *CNN* model exhibits consistent improvements in both KS and F1, with F1 score gains of up to 9.88 percentage points, highlighting its robustness to temporal adjustments in EF data. 3. Models such as *PERCEP* and *XNN* show significant sensitivity to the temporal shift, with marked fluctuations in both KS and F1. These results suggest potential challenges in handling shifted time-series data, possibly because of their reliance on specific temporal patterns. 4. Among the classical algorithms, *LDA* and *RIDGE* show modest but consistent improvements across scenarios, underscoring their adaptability to enriched datasets with temporal modifications. 5. The *FFNSP* model displays extreme variability in performance, particularly in the F1 scores, ranging from 4.07% to 64.39% across different scenarios. 6. Overall, the COVID MOV experiment underscores the critical influence of temporal changes in external factors. This experiment demonstrates that variables derived from the same EF but influenced by their temporal dynamics must be carefully considered in the experimental design to optimize the predictive performance of ML models. These findings highlight the nuanced relationship between temporal adjustments in EF data and model performance, emphasizing the importance of robust experimental frameworks for mitigating variability and maximizing predictive accuracy.

Table A8. Average KS and F1 across 10 folds of ML models with COVID MOV EF WOGB_CR, by evaluated scenario and relevance groups.

	KS (%)							F1 (%)						
	Scenarios				Results			Scenarios				Results		
	SFE	**P**	**F + P**	**F**	**P - SFE**	**F + P - SFE**	**F - SFE**	**SFE**	**P**	**F + P**	**F**	**P - SFE**	**F + P - SFE**	**F - SFE**
WOGB_CR														
1. D100														
cnn	36.11	37.36	38.52	37.33	1.25	2.42	1.23	75.97	82.01	82.19	81.77	6.05	6.22	5.80
ebm	47.01	48.82	49.10	48.98	1.82	2.10	1.97	83.33	83.96	84.02	84.02	0.63	0.68	0.69
ffn	35.81	33.75	35.86	35.51	−2.06	0.05	−0.30	75.18	71.71	74.16	74.06	−3.47	−1.02	−1.13
ffnlp	35.84	34.88	34.53	35.82	−0.95	−1.31	−0.02	72.43	73.53	75.64	67.77	1.10	3.21	−4.66
ffnsp	4.66	7.41	0.51	0.00	2.75	−4.16	−4.66	53.29	51.71	41.25	57.73	−1.58	−12.04	4.44
knn	24.89	24.34	23.91	23.98	−0.55	−0.98	−0.91	71.07	70.69	70.54	70.62	−0.38	−0.54	−0.46
lda	22.95	23.93	24.61	24.24	0.98	1.65	1.28	71.66	71.87	72.20	71.98	0.21	0.54	0.31
mlp	29.53	30.31	32.02	28.57	0.78	2.48	−0.96	76.57	76.42	69.61	62.90	−0.15	−6.96	−13.66
percep	10.32	9.32	9.54	12.82	−1.00	−0.78	2.50	55.33	65.60	46.28	61.28	10.27	−9.05	5.95
qda	25.55	23.93	21.35	24.43	−1.62	−4.20	−1.12	51.06	54.99	55.98	53.96	3.92	4.92	2.89
rf	45.86	45.21	43.93	44.97	−0.65	−1.93	−0.89	82.69	83.42	83.36	83.39	0.73	0.67	0.70
ridge	22.96	23.93	24.58	24.21	0.97	1.62	1.25	32.93	33.47	34.10	33.69	0.54	1.16	0.76
xgb	47.32	47.67	47.64	47.48	0.34	0.32	0.16	83.32	83.69	83.73	83.82	0.37	0.40	0.50
xnn	1.02	0.02	0.00	0.17	−1.00	−1.02	−0.85	43.91	66.01	74.25	65.97	22.10	30.34	22.06
2. D99														
cnn	36.49	36.46	37.86	37.35	−0.03	1.37	0.85	74.69	81.74	82.01	81.96	7.06	7.32	7.27
ebm	46.58	48.93	49.10	48.74	2.35	2.52	2.16	83.35	83.98	84.03	84.01	0.63	0.68	0.66
ffn	35.46	35.01	35.52	35.71	−0.45	0.07	0.26	73.04	68.68	75.97	75.23	−4.36	2.93	2.19
ffnlp	32.08	34.44	36.33	33.57	2.36	4.25	1.49	75.47	69.04	73.54	73.22	−6.43	−1.93	−2.25
ffnsp	0.51	2.38	4.64	4.67	1.87	4.12	4.15	41.37	37.07	35.62	45.06	−4.31	−5.76	3.68
knn	24.20	23.94	23.87	23.68	−0.27	−0.33	−0.52	70.93	70.47	70.60	70.55	−0.46	−0.32	−0.37
lda	22.99	23.78	24.43	24.27	0.79	1.44	1.28	71.70	71.89	72.26	71.94	0.19	0.56	0.24
mlp	29.08	29.66	30.50	30.28	0.58	1.43	1.20	70.32	53.49	68.32	73.70	−16.83	−1.99	3.39
percep	10.41	11.63	12.06	10.78	1.22	1.65	0.37	58.58	62.03	61.89	66.68	3.45	3.31	8.10
qda	25.43	23.97	21.15	24.20	−1.46	−4.28	−1.24	50.97	55.01	55.74	53.51	4.04	4.77	2.54
rf	45.52	45.09	44.23	45.05	−0.43	−1.29	−0.47	82.81	83.42	83.36	83.35	0.60	0.55	0.54
ridge	22.98	23.72	24.43	24.25	0.74	1.44	1.26	32.91	33.41	34.25	33.94	0.51	1.35	1.03
xgb	47.13	47.66	47.74	47.50	0.53	0.62	0.37	83.48	83.82	83.74	83.75	0.34	0.26	0.27
xnn	0.82	0.00	0.12	0.00	−0.81	−0.69	−0.82	66.48	74.30	82.47	82.53	7.82	15.99	16.05
3. D80														
cnn	35.50	36.04	36.84	36.65	0.53	1.34	1.15	73.13	82.62	82.55	81.95	9.49	9.42	8.83
ebm	45.55	47.97	48.27	47.98	2.43	2.72	2.43	83.62	84.43	84.46	84.40	0.81	0.84	0.79
ffn	34.42	34.08	34.11	34.17	−0.34	−0.31	−0.25	76.33	69.56	74.97	67.61	−6.76	−1.35	−8.71
ffnlp	33.67	32.17	35.10	34.43	−1.49	1.43	0.76	71.06	72.67	70.85	74.39	1.60	−0.21	3.32
ffnsp	2.26	0.01	9.65	5.17	−2.25	7.40	2.91	37.70	16.66	50.71	55.02	−21.04	13.01	17.31
knn	23.42	23.39	23.37	23.41	−0.03	−0.05	−0.01	70.84	70.82	70.81	70.84	−0.02	−0.03	0.00
lda	22.59	23.27	23.73	23.45	0.68	1.14	0.86	72.00	72.19	72.41	72.40	0.19	0.41	0.40
mlp	30.42	32.02	31.34	29.19	1.60	0.92	−1.23	64.56	81.48	61.72	71.11	16.92	−2.84	6.55
percep	10.39	11.01	12.65	14.68	0.63	2.26	4.29	50.31	58.93	50.53	63.93	8.61	0.22	13.62
qda	24.95	22.50	19.05	22.58	−2.45	−5.90	−2.37	52.42	56.95	55.20	56.38	4.53	2.77	3.96
rf	44.14	44.19	43.62	43.98	0.05	−0.52	−0.16	83.05	83.90	83.85	83.86	0.85	0.80	0.82
ridge	22.57	23.26	23.74	23.45	0.68	1.17	0.87	33.39	34.27	34.78	34.40	0.88	1.40	1.01
xgb	45.79	46.94	47.07	47.15	1.15	1.28	1.36	83.57	84.22	84.28	84.18	0.66	0.71	0.62
xnn	0.25	0.00	0.01	0.05	−0.25	−0.25	−0.20	66.50	74.99	66.67	66.62	8.49	0.18	0.12

Note: This table presents the average KS and F1 scores across ten folds for binary credit classification (good vs. bad payers) in datasets without group balancing (WOGB_CR) under the COVID MOV scenario, where the time series for fatalities was shifted by one period. Key findings include the following: 1. Models like *EBM*, *RF*, and *XGB* maintain robust performance, with KS values consistently above 45% and F1 scores exceeding 83% across all scenarios. The temporal shift in EF data has minimal impact on these models, demonstrating their stability. 2. The *FFNSP* model exhibit extreme performance fluctuations, particularly in D80, with KS ranging from 0.01% to 9.65% and F1 from 16.66% to 55.02%. This suggests sensitivity to both temporal shifts and dataset characteristics. 3. The *XNN* model shows significant variability in D100 and D99, with F1 scores ranging from 43.91% to 74.25% but stabilizing in D80. This indicates that its performance is highly dependent on the temporal alignment of EF data. 4. The *LDA* and *RIDGE* exhibits modest but consistent improvements, with F1 scores increasing by up to 1.40 percentage points when EF data are included, highlighting their reliability in unbalanced datasets. 5. The *MLP* model shows significant declines in D99 (F1 drop of 16.83 percentage points in P - SFE), suggesting potential overfitting or sensitivity to specific temporal configurations. 6. The WOGB_CR results reveal more pronounced variability compared to WGB_CR, particularly for models like *PERCEP* and *FFNSP*, underscoring the challenges of unbalanced datasets in temporal shift scenarios. These observations emphasize the importance of dataset balance and temporal alignment in model performance, particularly for algorithms sensitive to data distribution shifts. The results advocate for careful model selection and scenario testing in credit risk applications involving temporal EF adjustments.

Table A9. Average AUC across 10 folds of ML models with SOCIAL UNREST and CLIMATE CHANGE EF WGB_CR, by evaluated scenario and relevance groups.

AUC (%)															
	Scenarios								Results						
	SFE	**T**	**T + W**	**U**	**U + T**	**U + T + W**	**U + W**	**W**	**T - SFE**	**T + W - SFE**	**U - SFE**	**U + T - SFE**	**U + T + W - SFE**	**U + W - SFE**	**W - SFE**
WGB_CR															
1. D100															
cnn	62.05	64.79	59.93	64.11	60.59	62.64	60.90	64.56	2.74	−2.12	2.05	− 1.46	0.59	−1.15	2.51
ebm	75.08	74.99	75.07	75.01	75.06	75.19	75.15	75.03	−0.10	−0.01	−0.07	− 0.02	0.11	0.07	−0.05
ffn	67.42	67.39	67.07	67.33	67.55	67.95	67.59	67.12	−0.03	−0.35	−0.09	0.13	0.53	0.17	−0.30
ffnlp	67.66	67.42	67.95	67.85	67.47	67.89	67.58	68.24	−0.24	0.29	0.20	− 0.19	0.23	−0.08	0.58
ffnsp	67.45	68.17	66.83	68.22	67.21	67.11	68.17	67.53	0.72	−0.62	0.77	− 0.24	− 0.33	0.72	0.08
knn	61.08	61.22	61.24	61.25	61.25	61.24	61.24	61.26	0.14	0.17	0.18	0.17	0.16	0.17	0.18
lda	65.55	65.84	65.92	65.93	65.98	66.03	65.99	65.83	0.30	0.37	0.39	0.43	0.49	0.44	0.28
mlp	65.40	64.90	64.39	63.91	64.54	65.01	64.79	64.58	−0.51	−1.01	−1.49	− 0.86	− 0.39	−0.62	−0.82
percep	54.89	56.45	55.82	53.65	52.79	55.17	53.03	55.26	1.56	0.93	−1.24	− 2.09	0.28	−1.86	0.37
qda	67.21	67.14	66.58	66.42	66.20	64.98	65.55	67.07	−0.07	−0.63	−0.79	− 1.01	− 2.23	−1.67	−0.15
rf	76.83	77.02	76.67	76.73	76.61	76.32	76.54	76.81	0.19	−0.16	−0.10	− 0.21	− 0.50	−0.28	−0.02
ridge	64.30	64.61	64.71	64.77	64.82	64.89	64.85	64.62	0.30	0.40	0.47	0.52	0.59	0.54	0.32
xgb	74.21	75.31	75.54	75.24	75.36	75.51	75.52	75.47	1.09	1.33	1.03	1.15	1.30	1.31	1.26
xnn	50.27	51.07	49.37	50.71	51.74	52.51	50.07	50.35	0.79	−0.90	0.44	1.47	2.24	−0.20	0.08
2. D99															
cnn	62.51	63.71	61.43	62.88	61.68	61.32	61.35	64.21	1.20	−1.08	0.38	− 0.83	− 1.19	−1.16	1.71
ebm	74.81	74.90	74.99	74.98	75.01	75.08	75.05	74.97	0.09	0.18	0.16	0.19	0.26	0.23	0.16
ffn	68.01	68.42	68.15	67.85	67.56	67.83	67.73	67.59	0.41	0.14	−0.17	− 0.45	− 0.18	−0.29	−0.42
ffnlp	67.88	67.35	67.83	66.78	68.27	68.40	67.34	68.33	−0.53	−0.05	−1.10	0.39	0.52	−0.54	0.45
ffnsp	67.71	66.85	68.13	67.99	67.83	68.08	67.12	67.98	−0.86	0.41	0.27	0.12	0.36	−0.59	0.26
knn	60.92	61.12	61.01	61.03	61.03	61.02	61.02	61.03	0.20	0.10	0.12	0.11	0.10	0.10	0.11
lda	65.54	65.88	65.96	65.96	66.00	66.07	66.03	65.87	0.34	0.42	0.42	0.47	0.53	0.49	0.33
mlp	64.63	63.56	64.81	64.26	65.57	66.02	62.23	64.00	−1.07	0.18	−0.36	0.94	1.39	−2.40	−0.63
percep	53.08	57.12	53.78	55.46	56.71	55.10	55.56	56.06	4.04	0.70	2.38	3.62	2.01	2.48	2.98
qda	67.19	67.13	66.62	66.42	66.21	65.01	65.56	67.08	−0.06	−0.57	−0.77	− 0.98	− 2.18	−1.63	−0.11
rf	76.65	76.88	76.85	76.91	76.75	76.51	76.76	76.89	0.23	0.20	0.26	0.09	− 0.14	0.11	0.24
ridge	64.30	64.62	64.73	64.78	64.83	64.91	64.87	64.64	0.32	0.43	0.47	0.52	0.61	0.56	0.34
xgb	74.21	75.16	75.63	75.43	75.43	75.51	75.37	75.36	0.95	1.42	1.23	1.22	1.30	1.17	1.16
xnn	50.78	51.49	51.62	50.00	50.70	50.88	50.32	50.51	0.72	0.85	−0.77	− 0.08	0.11	−0.45	−0.27
3. D80															
cnn	62.23	63.29	59.66	64.72	61.58	61.98	61.48	63.53	1.05	−2.57	2.49	− 0.66	− 0.25	−0.75	1.30
ebm	74.70	74.90	75.10	75.02	75.05	75.15	75.17	75.03	0.19	0.40	0.31	0.35	0.45	0.46	0.33
ffn	66.87	66.35	67.69	68.23	67.42	67.73	68.18	67.09	−0.53	0.82	1.36	0.55	0.86	1.31	0.22
ffnlp	68.52	68.06	67.98	67.68	68.10	67.98	67.17	67.56	−0.46	−0.54	−0.83	− 0.42	− 0.53	−1.34	−0.95
ffnsp	67.74	67.31	67.34	67.70	67.49	68.52	67.82	67.37	−0.43	−0.40	−0.04	− 0.25	0.78	0.08	−0.37
knn	61.05	61.04	61.03	61.04	61.04	61.04	61.05	61.05	−0.01	−0.02	−0.01	− 0.01	− 0.02	−0.00	−0.00
lda	66.18	66.29	66.35	66.39	66.43	66.48	66.42	66.26	0.11	0.17	0.21	0.24	0.29	0.24	0.08
mlp	64.36	63.72	65.60	64.22	64.19	65.11	64.71	65.86	−0.64	1.24	−0.14	− 0.17	0.75	0.35	1.50
percep	49.78	53.42	52.71	51.46	51.65	53.65	53.89	50.50	3.64	2.93	1.68	1.87	3.87	4.11	0.72
qda	67.38	67.27	66.78	66.71	66.54	65.54	66.16	67.33	−0.10	−0.59	−0.66	− 0.84	− 1.84	−1.22	−0.05
rf	76.50	76.80	76.75	76.83	76.79	76.44	76.60	76.85	0.30	0.25	0.34	0.29	− 0.06	0.10	0.36
ridge	64.89	65.01	65.13	65.23	65.27	65.33	65.27	65.04	0.12	0.23	0.33	0.38	0.44	0.38	0.15
xgb	74.18	75.14	75.64	75.37	75.50	75.59	75.35	75.50	0.96	1.46	1.19	1.32	1.41	1.17	1.32
xnn	51.39	50.00	50.03	50.48	50.00	50.00	50.82	49.76	−1.39	−1.37	−0.92	− 1.40	− 1.39	−0.58	−1.64

Note: This table reports the average AUC across ten folds for binary credit classification (good vs. bad payers) in datasets enriched with EF related to social unrest and climate change under scenarios that include government-backed guarantees (WGB_CR). Key findings include the following: 1. Models such as *EBM*, *RF*, and *XGB* consistently achieve AUC values greater than 75%, showing robust predictive performance. 2. Neural network models, such as *CNN* benefit significantly from EF enrichment, with gains of up to 2.74 percentage points. 3. Classical models, such as *LDA* and *RIDGE* demonstrate resilience to EF variability, with minimal AUC fluctuation. 4. Some models (*MLP*, *XNN*, *QDA*) show sensitivity to EF shifts, highlighting potential challenges with scenario-specific performance. 5. Including WGB_CR data appears to stabilize the performance, with tree-based and neural network models benefiting the most.

Table A10. Average ACC across 10 folds of ML models with SOCIAL UNREST and CLIMATE CHANGE EF WGB_CR, by evaluated scenario and relevance groups.

ACC (%)															
	Scenarios								Results						
	SFE	**T**	**T + W**	**U**	**U + T**	**U + T + W**	**U + W**	**W**	**T - SFE**	**T + W - SFE**	**U - SFE**	**U + T - SFE**	**U + T + W - SFE**	**U + W - SFE**	**W - SFE**
WGB_CR															
1. D100															
cnn	52.61	65.41	84.94	83.75	84.38	85.01	85.11	83.95	12.80	32.32	31.14	31.77	32.40	32.50	31.34
ebm	85.17	85.04	85.00	84.99	85.05	85.00	85.01	84.98	−0.13	−0.16	−0.17	−0.12	− 0.17	−0.16	−0.19
ffn	64.44	69.23	57.39	62.26	59.06	68.61	63.64	66.31	4.79	−7.05	−2.18	−5.38	4.18	−0.80	1.88
ffnlp	67.89	71.13	66.22	62.92	64.40	68.43	67.28	57.41	3.24	−1.67	−4.97	−3.49	0.54	−0.61	−10.48
ffnsp	63.33	65.69	72.22	65.10	62.52	67.30	65.66	64.66	2.36	8.89	1.77	−0.81	3.97	2.33	1.33
knn	64.28	64.32	64.26	64.28	64.27	64.26	64.25	64.26	0.04	−0.02	0.00	−0.01	− 0.02	−0.03	−0.01
lda	66.62	66.80	66.99	66.85	66.94	67.08	67.03	66.84	0.18	0.37	0.22	0.32	0.46	0.40	0.22
mlp	64.75	61.55	65.94	70.37	64.75	62.93	60.47	55.55	−3.20	1.19	5.62	−0.00	− 1.82	−4.28	−9.20
percep	47.76	45.10	45.35	55.49	64.67	46.97	55.80	46.33	−2.66	−2.42	7.73	16.91	− 0.79	8.03	−1.43
qda	63.18	64.96	66.35	64.06	64.26	63.05	63.83	66.24	1.78	3.16	0.88	1.07	− 0.13	0.65	3.06
rf	83.93	85.34	85.66	85.55	85.65	85.66	85.67	85.62	1.41	1.72	1.61	1.72	1.73	1.74	1.68
ridge	65.52	65.74	65.94	65.94	66.01	66.05	66.01	65.97	0.22	0.43	0.43	0.50	0.54	0.49	0.45
xgb	83.75	84.78	84.94	84.92	84.93	84.99	85.06	85.01	1.03	1.19	1.17	1.17	1.24	1.30	1.26
xnn	69.59	71.18	76.49	74.46	77.08	56.18	62.83	74.82	1.59	6.90	4.87	7.49	−13.41	−6.76	5.23
2. D99															
cnn	54.37	61.27	84.98	80.98	83.86	85.05	84.98	83.42	6.90	30.61	26.61	29.48	30.68	30.61	29.05
ebm	85.18	85.01	85.00	84.96	84.99	84.99	84.98	84.98	−0.17	−0.18	−0.22	−0.19	− 0.19	−0.20	−0.20
ffn	64.94	63.78	60.07	65.04	57.60	64.29	62.59	69.11	−1.16	−4.87	0.10	−7.34	− 0.65	−2.35	4.16
ffnlp	63.67	61.15	65.34	67.51	64.26	64.08	66.43	66.05	−2.52	1.67	3.83	0.59	0.41	2.76	2.37
ffnsp	60.19	64.86	62.77	64.39	65.27	69.71	65.17	62.92	4.67	2.59	4.20	5.08	9.52	4.98	2.73
knn	64.33	64.25	64.11	64.14	64.12	64.11	64.12	64.13	−0.09	−0.22	−0.19	−0.21	− 0.22	−0.21	−0.21
lda	66.63	66.85	66.94	66.82	66.87	67.04	66.98	66.83	0.22	0.30	0.19	0.24	0.41	0.35	0.19
mlp	50.94	71.51	70.31	65.50	64.57	64.02	68.44	67.00	20.57	19.37	14.56	13.63	13.08	17.51	16.06
percep	56.28	37.18	58.04	49.47	42.86	45.73	47.04	46.14	−19.10	1.76	−6.80	−13.41	−10.54	−9.24	−10.13
qda	65.63	64.66	66.72	64.47	64.64	63.38	64.14	66.64	−0.97	1.09	−1.17	−0.99	− 2.25	−1.49	1.01
rf	83.77	85.35	85.60	85.43	85.61	85.63	85.61	85.55	1.59	1.84	1.67	1.84	1.87	1.85	1.78
ridge	65.56	65.72	65.92	65.92	66.01	66.06	66.03	65.96	0.16	0.36	0.36	0.45	0.50	0.47	0.40
xgb	83.71	84.77	85.01	84.88	84.97	84.95	84.93	84.98	1.06	1.30	1.18	1.26	1.24	1.22	1.27
xnn	72.46	78.30	75.69	78.10	80.99	73.49	80.06	71.88	5.84	3.22	5.64	8.53	1.03	7.60	−0.59
3. D80															
cnn	55.07	60.32	84.85	82.75	85.06	83.21	85.13	81.13	5.26	29.78	27.68	30.00	28.14	30.06	26.07
ebm	84.92	85.02	85.07	85.02	85.05	85.08	85.07	85.06	0.10	0.15	0.11	0.13	0.16	0.15	0.14
ffn	64.80	66.60	67.44	67.15	68.37	58.27	66.21	65.29	1.80	2.64	2.35	3.57	− 6.53	1.41	0.49
ffnlp	61.26	62.81	61.66	69.24	63.29	63.74	63.71	65.19	1.55	0.40	7.98	2.04	2.48	2.45	3.93
ffnsp	62.56	70.94	71.45	64.26	66.04	61.79	64.64	62.41	8.38	8.89	1.70	3.48	− 0.77	2.08	−0.15
knn	64.38	64.37	64.35	64.37	64.36	64.34	64.35	64.37	−0.01	−0.04	−0.02	−0.02	− 0.04	−0.03	−0.01
lda	66.82	67.06	67.14	67.10	67.16	67.33	67.22	67.02	0.24	0.32	0.27	0.34	0.50	0.40	0.20
mlp	74.21	61.59	58.76	61.62	60.92	59.31	62.27	55.87	−12.62	−15.45	−12.60	−13.30	−14.90	−11.94	−18.34
percep	75.06	53.34	54.05	64.33	70.05	57.29	61.29	65.05	−21.72	−21.01	−10.72	−5.01	−17.77	−13.77	−10.01
qda	64.38	65.24	66.14	62.86	63.23	62.24	62.99	66.21	0.86	1.76	−1.53	−1.16	− 2.15	−1.39	1.83
rf	83.56	85.23	85.57	85.47	85.59	85.58	85.56	85.48	1.67	2.01	1.92	2.04	2.02	2.01	1.93
ridge	65.63	65.95	66.14	66.14	66.19	66.36	66.22	66.13	0.31	0.50	0.50	0.56	0.73	0.59	0.50
xgb	83.57	84.75	85.03	84.90	85.01	84.99	85.04	84.96	1.17	1.46	1.32	1.44	1.42	1.46	1.38
xnn	54.64	78.16	78.20	79.41	85.19	71.12	81.61	61.07	23.52	23.56	24.77	30.55	16.48	26.97	6.43

Note: This table reports the average ACC across ten folds for binary credit classification (good vs. bad payers) in datasets enriched with EF related to social unrest and climate change under scenarios including government-backed guarantees (WGB_CR). Key findings include the following: 1. Tree-based models (*EBM*, *RF*, *XGB*) achieve accuracy values exceeding 84%, demonstrating stability and adaptability. 2. Neural networks such as *CNN* exhibit significant gains, achieving up to 32.50 percentage points improvement in specific scenarios. 3. Classical models (*LDA*, *RIDGE*) show modest but consistent improvements, with *KNN* maintaining a stable performance. 4. The *PERCEP* model displays variability, with gains of up to 8.03% but declines in the others. 5. Scenarios combining social unrest and climate change EFs led to notable gains for specific models (*CNN*, *XNN*). 6. Including WGB_CR enhances stability and consistency, particularly by benefiting neural networks and tree-based models.

Table A11. Average KS across 10 folds of ML models with SOCIAL UNREST and CLIMATE CHANGE EF WGB_CR, by evaluated scenario and relevance groups.

KS (%)															
	Scenarios								Results						
	SFE	**T**	**T + W**	**U**	**U + T**	**U + T + W**	**U + W**	**W**	**T - SFE**	**T + W - SFE**	**U - SFE**	**U + T - SFE**	**U + T + W - SFE**	**U + W - SFE**	**W - SFE**
WGB_CR															
1. D100															
cnn	22.71	22.47	15.19	23.10	17.07	18.76	20.09	22.50	−0.24	−7.51	0.39	−5.64	−3.94	−2.62	−0.21
ebm	34.93	37.07	36.90	36.45	36.76	36.95	36.87	36.67	2.14	1.97	1.51	1.83	2.02	1.94	1.74
ffn	26.42	27.04	25.31	28.07	27.66	27.69	27.85	26.73	0.61	−1.11	1.65	1.24	1.27	1.43	0.31
ffnlp	26.41	25.89	27.33	26.49	26.75	26.83	27.94	27.92	−0.53	0.91	0.08	0.34	0.42	1.52	1.50
ffnsp	17.91	12.90	14.64	13.02	15.15	11.07	11.49	14.58	−5.01	−3.27	−4.89	−2.76	−6.84	−6.42	−3.33
knn	17.11	17.14	17.08	17.09	17.09	17.06	17.08	17.14	0.03	−0.03	−0.01	−0.02	−0.05	−0.02	0.03
lda	22.76	23.29	23.29	23.34	23.47	23.44	23.38	23.25	0.52	0.52	0.58	0.71	0.68	0.62	0.48
mlp	22.73	24.17	21.95	21.93	24.44	23.02	23.96	23.37	1.44	−0.78	−0.80	1.71	0.29	1.23	0.65
percep	13.18	16.46	13.99	12.88	12.22	14.12	10.69	14.19	3.28	0.81	−0.30	−0.96	0.94	−2.49	1.01
qda	27.35	27.49	25.77	26.21	25.75	23.31	24.45	26.94	0.14	−1.58	−1.14	−1.59	−4.04	−2.90	−0.41
rf	40.39	40.76	39.54	39.89	39.88	39.57	39.37	40.18	0.37	−0.85	−0.50	−0.51	−0.82	−1.02	−0.21
ridge	22.74	23.26	23.24	23.28	23.40	23.43	23.33	23.18	0.52	0.50	0.54	0.66	0.69	0.59	0.44
xgb	37.09	38.47	38.69	37.99	38.29	38.72	38.58	38.47	1.39	1.60	0.90	1.20	1.63	1.49	1.38
xnn	2.41	0.04	3.91	1.41	0.75	2.46	0.16	2.97	−2.37	1.50	−1.00	−1.66	0.05	−2.25	0.55
2. D99															
cnn	19.08	22.90	11.33	19.51	13.78	14.81	13.94	17.88	3.82	−7.75	0.43	−5.29	−4.27	−5.14	−1.19
ebm	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ffn	26.16	24.97	24.29	25.94	24.89	25.94	24.71	26.08	−1.18	−1.86	−0.22	−1.27	−0.22	−1.44	−0.08
ffnlp	27.18	27.71	26.96	27.28	27.57	27.14	27.01	26.91	0.52	−0.22	0.09	0.38	−0.05	−0.17	−0.27
ffnsp	9.36	13.27	16.96	11.06	14.27	14.70	10.53	16.62	3.91	7.60	1.70	4.91	5.34	1.17	7.26
knn	14.05	13.54	13.27	13.29	13.23	13.22	13.25	13.28	−0.51	−0.78	−0.76	−0.82	−0.83	−0.80	−0.77
lda	21.33	21.73	22.23	22.02	22.19	22.41	22.29	22.13	0.40	0.90	0.69	0.87	1.08	0.96	0.80
mlp	20.87	21.98	24.80	22.33	22.53	21.69	22.91	23.03	1.11	3.93	1.46	1.66	0.83	2.04	2.16
percep	9.80	10.01	9.57	9.37	7.64	10.42	11.59	10.76	0.21	−0.23	−0.43	−2.16	0.62	1.79	0.96
qda	26.47	26.40	24.97	25.82	25.12	22.13	23.43	26.14	−0.06	−1.49	−0.65	−1.34	−4.34	−3.04	−0.33
rf	37.84	37.81	36.71	37.16	36.98	36.39	36.69	37.31	−0.03	−1.13	−0.67	−0.86	−1.44	−1.15	−0.53
ridge	21.31	21.70	22.21	22.00	22.17	22.39	22.27	22.12	0.39	0.90	0.69	0.87	1.09	0.97	0.82
xgb	34.58	36.33	35.93	36.26	36.36	36.62	36.46	36.59	1.76	1.35	1.68	1.78	2.04	1.88	2.01
xnn	2.52	0.00	2.86	0.02	1.77	0.98	0.41	1.09	−2.52	0.34	−2.50	−0.76	−1.54	−2.11	−1.43
3. D80															
cnn	22.20	21.68	14.44	22.97	17.89	17.16	18.40	20.04	−0.52	−7.76	0.77	−4.31	−5.04	−3.80	−2.16
ebm	35.08	36.81	36.88	36.98	37.02	37.14	37.07	36.76	1.73	1.80	1.89	1.93	2.06	1.99	1.68
ffn	27.38	27.80	27.50	26.89	26.50	28.07	28.20	27.70	0.41	0.12	−0.49	−0.88	0.69	0.82	0.31
ffnlp	27.38	26.81	26.45	27.89	27.28	27.82	28.31	26.94	−0.57	−0.93	0.51	−0.10	0.44	0.93	−0.44
ffnsp	13.61	16.83	17.36	14.58	13.18	16.98	14.93	16.65	3.22	3.75	0.97	−0.43	3.37	1.32	3.03
knn	17.35	17.35	17.32	17.33	17.33	17.30	17.32	17.33	−0.00	−0.03	−0.03	−0.02	−0.05	−0.04	−0.02
lda	23.84	24.07	24.08	24.29	24.45	24.49	24.46	24.00	0.23	0.25	0.45	0.61	0.65	0.63	0.16
mlp	23.88	23.83	21.46	23.34	23.32	24.26	21.67	24.16	−0.05	−2.42	−0.54	−0.57	0.38	−2.21	0.28
percep	10.79	12.50	12.73	10.46	11.96	13.20	10.32	12.19	1.71	1.93	−0.33	1.17	2.40	−0.48	1.39
qda	28.14	27.83	26.40	26.58	26.34	24.63	25.60	27.61	−0.31	−1.74	−1.56	−1.80	−3.50	−2.54	−0.52
rf	39.96	39.91	40.13	40.51	39.88	39.56	39.82	40.59	−0.04	0.18	0.55	−0.07	−0.39	−0.14	0.63
ridge	23.79	23.98	24.05	24.22	24.38	24.42	24.43	23.92	0.19	0.26	0.43	0.59	0.63	0.64	0.13
xgb	36.74	37.85	38.69	38.45	38.77	38.91	38.29	38.57	1.12	1.95	1.71	2.04	2.17	1.56	1.83
xnn	4.45	3.60	1.65	0.37	0.04	0.00	0.01	1.55	−0.85	−2.80	−4.08	−4.41	−4.44	−4.44	−2.90

Note: This table reports the average KS across ten folds for binary credit classification (good vs. bad payers) in datasets enriched with EF related to social unrest and climate change under scenarios that include government-backed guarantees (WGB_CR). Key findings include the following: 1. Models such as *RF*, *XGB*, and *EBM* consistently achieve high KS values, showing robust predictive performance. 2. The *RF* model shows the highest KS scores across most scenarios, with values around 40%. 3. Classical models, such as *LDA* and *RIDGE* demonstrate stable performance across different scenarios. 4. Some models (*CNN*, *MLP*, and *XNN*) show significant variability across scenarios, highlighting potential sensitivity to EF shifts. 5. Including WGB_CR data appears to stabilize the performance for some models, while others show large fluctuations between scenarios.

Table A12. Average F1 across 10 folds of ML models with SOCIAL UNREST and CLIMATE CHANGE EF WGB_CR, by evaluated scenario and relevance groups.

F1 (%)															
	Scenarios								Results						
	SFE	**T**	**T + W**	**U**	**U + T**	**U + T + W**	**U + W**	**W**	**T - SFE**	**T + W - SFE**	**U - SFE**	**U + T - SFE**	**U + T + W - SFE**	**U + W - SFE**	**W - SFE**
WGB_CR															
1. D100															
cnn	65.60	67.92	91.94	90.86	83.99	91.94	91.67	91.07	2.32	26.34	25.26	18.39	26.34	26.07	25.47
ebm	91.22	91.67	91.70	91.64	91.67	91.68	91.66	91.67	0.44	0.47	0.42	0.45	0.46	0.44	0.45
ffn	77.10	72.85	76.60	70.25	77.98	75.85	75.64	69.85	−4.25	−0.50	−6.85	0.88	−1.25	−1.46	−7.25
ffnlp	69.97	79.18	71.61	75.03	73.50	80.86	76.05	71.79	9.21	1.64	5.06	3.53	10.89	6.08	1.82
ffnsp	48.74	44.69	42.64	64.14	51.68	56.75	47.90	40.48	−4.05	−6.10	15.40	2.94	8.01	−0.83	−8.25
knn	76.16	76.14	76.08	76.10	76.09	76.08	76.07	76.09	−0.01	−0.07	−0.05	−0.07	−0.08	−0.08	−0.07
lda	77.06	77.19	77.35	77.37	77.42	77.45	77.42	77.40	0.13	0.29	0.31	0.36	0.39	0.36	0.34
mlp	66.55	71.50	83.00	79.79	68.07	64.94	72.57	69.97	4.95	16.44	13.24	1.52	−1.61	6.01	3.42
percep	50.71	43.53	46.50	61.18	71.86	50.13	60.29	46.99	−7.18	−4.21	10.47	21.15	−0.58	9.58	−3.72
qda	74.51	76.13	77.46	75.36	75.57	74.58	75.23	77.31	1.62	2.95	0.85	1.06	0.07	0.72	2.80
rf	90.87	91.94	92.12	92.04	92.15	92.15	92.15	92.10	1.06	1.25	1.16	1.27	1.28	1.27	1.22
ridge	43.83	44.59	44.90	44.82	45.15	45.42	45.08	44.59	0.76	1.08	0.99	1.33	1.59	1.25	0.76
xgb	90.90	91.62	91.73	91.71	91.72	91.76	91.80	91.77	0.73	0.84	0.81	0.82	0.86	0.90	0.87
xnn	79.95	91.89	65.74	86.19	90.54	78.73	91.04	66.31	11.94	−14.22	6.24	10.59	−1.22	11.09	−13.64
2. D99															
cnn	65.59	70.45	93.35	91.64	92.45	93.37	93.30	91.60	4.86	27.75	26.04	26.86	27.78	27.71	26.01
ebm	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ffn	75.92	77.65	78.86	75.96	75.82	80.24	73.25	74.40	1.73	2.94	0.04	−0.10	4.32	−2.66	−1.52
ffnlp	66.46	77.13	74.14	72.20	76.00	78.50	68.27	78.54	10.67	7.67	5.74	9.54	12.04	1.81	12.07
ffnsp	52.73	33.01	54.69	48.28	50.96	47.14	59.38	42.15	−19.72	1.97	−4.45	−1.77	−5.59	6.65	−10.58
knn	77.62	77.23	77.20	77.21	77.20	77.19	77.20	77.22	−0.39	−0.42	−0.41	−0.42	−0.43	−0.42	−0.40
lda	78.03	78.18	78.35	78.30	78.37	78.45	78.38	78.23	0.15	0.32	0.27	0.35	0.43	0.36	0.20
mlp	72.53	58.47	73.74	70.61	83.61	73.28	75.22	66.40	−14.06	1.22	−1.92	11.08	0.76	2.69	−6.13
percep	54.99	67.66	76.26	74.05	79.16	67.38	60.35	63.75	12.66	21.26	19.06	24.17	12.39	5.35	8.76
qda	75.22	76.07	76.52	74.35	74.57	72.86	73.44	76.07	0.85	1.30	−0.86	−0.65	−2.36	−1.78	0.85
rf	92.15	93.28	93.40	93.32	93.39	93.40	93.38	93.35	1.13	1.26	1.18	1.24	1.26	1.24	1.21
ridge	44.18	44.46	44.90	44.78	44.93	45.34	45.12	44.67	0.29	0.73	0.60	0.75	1.16	0.95	0.49
xgb	92.48	93.18	93.28	93.22	93.30	93.30	93.25	93.27	0.70	0.81	0.74	0.83	0.82	0.77	0.79
xnn	73.03	92.03	76.34	73.55	88.17	86.03	90.93	86.01	19.00	3.31	0.52	15.14	13.00	17.90	12.98
3. D80															
cnn	65.03	67.74	91.86	91.36	92.00	92.01	92.00	88.43	2.71	26.83	26.34	26.97	26.98	26.98	23.41
ebm	91.09	91.60	91.70	91.64	91.68	91.72	91.71	91.69	0.51	0.61	0.55	0.59	0.63	0.62	0.60
ffn	78.15	78.28	71.12	73.95	76.17	74.26	74.60	75.33	0.12	−7.03	−4.20	−1.98	−3.90	−3.55	−2.83
ffnlp	73.69	74.93	78.64	74.90	75.63	79.30	75.33	73.38	1.24	4.94	1.20	1.94	5.61	1.64	−0.31
ffnsp	41.02	44.83	54.21	57.41	36.19	55.66	49.21	43.53	3.81	13.20	16.39	−4.83	14.65	8.19	2.51
knn	76.22	76.21	76.19	76.20	76.20	76.18	76.19	76.21	−0.01	−0.03	−0.01	−0.02	−0.03	−0.02	−0.01
lda	77.09	77.31	77.49	77.50	77.54	77.67	77.58	77.50	0.23	0.40	0.42	0.45	0.59	0.49	0.42
mlp	57.55	69.54	77.77	67.28	80.29	62.86	73.73	66.26	12.00	20.22	9.73	22.75	5.32	16.19	8.71
percep	84.49	57.76	61.38	70.60	77.07	66.03	70.12	71.36	−26.73	−23.11	−13.89	−7.42	−18.46	−14.37	−13.13
qda	75.56	76.37	77.23	74.19	74.56	73.76	74.41	77.24	0.81	1.68	−1.37	−0.99	−1.80	−1.15	1.68
rf	90.64	91.86	92.09	92.01	92.09	92.14	92.10	92.04	1.22	1.45	1.38	1.45	1.50	1.47	1.40
ridge	44.35	44.98	45.29	45.22	45.56	45.76	45.40	44.91	0.64	0.94	0.88	1.21	1.41	1.06	0.57
xgb	90.78	91.59	91.78	91.69	91.77	91.76	91.78	91.73	0.81	1.00	0.91	0.98	0.98	1.00	0.95
xnn	80.64	72.23	85.18	81.25	73.62	92.01	73.60	61.87	−8.41	4.54	0.61	−7.02	11.37	−7.04	−18.77

Note: This table reports the average F1 across ten folds for binary credit classification (good vs. bad payers) in datasets enriched with EF related to social unrest and climate change under scenarios including government-backed guarantees (WGB_CR). Key findings include the following: 1. The *CNN* model shows dramatic improvements in F1 score (up to +27.75 percentage points) in certain scenarios, particularly those combining multiple EF types. 2. Models like *EBM* and *RF* maintain consistently high F1 scores (above 90%) across most scenarios. 3. The *FFNSP* model exhibits the most variability, with improvements up to +16.39 in some scenarios but declines in others. 4. Classical models (*LDA*, *RIDGE*) show modest but stable performance across scenarios. 5. Some models (*PERCEP*, *XNN*) show extreme variability, with both large gains and losses depending on the scenario. 6. The inclusion of WGB_CR appears to benefit neural network models particularly, while tree-based models maintain their high performance.

Table A13. Average AUC across 10 folds of ML models with SOCIAL UNREST and CLIMATE CHANGE EF WOGB_CR, by evaluated scenario and relevance groups.

AUC (%)															
	Scenarios								Results						
	SFE	**T**	**T + W**	**U**	**U + T**	**U + T + W**	**U + W**	**W**	**T - SFE**	**T + W - SFE**	**U - SFE**	**U + T - SFE**	**U + T + W - SFE**	**U + W - SFE**	**W - SFE**
WOGB_CR															
1. D100															
cnn	62.35	61.96	57.87	62.34	58.89	59.00	58.34	61.89	−0.39	−4.48	−0.01	− 3.46	−3.35	−4.00	−0.46
ebm	73.73	73.71	73.92	73.87	73.91	74.03	74.02	73.89	−0.03	0.19	0.13	0.18	0.29	0.29	0.15
ffn	65.16	66.25	66.50	66.81	66.61	66.50	66.49	66.62	1.09	1.34	1.65	1.45	1.35	1.34	1.46
ffnlp	66.45	66.06	65.46	66.77	66.60	66.74	65.99	66.40	−0.39	−0.99	0.32	0.14	0.28	−0.46	−0.06
ffnsp	66.26	66.27	66.90	67.05	65.92	67.00	66.82	66.19	0.01	0.65	0.79	− 0.34	0.74	0.57	−0.07
knn	58.68	58.54	58.63	58.62	58.62	58.61	58.60	58.63	−0.14	−0.05	−0.06	− 0.06	−0.07	−0.08	−0.05
lda	64.60	64.84	64.96	65.00	65.02	65.06	65.04	64.89	0.24	0.35	0.39	0.41	0.46	0.43	0.29
mlp	63.72	61.49	64.36	64.96	64.06	64.77	61.77	64.21	−2.23	0.64	1.24	0.34	1.05	−1.95	0.49
percep	55.32	54.51	53.47	55.14	54.96	56.40	53.95	56.10	−0.81	−1.85	−0.18	− 0.36	1.07	−1.37	0.77
qda	66.35	66.39	65.74	65.80	65.52	64.16	64.76	66.21	0.04	−0.60	−0.54	− 0.83	−2.19	−1.58	−0.14
rf	74.68	74.83	74.47	74.72	74.70	74.38	74.47	74.77	0.14	−0.22	0.03	0.01	−0.31	−0.22	0.09
ridge	63.40	63.65	63.80	63.90	63.93	63.98	63.96	63.74	0.26	0.41	0.50	0.53	0.59	0.56	0.34
xgb	72.53	73.97	73.97	73.72	73.83	73.90	74.02	74.01	1.44	1.44	1.18	1.30	1.37	1.48	1.48
xnn	51.66	50.13	49.93	50.52	50.55	51.25	49.96	50.32	−1.54	−1.73	−1.15	− 1.11	−0.41	−1.71	−1.34
2. D99															
cnn	62.49	62.40	58.08	61.81	59.52	59.19	59.23	62.85	−0.10	−4.41	−0.68	− 2.97	−3.31	−3.26	0.35
ebm	73.68	73.85	73.64	73.58	73.68	73.80	73.70	73.60	0.17	−0.04	−0.10	− 0.00	0.12	0.03	−0.08
ffn	66.85	66.06	66.54	65.76	66.16	66.09	65.66	66.50	−0.79	−0.31	−1.09	− 0.69	−0.76	−1.19	−0.35
ffnlp	66.07	66.53	66.58	65.89	66.35	66.94	66.20	66.11	0.47	0.51	−0.18	0.28	0.87	0.13	0.04
ffnsp	66.51	66.78	66.49	66.63	67.09	66.90	66.45	66.19	0.28	−0.02	0.13	0.58	0.39	−0.06	−0.31
knn	58.65	58.45	58.35	58.36	58.35	58.35	58.35	58.36	−0.20	−0.29	−0.28	− 0.30	−0.30	−0.29	−0.29
lda	64.59	64.80	64.94	64.98	65.00	65.04	65.02	64.87	0.22	0.35	0.39	0.41	0.46	0.43	0.28
mlp	63.49	64.05	64.20	62.84	64.36	64.80	62.98	63.31	0.55	0.71	−0.65	0.86	1.30	−0.52	−0.18
percep	53.15	51.87	52.01	51.72	50.70	53.44	54.65	53.33	−1.28	−1.14	−1.43	− 2.45	0.29	1.50	0.18
qda	66.34	66.28	65.63	65.67	65.40	64.02	64.62	66.09	−0.06	−0.71	−0.67	− 0.94	−2.32	−1.72	−0.25
rf	74.54	75.02	74.71	74.83	74.61	74.47	74.49	74.83	0.48	0.17	0.29	0.08	−0.07	−0.05	0.29
ridge	63.38	63.61	63.78	63.89	63.92	63.96	63.94	63.71	0.23	0.40	0.51	0.54	0.58	0.56	0.34
xgb	72.60	73.73	73.74	73.67	73.89	73.92	73.91	73.93	1.13	1.14	1.08	1.29	1.32	1.31	1.33
xnn	49.59	51.89	51.69	50.47	50.00	50.29	52.04	50.65	2.29	2.10	0.88	0.41	0.70	2.45	1.05
3. D80															
cnn	61.73	60.58	58.10	60.91	59.23	58.89	58.88	61.34	−1.14	−3.62	−0.82	− 2.50	−2.84	−2.85	−0.39
ebm	73.49	73.69	73.81	73.74	73.85	73.89	73.89	73.80	0.19	0.32	0.25	0.36	0.40	0.40	0.31
ffn	66.74	66.58	66.16	65.41	66.59	66.38	66.47	65.86	−0.16	−0.58	−1.32	− 0.15	−0.36	−0.26	−0.88
ffnlp	66.29	66.37	66.13	66.85	66.49	66.61	67.25	65.60	0.09	−0.16	0.57	0.21	0.32	0.96	−0.69
ffnsp	66.90	66.42	66.55	66.96	65.59	66.67	66.76	65.94	−0.48	−0.36	0.06	− 1.31	−0.24	−0.14	−0.97
knn	58.56	58.55	58.56	58.54	58.54	58.55	58.55	58.57	−0.01	0.00	−0.02	− 0.02	−0.01	−0.01	0.01
lda	65.11	65.21	65.28	65.35	65.37	65.40	65.37	65.22	0.09	0.17	0.23	0.26	0.29	0.26	0.11
mlp	64.27	64.98	62.54	63.20	65.27	65.21	65.10	63.60	0.71	−1.73	−1.07	1.00	0.94	0.83	−0.67
percep	51.76	51.39	53.77	54.19	50.55	53.06	53.66	52.41	−0.37	2.01	2.43	−1.21	1.31	1.90	0.65
qda	66.49	66.36	65.84	66.02	65.75	64.61	65.25	66.33	−0.14	−0.65	−0.48	−0.74	−1.88	−1.25	−0.16
rf	74.30	74.70	74.49	74.58	74.60	74.21	74.22	74.66	0.41	0.20	0.29	0.31	−0.09	−0.07	0.37
ridge	63.90	64.01	64.14	64.27	64.31	64.35	64.30	64.07	0.11	0.24	0.37	0.41	0.45	0.40	0.17
xgb	72.43	73.64	74.01	73.71	73.76	74.05	73.86	74.05	1.21	1.58	1.28	1.33	1.62	1.43	1.62
xnn	50.40	50.55	50.00	49.97	50.00	50.00	49.94	50.66	0.15	−0.40	−0.43	−0.40	−0.40	−0.47	0.25

Note: This table reports the average AUC across ten folds for binary credit classification (good vs. bad payers) in datasets enriched with EF related to social unrest and climate change, excluding credits with government-backed guarantees (WOGB_CR). Key findings include the following: 1. Tree-based models (*EBM*, *RF*, and *XGB*) demonstrate strong and consistent performance, with AUC values above 73%. 2. Neural networks (*CNN*, *XNN*) show greater sensitivity to the exclusion of government-backed credits, with notable performance declines in some scenarios. 3. Classical models (*LDA*, *RIDGE*) show modest but stable improvements, whereas *KNN* maintains consistent performance. 4. Scenarios including social unrest and climate change factors (U, T + W) highlight the adaptability of tree-based models, while neural networks exhibit higher variability. 5. The exclusion of government-backed credit data affects the predictive performance of some models, particularly neural network-based models, underscoring the value of enriched datasets in maintaining predictive accuracy.

Table A14. Average ACC across 10 folds of ML models with SOCIAL UNREST and CLIMATE CHANGE EF WOGB_CR, by evaluated scenario and relevance groups.

ACC (%)															
	Scenarios								Results						
	SFE	**T**	**T + W**	**U**	**U + T**	**U + T + W**	**U + W**	**W**	**T - SFE**	**T + W - SFE**	**U - SFE**	**U + T - SFE**	**U + T + W - SFE**	**U + W - SFE**	**W - SFE**
WOGB_CR															
1. D100															
cnn	56.63	60.47	87.65	84.37	87.69	87.67	82.41	84.51	3.84	31.02	27.74	31.06	31.04	25.78	27.88
ebm	87.78	87.69	87.65	87.63	87.64	87.64	87.65	87.65	−0.09	−0.13	−0.15	−0.14	−0.14	−0.13	−0.13
ffn	70.07	70.60	67.96	64.99	66.82	68.50	68.26	68.25	0.53	−2.11	−5.08	−3.24	−1.56	−1.81	−1.82
ffnlp	64.10	65.67	72.32	67.66	66.18	65.77	65.80	66.70	1.57	8.22	3.56	2.08	1.66	1.69	2.59
ffnsp	67.21	63.53	60.93	69.04	67.53	60.97	66.57	69.85	−3.68	−6.28	1.82	0.31	−6.24	−0.64	2.63
knn	65.24	65.20	64.98	65.00	65.00	64.95	64.96	64.99	−0.03	−0.26	−0.24	−0.24	−0.28	−0.27	−0.24
lda	67.24	67.31	67.52	67.43	67.45	67.64	67.55	67.49	0.07	0.28	0.19	0.21	0.40	0.30	0.25
mlp	58.58	70.53	56.07	66.44	67.42	56.94	63.93	56.47	11.95	−2.50	7.87	8.84	−1.63	5.36	−2.11
percep	48.84	46.38	62.42	36.78	63.48	40.90	56.64	42.35	−2.46	13.58	−12.06	14.64	−7.94	7.80	−6.49
qda	62.81	64.24	65.37	63.18	63.41	61.37	62.09	65.05	1.43	2.55	0.36	0.59	−1.44	−0.73	2.23
rf	85.72	87.44	87.63	87.53	87.68	87.67	87.65	87.63	1.73	1.91	1.82	1.96	1.95	1.93	1.91
ridge	66.18	66.31	66.60	66.58	66.67	66.77	66.63	66.50	0.13	0.43	0.40	0.49	0.59	0.45	0.33
xgb	86.18	87.29	87.50	87.41	87.49	87.49	87.49	87.51	1.11	1.32	1.23	1.31	1.30	1.31	1.32
xnn	72.01	56.05	87.57	81.71	77.36	80.53	79.95	77.88	−15.96	15.56	9.70	5.36	8.52	7.94	5.88
2. D99															
cnn	66.29	54.66	87.59	83.57	87.67	87.70	87.53	87.04	−11.63	21.30	17.28	21.38	21.41	21.24	20.76
ebm	87.82	87.69	87.62	87.59	87.62	87.64	87.62	87.63	−0.13	−0.20	−0.22	−0.20	−0.18	−0.20	−0.19
ffn	67.13	70.09	62.29	64.60	68.45	67.81	71.27	67.15	2.96	−4.84	−2.53	1.32	0.68	4.14	0.02
ffnlp	58.27	68.32	65.68	65.02	67.64	65.62	60.54	69.23	10.05	7.41	6.76	9.37	7.35	2.27	10.96
ffnsp	64.10	67.73	67.50	66.28	66.38	66.31	69.18	65.03	3.64	3.40	2.18	2.29	2.21	5.08	0.93
knn	65.38	64.92	64.89	64.89	64.89	64.87	64.89	64.90	−0.45	−0.49	−0.48	−0.49	−0.50	−0.49	−0.47
lda	67.24	67.40	67.50	67.42	67.41	67.54	67.50	67.40	0.16	0.26	0.18	0.17	0.30	0.26	0.16
mlp	64.91	61.26	68.51	66.46	58.05	67.96	64.92	65.38	−3.65	3.59	1.55	−6.86	3.04	0.01	0.46
percep	51.55	61.77	69.54	66.65	72.93	60.93	53.92	57.20	10.22	17.99	15.10	21.38	9.38	2.37	5.64
qda	63.27	64.28	64.73	62.34	62.54	60.49	61.20	64.28	1.01	1.46	−0.93	−0.73	−2.77	−2.07	1.02
rf	85.73	87.51	87.64	87.59	87.66	87.68	87.66	87.62	1.78	1.92	1.86	1.93	1.96	1.94	1.89
ridge	66.16	66.38	66.61	66.55	66.64	66.74	66.66	66.44	0.22	0.45	0.39	0.48	0.58	0.50	0.28
xgb	86.20	87.32	87.48	87.38	87.52	87.51	87.42	87.46	1.12	1.27	1.18	1.32	1.31	1.21	1.26
xnn	63.78	68.47	62.79	81.82	87.40	84.20	64.94	82.32	4.69	−0.99	18.04	23.62	20.43	1.16	18.54
3. D80															
cnn	62.36	58.85	87.57	84.52	87.42	87.68	87.71	82.55	−3.50	25.22	22.16	25.06	25.32	25.35	20.20
ebm	87.51	87.58	87.64	87.62	87.62	87.65	87.63	87.64	0.07	0.13	0.10	0.11	0.13	0.12	0.12
ffn	63.83	58.29	69.62	66.20	69.87	71.70	62.72	62.91	−5.54	5.79	2.37	6.04	7.87	−1.12	−0.92
ffnlp	58.46	61.38	71.01	70.00	64.74	66.28	62.12	63.30	2.92	12.55	11.54	6.28	7.82	3.66	4.84
ffnsp	65.89	62.61	63.44	63.93	72.89	62.83	65.32	66.60	−3.28	−2.45	−1.96	7.00	−3.06	−0.57	0.71
knn	65.07	65.06	65.03	65.05	65.04	65.02	65.03	65.06	−0.00	−0.03	−0.02	−0.03	−0.05	−0.03	−0.01
lda	67.33	67.59	67.76	67.58	67.69	67.85	67.74	67.56	0.26	0.43	0.25	0.36	0.52	0.41	0.23
mlp	67.93	65.93	72.31	68.12	63.21	66.36	59.38	59.67	−2.00	4.38	0.19	−4.72	−1.57	−8.55	−8.26
percep	60.08	62.61	50.07	53.04	62.32	63.72	53.02	64.28	2.53	−10.01	−7.03	2.24	3.64	−7.06	4.20
qda	62.47	63.65	64.27	60.94	61.33	59.35	60.28	63.88	1.18	1.81	−1.53	−1.14	−3.12	−2.19	1.41
rf	85.45	87.36	87.68	87.58	87.70	87.70	87.65	87.60	1.91	2.22	2.13	2.25	2.25	2.20	2.14
ridge	66.32	66.58	66.73	66.72	66.84	66.93	66.78	66.64	0.26	0.41	0.40	0.52	0.61	0.46	0.31
xgb	85.95	87.21	87.51	87.38	87.46	87.51	87.49	87.47	1.26	1.55	1.42	1.51	1.56	1.54	1.52
xnn	66.33	81.92	87.74	80.03	87.74	80.19	79.88	69.28	15.59	21.41	13.69	21.41	13.86	13.55	2.95

Note: This table reports the average ACC across ten folds for binary credit classification (good vs. bad payers) in datasets enriched with EF related to social unrest and climate change, excluding credits with government-backed guarantees (WOGB_CR). Key findings include the following: 1. Tree-based models (*EBM*, *RF*, *XGB*) demonstrate robust and consistent performance, with ACC values above 87% across most scenarios. 2. Neural networks (*CNN*, *XNN*, *MLP*) exhibit greater sensitivity to the exclusion of government-backed credits, with significant performance variations in combined scenarios (U + T + W), particularly for *XNN* (ACC reductions of up to 21% in some scenarios). 3. Classical models (*LDA*, *RIDGE*) maintain stable performance with minimal changes (±0.2%), achieving consistent ACC values of approximately 67%. 4. Perceptron-based models show notable performance reductions in specific scenarios, such as U + T- SFE, with ACC dropping to 48.84% (−2.46% compared with datasets including government-backed credits). 5. Scenarios enriched with combined external factors (U + T + W) highlight the adaptability and stability of tree-based models, whereas neural network models show greater variability. 6. The exclusion of government-backed credit data significantly affects the predictive performance of certain models, particularly neural networks, emphasizing the importance of enriched datasets in maintaining predictive accuracy.

Table A15. Average KS across 10 folds of ML models with SOCIAL UNREST and CLIMATE CHANGE EF WOGB_CR, by evaluated scenario and relevance groups.

KS (%)															
	Scenarios								Results						
	SFE	**T**	**T + W**	**U**	**U + T**	**U + T + W**	**U + W**	**W**	**T - SFE**	**T + W - SFE**	**U - SFE**	**U + T - SFE**	**U + T + W - SFE**	**U + W - SFE**	**W - SFE**
WOGB_CR															
1. D100															
cnn	18.94	19.98	12.10	19.87	12.88	14.37	13.07	17.53	1.04	−6.84	0.93	−6.06	−4.57	−5.87	−1.41
ebm	33.22	34.51	34.74	34.83	34.90	35.00	34.98	34.56	1.29	1.52	1.61	1.68	1.78	1.76	1.35
ffn	25.68	25.89	26.54	26.06	25.95	25.99	24.64	24.54	0.21	0.87	0.38	0.27	0.31	−1.04	−1.14
ffnlp	25.04	25.92	25.95	26.42	25.63	26.39	25.59	25.24	0.88	0.91	1.38	0.60	1.35	0.55	0.20
ffnsp	9.84	11.59	4.96	10.07	11.05	13.18	9.46	8.22	1.75	−4.88	0.23	1.21	3.34	−0.38	−1.62
knn	14.22	13.64	14.05	14.06	14.05	14.03	14.04	14.07	−0.58	−0.17	−0.16	−0.17	−0.19	−0.18	−0.15
lda	21.34	21.68	22.05	22.26	22.36	22.35	22.28	21.96	0.34	0.71	0.92	1.03	1.01	0.94	0.62
mlp	21.86	22.19	22.80	20.70	23.36	22.86	20.86	21.43	0.33	0.94	−1.17	1.49	1.00	−1.01	−0.44
percep	12.60	11.91	10.31	11.87	12.76	13.63	11.64	13.77	−0.70	−2.29	−0.74	0.15	1.02	−0.96	1.16
qda	26.31	26.61	25.04	25.71	25.15	22.29	23.59	26.49	0.30	−1.27	−0.60	−1.16	−4.02	−2.72	0.18
rf	37.89	37.84	36.79	37.37	37.11	36.32	36.64	37.27	−0.05	−1.10	−0.52	−0.78	−1.58	−1.25	−0.63
ridge	21.32	21.65	22.02	22.25	22.33	22.33	22.24	21.96	0.33	0.70	0.93	1.01	1.01	0.92	0.64
xgb	34.76	36.80	36.94	36.15	36.25	36.53	36.58	36.63	2.04	2.19	1.40	1.49	1.77	1.83	1.87
xnn	3.37	0.26	3.18	1.09	2.80	0.03	1.04	0.00	−3.11	−0.20	−2.29	−0.58	−3.34	−2.33	−3.37
2. D99															
cnn	19.08	22.90	11.33	19.51	13.78	14.81	13.94	17.88	3.82	−7.75	0.43	−5.29	−4.27	−5.14	−1.19
ebm	33.10	34.52	35.30	35.03	35.27	35.58	35.52	34.90	1.42	2.20	1.93	2.18	2.48	2.42	1.81
ffn	26.16	24.97	24.29	25.94	24.89	25.94	24.71	26.08	−1.18	−1.86	−0.22	−1.27	−0.22	−1.44	−0.08
ffnlp	26.25	24.57	26.45	25.67	25.72	25.97	26.49	25.55	−1.68	0.20	−0.58	−0.53	−0.28	0.24	−0.70
ffnsp	9.60	9.95	12.99	9.68	13.82	13.25	8.60	6.58	0.35	3.39	0.08	4.22	3.65	−1.00	−3.02
knn	14.05	13.54	13.27	13.29	13.23	13.22	13.25	13.28	−0.51	−0.78	−0.76	−0.82	−0.83	−0.80	−0.77
lda	21.33	21.73	22.23	22.02	22.19	22.41	22.29	22.13	0.40	0.90	0.69	0.87	1.08	0.96	0.80
mlp	20.87	21.98	24.80	22.33	22.53	21.69	22.91	23.03	1.11	3.93	1.46	1.66	0.83	2.04	2.16
percep	9.80	10.01	9.57	9.37	7.64	10.42	11.59	10.76	0.21	−0.23	−0.43	−2.16	0.62	1.79	0.96
qda	26.47	26.40	24.97	25.82	25.12	22.13	23.43	26.14	−0.06	−1.49	−0.65	−1.34	−4.34	−3.04	−0.33
rf	37.84	37.81	36.71	37.16	36.98	36.39	36.69	37.31	−0.03	−1.13	−0.67	−0.86	−1.44	−1.15	−0.53
ridge	21.31	21.70	22.21	22.00	22.17	22.39	22.27	22.12	0.39	0.90	0.69	0.87	1.09	0.97	0.82
xgb	34.58	36.33	35.93	36.26	36.36	36.62	36.46	36.59	1.76	1.35	1.68	1.78	2.04	1.88	2.01
xnn	2.67	0.92	0.03	1.05	3.53	1.61	0.00	1.08	−1.75	−2.64	−1.62	0.86	−1.06	−2.67	−1.60
3. D80															
cnn	19.50	21.15	12.74	20.39	14.50	14.56	13.93	19.92	1.66	−6.76	0.90	−5.00	−4.93	−5.57	0.42
ebm	33.07	34.47	34.92	35.00	35.11	35.01	35.13	34.92	1.39	1.85	1.93	2.04	1.94	2.06	1.84
ffn	24.92	25.45	25.35	25.44	25.50	26.16	26.50	24.64	0.53	0.43	0.52	0.58	1.24	1.58	−0.28
ffnlp	25.54	25.91	24.81	25.28	26.58	26.37	26.40	24.38	0.37	−0.73	−0.26	1.04	0.83	0.86	−1.16
ffnsp	13.65	15.66	10.11	10.36	11.86	12.60	11.65	11.45	2.02	−3.54	−3.29	−1.79	−1.05	−2.00	−2.20
knn	13.78	13.78	13.82	13.78	13.77	13.79	13.84	13.86	−0.01	0.04	0.00	−0.01	0.00	0.06	0.08
lda	22.24	22.45	22.74	22.95	23.11	23.00	22.89	22.60	0.21	0.50	0.71	0.87	0.76	0.65	0.36
mlp	23.25	21.95	21.19	19.73	21.17	22.74	20.36	22.79	−1.30	−2.05	−3.52	−2.08	−0.51	−2.89	−0.46
percep	10.56	10.86	11.48	11.94	11.49	12.40	12.16	10.83	0.29	0.92	1.38	0.93	1.84	1.60	0.27
qda	27.14	26.85	25.42	26.06	25.79	23.22	24.15	26.57	−0.29	−1.72	−1.08	−1.35	−3.92	−2.99	−0.57
rf	37.21	37.14	36.89	37.36	37.37	36.46	36.62	36.98	−0.07	−0.32	0.15	0.15	−0.76	−0.60	−0.23
ridge	22.22	22.45	22.75	22.94	23.08	22.99	22.86	22.58	0.24	0.53	0.73	0.86	0.77	0.64	0.36
xgb	34.35	36.15	36.51	36.19	36.29	36.89	36.24	36.87	1.80	2.16	1.84	1.94	2.54	1.89	2.52
xnn	3.96	0.04	0.00	4.34	0.80	1.05	2.69	0.01	−3.92	−3.96	0.39	−3.16	−2.91	−1.27	−3.94

Note: This table reports the average KS across ten folds for binary credit classification (good vs. bad payers) in datasets enriched with EF related to social unrest and climate change, excluding credits with government-backed guarantees (WOGB_CR). Key findings include the following: 1. Tree-based models (*EBM*, *RF*, and *XGB*) demonstrate strong and consistent performance, with KS values above 33%. 2. Neural networks (*CNN*, *XNN*) show greater sensitivity to the exclusion of government-backed credits, with notable performance declines in some scenarios. 3. Classical models (*LDA*, *RIDGE*) show modest but stable improvements, whereas *KNN* maintains consistent performance. 4. Scenarios including social unrest and climate change factors (U, T + W) highlight the adaptability of tree-based models, while neural networks exhibit higher variability. 5. The exclusion of government-backed credit data affects the predictive performance of some models, particularly neural network-based models, underscoring the value of enriched datasets in maintaining predictive accuracy.

Table A16. Average F1 across 10 folds of ML models with social unrest and climate change EF WOGB_CR, by evaluated scenario and relevance groups.

F1 (%)															
	Scenarios								Results						
	SFE	**T**	**T + W**	**U**	**U + T**	**U + T + W**	**U + W**	**W**	**T - SFE**	**T + W - SFE**	**U - SFE**	**U + T - SFE**	**U + T + W - SFE**	**U + W - SFE**	**W - SFE**
WOGB_CR															
1. D100															
cnn	74.52	72.22	93.43	92.93	93.31	93.44	93.18	91.34	−2.29	18.91	18.42	18.80	18.93	18.66	16.83
ebm	92.94	93.31	93.32	93.29	93.31	93.32	93.32	93.31	0.38	0.38	0.35	0.37	0.38	0.38	0.37
ffn	78.15	77.34	71.58	71.25	80.09	73.72	79.59	80.88	−0.81	−6.57	−6.90	1.94	−4.43	1.44	2.73
ffnlp	78.39	78.26	78.67	76.47	77.80	73.82	80.79	83.01	−0.12	0.28	−1.92	−0.59	−4.56	2.41	4.62
ffnsp	55.25	35.46	36.63	40.18	48.29	47.66	36.39	53.66	−19.80	−18.62	−15.07	−6.96	−7.59	−18.86	−1.59
knn	77.50	77.44	77.27	77.29	77.29	77.25	77.26	77.29	−0.06	−0.23	−0.21	−0.21	−0.25	−0.24	−0.22
lda	78.03	78.11	78.35	78.34	78.40	78.48	78.36	78.26	0.08	0.32	0.31	0.37	0.45	0.33	0.24
mlp	76.41	64.90	72.26	76.02	70.03	63.77	70.71	72.75	−11.51	−4.16	−0.39	−6.38	−12.65	−5.71	−3.67
percep	54.08	51.78	66.30	38.40	68.94	44.24	65.55	44.69	−2.31	12.22	−15.68	14.85	−9.84	11.47	−9.39
qda	74.80	76.03	77.08	75.11	75.35	73.66	74.25	76.73	1.23	2.28	0.31	0.55	−1.14	−0.55	1.93
rf	92.15	93.27	93.38	93.34	93.40	93.40	93.38	93.37	1.12	1.23	1.19	1.25	1.25	1.23	1.22
ridge	44.22	44.42	45.05	44.87	45.07	45.55	45.26	44.85	0.20	0.83	0.65	0.85	1.33	1.04	0.63
xgb	92.46	93.16	93.29	93.23	93.28	93.29	93.29	93.30	0.70	0.83	0.77	0.82	0.83	0.82	0.83
xnn	60.57	73.92	81.69	83.04	82.61	83.99	78.51	84.10	13.35	21.13	22.47	22.05	23.42	17.94	23.54
2. D99															
cnn	65.59	70.45	93.35	91.64	92.45	93.37	93.30	91.60	4.86	27.75	26.05	26.86	27.78	27.71	26.01
ebm	92.94	93.28	93.32	93.30	93.32	93.33	93.34	93.31	0.33	0.38	0.36	0.38	0.39	0.40	0.37
ffn	75.92	77.65	78.86	75.96	75.82	80.24	73.25	74.40	1.73	2.94	0.04	−0.10	4.32	−2.66	−1.52
ffnlp	73.83	80.34	75.97	79.03	76.65	74.48	74.79	74.32	6.51	2.14	5.20	2.82	0.65	0.96	0.49
ffnsp	61.05	37.53	59.15	39.26	55.64	36.89	51.31	27.41	−23.52	−1.90	−21.80	−5.42	−24.17	−9.74	−33.64
knn	77.62	77.23	77.20	77.21	77.20	77.19	77.20	77.22	−0.39	−0.42	−0.41	−0.42	−0.43	−0.42	−0.40
lda	78.03	78.18	78.35	78.30	78.37	78.45	78.38	78.23	0.15	0.32	0.27	0.35	0.43	0.36	0.20
mlp	72.53	58.47	73.74	70.61	83.61	73.28	75.22	66.40	−14.06	1.22	−1.92	11.08	0.76	2.69	−6.13
percep	54.99	67.66	76.26	74.05	79.16	67.38	60.35	63.75	12.66	21.26	19.06	24.17	12.39	5.35	8.76
qda	75.22	76.07	76.52	74.35	74.57	72.86	73.44	76.07	0.85	1.30	−0.86	−0.65	−2.36	−1.78	0.85
rf	92.15	93.28	93.40	93.32	93.39	93.40	93.38	93.35	1.13	1.25	1.18	1.24	1.26	1.23	1.20
ridge	44.18	44.46	44.90	44.78	44.93	45.34	45.12	44.67	0.29	0.73	0.60	0.75	1.16	0.95	0.49
xgb	92.48	93.18	93.28	93.22	93.30	93.30	93.25	93.27	0.70	0.81	0.74	0.83	0.82	0.77	0.79
xnn	87.58	87.39	84.09	83.05	73.30	69.05	93.45	87.28	−0.19	−3.49	−4.53	−14.28	−18.53	5.87	−0.30
3. D80															
cnn	59.66	73.36	93.35	92.35	93.29	93.37	93.40	90.92	13.70	33.70	32.70	33.63	33.71	33.74	31.26
ebm	92.85	93.25	93.34	93.32	93.33	93.34	93.35	93.33	0.39	0.48	0.46	0.47	0.49	0.49	0.47
ffn	74.53	77.09	81.31	79.85	79.37	76.66	74.82	75.94	2.55	6.78	5.32	4.83	2.13	0.29	1.41
ffnlp	74.58	75.44	76.28	76.44	71.20	76.51	76.87	76.48	0.87	1.70	1.86	−3.38	1.93	2.29	1.90
ffnsp	34.07	44.07	46.46	67.61	32.96	51.48	33.21	46.91	10.00	12.39	33.53	−1.12	17.40	−0.86	12.83
knn	77.34	77.33	77.31	77.32	77.31	77.30	77.31	77.33	−0.00	−0.03	−0.01	−0.02	−0.04	−0.03	−0.01
lda	78.11	78.32	78.45	78.44	78.52	78.57	78.47	78.36	0.20	0.34	0.32	0.41	0.46	0.36	0.24
mlp	61.39	71.07	81.42	87.36	77.18	68.34	78.18	77.46	9.67	20.03	25.97	15.78	6.95	16.79	16.07
percep	63.41	66.38	54.50	60.03	65.79	72.48	58.12	69.26	2.97	−8.91	−3.38	2.38	9.06	−5.29	5.85
qda	74.42	75.49	76.12	73.08	73.46	71.77	72.56	75.71	1.07	1.70	−1.34	−0.96	−2.65	−1.86	1.29
rf	91.96	93.20	93.39	93.35	93.40	93.39	93.38	93.36	1.24	1.43	1.39	1.44	1.44	1.42	1.40
ridge	44.91	45.21	45.47	45.52	45.66	45.97	45.74	45.27	0.30	0.56	0.61	0.75	1.06	0.83	0.36
xgb	92.32	93.12	93.30	93.21	93.27	93.30	93.28	93.27	0.80	0.98	0.90	0.95	0.98	0.97	0.95
xnn	76.29	93.19	93.47	83.06	77.56	78.10	84.41	84.15	16.90	17.17	6.77	1.26	1.81	8.12	7.86

Note: This table reports the average F1 across ten folds for binary credit classification (good vs. bad payers) in datasets enriched with EF related to social unrest and climate change, excluding credits with government-backed guarantees (WOGB_CR). Key findings include the following: 1. Tree-based models (*EBM*, *RF*, *XGB*) demonstrate robust and consistent performance, with F1 values above 92% across most scenarios. 2. Neural networks (*CNN*, *XNN*, *MLP*) exhibit greater sensitivity to the exclusion of government-backed credits, with significant performance variations in combined scenarios (U + T + W). 3. Classical models (*LDA*, *RIDGE*) maintain stable performance with minimal changes, achieving consistent F1 values. 4. Perceptron-based models show notable performance reductions in specific scenarios, such as U + T- SFE. 5. Scenarios enriched with combined external factors (U + T + W) highlight the adaptability and stability of tree-based models, whereas neural network models show greater variability. 6. The exclusion of government-backed credit data significantly affects the predictive performance of certain models, particularly neural networks, emphasizing the importance of enriched datasets in maintaining predictive accuracy.

Table A17. Models with statistically significant improvements for AUC, ACC, KS, and F1 metrics with COVID EF, considering WGB_CR and WOGB_CR, by evaluated scenario and relevance groups—Scenarios.

	WGB_CR													WOGB_CR											
	AUC			ACC			KS			F1			TOTAL	AUC			ACC			KS			F1			TOTAL
	F	F + P	P	F	F + P	P	F	F + P	P	F	F + P	P		F	F + P	P	F	F + P	P	F	F + P	P	F	F + P	P	
1. D100																										
cnn	Y	Y	-	Y	Y	Y	Y	Y	-	Y	Y	Y	8	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
ebm	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
ffn	-	-	-	-	-	Y	-	-	-	-	-	-	1	-	-	-	-	Y	Y	-	-	-	-	-	-	2
ffnlp	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	0
ffnsp	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	-	-	Y	Y	-	-	-	-	-	-	2
lda	Y	Y	Y	-	Y	-	Y	Y	Y	Y	Y	Y	9	Y	Y	-	Y	Y	-	Y	Y	-	Y	Y	-	7
mlp	-	-	-	-	-	Y	-	-	-	-	-	-	1	Y	-	-	-	-	-	-	-	-	-	-	-	1
percep	-	-	-	-	Y	Y	-	-	-	-	Y	Y	4	-	-	-	-	-	-	-	-	-	-	-	-	0
qda	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6
rf	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6
ridge	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	Y	Y	-	Y	Y	-	Y	Y	-	Y	Y	-	8
xgb	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6
xnn	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	-	-	Y	-	-	-	-	-	Y	-	1
2. D99																										
cnn	-	-	-	Y	-	Y	-	-	-	Y	-	Y	3	Y	-	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	9
ebm	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
ffn	-	-	-	-	-	Y	-	-	-	-	-	-	1	-	-	-	-	-	Y	-	-	-	-	-	-	1
ffnlp	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	0
ffnsp	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	Y	-	-	-	-	-	-	-	-	-	1
lda	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	10	Y	Y	-	Y	Y	-	Y	Y	-	Y	Y	-	7
percep	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	0
qda	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6
rf	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6
ridge	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	Y	Y	-	Y	Y	-	Y	Y	-	Y	Y	-	8
xgb	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
xnn	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	0

Table A17. *Cont.*

	WGB_CR													WOGB_CR											
	AUC			ACC			KS			F1			TOTAL	AUC			ACC			KS			F1			TOTAL
	F	F + P	P	F	F + P	P	F	F + P	P	F	F + P	P		F	F + P	P	F	F + P	P	F	F + P	P	F	F + P	P	
3. D80																											
cnn	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	10	Y	Y	-	Y	Y	Y	Y	Y	-	Y	Y	Y	9
ebm	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
ffn	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	-	-	-	-	-	-	-	-	-	-	0
ffnlp	-	Y	-	-	Y	-	-	Y	-	-	Y	-	3	-	-	-	-	Y	-	-	-	-	-	Y	-	1
ffnsp	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	Y	-	-	-	-	-	-	-	-	-	1
lda	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
mlp	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	-	Y	-	Y	-	-	-	-	-	-	2
percep	-	Y	-	-	-	-	-	Y	-	-	-	-	1	-	Y	-	-	-	-	-	Y	-	-	-	-	1
qda	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6
rf	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6
ridge	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
xnn	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6	-	-	-	-	Y	-	-	-	-	-	Y	-	1

Note: This table reports the machine learning models (*CNN*, *EBM*, *FFN*, etc.) that exhibited statistically significant improvements in their classification metrics (AUC, ACC, KS, and F1) when EF related to COVID-19 was included. The evaluation considers scenarios with and without government-backed guarantees (*WGB_CR* and *WOGB_CR*). Key findings include the following: 1. Tree-based models (*XGB*, *RF*) demonstrate consistent performance, with statistical significance in most scenarios, particularly in F1 and ACC metrics. 2. Neural networks (*CNN*, *FFN*) show variability, excelling in *WGB_CR* but with mixed results in *WOGB_CR*. 3. Classical models (*LDA*, *RIDGE*) maintain robust performance across all metrics, especially in *WGB_CR*. 4. Model-specific trends are as follows: - *EBM* achieves perfect significance in all metrics for both groups. - *PERCEP* and *MLP* are scenario dependent, with limited improvements. 5. Dataset impact: *WOGB_CR* introduces higher variability, particularly for KS and F1 metrics.

Table A18. Models with statistically significant improvements for AUC, ACC, KS, and F1 metrics with COVID MOV EF, considering WGB_CR and WOGB_CR, by evaluated scenario and relevance groups—Scenarios.

	WGB_CR													WOGB_CR											
	AUC			ACC			KS			F1			TOTAL	AUC			ACC			KS			F1			TOTAL
	F	F + P	P	F	F + P	P	F	F + P	P	F	F + P	P		F	F + P	P	F	F + P	P	F	F + P	P	F	F + P	P	
1. D100																											
cnn	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	10	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
ebm	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	6	-	-	-	-	-	-	-	-	-	-	-	-	0
lda	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	11	Y	Y	-	Y	Y	-	Y	Y	-	Y	Y	-	7
percep	-	-	-	-	-	Y	-	-	-	-	-	Y	2	-	-	-	-	-	-	-	-	-	-	-	-	0
qda	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6
rf	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6
ridge	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	Y	Y	-	Y	Y	-	Y	Y	-	Y	Y	-	8
xgb	Y	-	-	Y	-	Y	-	-	-	Y	Y	Y	5	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6

Table A18. *Cont.*

	WGB_CR													WOGB_CR											
	AUC			ACC			KS			F1			TOTAL	AUC			ACC			KS			F1			TOTAL
	F	F + P	P	F	F + P	P	F	F + P	P	F	F + P	P		F	F + P	P	F	F + P	P	F	F + P	P	F	F + P	P	
2. D99																											
cnn	Y	Y	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	10	Y	Y	-	Y	Y	Y	Y	Y	-	Y	Y	Y	9
ebm	-	Y	-	-	-	-	Y	Y	Y	Y	Y	Y	6	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
lda	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	11	Y	Y	-	Y	Y	-	Y	Y	-	Y	Y	-	7
mlp	-	Y	-	-	-	-	-	Y	-	-	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	0
percep	Y	-	-	-	-	-	Y	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	0
qda	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6
rf	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6
ridge	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	Y	Y	-	Y	Y	-	Y	Y	-	Y	Y	-	8
xgb	-	Y	-	-	Y	-	-	-	-	Y	Y	Y	4	-	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	9
ffn	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	-	Y	Y	-	-	Y	-	Y	Y	-	5
ffnlp	-	-	-	-	-	-	-	-	-	-	-	-	0	-	Y	-	-	-	-	-	Y	-	-	-	-	2
3. D80																											
cnn	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	-	Y	-	Y	Y	Y	-	Y	-	Y	Y	Y	7
ebm	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
ffnlp	-	Y	-	-	-	-	-	Y	-	-	-	-	2	-	Y	-	-	-	-	-	Y	-	-	-	-	2
lda	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
qda	-	-	-	Y	-	Y	-	-	-	Y	Y	Y	5	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6
rf	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6	-	-	-	Y	Y	Y	-	-	-	Y	Y	Y	6
ridge	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
mlp	-	-	-	-	-	-	-	-	-	-	-	-	0	-	-	-	-	-	Y	-	-	-	-	-	-	1
ffnsp	-	-	-	-	-	-	-	-	-	-	-	-	0	-	Y	-	-	-	-	-	Y	-	-	-	-	2

Note: This table reports the machine learning models (*CNN*, *EBM*, *FFN*, etc.) that exhibited statistically significant improvements in their classification metrics (AUC and ACC) when EF related to COVID-19 MOV were included. This analysis introduced a temporal shift in mortality data, considering scenarios with and without government-backed guarantees (*WGB_CR* and *WOGB_CR*). Key findings include the following: 1. Tree-based models (*XGB*, *RF*): These models demonstrate consistent and robust performance, achieving statistical significance in all relevance groups (F, F + P, P) under *WGB_CR* and most scenarios under *WOGB_CR*. The temporal shift appears to further stabilize performance. 2. Neural network models (*CNN*): *Cnn* consistently achieves statistical significance across all scenarios under both datasets (*WGB_CR* and *WOGB_CR*), underscoring its adaptability to enriched datasets and the temporal shift introduced by the COVID MOV. 3. Classical models (*LDA*, *RIDGE*): These models exhibit stable and statistically significant improvements, particularly in scenarios enriched with external factors. Their performance remains robust across most scenarios, with or without government-backed guarantees. 4. Scenario-specific improvements: Models such as *FFN*, *FFNLP*, and *MLP* exhibit improvements in specific scenarios, indicating higher sensitivity to dataset composition and relevance group combinations. 5. Impact of dataset composition: The exclusion of government-backed guarantees (*WOGB_CR*) amplifies the performance variability across models, particularly for neural networks and classical methods, suggesting a nuanced dependency on data characteristics to achieve statistical significance. 6. Temporal shift effects: The inclusion of a temporal shift in the COVID mortality data (*MOV*) enhances the ability of certain models (*CNN*, *XGB*) to identify statistically significant improvements, highlighting the importance of temporally adjusted external factors in predictive modeling. Overall, the findings underscore the critical role of enriched datasets and temporal adjustments in improving predictive accuracy, particularly for neural networks and tree-based models, while highlighting model-specific sensitivities to data composition and experimental design.

Table A19. Models with statistically significant improvements for AUC and ACC metrics with SOCIAL UNREST and CLIMATE CHANGE EF, considering WGB_CR, by evaluated scenario and relevance groups—Scenarios.

	WGB_CR														
	AUC							ACC							TOTAL
	T	**T + W**	**U**	**U + T**	**U + T + W**	**U + W**	**W**	**T**	**T + W**	**U**	**U + T**	**U + T + W**	**U + W**	**W**	
1. D100															
cnn	Y	-	Y	-	-	-	Y	Y	Y	Y	Y	Y	Y	Y	10
ffn	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	1
ffnsp	-	-	Y	-	-	Y	-	-	Y	-	-	-	-	-	3
lda	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
percep	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	1
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	Y	3
rf	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	7
ridge	-	-	-	-	-	-	-	-	-	-	Y	Y	Y	-	3
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xnn	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	1
2. D99															
cnn	-	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	6
ffnlp	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
ffnsp	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	1
lda	-	-	-	-	-	-	-	-	-	-	-	Y	Y	-	2
mlp	-	-	-	-	Y	-	-	Y	Y	-	-	-	Y	Y	5
percep	Y	-	-	Y	-	Y	-	-	-	-	-	-	-	-	3
qda	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
rf	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	7
ridge	-	-	-	-	Y	-	-	-	Y	Y	Y	Y	Y	Y	7
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xnn	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
3. D80															
cnn	-	-	Y	-	-	-	-	-	Y	Y	Y	Y	Y	Y	7
ebm	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
ffn	-	-	Y	-	-	Y	-	-	-	-	-	-	-	-	2
ffnlp	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	1
ffnsp	-	-	-	-	-	-	-	Y	Y	-	-	-	-	-	2
lda	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
mlp	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
percep	Y	-	-	-	Y	-	-	-	-	-	-	-	-	-	2
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	Y	3
rf	Y	Y	Y	Y	-	-	Y	Y	Y	Y	Y	Y	Y	Y	12
ridge	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xnn	-	-	-	-	-	-	-	-	Y	Y	Y	-	Y	-	4

Note: This table reports the machine learning models (*CNN*, *EBM*, *FFN*, etc.) that exhibited statistically significant improvements in their classification metrics (AUC and ACC) when EF related to social unrest and climate change were included. The evaluation considered scenarios in the presence of government-backed guarantees (*WGB_CR*). Key findings include the following: 1. Tree-based models (*XGB*, *RF*): These models demonstrate consistent and robust performance across most scenarios, achieving statistical significance in all relevance groups (T, U, W, and their combinations). Notably, *XGB* achieved statistical significance in all the evaluated scenarios, highlighting its adaptability to diverse external factors. 2. Neural network models (*CNN*): *Cnn* achieved statistically significant improvements across multiple scenarios, particularly in combinations of temperature (T) and other external factors. Its performance underscores its sensitivity to enriched datasets and potential to adapt to complex external factor scenarios. 3. Classical models (*LDA*, *RIDGE*): While *LDA* showed limited significance in this analysis, *RIDGE* exhibited consistent improvements in scenarios that included climate-related factors (T + W, U + T + W) and achieved statistical significance across all combinations, demonstrating its robustness when government-backed guarantees were present. 4. Scenario-specific improvements: Models such as *EBM*, *FFN*, and *FFNLP* demonstrate significance in the selected scenarios. For instance, *EBM* consistently achieved significance in scenarios involving a combination of social unrest and climate factors, suggesting context-dependent performance. 5. Impact of external factors (W, T): The inclusion of climate-related factors (T) and social unrest (U) enhanced the predictive capabilities of several models, particularly tree-based and neural network approaches. Models that combined these factors (U + T + W) exhibited notable improvements in statistical performance. 6. Comprehensive performance of *XGB* and *RIDGE*: Both models consistently demonstrated robust results across all relevance groups and scenarios, making them highly reliable for predictive tasks involving external factors such as social unrest and climate change. These findings highlight the importance of incorporating external factors into predictive modeling frameworks. In particular, tree-based models and neural networks benefit from these enriched datasets, achieving significant improvements in their predictive accuracy and generalization performance under complex scenarios.

Table A20. Models with statistically significant improvements for AUC and ACC metrics with SOCIAL UNREST and CLIMATE CHANGE EF, considering WOGB_CR, by evaluated scenario and relevance groups—Scenarios.

	WOGB_CR														
	AUC							ACC							TOTAL
	T	T + W	U	U + T	U + T + W	U + W	W	T	T + W	U	U + T	U + T + W	U + W	W	
1. D100															
cnn	-	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	6
ffn	-	-	Y	-	-	-	-	-	-	-	-	-	-	-	1
ffnsp	-	Y	Y	-	-	-	-	-	-	-	-	-	-	-	2
lda	-	Y	Y	Y	Y	Y	-	-	-	-	-	Y	-	-	6
percep	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	Y	3
rf	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	7
ridge	-	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	12
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xnn	-	-	-	-	-	-	-	-	Y	-	-	-	-	-	1
2. D99															
cnn	-	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	6
ffnlp	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	-	Y	6
ffnsp	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
lda	-	-	-	-	-	-	-	-	Y	-	-	Y	-	-	2
mlp	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
percep	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	1
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	Y	3
rf	Y	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	8
ridge	-	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	-	5
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xnn	-	Y	-	-	-	Y	-	-	-	Y	Y	-	-	-	4
3. D80															
cnn	-	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	6
ebm	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
ffn	-	-	-	-	-	-	-	-	Y	-	-	Y	-	-	2
ffnlp	-	-	-	-	-	Y	-	-	Y	Y	-	-	-	-	3
ffnsp	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
lda	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
mlp	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	1
percep	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	Y	3
rf	Y	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	Y	9
ridge	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xnn	-	-	-	-	-	-	-	-	Y	-	Y	-	-	-	2

Note: This table reports the machine learning models (*CNN*, *EBM*, *FFN*, etc.) that exhibited statistically significant improvements in their classification metrics (AUC and ACC) when EF related to social unrest and climate change were included. The evaluation considers scenarios without government-backed guarantees (*WOGB_CR*). Key findings include the following: 1. Tree-based models (*XGB*, *RF*): These models demonstrate consistent and robust performance across most scenarios, achieving statistical significance in all relevance groups (T, U, W, and their combinations). Notably, *XGB* achieved statistical significance in all the evaluated scenarios, further confirming its adaptability to diverse external factors. 2. Neural network models (*CNN*): *Cnn* achieve statistically significant improvements in scenarios incorporating temperature-related factors (T) and their combinations, particularly in the presence of external factors such as social unrest (U). However, its overall performance slightly decreases compared to the *WGB_CR* dataset. 3. Classical models (*LDA*, *RIDGE*): - *Ridge* continues to show consistent improvements across most scenarios, achieving statistical significance in a variety of combinations, especially those involving temperature (T + W, U + T + W). - *Lda* exhibits moderate performance under *WOGB_CR* scenarios, with improvements observed primarily in social unrest and climate change combinations (U + T + W). 4. Scenario-specific improvements: - Models such as *EBM*, *FFN*, and *FFNLP* demonstrate limited statistical significance, with their performance varying considerably across different scenarios. This suggests that these models are highly dependent on the context and specific dataset composition. 5. Impact of external factors (W, T, U): The inclusion of combined factors, particularly temperature (T) and social unrest (U), significantly enhance the predictive capabilities of models, such as *XGB* and *RF*. The *WOGB_CR* dataset amplifies the variability of improvements, particularly for neural networks and classical models. 6. Superior performance of *XGB* and *RIDGE*: Both models are the most reliable, achieving statistical significance across all relevance groups and scenarios. Their robustness and adaptability make them excellent candidates for predictive tasks in datasets without government-backed guarantees. These findings underscore the importance of incorporating external factors into predictive modeling frameworks, particularly in datasets with complex characteristics such as the absence of government-backed guarantees. Tree-based models and neural networks demonstrate the most significant benefits from these enriched datasets, further emphasizing their value for addressing challenging prediction tasks.

Table A21. Models with statistically significant improvements for KS and F1 metrics with climate change EF, considering WGB_CR, by evaluated scenario and relevance groups.

	WGB_CR															
	KS								F1							TOTAL
	T	T + W	U	U + T	U + T + W	U + W	W	T	T + W	U	U + T	U + T + W	U + W	W	
1. D100																
cnn	-	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	6
ebm	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
ffnlp	-	-	-	-	-	Y	Y	Y	-	-	-	Y	Y	-	5
ffnsp	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	1
mlp	-	-	-	-	-	-	-	-	Y	Y	-	-	-	-	2
percep	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	1
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	Y	3
rf	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	7
ridge	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	7
xgb	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	12
2. D99																
cnn	Y	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	7
ffnlp	-	-	-	-	-	-	-	Y	-	-	Y	Y	-	Y	4
ffnsp	-	Y	-	-	-	-	-	-	-	-	-	-	-	-	1
lda	-	-	-	-	Y	Y	-	-	Y	-	Y	Y	Y	-	6
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	-	2
rf	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	7
ridge	-	-	-	-	Y	Y	-	-	Y	Y	Y	Y	Y	Y	8
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
3. D80																
cnn	-	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	6
ebm	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
ffnlp	-	-	-	-	-	Y	-	-	-	-	-	-	-	-	1
lda	-	-	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	11
mlp	-	-	-	-	-	-	-	-	Y	-	Y	-	-	-	2
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	Y	3
rf	-	-	Y	-	-	-	Y	Y	Y	Y	Y	Y	Y	Y	9
ridge	-	-	-	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	10
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14

Note: This table reports the machine learning models that exhibited statistically significant improvements in KS and F1 metrics when climate change external factors were included. Key findings include the following: 1. Top Performers: - *EBM* and *XGB* achieve perfect scores (14/14) in the 3. D80 scenario. - *RIDGE* and *LDA* show strong consistency across scenarios. 2. Scenario Variations: - 1. D100: *EBM* dominates (14/14), while *CNN* only shows F1 improvements. - 2. D99: *XGB* maintains perfect performance, while *FFNLP* shows scenario-specific significance. - 3. D80: Most models perform better, especially in F1 metrics. 3. Metric Differences: - KS improvements are less common than F1. - *XGB* shows the most balanced performance across both metrics. 4. New Models: - Added *FFNLP* and *FFNSP* show specific scenario significance. - Included *MLP* demonstrates limited but notable improvements.

Table A22. Models with statistically significant improvements for KS and F1 metrics with climate change EF, considering WOGB_CR, by evaluated scenario and relevance groups.

	WOGB_CR															
	KS								F1							TOTAL
	T	T + W	U	U + T	U + T + W	U + W	W	T	T + W	U	U + T	U + T + W	U + W	W	
1. D100																
cnn	-	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	6
ebm	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
lda	-	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	-	11
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	Y	3
rf	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	7
ridge	-	Y	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y	12
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
2. D99																
cnn	Y	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	7
ebm	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
lda	-	-	-	-	Y	Y	-	-	Y	-	Y	Y	Y	-	6
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	-	2
rf	Y	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	8
ridge	-	-	-	-	Y	Y	-	-	Y	Y	Y	Y	Y	Y	8
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
ffn	-	-	-	-	-	-	-	Y	Y	-	-	Y	-	-	3

Table A22. *Cont.*

	WOGB_CR														
	KS							F1							TOTAL
	T	T + W	U	U + T	U + T + W	U + W	W	T	T + W	U	U + T	U + T + W	U + W	W	
3. D80															
cnn	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	7
ebm	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
ffnlp	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	1
ffnsp	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	1
knn	-	-	-	-	-	Y	Y	-	-	-	-	-	-	-	2
lda	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
mlp	-	-	-	-	-	-	-	-	Y	Y	-	-	Y	Y	4
qda	-	-	-	-	-	-	-	Y	Y	-	-	-	-	Y	3
rf	-	-	-	-	-	-	-	Y	Y	Y	Y	Y	Y	Y	7
ridge	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xgb	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	14
xnn	-	-	-	-	-	-	-	Y	Y	-	-	-	-	-	2

Note: This table reports the machine learning models that exhibited statistically significant improvements in KS and F1 metrics when climate change external factors were included, without government-backed guarantees. Key observations are the following: 1. Consistent Top Performers: - *EBM*, *XGB*, and *RIDGE* maintain perfect scores (14/14) in multiple scenarios. - *LDA* shows particularly strong performance in the 3. D80 scenario. 2. Dataset Impact: - WOGB_CR shows more variability in model performance compared to WGB_CR. - *KNN* and *XNN* appear only in WOGB_CR scenarios. 3. Metric Patterns: - F1 improvements are more frequent than KS across all models. - *RF* shows better F1 than KS performance consistently. 4. New Findings: - *FFNLP* and *FFNSP* show scenario-specific significance. - *MLP* demonstrates limited but notable improvements in 3. D80. - *KNN* appears as a new model with specific scenario significance. 5. Scenario Analysis: - 3. D80 shows the most consistent improvements across models. - 1. D100 has the fewest KS improvements overall. - 2. D99 reveals interesting variations in model performance.

Appendix B. Categorical Variable Mappings

Table A23. Mapping for credit portfolio segment.

Code	Description	Count	%
3	Consumer: Individual retail clients with personal loans	254,694	69.33
1	Microenterprise: Clients with small business activities	77,437	21.08
2	Small enterprise: SMEs with higher credit exposure than microenterprises	35,106	9.56
4	Corporate: Large-scale borrowers or commercial entities	109	0.03

Table A24. Mapping for gender.

Code	Description	Count	%
2	Female	225,489	61.38
1	Male	141,857	38.62

Table A25. Mapping for occupation code.

Code	Description	Count	%
1	Primary self-employed occupation (e.g., informal or freelance worker)	201,792	54.93
4	Employee in private or public sector	105,709	28.78
12	Independent professional (e.g., technician and artisan)	39,932	10.87
11	Merchant or small business owner	11,945	3.25
–	Other values (codes 2, 3, 5, 6, 7, 8, 9, 10)	8968	2.44

Table A26. Mapping for education level.

Code	Description	Count	%
1	Data not available	209,809	57.11
5	Complete primary education	55,491	15.11
8	Complete secondary education	53,552	14.58
6	Incomplete primary education	17,122	4.66
4	Non-university technical education	14,686	4.00
3	Incomplete university education	14,665	3.99
7	Complete university education	1015	0.28
2	Incomplete secondary education	1006	0.27

Table A27. Mapping for employment code.

Code	Description	Count	%
15	Merchant/Shopkeeper	267,270	72.76
19	Technician/Technologist	22,470	6.12
16	Street Vendor or Salesperson	17,692	4.82
31	Gas Fitter/Plumber	11,628	3.17
80	Upholsterer	6957	1.89
66	Employee (general services)	4889	1.33
–	Other values (codes 1–94)	36,440	9.91

Table A28. Mapping for residence ownership code.

Code	Description	Count	%
3	Rented property	247,847	67.47
2	Owned home	95,633	26.03
4	Family-owned residence (shared)	23,863	6.50
1	Missing or unclassified	3	0.00

Table A29. Mapping for residence type code.

Code	Description	Count	%
9	Multifamily housing or apartment	142,378	38.76
5	Detached or single-family house	121,216	33.00
10	Shared family housing	55,366	15.07
7	Informal housing or makeshift dwelling	23,913	6.51
11	Non-residential property used as home	23,143	6.30
–	Other values (codes 1, 2, 3, 4, 6, 8)	1330	0.36

References

1. Liu, C.; Ming, Y.; Xiao, Y.; Zheng, W.; Hsu, C. Finding the next interesting loan for investors on a peer-to-peer lending platform. *IEEE Access* **2021**, *9*, 111293–111304. [CrossRef]
2. Lombardo, G.; Pellegrino, M.; Adosoglou, G.; Cagnoni, S.; Pardalos, P.M.; Poggi, A. Machine learning for bankruptcy prediction in the American stock market: Dataset and benchmarks. *Future Internet* **2022**, *14*, 244. [CrossRef]
3. Wen, C.; Yang, J.; Gan, L.; Pan, Y. Big data driven internet of things for credit evaluation and early warning in finance. *Future Gener. Comput. Syst.* **2021**, *124*, 295–307. [CrossRef]
4. Shih, D.H.; Wu, T.W.; Shih, P.Y.; Lu, N.A.; Shih, M.H. A framework of global credit-scoring modeling using outlier detection and machine learning in a P2P lending platform. *Mathematics* **2022**, *10*, 2282. [CrossRef]
5. Mousavi, M.M.; Lin, J. The application of PROMETHEE multi-criteria decision aid in financial decision making: Case of distress prediction models evaluation. *Expert Syst. Appl.* **2020**, *159*, 113438. [CrossRef]
6. Hani, U.; Wickramasinghe, A.; Kattiyapornpong, U.; Shahriar, S. The future of data-driven relationship innovation in the microfinance industry. *Ann. Oper. Res.* **2022**, *333*, 971–997. [CrossRef]

7. Chen, Z.; Chen, W.; Shi, Y. Ensemble learning with label proportions for bankruptcy prediction. *Expert Syst. Appl.* **2020**, *146*, 113155. [CrossRef]
8. Orlova, E.V. Decision-making techniques for credit resource management using machine learning and optimization. *Information* **2020**, *11*, 144. [CrossRef]
9. Sun, M.; Li, Y. Credit risk simulation of enterprise financial management based on machine learning algorithm. *Mob. Inf. Syst.* **2022**, *2022*, 9007140. [CrossRef]
10. Chen, S.F.; Chakraborty, G.; Li, L.H. Feature selection on credit risk prediction for peer-to-peer lending. In *New Frontiers in Artificial Intelligence: JSAI-isAI 2018 Workshops, JURISIN, AI-Biz, SKL, LENLS, IDAA, Yokohama, Japan, 12–14 November 2018, Revised Selected Papers*; Springer: Cham, Switzerland, 2019; pp. 5–18.
11. Telg, S.; Dubinova, A.; Lucas, A. COVID-19, credit risk management modeling, and government support. *J. Bank. Financ.* **2023**, *147*, 106638. [CrossRef]
12. Ahmed, I. Essays on Climate-Related Financial Risks. Ph.D. Dissertation, University of Otago, Dunedin, New Zealand, 2023.
13. Cho, S.H.; Shin, K. Feature-weighted counterfactual-based explanation for bankruptcy prediction. *Expert Syst. Appl.* **2023**, *216*, 119390. [CrossRef]
14. Mancisidor, R.A.; Kampffmeyer, M.; Aas, K.; Jenssen, R. Learning latent representations of bank customers with the variational autoencoder. *Expert Syst. Appl.* **2021**, *164*, 114020. [CrossRef]
15. Wang, T.; Liu, R.; Qi, G. Multi-classification assessment of bank personal credit risk based on multi-source information fusion. *Expert Syst. Appl.* **2022**, *191*, 116236. [CrossRef]
16. Chen, C.; Lin, K.; Rudin, C.; Shaposhnik, Y.; Wang, S.; Wang, T. A holistic approach to interpretability in financial lending: Models, visualizations, and summary-explanations. *Decis. Support Syst.* **2022**, *152*, 113647. [CrossRef]
17. Pavitha, N.; Sugave, S. Explainable Multistage Ensemble 1D Convolutional Neural Network for Trust Worthy Credit Decision. *Int. J. Adv. Comput. Sci. Appl.* **2024**, *15*, 351–358.
18. Nallakaruppan, M.K.; Chaturvedi, H.; Grover, V.; Balusamy, B.; Jaraut, P.; Bahadur, J.; Meena, V.P.; Hameed, I.A. Credit Risk Assessment and Financial Decision Support Using Explainable Artificial Intelligence. *Risks* **2024**, *12*, 164. [CrossRef]
19. Bastos, J.A.; Matos, S.M. Explainable models of credit losses. *Eur. J. Oper. Res.* **2022**, *301*, 386–394. [CrossRef]
20. BBC News Mundo. Muertes por COVID-19: El Gráfico que Muestra los 10 Primeros Países del Mundo en el Ranking de Fallecimientos per Cápita (y Cuáles son de América Latina), BBC News Mundo. 2020. Available online: https://www.bbc.com/mundo/noticias-54358383 (accessed on 26 December 2024).
21. CNN en Español, Protestas Violentas en Perú tras Rechazo de Adelantar Elecciones. *CNN en Español*, Video. 2024. Available online: https://cnnespanol.cnn.com/video/protestas-violenta-lima-congreso-rechazo-elecciones-crisis-politica-jimena-de-la-quintana-mirador-mundial (accessed on 26 December 2024).
22. Ye, Y.F.; Jiang, Y.X.; Shao, Y.H.; Li, C.N. Financial conditions index construction through weighted Lp-norm support vector regression. *J. Adv. Comput. Intell. Intell. Inform.* **2015**, *19*, 397–406. [CrossRef]
23. Noriega, J.P.; Rivera, L.A.; Herrera, J.A. Machine learning for credit risk prediction: A systematic literature review. *Data* **2023**, *8*, 169. [CrossRef]
24. Aranha, M.; Bolar, K. Efficacies of artificial neural networks ushering improvement in the prediction of extant credit risk models. *Cogent Econ. Financ.* **2023**, *11*, 2210916. [CrossRef]
25. Merćep, A.; Mrčela, L.; Birov, M.; Kostanjčar, Z. Deep neural networks for behavioral credit rating. *Entropy* **2020**, *23*, 27. [CrossRef]
26. Fan, S.; Shen, Y.; Peng, S. Improved ML-based technique for credit card scoring in internet financial risk control. *Complexity* **2020**, *2020*, 8706285. [CrossRef]
27. Geraldo-Campos, L.A.; Soria, J.J.; Pando-Ezcurra, T. Machine learning for credit risk in the Reactiva Perú Program: A comparison of the Lasso and Ridge regression models. *Economies* **2022**, *10*, 188. [CrossRef]
28. Quintanilla Llanos, M.A. *Efecto del desempleo en la morosidad en portafolios de tarjetas de crédito ante el impacto de la pandemia (2011–2023)*; Universidad Peruana de Ciencias Aplicadas (UPC): Santiago de Surco, Peru, 2024.
29. Ontón, R.; Reilly, B. *Dynamic Between Business and Personal Debt in a Group of Microenterprises in Peru, in the Context of the COVID-19 Pandemic*; Superintendencia de Banca, Seguros y AFP del Perú (SBS): San Isidro, Peru, 2024.
30. Flori, A.; Pammolli, F.; Spelta, A. Commodity prices co-movements and financial stability: A multidimensional visibility nexus with climate conditions. *J. Financ. Stab.* **2021**, *54*, 100876. [CrossRef]
31. Capasso, G.; Gianfrate, G.; Spinelli, M. Climate change and credit risk. *J. Clean. Prod.* **2020**, *266*, 121634. [CrossRef]
32. Nieto, M.J. Banks, climate risk and financial stability. *J. Financ. Regul. Compliance* **2019**, *27*, 243–262. [CrossRef]
33. Fayyaz, M.R.; Rasouli, M.R.; Amiri, B. A data-driven and network-aware approach for credit risk prediction in supply chain finance. *Ind. Manag. Data Syst.* **2020**, *121*, 785–808. [CrossRef]
34. Acharya, V.V.; Berner, R.; Engle, R.; Jung, H.; Stroebel, J.; Zeng, X.; Zhao, Y. Climate stress testing. *Annu. Rev. Financ. Econ.* **2023**, *15*, 291–326. [CrossRef]

35. Calsin Enriquez, H.O.; Vargas Salazar, I.Y. Monetary policy and microfinance credit risk: The case of Peru, period 2003–2021. In *Pensamiento Crítico*; Facultad de Ciencias Económicas, Universidad Nacional Mayor de San Marcos: Lima, Peru, 2023.
36. Madeira, C. The double impact of deep social unrest and a pandemic: Evidence from Chile. *Can. J. Econ. Can. d'économique* **2022**, *55*, 135–171. [CrossRef]
37. Espinosa-Méndez, C. Civil unrest and herding behavior: Evidence in an emerging market. *Econ.-Res.-Ekon. Istraz.* **2022**, *35*, 1243–1261. [CrossRef]
38. Dendramis, Y.; Tzavalis, E.; Adraktas, G. Credit risk modelling under recessionary and financially distressed conditions. *J. Bank. Financ.* **2018**, *91*, 160–175. [CrossRef]
39. Elkhayat, N.; ElBannan, M.A. State divestitures and bank performance: Empirical evidence from the Middle East and North Africa region. *Asian Econ. Financ. Rev.* **2018**, *8*, 145. [CrossRef]
40. Kemp, L.; Xu, C.; Depledge, J.; Ebi, K.L.; Gibbins, G.; Kohler, T.A.; Rockström, J.; Scheffer, M.; Schellnhuber, H.J.; Steffen, W.; et al. Climate endgame: Exploring catastrophic climate change scenarios. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2108146119. [CrossRef] [PubMed]
41. Aguilar-Valenzuela, G.R.; Vilca, E.M. Algoritmos para Machine Learning utilizados en la Gestión de Riesgo Crediticio en Perú. *Micaela Rev. De Investig.-UNAMBA* **2024**, *5*, 30–35. [CrossRef]
42. Alzamora, G.S.; Aceituno-Rojo, M.R.; Condori-Alejo, H.I. An assertive machine learning model for rural micro credit assessment in Peru. *Procedia Comput. Sci.* **2022**, *202*, 301–306. [CrossRef]
43. Marbán, O.; Mariscal, G.; Segovia, J. A Data Mining & Knowledge Discovery Process Model. In *Data Mining and Knowledge Discovery in Real Life Applications*; Ponce, J., Karahoca, A., Eds.; IntechOpen: Rijeka, Croatia, 2009; Chapter 1. [CrossRef]
44. Mann, S.C.; Logeswaran, R. Data analytics in improved bankruptcy prediction with industrial risk. In Proceedings of the 2021 14th International Conference on Developments in eSystems Engineering (DeSE), Sharjah, United Arab Emirates, 7–10 December 2021; pp. 23–26.
45. Sahiq, A.N.M.; Ismail, S.; Nor, S.H.S.; Ul-Saufie, A.Z.; Yaacob, W.F.W. Application of logistic regression model on imbalanced data in personal bankruptcy prediction. In Proceedings of the 2022 3rd International Conference on Artificial Intelligence and Data Sciences (AiDAS), Ipoh, Malaysia, 7–8 September 2022; pp. 120–125.
46. Liu, J.; Zhang, S.; Fan, H. A two-stage hybrid credit risk prediction model based on XGBoost and graph-based deep neural network. *Expert Syst. Appl.* **2022**, *195*, 116624. [CrossRef]
47. Awad, I.M.; Karaki, M.S.A. The impact of bank lending on Palestine economic growth: An econometric analysis of time series data. *Financ. Innov.* **2019**, *5*, 14. [CrossRef]
48. Frazier, P.I. A tutorial on Bayesian optimization. *arXiv* **2018**, arXiv:1807.02811.
49. Pandey, M.K.; Mittal, M.; Subbiah, K. Optimal balancing & efficient feature ranking approach to minimize credit risk. *Int. J. Inf. Manag. Data Insights* **2021**, *1*, 100037.
50. Biswas, N.; Mondal, A.S.; Kusumastuti, A.; Saha, S.; Mondal, K.C. Automated credit assessment framework using ETL process and machine learning. *Innov. Syst. Softw. Eng.* **2022**, *2022*, 257–270. [CrossRef]
51. Bosker, J.; Gürtler, M.; Zöllner, M. Machine learning-based variable selection for clustered credit risk modeling. *J. Bus. Econ.* **2024**, *Online First*, 1–36. [CrossRef]
52. Nwafor, C.N.; Nwafor, O.; Brahma, S. Enhancing transparency and fairness in automated credit decisions: An explainable novel hybrid machine learning approach. *Sci. Rep.* **2024**, *14*, 25174. [CrossRef] [PubMed]
53. Yu, X.; Yang, Q.; Wang, R.; Fang, R.; Deng, M. Data cleaning for personal credit scoring by utilizing social media data: An empirical study. *IEEE Intell. Syst.* **2020**, *35*, 7–15. [CrossRef]
54. Alam, T.M.; Shaukat, K.; Hameed, I.A.; Luo, S.; Sarwar, M.U.; Shabbir, S.; Li, J.; Khushi, M. An investigation of credit card default prediction in the imbalanced datasets. *IEEE Access* **2020**, *8*, 201173–201198. [CrossRef]
55. Dumitrescu, E.; Hué, S.; Hurlin, C.; Tokpavi, S. Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. *Eur. J. Oper. Res.* **2022**, *297*, 1178–1192. [CrossRef]
56. Zheng, B. Financial default payment predictions using a hybrid of simulated annealing heuristics and extreme gradient boosting machines. *Int. J. Internet Technol. Secur. Trans.* **2019**, *9*, 404–425. [CrossRef]
57. Ma, Y.; Zhang, P.; Duan, S.; Zhang, T. Credit default prediction of Chinese real estate listed companies based on explainable machine learning. *Financ. Res. Lett.* **2023**, *58*, 104305. [CrossRef]
58. Chen, C.-C.; Liu, Z.; Yang, G.; Wu, C.-C.; Ye, Q. An Improved Fault Diagnosis Using 1D-Convolutional Neural Network Model. *Electronics* **2021**, *10*, 59. [CrossRef]
59. Lee, J.Y.; Yang, J.; Anderson, E. Who benefits from alternative data for credit scoring? Evidence from Peru. *SSRN* **2024**. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.