



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

**Modelo de explicabilidad para la deserción
universitaria a través de aprendizaje automático**

TESIS

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

AUTOR

Mauricio Alexander QUIMIZ MOREIRA

ASESOR

Dr. Hugo David CALDERÓN VILCA

Lima, Perú

2026



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

Quimiz, M. (2026). *Modelo de explicabilidad para la deserción universitaria a través de aprendizaje automático*. [Tesis de 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	Mauricio Alexander Quimiz Moreira
Tipo de documento de identidad	Cédula de identidad
Número de documento de identidad	EC/ 1312080755
URL de ORCID	https://orcid.org/0000-0002-5430-0215
Datos de asesor	
Nombres y apellidos	Dr. Hugo David Calderón Vilca
Tipo de documento de identidad	DNI
Número de documento de identidad	01317748
URL de ORCID	https://orcid.org/0000-0002-1177-4947
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Dr. Ciro Rodriguez Rodriguez
Tipo de documento	DNI
Número de documento de identidad	06020241
Miembro del jurado 1	
Nombres y apellidos	Dr. Pedro Martin Lezama Gonzales
Tipo de documento	DNI
Número de documento de identidad	09656793
Miembro del jurado 2	
Nombres y apellidos	Dr. Ivan Carlo Petrlik Azabache
Tipo de documento	DNI
Número de documento de identidad	10140461
Miembro del jurado 3	
Nombres y apellidos	Dr. Hugo David Calderón Vilca
Tipo de documento	DNI

Número de documento de identidad	01317748
Datos de investigación	
Línea de investigación	C.0.5.8. Inteligencia Artificial
Grupo de investigación	No aplica
Agencia de financiamiento	Sin financiamiento.
Ubicación geográfica de la investigación	País: Ecuador Provincia: Manabí Ciudad: Portoviejo Latitud: -1.051560 Longitud: -80.454189
Año o rango de años en que se realizó la investigación	2022 - 2025
URL de disciplinas OCDE	Ingeniería de Sistemas y comunicaciones https://purl.org/pe-repo/ocde/ford#2.02.04 Ciencias de la computación https://purl.org/pe-repo/ocde/ford#1.02.01



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 quince (15) días del mes de mayo de 2026, siendo las 11:00 am., se reunieron en el Auditorio Alfredo Celso Alva Bravo, el Jurado de Tesis conformado por los siguientes docentes:

Dr. Ciro Rodriguez Rodriguez (Presidente)
Dr. Ivan Carlo Petrlik Azabache (Miembro)
Dr. Pedro Martin Lezama Gonzales (Miembro)
Dr. Hugo David Calderón Vilca (Miembro Asesor)

Se inició la Sustentación invitando al candidato a Doctor **MAURICIO ALEXANDER QUIMIZ MOREIRA**, 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:

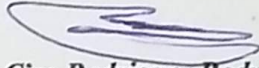
**“MODELO DE EXPLICABILIDAD PARA LA DESERCIÓN
UNIVERSITARIA A TRAVÉS DE APRENDIZAJE AUTOMÁTICO”**

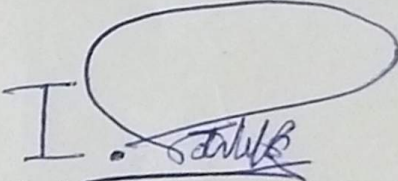
Concluida la exposición, los miembros del Jurado de Tesis procedieron a formular sus preguntas que fueron absueltas por el graduando; acto seguido se procedió a la evaluación correspondiente, habiendo obtenido la siguiente calificación:

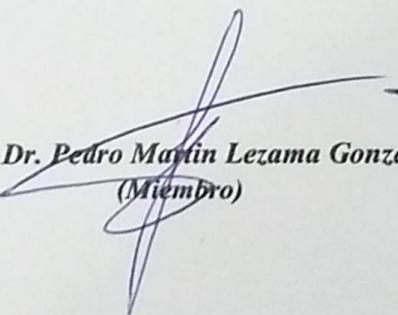
Dieciocho (18) *Muy Bueno*

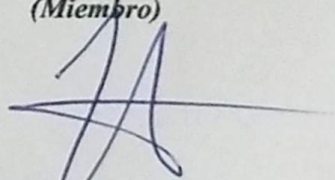
Por tanto, el presidente del Jurado, de acuerdo con el Reglamento General de Estudios de Posgrado, otorga al Magister **MAURICIO ALEXANDER QUIMIZ MOREIRA** el Grado Académico de Doctor en Ingeniería de Sistemas e Informática.

Siendo las *2:16* horas, el presidente del Jurado de Tesis, da por concluido el acto académico de Sustentación de Tesis.


Dr. Ciro Rodriguez Rodriguez
(Presidente)


Dr. Ivan Carlo Petrlik Azabache
(Miembro)


Dr. Pedro Martin Lezama Gonzales
(Miembro)


Dr. Hugo David Calderón Vilca
(Miembro Asesor)



Anexo 1

CERTIFICADO DE SIMILITUD

Yo, Hugo David Calderón Vilca en mi condición de asesor acreditado mediante Dictamen N° 000524-2025- UPG-VDIP-FISI/UNMSM de la tesis, cuyo título es **Modelo de explicabilidad para la deserción universitaria a través de aprendizaje automático**, presentado por el magíster Mauricio Alexander QUIMIZ MOREIRA para optar el grado Académico 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 4% 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 Académico de Doctor en Ingeniería de Sistemas e Informática

Firma del Asesor _____

Lima, 19 de marzo de 2026

Nombres y Apellidos del asesor: Hugo David Calderón Vilca

DNI: 01317748

* Según corresponda: bachiller, magíster, doctor; título profesional o título de segunda especialidad

La vigencia del presente certificado corresponde a un período de noventa (90) días desde la fecha de su emisión; siempre que no haya observaciones en el texto durante el proceso de revisión por el jurado informante o surja alguna contingencia.

DEDICATORIA

A Dios, por acompañarme en cada etapa de este proceso, por ser la luz que orientó mis decisiones y la fuente inagotable de fortaleza y sabiduría que sostuvo mi perseverancia hasta alcanzar esta meta. Sin su gracia y guía constante, este logro no habría sido posible.

A mi querida esposa, María Eugenia, por su amor incondicional, su apoyo constante y, sobre todo, por la paciencia con la que me acompañó durante todo este recorrido. Gracias por impulsarme cada día a no rendirme, por creer en este sueño y por caminar a mi lado como compañera de vida.

A mis hijos, Gabriel y Rafael, porque este logro también les pertenece; deseo que siempre encuentren en él una inspiración para perseguir sus sueños con determinación y valentía, recordando que no existen límites cuando se trabaja con fe y constancia.

El camino hacia la culminación de esta tesis no estuvo exento de desafíos, aciertos y aprendizajes. No fue un trayecto sencillo, pero gracias al apoyo de todos ustedes fue posible transformar cada obstáculo en un impulso y, finalmente, alcanzar este anhelado objetivo.

AGRADECIMIENTOS

Deseo expresar mi más sincero agradecimiento a todas las personas e instituciones que han sido parte esencial en la culminación de esta tesis doctoral.

A mi director de tesis, el Dr. Hugo Calderón, por su orientación, paciencia constante y compromiso académico. Su acompañamiento fue determinante no solo en la estructuración y finalización de esta investigación doctoral, sino también en el desarrollo de los artículos científicos derivados de ella. Agradezco profundamente su confianza en mi capacidad y la generosidad con la que compartió su experiencia y sus conocimientos. Extiendo igualmente mi reconocimiento a cada uno de los docentes que formaron parte de este proceso formativo. En cada asignatura encontré no solo nuevos saberes, sino también perspectivas críticas y experiencias académicas que fortalecieron mi formación investigativa. De manera especial, agradezco al Dr. Ciro Rodríguez, al Dr. David Mauricio Sánchez y al Dr. Iván Túpac Valdivia, cuyas enseñanzas contribuyeron de manera significativa al desarrollo de este trabajo.

De igual manera, mi agradecimiento al personal administrativo, que siempre estuvo presente durante el proceso de formación, brindando su apoyo constante y asesoramiento continuo.

A la Universidad Nacional Mayor de San Marcos, por ofrecerme un entorno académico exigente y estimulante, que promueve la excelencia, el pensamiento crítico y el crecimiento profesional. Haber formado parte de esta institución representa un honor y un privilegio en mi trayectoria académica.

ÍNDICE GENERAL

AGRADECIMIENTOS	III
LISTA DE TABLAS.....	VII
LISTA DE FIGURAS.....	VIII
RESUMEN	IX
ABSTRACT	XI
CAPÍTULO I. INTRODUCCIÓN	12
1.1. Situación Problemática	12
1.2. Formulación del Problema	15
1.2.1. Problema general.....	15
1.2.2. Problemas específicos.....	16
1.3 Justificación Teórica	16
1.4 Justificación Práctica	17
1.5. Objetivos de la Investigación	20
1.5.1. Objetivo General	20
1.5.2. Objetivos Específicos.....	20
1.6. Hipótesis e identificación de variables	21
1.6.1 Hipótesis general	21
1.6.2 Hipótesis específicas	21
CAPÍTULO II. MARCO TEORICO	22
2.1. Marco Epistemológico.....	22
2.2. Antecedentes de Investigación.....	23
2.3. Bases Teóricas.....	25
2.3.1 Deserción Estudiantil.....	26
2.3.2 Tipo de Deserción Estudiantil.....	28
2.3.3 Factores de deserción universitaria	29
2.3.4 Aprendizaje automático	30
2.3.5 Técnicas de Aprendizaje Automático.....	31
2.3.6 Predicción.....	32
2.3.7 Explicabilidad.....	33
2.4. Marcos Conceptuales o Glosario.....	36
2.4.1 Deserción universitaria	36
2.4.2 Retención estudiantil.....	37
CAPÍTULO III. METODOLOGÍA DE LA INVESTIGACIÓN.....	39
3.1. Tipo de investigación.....	39
3.1.1 Nivel de investigación.....	39
3.1.2 Tipo de investigación.....	40
3.1.3 Diseño de la investigación.....	40
3.1.4 Método de la investigación.	42

3.1.5	Unidad de análisis	43
3.1.6	Población de estudio.....	44
3.1.7	Selección de muestra	45
3.2.	Dataset y partición.....	46
3.2.1	Dataset	46
3.2.2	Partición	49
3.3	Operacionalización de variables.	49
3.3.1.	Variable independiente	49
3.3.2.	Variable dependiente	50
3.3.3.	Matriz de operacionalización de variables.....	50
3.4.	Técnicas e instrumentos de recolección de datos.....	52
3.4.1	Técnicas.....	52
3.4.2	Procedimientos de análisis y modelado.....	52
3.4.3	Validación y confiabilidad de los modelos de ML.....	54
3.4.4	Procedimientos	56
3.5.	Matriz De Consistencia	63
CAPÍTULO IV. FASE EXPERIMENTAL.....		65
4.1.	Análisis de dataset	65
4.1.1	Ventajas del dataset seleccionado	66
4.1.2	Limitaciones del dataset.....	66
4.1.3	Preparación y enriquecimiento del dataset.....	67
4.2.	Modelo Propuesto.....	70
4.2.1.	Investigación científica.....	71
4.2.2.	Construcción del dataset.....	73
4.2.3	Modelado predictivo	75
4.2.4	Explicabilidad del modelo	79
4.3.	Implementación de Modelos.....	80
4.4.	Métricas de evaluación del modelo	84
4.5.	Hiperparámetros de los modelos propuestos	86
CAPÍTULO V. RESULTADOS Y DISCUSIÓN		88
5.1.	Resultados de construcción de un conjunto de datos	88
5.1.1.	Selección y adquisición de los datos.....	88
5.1.2.	Comprensión y Análisis Exploratorio de los Datos (EDA).....	91
5.1.3.	Preprocesamiento	101
5.1.4.	Selección de características	109
5.1.5.	Hipótesis Específica 1	130
5.2.	Resultados del entrenar y evaluar algoritmos de predicción	132
5.2.1	División del conjunto de datos.....	133
5.2.2	Ejecución de algoritmos de predicción.....	134
5.2.3	Optimización del modelo predictivo con oversampling.....	156
5.2.4	Optimización mediante hiperparámetros	161
5.2.5.	Hipótesis Específica 2.....	168
5.3.	Resultados de diseño de modelo de explicabilidad	170
5.3.1	Explicabilidad Global	170
5.3.2	Explicabilidad Local	175

5.3.3. Hipótesis Específica 3	187
5.4. Discusión.....	189
CONCLUSIONES	197
RECOMENDACIONES	199
REFERENCIAS BIBLIOGRÁFICAS	200

LISTA DE TABLAS

Tabla 1	44
Tabla 2	46
Tabla 3	47
Tabla 4	51
Tabla 5	55
Tabla 6	64
Tabla 7	86
Tabla 8	88
Tabla 9	93
Tabla 10	99
Tabla 11	102
Tabla 12	105
Tabla 13	107
Tabla 14	114
Tabla 15	129
Tabla 16	134
Tabla 17	136
Tabla 18	138
Tabla 19	140
Tabla 20	140
Tabla 21	142
Tabla 22	143
Tabla 23	145
Tabla 24	147
Tabla 25	148
Tabla 26	150
Tabla 27	152
Tabla 28	154
Tabla 29	155
Tabla 30	157
Tabla 31	158
Tabla 32	160
Tabla 33	162
Tabla 34	163
Tabla 35	164
Tabla 36	175

LISTA DE FIGURAS

Figura 1.....	70
Figura 2.....	95
Figura 3.....	101
Figura 4.....	104
Figura 5.....	106
Figura 6.....	111
Figura 7.....	112
Figura 8.....	117
Figura 9.....	118
Figura 10.....	119
Figura 11.....	121
Figura 12.....	122
Figura 13.....	124
Figura 14.....	125
Figura 15.....	127
Figura 16.....	128
Figura 17.....	132
Figura 19.....	165
Figura 20.....	166
Figura 21.....	171
Figura 22.....	172
Figura 23.....	173
Figura 24.....	174
Figura 25.....	177
Figura 26.....	178
Figura 27.....	179
Figura 28.....	180
Figura 29.....	182
Figura 30.....	183
Figura 31.....	185
Figura 32.....	187

RESUMEN

La deserción estudiantil universitaria constituye una problemática compleja y persistente que afecta la eficiencia de los sistemas de educación superior, interrumpe las trayectorias formativas de los estudiantes y genera repercusiones académicas, sociales e institucionales. En este contexto, desarrolla un modelo de explicabilidad basado en el aprendizaje automático para predecir e interpretar los factores asociados a la deserción estudiantil universitaria. La investigación se desarrolló a partir de la información del sistema institucional de gestión académica. El conjunto de datos consolidado se conformó por 17 183 registros y, tras la depuración y validación de calidad, se trabajó con 36 variables sociodemográficas, económicas, institucionales, personales y académicas. Posteriormente, se aplican técnicas de preprocesamiento, tratamiento del desequilibrio de clases, selección de características, entrenamiento de modelos, optimización de hiperparámetros y evaluación mediante métricas robustas.

Los resultados evidenciaron que los modelos de aprendizaje automático permiten identificar con alta capacidad discriminativa a los estudiantes en riesgo de abandono. Entre los algoritmos evaluados, Extra Trees presentó el mejor desempeño global, con F1-score de 0.892, recall de 1.000, precision de 0.971, accuracy de 0.998 y ROC-AUC cercano a 1.000, superando a Random Forest y Logistic Regression en el equilibrio entre sensibilidad y precisión para la clasificación de la clase minoritaria. Estos hallazgos confirman que los métodos de ensamble basados en árboles constituyen una alternativa especialmente robusta para el análisis de la deserción en escenarios con datos desbalanceados.

El estudio concluye que el uso del aprendizaje automático, articulado con técnicas de explicabilidad, contribuye significativamente a la comprensión y la predicción de la deserción estudiantil universitaria, permitiendo identificar patrones y factores de riesgo

que no son evidentes a simple vista y resultan fundamentales para orientar acciones de intervención y de permanencia estudiantil. Este enfoque fue desarrollado en el contexto de la educación superior, lo que demuestra su utilidad para apoyar la toma de decisiones institucionales con mayor precisión, transparencia y sustento analítico. Asimismo, su estructura metodológica puede extenderse a otras facultades, carreras e incluso a distintas instituciones de educación superior, incorporando nuevas variables significativas que continúen fortaleciendo y optimizando la capacidad predictiva y explicativa del modelo.

Palabras clave: deserción estudiantil universitaria, aprendizaje automático, inteligencia artificial explicable, predicción, educación superior, permanencia estudiantil.

ABSTRACT

University student dropout is a complex and persistent problem that affects the efficiency of higher education systems, interrupts students' educational trajectories, and has academic, social, and institutional repercussions. In this context, a model of explainability based on machine learning is developed to predict and interpret the factors associated with university student dropout. The research was developed using information from the institutional academic management system. The consolidated dataset consisted of 17,183 records and, after cleaning and quality validation, 36 sociodemographic, economic, institutional, personal, and academic variables were used. Subsequently, pre-processing techniques, class imbalance treatment, feature selection, model training, hyperparameter optimization, and evaluation using robust metrics were performed. The results showed that machine learning models can identify students at risk of dropping out with high discriminatory power.

The study concludes that the use of machine learning, combined with explainability techniques, significantly contributes to understanding and predicting university student dropout rates, enabling the identification of patterns and risk factors that are not immediately apparent and are essential for guiding intervention and student retention actions. This approach was developed in the context of higher education, demonstrating its usefulness in supporting institutional decision-making with greater precision, transparency, and analytical support. Likewise, its methodological structure can be extended to other faculties, degree programs, and even other higher education institutions, incorporating new significant variables that continue to strengthen and optimize the model's predictive and explanatory capacity.

Keywords: university student dropout, machine learning, explainable artificial intelligence, prediction, higher education, student retention.

CAPÍTULO I. INTRODUCCIÓN

1.1. Situación Problemática

La deserción estudiantil en las instituciones de educación superior (IES) públicas es un fenómeno persistente que afecta de manera significativa tanto a las instituciones educativas como a la sociedad en general, con un impacto social y económico considerable. En los últimos años, su incidencia ha aumentado, lo que lo convierte en un asunto clave tanto para la investigación educativa como para el diseño de políticas públicas. Esta inquietud se afianza ante los niveles persistentemente elevados de abandono en las IES, lo que lleva a distintos actores académicos a buscar respuestas más eficaces y ajustadas a cada contexto. Como señalan (Rodríguez-Muñiz et al., 2019), la deserción sigue siendo uno de los mayores retos que las IES deben sortear para asegurar que los estudiantes permanezcan y finalicen con éxito su formación.

Aunque los primeros trabajos sobre el tema aparecieron en la década de 1970, (Kabathova & Drlik, 2021), sostienen que fue en los años noventa cuando el problema cobró verdadera urgencia para las casas de estudio, impulsado por el crecimiento sostenido de la matrícula y por las repercusiones negativas que conlleva el abandono. En ese sentido, la (UNESCO, 2023) indica que existen 264 millones de estudiantes matriculados en las IES en el mundo, lo que representa el doble de los inicios del siglo XXI. Aunque esto indica un crecimiento notable en el acceso a la educación superior, también plantea nuevos retos en torno a la equidad, la calidad y la permanencia estudiantil.

Desde una óptica teórica, (Villegas & Núñez Lira, 2024) argumentan que la deserción no puede analizarse solo en términos académicos, sino que requiere considerar factores sociales, emocionales e institucionales que determinan si un estudiante permanece o se

va. Para ellos, el abandono surge de la interacción activa entre las características personales del individuo y las condiciones que ofrece la universidad.

En este marco, (Gallego et al., 2021) identifican el riesgo académico, entendido como déficit de rendimiento o de vinculación con el estudio, como uno de los indicadores más confiables de abandono. No obstante, los autores lamentan que este asunto aún no cuente con la investigación exhaustiva que su relevancia exige. En línea con esta crítica, (Cortés Cáceres et al., 2019) señalan que incluso quienes mantienen buenas calificaciones pueden desertar si las metodologías docentes resultan aburridas o desactualizadas, ya que ello mina su motivación y los aleja del proceso formativo.

Según (Henriquez Cabezas & Vargas Escobar, 2022) , si bien han surgido numerosos estudios y modelos encaminados a prever la deserción estudiantil, la mayoría de ellos alcanza una precisión que no supera el 80 %. Según ellos, dicha limitación radica en el manejo deficiente de las variables que alimentan el análisis, lo que repercute de inmediato en la fiabilidad del pronóstico.

Ante esta situación, (Padmasiri & Kasthuriarachchi, 2024) defienden que las técnicas de aprendizaje automático (ML) constituyen una herramienta más sólida para explorar grandes conjuntos de datos y descubrir patrones ocultos que los métodos estadísticos clásicos pasarían por alto. Gracias a esos algoritmos, es posible no solo estimar con mayor certeza qué alumnos están en riesgo de abandonar, sino también desglosar los factores involucrados: rendimiento académico, condiciones económicas, integración social, ambiente familiar y grado de participación en actividades externas.

En un estudio reciente (Nnadi et al., 2024) se subraya la necesidad de que los modelos predictivos utilizados en la educación sean explicables. Los autores indican que técnicas tradicionales como la regresión logística (LR) o los árboles de decisión, así como los

métodos SHAP (SHapley Additive exPlanations), ofrecen a los gestores académicos una visión clara de qué variables influyen en la deserción. Esa transparencia permite diseñar intervenciones ajustadas a cada contexto y respaldadas por datos concretos.

Aunque la deserción universitaria ha sido objeto de estudio en múltiples contextos, numerosos investigadores coinciden en que aún se requiere con urgencia la construcción de modelos integrales que abarquen todas las variables involucradas. En esta línea, (Coleman et al., 2024) argumentan que la mayoría de los enfoques tradicionales se han concentrado en aspectos individuales, por ejemplo, en el rendimiento académico o en la situación económica, sin articularlos con dimensiones psicológicas, motivacionales, institucionales y contextuales que también influyen en el abandono. Esa mirada fragmentada impide alcanzar una comprensión integral del fenómeno y, en consecuencia, debilita la eficacia de las políticas diseñadas para retener a los estudiantes.

En consonancia con esta preocupación, (Zeineddine et al., 2021) indican que la carencia de modelos amplios y claros limita la toma de decisiones estratégicas en las IES públicas. De ahí que las intervenciones implementadas suelen ser genéricas, descontextualizadas o reactivas, lo que reduce su capacidad de incidir en las tasas de retención y en el éxito académico. Así, la falta de una base empírica sólida que vincule las causas del abandono con acciones concretas perpetúa la brecha entre el conocimiento académico y la práctica en las instituciones.

Como recuerdan (Bowers et al., 2013), saber predecir un fenómeno no basta si el pronóstico permanece oculto. Los sistemas que identifican a los alumnos en riesgo de abandono deben emplear un lenguaje accesible para profesores, tutores y gestores académicos. Sin esa lucidez, aunque el algoritmo acierte, se convierte en una caja negra que bloquea acciones educativas informadas y a medida.

La presente investigación aborda esta necesidad mediante un marco claro que explique cuáles son y cómo interactúan los factores que llevan a un estudiante a abandonar la universidad pública. En tal contexto, el trabajo de (Oqaidi et al., 2022a) diseña herramientas que, de manera integrada, predicen y explican la DEU, de modo que las instituciones no solo avisen a tiempo, sino que también puedan gestar respuestas proactivas, inclusivas y ajustadas a cada trayectoria estudiantil.

Con este objetivo, se plantea desarrollar una herramienta basada en ML que detecte con alta precisión a los estudiantes de alto riesgo de abandono y, además, explique de forma comprensible los factores que sustentan cada predicción. En este contexto, (S. Lundberg & Lee, 2016) proponen SHAP como técnica de explicabilidad, ya que permite descomponer las decisiones de modelos complejos en contribuciones individuales de cada variable, lo que ofrece una interpretación transparente. De forma complementaria, el método LIME (Local Interpretable Model-Agnostic Explanations) usado por (Ribeiro et al., 2016) explica de manera individual cada predicción mediante un modelo local interpretable, que facilita comprender las causales detrás de cada decisión

La integración de ambas técnicas, a nivel global y local, fortalece tanto la interpretabilidad del conjunto de datos respecto del modelo como la transparencia de cada caso, lo que habilita a las instituciones de educación superior para tomar decisiones informadas e implementar intervenciones tempranas, personalizadas y efectivas.

1.2. Formulación del Problema

1.2.1. Problema general

¿Cómo pueden los modelos de explicabilidad basados en el aprendizaje automático contribuir a predecir e interpretar los factores asociados a la deserción estudiantil universitaria?

1.2.2. Problemas específicos

- ¿Cuáles son los factores más relevantes que deben integrarse en un conjunto de datos para modelar adecuadamente la deserción estudiantil universitaria?
- ¿Qué modelos de aprendizaje automático presentan la mayor capacidad predictiva y la mayor robustez para identificar a estudiantes en riesgo de deserción universitaria?
- ¿Cómo pueden interpretar los modelos de explicabilidad la contribución y las interacciones entre los factores asociados a la deserción estudiantil universitaria?

Justificación Teórica

En el campo del análisis de datos educativos, el uso de modelos predictivos basados en técnicas avanzadas ha permitido avanzar notablemente en la identificación de patrones asociados al rendimiento académico y al abandono estudiantil. No obstante, una limitación recurrente en la literatura científica es que muchos de estos modelos se centran en mejorar la capacidad predictiva, sin considerar la explicabilidad de los resultados. En consecuencia, aunque estos enfoques permiten identificar estudiantes en riesgo de deserción, con frecuencia presentan dificultades para explicar cómo inciden las diferentes variables en las predicciones obtenidas. Este contexto limita la generación de conocimiento explicativo sobre el fenómeno de la deserción y reduce la posibilidad de que los resultados sean comprendidos y utilizados en la toma de decisiones en las IES (Murdoch et al., 2019; Ribeiro et al., 2016)

Además, diversos estudios han comenzado a destacar la importancia de incorporar mecanismos que permitan interpretar el comportamiento de los modelos predictivos. La interpretabilidad de los modelos facilita analizar la contribución de las variables a la generación de predicciones y permite comprender los patrones presentes en los datos. En

el ámbito educativo, esta capacidad resulta especialmente relevante, ya que no solo implica anticipar el riesgo de abandono, sino también comprender los factores que inciden en la DEU. Esta posibilidad de explicar las predicciones fortalece la transparencia de los modelos predictivos y contribuye a transformar el conocimiento adecuado en la investigación educativa y en la gestión institucional (Adadi & Berrada, 2018; S. M. Lundberg & Lee, 2017)

La presente investigación plantea desarrollar un modelo orientado a explicar las predicciones del riesgo de la DEU y a analizar la influencia de las variables involucradas. A diferencia de los enfoques tradicionales centrados exclusivamente en la predicción, el modelo planteado incorpora procedimientos que permiten identificar la importancia relativa de las variables y comprender cómo determinados factores académicos o institucionales influyen en la probabilidad de abandono estudiantil.

En consecuencia, la contribución teórica de esta investigación radica en integrar la capacidad predictiva de los modelos de análisis de datos con mecanismos de interpretación que permiten comprender el comportamiento de las predicciones generadas. Esto contribuye al desarrollo de la analítica educativa al proporcionar un marco metodológico que favorece la generación de conocimiento explicativo sobre los factores que inciden en la deserción estudiantil. Asimismo, permite fortalecer el vínculo entre los métodos de análisis predictivo y la investigación educativa, facilitando la interpretación de los resultados y apoyando la formulación de estrategias institucionales orientadas a mejorar la permanencia estudiantil.

Justificación Práctica

Desde una perspectiva económica, el incremento del nivel educativo de un país suele ir acompañado de un aumento significativo del gasto público por estudiante. De acuerdo

con el informe de la (OECD, 2019) , algunos países destinan importantes recursos a la educación superior; Estados Unidos y Noruega invierten más de 17 000 USD anuales por estudiante universitario, mientras que Suecia y los Países Bajos superan los 13 000 USD. Sin embargo, una parte considerable de estos recursos se desaprovecha cuando los estudiantes abandonan prematuramente sus estudios, lo que genera pérdidas económicas derivadas de la interrupción de los procesos formativos y reduce la eficiencia de los sistemas educativos.

Las consecuencias económicas de la DEU son particularmente evidentes en América Latina. Diversos estudios han señalado que una proporción importante de los recursos destinados a la formación universitaria se pierde cuando los estudiantes no concluyen sus programas académicos. El estudio de (Lykourantzou et al., 2009) estima que aproximadamente el 26 % del presupuesto universitario destinado a la matrícula y al apoyo estudiantil puede perderse como consecuencia del abandono académico. Asimismo, investigaciones recientes indican que en países como Argentina y México las tasas de deserción universitaria pueden superar el 50 %, mientras que en Chile y Perú se sitúan alrededor del 30 %, lo que refleja la magnitud del problema en la región (Núñez-Naranjo, 2024).

Este fenómeno no se limita al contexto latinoamericano, sino que constituye un desafío de alcance global. En España, estudios recientes evidencian que la deserción acumulada al finalizar los estudios universitarios puede alcanzar el 50 % entre los estudiantes que ingresan inicialmente a la educación superior (Berral-Ortiz et al., 2022; Santos-Villalba et al., 2023). De igual forma, en Colombia se ha documentado que cerca del 50 % de los estudiantes que ingresan a programas de pregrado abandonan sus estudios antes de obtener el título, mientras que en modalidades de educación a distancia esta cifra puede superar el 60 % (Barbosa-Camargo et al., 2021). Asimismo, investigaciones realizadas en

Brasil evidencian que el riesgo de abandono es especialmente elevado durante los primeros años de la formación universitaria (Klitzke & Carvalhaes, 2023).

El problema se intensifica cuando el financiamiento institucional se basa principalmente en el número de estudiantes matriculados y no en la eficiencia de los resultados académicos. Por lo tanto, algunos estudios señalan que estos esquemas pueden generar incentivos institucionales inadecuados que dificultan el desarrollo de estrategias de acompañamiento estudiantil efectivas, lo que contribuye a perpetuar los niveles de deserción en la educación superior (Yin et al., 2020). De forma complementaria, (Espinoza et al., 2024) destacan que la reducción de la DEU no solo contribuye a mejorar la equidad en el acceso y la permanencia en la educación superior, sino que también permite optimizar la inversión pública en educación.

Ante este escenario, las IES requieren herramientas analíticas para identificar a los estudiantes con mayor probabilidad de abandonar sus estudios y diseñar estrategias de intervención orientadas a mejorar su permanencia académica. De este modo, el uso de modelos predictivos basados en el análisis avanzado de datos ha demostrado ser una alternativa para apoyar la gestión académica y fortalecer los procesos de seguimiento estudiantil. Estos modelos permiten analizar grandes volúmenes de información académica e institucional con el propósito de identificar patrones asociados al riesgo de deserción y facilitar la implementación de acciones preventivas dirigidas a los estudiantes en situación de vulnerabilidad académica (Romero & Ventura, 2020).

No obstante, más allá de la capacidad predictiva, es fundamental que los resultados puedan interpretarse con claridad, lo que permite analizar los factores que influyen en las predicciones y comprender qué variables académicas, sociales o institucionales se asocian con el riesgo de abandono estudiantil. Desde esta perspectiva, la interpretación de los factores asociados a la deserción estudiantil contribuye a mejorar la toma de decisiones

basada en evidencia en las IES. Mediante herramientas que permiten identificar patrones de riesgo y comprender las variables que inciden en el abandono académico.

En consecuencia, el desarrollo de modelos que permitan no solo anticipar el riesgo de abandono, sino también comprender los factores que influyen en la DEU. Este tipo de enfoques facilita la generación de información analítica que puede utilizarse para fortalecer los sistemas de seguimiento académico, mejorar los programas de acompañamiento estudiantil y orientar la formulación de políticas institucionales orientadas a incrementar la permanencia y el éxito académico de los estudiantes.

1.3.Objetivos de la Investigación

1.3.1. Objetivo General

- Desarrollar un modelo de explicabilidad basado en el aprendizaje automático que permita predecir e interpretar los factores asociados a la deserción estudiantil universitaria.

1.3.2. Objetivos Específicos

- Construir un conjunto de datos con los factores relevantes que influyan en la predicción de la deserción universitaria
- Entrenar y evaluar algoritmos de aprendizaje automático para predecir la deserción estudiantil universitaria, considerando criterios de desempeño y de robustez del modelo.
- Diseñar un modelo explicativo para interpretar la contribución y las relaciones entre los factores que inciden en la deserción estudiantil universitaria.

1.4. Hipótesis e identificación de variables

1.4.1 Hipótesis general

El desarrollo de un modelo de explicabilidad basado en el aprendizaje automático permite interpretar de manera efectiva los factores que influyen en la deserción universitaria.

1.4.2 Hipótesis específicas

- La construcción de un conjunto de datos integrado por variables relevantes permite modelar adecuadamente la deserción estudiantil universitaria.
- El entrenamiento y la evaluación de algoritmos de aprendizaje automático permiten obtener modelos predictivos para identificar la deserción estudiantil universitaria.
- El diseño de un modelo explicativo permite interpretar la contribución de los factores asociados a la deserción estudiantil universitaria, así como las relaciones entre ellos.

CAPÍTULO II. MARCO TEORICO

2.1. Marco Epistemológico

Desde una óptica epistemológica, la producción de conocimiento científico exige reconocer las razones que legitiman los métodos, los enfoques y los resultados de una investigación. En ese sentido, (Klitzke & Carvalhaes, 2023) sostienen que la epistemología, además de presentar un recorrido histórico sobre cómo se construye el saber, sitúa cada etapa dentro de una lógica que la comunidad científica ha validado, otorgando así coherencia, reconocimiento y valor a las conclusiones sobre fenómenos concretos. Gracias a ello, comprendemos que, mediante la observación, la experimentación y el análisis sistemático, el ser humano convierte lo que le rodea en conocimiento comprobable, estableciendo lazos de causalidad entre hechos que, a primera vista, parecen dispersos (Imai et al., 2013).

En el ámbito de la informática, el auge del ML en el siglo XX marcó una profunda ruptura en nuestra manera de entender el conocimiento científico. Como indica (Mitchell, 1997), los algoritmos de ML se basan en la capacidad de descubrir patrones en conjuntos de datos y, a partir de esa experiencia, anticipar lo que ocurrirá en situaciones nuevas o desconocidas. Consecuentemente, el conocimiento generado no se limita al razonamiento deductivo, sino que también presenta un carácter inductivo y probabilístico, basado en la evidencia empírica computacional. Además, (Liu et al., 2025) subrayan que los sistemas actuales de inteligencia artificial combinan deducciones estadísticas con procesos de aprendizaje, dando forma a un tipo mixto de saber que integra el juicio lógico, la práctica y la anticipación.

Desde esta perspectiva (Desai et al., 2022), los sistemas de ML se basan en unidades elementales de información, conocidas como hechos atómicos, que se consideran piezas autónomas con valor por sí mismas. Esta concepción se contrapone al

modelo lineal tradicional y abre paso a un terreno en el que el conocimiento surge del análisis de grandes volúmenes de datos, sin que una jerarquía previa determine qué es importante.

La presente investigación se basa, por tanto, en un enfoque que parte de la ciencia educativa como construcción analítica de la información académica mediante herramientas computacionales. A través de este procedimiento no solo es posible prever fenómenos complejos como la deserción universitaria, sino que también se pueden desmenuzar sus causas mediante un modelo lógico-probabilístico que sigue aprendiendo con cada nuevo dato.

2.2. Antecedentes de Investigación

La deserción universitaria sigue siendo un desafío estructural para las IES, sobre todo en América Latina y el Caribe. (Shafiq et al., 2022) Señalan que, según informes del Banco Mundial y de la UNESCO, casi el 40 % de los alumnos en esas regiones abandona la escuela antes de terminarla, lo que evidencia una crisis profunda en la capacidad de las escuelas para retener a sus estudiantes. En el caso de Colombia, la mitad de los estudiantes que ingresan a la educación superior la abandona y no logra titularse, lo que evidencia una crisis de retención académica (Arias et al., 2024). La Universidad de Chile realizó un estudio en el que la deserción alcanzó el 44% en el primer año de estudio, lo que evidencia una brecha significativa entre el ingreso y la permanencia y subraya la urgencia de intervenciones a nivel institucional (Mellado et al., 2019).

Sin embargo, el problema no solo está presente en el contexto latinoamericano. En Europa, países como Portugal también registran datos alarmantes. El informe de la Direção-Geral de Estatística da Educação e Ciência (DGEEC) indica que tres de cada diez estudiantes en carreras de pregrado ya se encontraban matriculados y no habían podido titularse después de cuatro años de haber ingresado; también registra cifras

alarmantes: la DEU alcanzó un 31,8% de abandono universitario, un valor mucho más alto que el promedio de los países de la OCDE (Ibáñez-Cubillas et al., 2023). Italia tampoco es un caso aislado: la deserción cobra relevancia; el 24% de los estudiantes no continuó con sus estudios, y se destaca que ocurre en los dos primeros años de estudios, lo que evidencia que ni siquiera los sistemas educativos tradicionales cuentan con una estrategia clara para garantizar la retención estudiantil (Delogu et al., 2024).

El trabajo de (Apumayta et al., 2024) afirma que dejar la universidad no solo implica perder dinero, becas, fondos institucionales o matrículas que nunca se usan, sino que también golpea a la sociedad y a las propias instituciones, al debilitar la formación profesional, restringir la movilidad social y, al final, frenar el crecimiento humano de las comunidades.

Desde una perspectiva más amplia, (Namoun & Alshantiti, 2020) sostienen que los efectos de la DEU se manifiestan en tres niveles: personal, institucional y social. A nivel individual, el abandono puede dejar huellas psicológicas, como el fracaso y un bajón de la autoestima, y, por otra parte, el escenario económico implica gastos que nunca se recuperan ni para el estudiante ni para la universidad. Para las IES, la fuga de estudiantes genera problemas académicos, recortes presupuestarios y trastornos organizativos que se prolongan a largo plazo.

Por ello, como señala (Villegas-Ch et al., 2023), numerosas IES han empezado a orientar sus recursos hacia campañas y planes que impulsen la permanencia, ofreciendo caminos alternativos que frenen el retiro en las distintas carreras. Aun así, el reto más grande sigue siendo identificar con precisión las razones que llevan a un joven a abandonar los estudios. Según Namoun y Alshantiti (2020), esta tarea no es sencilla porque cada alumno vive en su propio contexto, donde factores personales, sociales y

académicos influyen de manera distinta en la decisión de seguir adelante o dar un paso al costado.

En el contexto internacional, (Reinoso et al., 2024) destacan que las tasas de la DEU indican diferencias notables entre cada país como Estados Unidos, España y Francia, donde se sitúan entre 30 % y el 50 %; sin embargo, en Suiza, Finlandia y Alemania los números son mucho más bajos-alrededor del 7 % al 30 % en Suiza, el 10 % en Finlandia y entre el 20 % y 25 % en Alemania, que indica que las políticas educativas y la red de apoyo académico difieren de manera sustancial en cada país (OECD, 2023).

En América Latina, de acuerdo a datos del Banco Mundial , casi el 50 % de los jóvenes entre 25 y 29 años que inician la universidad no la finaliza; además, alrededor del 30 % de esos abandonos se produce en el primer año de estudios, lo que pone de relieve la urgente necesidad de contar con sistemas de alerta temprana que permitan intervenir antes de que los estudiantes se alejen definitivamente (Chiarino et al., 2024). Además, en el trabajo de (Espíndola & León, 2002) muestran que Guatemala (82 %), Bolivia (73 %) y Uruguay (72 %) lideran el continente en temas de la DEU, seguidos por Brasil (59 %), Chile y Costa Rica (54 % cada uno), lo que demanda esfuerzos urgentes y mejores planes de acción.

Ante este panorama, (Goren et al., 2024) subrayan que recolectar, procesar y analizar datos con precisión permite crear modelos predictivos que detecten a tiempo a quienes corren el riesgo de desertar. Así, las instituciones de educación superior pueden intervenir a tiempo y los sistemas educativos funcionan de manera más ágil y efectiva.

2.3. Bases Teóricas

Predecir si un estudiante abandonará la universidad exige construir un modelo sólido que no solo calcule con precisión el riesgo, sino que también explique qué factores lo impulsan. Para lograrlo, hay que definir con claridad qué datos se proporcionan al

algoritmo (variables independientes) y cuál es el resultado esperado (variable objetivo), de modo que el sistema aprenda a vincular el pasado con el futuro.

Dentro de este esquema, el modelo se comporta como un mapa que intenta resumir la realidad, captando los patrones complejos que anidan en las notas, las interacciones sociales y las circunstancias particulares de cada alumno. Gracias a técnicas de ML, esa abstracción es capaz de señalar, antes de que ocurra, a los estudiantes que corren un alto riesgo de desertar, lo cual es clave para activar a tiempo las medidas de apoyo.

Una de las ventajas más valiosas de estos modelos radica en su capacidad para emitir alertas anticipadas, por lo que se convierten en recursos de vigilancia y de apoyo institucional. Diversas investigaciones coinciden en que su impacto aumenta cuando se consulta temprano y se integra a redes de acompañamiento académico y emocional, pues así se pueden diseñar intervenciones a medida para cada persona.

2.3.1 Deserción Estudiantil

Dejar la universidad no es algo sencillo ni aislado; se debe a que confluyen muchas circunstancias y el impacto se siente tanto en la persona que se va como en la institución que pierde a ese alumno. En (Alban & Mauricio, 2019a) se explica que la deserción no ocurre de la noche a la mañana, sino que es un proceso largo que comienza con la acumulación de problemas académicos, personales y sociales hasta que el estudiante decide marcharse del sistema.

De este modo, comprender la DEU permite vincularla con otros indicios previos, como el rezago académico, la baja motivación y la pérdida de compromiso institucional. Estos signos tempranos evidencian una desconexión entre el estudiante y su entorno educativo, lo que se refleja en un rendimiento bajo, en irregularidad en la asistencia o en falta de participación en las actividades formativas.

En América Latina, el abandono estudiantil no solo afecta las tasas de eficiencia terminal y de retención institucional, sino que también limita la movilidad social y perpetúa las brechas de desigualdad (Apumayta et al., 2024; Arias et al., 2024). Diversos estudios confirman que el rezago académico es el principal predictor del abandono en los primeros años y que su detección temprana, mediante modelos predictivos y explicables, puede transformar la gestión educativa hacia un enfoque más preventivo e inclusivo (Coussement et al., 2020; Núñez-Naranjo, 2024).

Según (Rabuco, 2022) un estudiante se considera desertor cuando abandona la carrera sin obtener el título y no vuelve durante un tiempo prolongado. Esta definición permite distinguir entre un abandono definitivo y otras formas de desvinculación temporal, como una pausa académica o un cambio de carrera, diferenciaciones necesarias para medir adecuadamente el fenómeno y diseñar estrategias de retención.

Por otro lado, (López-Angulo et al., 2023) sostienen que el abandono puede ser elegido por el alumno o empujado por factores externos a la universidad, como problemas económicos, familiares, de salud o laborales. Por lo general, este tipo de salida ocurre en los tramos iniciales, justo cuando la persona intenta adaptarse a las exigencias universitarias. De acuerdo con (Mnyawami et al., 2022), desertar no solo depende de factores personales e institucionales, sino también de aspectos psicológicos y emocionales. La motivación personal, el bagaje escolar previo, el sentido de pertenencia y las habilidades personales son los factores que más destacan.

Además, (Vidal et al., 2022) señalan que, aunque el número de investigaciones sobre la DEU ha tenido un notable incremento, todavía no existe un consenso que describa o explique las causas del abandono de los estudiantes. Por lo tanto, la deserción sigue siendo vista como un problema de atención prioritaria, tanto social como económico, para

las instituciones de educación superior, y comprenderla a detalle sigue costando trabajo a investigadores, profesores y responsables de políticas.

Por eso mismo, el abandono universitario no puede verse desde un solo ángulo; en él coexisten aspectos relacionados con lo académico, lo emocional, lo que ofrece la institución y las condiciones de cada hogar. Atajarlo exige combinar análisis numéricos, modelos de adivinanza de riesgos y, sobre todo, un acompañamiento a medida que realmente ayude al estudiante a quedarse.

2.3.2 Tipo de Deserción Estudiantil

La DEU puede manifestarse de varias formas, según el momento en que ocurre, las razones que la impulsan y lo que el estudiante decida hacer después. En (Kocsis & Molnár, 2025) se ofrece una tipología que aclara el fenómeno de la deserción y, hasta la fecha, se utiliza como referencia en estudios contemporáneos sobre el abandono y facilita el diseño de respuestas específicas.

En primer lugar, hablan de la deserción total, que ocurre cuando el alumno abandona la universidad de forma definitiva y no vuelve. Este abandono suele deberse a problemas serios, como la falta de dinero, las dificultades familiares o la falta de orientación vocacional. La segunda categoría es la deserción por cambio de carrera, que se manifiesta cuando el estudiante renuncia a un programa para matricularse en otro, ya sea en la misma universidad o en otra. Aunque la ruta académica original se detiene, el vínculo con la educación superior puede mantenerse intacto, por lo que este tipo de deserción se asemeja más a una salida temporal que a un corte absoluto. Por último, (Kocsis & Molnár, 2025) señalan que la deserción en el primer semestre indica una adaptación deficiente del estudiante a la vida universitaria o al programa elegido. Este abandono revela diferencias marcadas entre las expectativas iniciales del alumno y las

exigencias reales del entorno académico, así como falencias en los sistemas de orientación y de nivelación.

No obstante, en trabajos recientes se señala que esta tipología debe actualizarse y adaptarse a contextos variables. Según (Apumayta et al., 2024), la deserción es un fenómeno multicausal y dinámico que requiere categorías que capten dimensiones como la voluntariedad, la continuación informal, la transferencia, la pausa temporal y el retorno potencial. En la misma línea, (Acevedo, 2021) resalta que los procedimientos de medición en América Latina aún presentan problemas de definición y comparabilidad que dificultan la aplicación uniforme de las tipologías tradicionales.

Dicha clasificación resulta esencial para la presente investigación, pues permite examinar el fenómeno desde diversas ópticas y ajustar los modelos predictivos según el tipo de deserción observado.

2.3.3 Factores de deserción universitaria

En (Khatiry & Abdallah, 2023) se afirma que el tránsito de la escuela secundaria a la universidad es una fase crítica en la que se combinan el contexto familiar, las condiciones económicas y el saber. Cuando esas fuerzas se suman a un entorno académico riguroso, empiezan a influir de manera decisiva en el rendimiento y pueden aumentar notablemente el riesgo de abandono.

En esta línea, investigaciones recientes (Martínez-Líbano & Yeomans-Cabrera, 2023) ratifican que las dificultades están relacionadas principalmente con el bajo apoyo familiar, la deficiente formación en la secundaria y las limitaciones socioeconómicas, que afectan tanto la permanencia como el sentido de pertenencia institucional. El trabajo de (Ramírez et al., 2024) sostiene que la escasa calidad de las aulas en muchos sistemas latinoamericanos, unida a presiones económicas y problemas psicológicos, eleva la probabilidad de deserción en los primeros semestres universitarios.

Además, (Li & Carroll, 2020) observan que los factores que incitan al abandono evolucionan a lo largo de la carrera: en los años iniciales predominan aspectos personales como la autoeficacia, la motivación y la integración social; en etapas avanzadas, adquieren mayor relevancia los componentes institucionales, en especial la calidad docente y la relación entre profesores y estudiantes. Esta visión dinámica sugiere que las intervenciones deben diferenciarse según el momento del recorrido formativo.

2.3.4 Aprendizaje automático

Los avances recientes en tecnologías de la información y la comunicación han impulsado el desarrollo de modelos predictivos que cada vez muestran mayor precisión y eficacia en el ámbito de la educación superior. Según Hasan et al. (2020), la fiabilidad de los modelos predictivos depende, sobre todo, de la calidad y la relevancia de los datos utilizados, por lo que resulta esencial dedicar el tiempo adecuado a las etapas de limpieza y normalización de los datos.

Por su parte, Márquez (2020) concibe el ML como una rama que busca patrones complejos mediante algoritmos que aprenden de la experiencia y generan pronósticos útiles para la toma de decisiones. En el aula, los enfoques más comunes son los árboles de decisión, los métodos bayesianos, las redes neuronales, las regresiones y los clústeres basados en instancias.

De manera complementaria, (Chang, 2025) señala que cada sistema de ML descansa en una representación, ya sea explícita o implícita, de los patrones que examina, y que dicha representación se perfecciona mediante un conjunto de datos históricos cuidadosamente etiquetados. Aplicada a la DEU, esta idea ha producido avances notables: los modelos no solo identifican de manera anticipada a los estudiantes más vulnerables, sino que también aportan pistas sobre las causas subyacentes del riesgo y permiten diseñar intervenciones tempranas y personalizadas.

2.3.5 Técnicas de Aprendizaje Automático

El ML, que forma parte de la inteligencia artificial, se dedica a desarrollar modelos para reconocer patrones y realizar pronósticos a partir de grandes volúmenes de datos. Gracias a su capacidad de autoaprendizaje y de ajuste continuo, estas técnicas a menudo superan a los modelos estadísticos clásicos. Según (Mnyawami et al., 2022), esta herramienta puede ser una respuesta efectiva frente a la DEU, siempre que se limpien adecuadamente las características y se elija el algoritmo que mejor se adapte a cada institución.

En la misma línea, (Christou et al., 2023) señalan que el ML sobresale en escenarios con una gran diversidad de información sobre estudiantes, asignaturas y carreras, ya que permite identificar a quienes están en riesgo de abandonar sus estudios. La fortaleza de estas herramientas radica en su capacidad para analizar grandes volúmenes de datos y descubrir relaciones que pasan desapercibidas con los métodos tradicionales.

Asimismo, estudios recientes indican que el uso de modelos supervisados y no supervisados en contextos universitarios permite identificar grupos vulnerables, caracterizar sus trayectorias académicas y anticipar con alta precisión las decisiones de abandono (Goren et al., 2024). De este modo, el ML no solo mejora la capacidad predictiva en las IES, sino que también fortalece su potencial explicativo, lo que contribuye a la formulación de políticas basadas en evidencia que promuevan la retención y la equidad educativa (Apumayta et al., 2024).

Además, (Kok et al., 2024) subrayan que el ML constituye un hito clave en el diseño de sistemas de alerta temprana para la educación superior, pues permite detectar a tiempo a estudiantes en riesgo de desertar. Esta anticipación brinda margen para preparar

intervenciones adecuadas que aumenten las probabilidades de que esos estudiantes continúen su trayectoria académica.

Los algoritmos más frecuentes en estas aplicaciones incluyen k vecinos más cercanos (KNN), máquinas de soporte vectorial (SVM), árboles de decisión (DT), redes neuronales artificiales (ANN), modelos de aprendizaje profundo, así como métodos de ensamblaje como Extra Trees (ET) y CatBoost, reconocidos por su capacidad para manejar datos de alta dimensionalidad y relaciones no lineales. En el estudio de (Sihare, 2024) se advierte que la selección del enfoque debe depender de la naturaleza de los datos y de los objetivos concretos que persiga el modelo predictivo.

2.3.6 Predicción

En el campo del ML, hacer una predicción implica construir un modelo que utilice datos históricos para predecir qué podría ocurrir en el futuro. Según (Al-Barrak & Al-Razgan, 2015), este procedimiento produce un pronóstico basado en el pasado y les proporciona a los responsables la información necesaria para actuar antes de que los problemas aparezcan, especialmente en un entorno tan cambiante como el de la educación superior.

Siguiendo esta línea, (Tsai et al., 2020) afirman que un sistema que predice la DEU constituye un marco ordenado capaz de identificar, con antelación, a quienes corren el riesgo de abandonar sus estudios, lo que, a su vez, permite el diseño de políticas de retención a medida. La utilidad real de dicho sistema, no obstante, está ligada a la limpieza de los datos y a la pertinencia de las características seleccionadas durante la fase de entrenamiento. Además, (Coleman et al., 2024) subrayan que la exactitud de la predicción se basa en dos factores fundamentales: seleccionar adecuadamente las variables de entrada (rendimiento académico, situación económica, asistencia a clases) y aplicar técnicas de procesamiento que pongan al modelo en su mejor estado. De ahí que la predicción no se limite a emitir un número que cuantifique el riesgo; se transforma, en cambio, en un recurso estratégico capaz

de descubrir pautas de abandono y guiar a las instituciones en la creación de respuestas efectivas.

2.3.7 Explicabilidad

En años recientes ha surgido un notable interés por la explicabilidad de los modelos de ML, sobre todo en contextos en los que las decisiones basadas en la inteligencia artificial impactan directamente a personas e instituciones, como ocurre en la educación superior. Esta notoriedad se debe a la necesidad de hacer transparentes los procesos de predicción y clasificación, de modo que los responsables institucionales puedan comprenderlos, auditarlos y confiar en los resultados generados por los algoritmos.

En tal línea, (Qing et al., 2022) sostienen que los métodos explicativos son fundamentales para promover la confianza y la interpretabilidad, pues permiten visualizar en qué medida cada característica influye en la decisión final del modelo. Afrontar la explicación de decisiones complejas resulta crítico en los modelos de caja negra, cuyas reglas internas no se revelan de forma natural a los usuarios. En esta línea, (Arrieta et al., 2020) y (Czerwinska, 2022) argumentan que un modelo explicable debe ofrecer no solo resultados acertados, sino también interpretaciones claras y accesibles que refuercen la confianza de quienes lo utilizan.

En tal sentido, (Nnadi et al., 2024) sostienen que una explicación integral del ML comienza por establecer vínculos cuantitativos claros entre las variables de entrada y de salida del modelo. Con este principio, el sistema gana credibilidad y trazabilidad en el modelo, lo que permite a los usuarios verificar que el comportamiento algorítmico es válido y coherente. Asimismo, (Vilone & Longo, 2021) señalan que la validez de las explicaciones, tanto a nivel local como global, es una condición fundamental para que los modelos de ML explicables sean fiables y éticamente auditables. En el ámbito educativo, esto implica que las IES deben asegurarse de que las predicciones del riesgo de abandono

se mantengan consistentes ante variaciones en los datos estudiantiles, evitando sesgos o interpretaciones erróneas. En conjunto, estas prácticas fortalecen la transparencia, la interpretabilidad y la confianza institucional en el uso de modelos de predicción académica explicables.

SHAP

El algoritmo SHAP (Shapley Additive Explanations) es un método de interpretación ampliamente utilizado en modelos de ML que permite explicar la contribución de cada variable a las predicciones del modelo. Este enfoque se basa en los valores de Shapley de la teoría de juegos cooperativos, que permiten distribuir de manera equitativa la contribución de cada característica al resultado de una predicción. En el contexto del ML, SHAP proporciona un marco teórico sólido para cuantificar el impacto de cada variable en la salida del modelo, lo que permite comprender cómo los distintos factores influyen en la predicción (Min et al., 2023).

Uno de los principales aportes de SHAP radica en su capacidad para ofrecer explicaciones tanto globales como locales del comportamiento del modelo. A nivel global, permite identificar qué variables tienen mayor influencia en las predicciones generales del modelo, lo que facilita el análisis de patrones relevantes en el conjunto de datos. A nivel local, SHAP permite analizar la contribución específica de cada variable a la predicción de un caso particular, lo que resulta especialmente útil para comprender las decisiones del modelo y mejorar la transparencia de los sistemas de inteligencia artificial (Freiesleben et al., 2022).

En los últimos años, SHAP se ha consolidado como una de las técnicas más utilizadas en el campo de la inteligencia artificial explicable (XAI), debido a su capacidad para proporcionar interpretaciones consistentes y fundamentadas de modelos complejos, incluidos algoritmos basados en árboles, redes neuronales y métodos de ensamble. Su

aplicación ha permitido mejorar la interpretabilidad de modelos predictivos en diversos dominios, incluidos la salud, las finanzas y la educación, donde la comprensión de los factores que influyen en las predicciones resulta esencial para la toma de decisiones informadas (Barredo Arrieta et al., 2020; Speith, 2022).

En el contexto de la analítica de datos educativos y de la predicción de la deserción universitaria, SHAP permite identificar con mayor precisión qué variables académicas, socioeconómicas o institucionales inciden en el riesgo de abandono estudiantil. De esta manera, la aplicación de SHAP facilita no solo el desarrollo de modelos predictivos precisos, sino también la generación de conocimiento interpretable que puede apoyar la toma de decisiones institucionales orientadas a mejorar la permanencia estudiantil y reducir los niveles de deserción en la educación superior.

LIME

LIME (Local Interpretable Model-Agnostic Explanations) es una técnica de interpretabilidad diseñada para explicar el comportamiento de modelos de ML complejos mediante aproximaciones locales interpretables. Este método se basa en la idea de que, aunque un modelo global pueda ser altamente complejo y difícil de interpretar, es posible comprender su comportamiento en torno a una predicción específica mediante la construcción de un modelo interpretable que lo aproxime en una región local del espacio de datos (Ribeiro et al., 2016).

El enfoque de LIME consiste en generar perturbaciones alrededor de una observación particular y evaluar cómo cambian las predicciones del modelo ante dichas perturbaciones. A partir de esta información, el método ajusta un modelo interpretable, generalmente lineal, que permite estimar la contribución de cada característica a la predicción en estudio. De esta manera, LIME proporciona explicaciones locales que facilitan la comprensión de las decisiones tomadas por modelos complejos, como redes

neuronales, métodos de ensamble o algoritmos de clasificación no lineales (Freiesleben et al., 2022).

En el campo de la XAI, LIME se ha consolidado como una de las técnicas más utilizadas para interpretar modelos de ML, especialmente en contextos en los que la transparencia y la interpretabilidad son fundamentales para la toma de decisiones. Su carácter model-agnostic permite aplicarlo a distintos tipos de algoritmos sin necesidad de modificar la estructura interna del modelo, lo que facilita su integración en diversos escenarios de análisis predictivo (Barredo Arrieta et al., 2020).

En el contexto de la analítica de datos educativos y de la predicción de la deserción universitaria, LIME permite analizar cómo las variables académicas, sociodemográficas e institucionales influyen en la predicción individual del riesgo de deserción estudiantil. Este enfoque resulta especialmente útil para interpretar casos específicos de estudiantes identificados como de alto riesgo, lo que permite comprender qué factores contribuyen en mayor medida a la predicción del modelo.

2.4. Marcos Conceptuales o Glosario

2.4.1 Deserción universitaria

La DEU se entiende como el proceso por el cual un estudiante abandona sus estudios superiores antes de completarlos y obtener un título. Según (Rahmani et al., 2024), la decisión de abandonar los estudios puede ser voluntaria o forzosa y se debe a múltiples causas, como factores económicos, académicos, institucionales o personales. Entre los factores más destacados se encuentran los problemas económicos, el bajo rendimiento escolar, las dificultades para adaptarse, la orientación vocacional insuficiente o la falta de ayuda familiar. Por lo tanto, la deserción no debe considerarse una mera elección individual, sino el resultado de una interacción compleja de circunstancias sociales, psicológicas e institucionales que inciden en la permanencia del estudiante.

Por lo tanto, la deserción en el sistema de educación superior debe considerarse un signo de vulnerabilidad estructural y no únicamente una elección individual. Según (Arias et al., 2024), la decisión de abandonar los estudios está relacionada con la capacidad de las IES para proporcionar apoyo, guía y un entorno socioemocional favorable que faciliten la continuidad. Además, (Martínez-Líbano & Yeomans-Cabrera, 2023) señalan que las variables psicológicas, como la depresión, el estrés y la sensación de soledad, aumentaron el riesgo de abandono tras la pandemia, lo que resalta la necesidad de desarrollar políticas universitarias que prioricen la salud mental y el bienestar de los estudiantes.

2.4.2 Retención estudiantil

La retención estudiantil es la capacidad de las IES para mantener a los estudiantes matriculados y activos en sus programas académicos hasta la culminación de sus estudios. Además, constituye uno de los principales indicadores del desempeño institucional, ya que refleja el grado en que las IES acompañan a los estudiantes a lo largo de su trayectoria formativa y favorecen su progreso académico. En términos operativos, la retención suele evaluarse mediante la proporción de estudiantes que permanecen matriculados en periodos académicos consecutivos o que continúan en la institución hasta completar su programa de estudios (Bukhatir et al., 2023).

En años recientes, la retención estudiantil se ha considerado un fenómeno multidimensional, influido por factores académicos, institucionales, sociales y personales que inciden en la permanencia de los estudiantes en las IES. Diversos estudios señalan que variables como el rendimiento académico, la integración social en la universidad, el bienestar estudiantil y el acceso a servicios de apoyo influyen de manera significativa en la probabilidad de que los estudiantes continúen su trayectoria universitaria (Pérez-Jorge et al., 2025).

Asimismo, investigaciones recientes destacan que mejorar la retención estudiantil se ha convertido en uno de los principales desafíos de las IES, debido al impacto de la deserción en la eficiencia del sistema educativo y en las oportunidades de desarrollo académico y profesional de los estudiantes. En este contexto, se han comenzado a implementar estrategias institucionales orientadas al seguimiento de la trayectoria académica, al desarrollo de programas de acompañamiento y a la implementación de sistemas de alerta temprana para identificar a estudiantes con alto riesgo de abandono (González-Morales et al., 2025).

Además, la analítica de datos educativos y las técnicas de ML han cobrado un papel relevante en el estudio de la retención estudiantil, ya que permiten analizar grandes volúmenes de datos académicos y sociodemográficos con el propósito de identificar patrones asociados al riesgo de abandono y facilitar la toma de decisiones institucionales orientadas a mejorar la permanencia estudiantil. Estas aproximaciones han permitido desarrollar modelos predictivos capaces de anticipar situaciones de riesgo y de apoyar el diseño de estrategias de intervención temprana que contribuyan a fortalecer la continuidad de los estudiantes en la educación superior (Sodagari, 2025).

En consecuencia, la retención estudiantil no debe entenderse únicamente como la permanencia de los estudiantes en una institución, sino como el resultado de un conjunto de políticas institucionales, estrategias pedagógicas y sistemas de apoyo académico orientados a garantizar la continuidad, el progreso y el éxito educativo en el sistema de educación superior.

CAPÍTULO III. METODOLOGÍA DE LA INVESTIGACIÓN

El presente estudio es una investigación aplicada, de carácter explicativo y con un enfoque cuantitativo. La investigación tiene como propósito diseñar un modelo de predicción, explicación y simulación de la deserción universitaria mediante ML que, además de pronosticar la deserción de los estudiantes, permita identificar e interpretar los factores más influyentes en ella y brinde sustento para la toma de decisiones institucionales.

3.1. Tipo de investigación.

3.1.1 Nivel de investigación.

La presente investigación es aplicada, pues tiene como finalidad generar conocimiento con un propósito práctico, enfocado en brindar soluciones a los problemas del sector educativo. El estudio se centra en la predicción de la DEU mediante técnicas de ML.

El objetivo es optimizar la precisión y la explicabilidad de los modelos de predicción del riesgo de abandono mediante la inclusión de variables que reflejen aspectos demográficos, socioeconómicos, personales, institucionales y académicos, y se emplearon enfoques basados en técnicas explicativas, como SHAP y LIME, para facilitar la interpretación de los resultados.

La investigación tiene como finalidad clasificar a los estudiantes en dos grupos: los con permanencia y los en riesgo de deserción, considerando sus patrones académicos, socioeconómicos y de participación institucional. A partir de esta clasificación, se plantea una metodología replicable y robusta que optimiza la gestión del abandono estudiantil en contextos de alta incertidumbre, lo que permite a las instituciones implementar intervenciones tempranas y focalizadas.

3.1.2 Tipo de investigación.

La presente investigación adopta un enfoque cuantitativo, basado en la recopilación, el procesamiento y el análisis de datos numéricos mediante técnicas estadísticas y de ML. Se emplean métodos cuantitativos para evaluar el desempeño de los modelos de predicción del riesgo de la DEU, analizando su precisión y capacidad discriminativa en distintos escenarios. La validación se basa en métricas de clasificación como el F1-score (F1), la exactitud (ACC) y el área bajo la curva ROC (AUC), destacando el F1, que permite medir con mayor rigor la efectividad del modelo en la identificación de estudiantes en riesgo de abandono.

De igual manera, la investigación incorpora algoritmos de ML para analizar la relación entre los distintos factores académicos, socioeconómicos e institucionales y la probabilidad de deserción, garantizando un proceso replicable y basado en evidencia empírica. Este enfoque permite evaluar objetivamente cómo cada variable contribuye al comportamiento del modelo y a la caracterización del fenómeno del abandono estudiantil.

3.1.3 Diseño de la investigación.

El diseño metodológico se enmarca en un enfoque cuantitativo, no experimental y de alcance explicativo-predictivo. Se trabajó con un conjunto de datos históricos provenientes de registros institucionales oficiales que integran factores académicos, socioeconómicos e institucionales asociados a la DEU, sin manipular directamente las variables independientes ni intervenir en la población. En consecuencia, el fenómeno fue analizado en su contexto natural mediante un esquema retrospectivo basado en datos observacionales.

El estudio presenta un carácter correlacional y longitudinal de corte retrospectivo, dado que examina la relación entre múltiples dimensiones del perfil estudiantil y la

probabilidad de abandono a lo largo de la trayectoria académica. En este sentido, no se persigue establecer una causalidad directa, sino identificar patrones consistentes y construir modelos predictivos robustos mediante técnicas avanzadas de aprendizaje automático.

Para el contraste de las hipótesis y la evaluación del desempeño predictivo, se adoptó una estrategia de experimentación computacional comparativa, en la que se evaluaron múltiples familias algorítmicas, incluyendo modelos lineales, probabilísticos y ensambles de árboles, en distintos escenarios de balanceo de clases. La validación del desempeño se realizó mediante validación cruzada estratificada (k-fold) y un conjunto de prueba independiente, lo que permitió estimar la capacidad de generalización de los modelos y reducir el sesgo asociado a particiones específicas del conjunto de datos. Asimismo, se incorporó una fase de optimización de hiperparámetros y de ajuste del umbral de decisión, con el propósito de controlar el sobreajuste y mejorar el equilibrio entre sensibilidad y precisión.

Finalmente, se integraron técnicas de explicabilidad, específicamente SHAP y LIME, con el fin de analizar no solo el rendimiento predictivo del modelo, sino también la lógica subyacente de sus decisiones. Esto fortaleció la transparencia, la auditabilidad y la aplicabilidad institucional del sistema propuesto.

3.1.4 Método de la investigación.

La presente investigación se desarrolló con un enfoque cuantitativo, de tipo aplicado y de nivel explicativo, orientado a analizar el impacto de diversos factores académicos, socioeconómicos e institucionales sobre la capacidad predictiva de los modelos de ML para predecir la DEU. Con este propósito, se diseñó un procedimiento metodológico riguroso para evaluar el desempeño de los algoritmos de clasificación supervisada en un conjunto de datos históricos provenientes de fuentes institucionales oficiales. El método propuesto se organiza en las siguientes fases:

- **Recolección y preparación de datos.** Incluye la obtención y consolidación de información académica, demográfica y socioeconómica, seguidas de un proceso de limpieza, codificación, integración y enriquecimiento para conformar un dataset tabular robusto y analíticamente consistente. Esta etapa garantiza la calidad y la pertinencia de los datos empleados en el modelado.
- **Análisis exploratorio.** Comprende la evaluación inicial de las relaciones entre variables mediante técnicas estadísticas y visualizaciones descriptivas, como matrices de correlación, análisis de distribuciones y pruebas de dependencia, con el fin de identificar patrones relevantes y posibles indicadores tempranos de deserción.
- **Modelado predictivo.** Consiste en el entrenamiento y la evaluación de distintos algoritmos de ML para estimar la probabilidad de abandono estudiantil. El desempeño de los modelos se analizó mediante métricas de clasificación ampliamente aceptadas, como la exactitud (ACC), el área bajo la curva ROC (AUC) y el F1-score; este último, especialmente útil debido a la naturaleza desbalanceada del fenómeno.

- **Validación y comparación.** Incluye la aplicación de técnicas de validación cruzada (10-fold cross-validation) y de análisis de sensibilidad para comparar el rendimiento de los modelos en distintas configuraciones. Esta fase permitió determinar el aporte real de cada conjunto de variables, así como la estabilidad y la capacidad de generalización del modelo propuesto.
- **Explicabilidad del modelo (XAI).** Finalmente, se incorporó un componente de explicabilidad, tanto global como local, para interpretar de manera transparente las decisiones del modelo predictivo. Para ello, se utilizaron las técnicas SHAP y LIME, ampliamente reconocidas en el campo de la inteligencia artificial explicable.

Desde una perspectiva global, SHAP permitió identificar las variables con la mayor contribución promedio al comportamiento del modelo, lo que proporcionó una comprensión estructural de los factores que explican la deserción estudiantil. En el plano local, SHAP y LIME facilitaron el análisis detallado de casos individuales, evidenciando cómo cada variable influye, positiva o negativamente, en la predicción de un estudiante en particular. La inclusión de este módulo de explicabilidad garantiza que el modelo no solo alcance un desempeño predictivo adecuado, sino que también sea transparente, interpretable y útil para la toma de decisiones institucionales, lo que permite diseñar estrategias de intervención temprana basadas en evidencia.

3.1.5 Unidad de análisis

Para la validación del modelo de explicabilidad de la DEU, se definió como unidad de análisis a los estudiantes matriculados en pregrado en una universidad pública de la provincia de Manabí, Ecuador. El estudio se realizó a nivel institucional, incorporando información académica, administrativa y sociodemográfica generada a lo largo de los

periodos académicos analizados, con el propósito de identificar, predecir y explicar los factores asociados a la DEU. En este contexto, se excluyeron del análisis los institutos, escuelas y colegios, así como las IES de carácter privado, a fin de mantener la coherencia normativa, estructural y académica del sistema universitario público ecuatoriano.

3.1.6 Población de estudio

La población de estudio está conformada por el conjunto de datos de estudiantes de la Facultad de Ciencias Informáticas de una universidad pública de la provincia de Manabí, Ecuador, que registran información demográfica, socioeconómica, académica e institucional, relevante para el análisis de la DEU. El estudio considera diez periodos académicos consecutivos, comprendidos entre 2019-2 y 2024-1, y abarca a los estudiantes de pregrado. En la Tabla 1 se presenta la distribución de los registros analizados por periodo académico, lo que permite contextualizar el volumen y la cobertura temporal de los datos utilizados en el modelo propuesto.

Tabla 1

Conjunto de datos de estudiantes iniciales

Periodo	Matriculados
OCTUBRE DEL 2019 HASTA: FEBRERO DEL 2020	1140
MAYO 2020 HASTA: OCTUBRE DEL 2020	1319
NOVIEMBRE DEL 2020 HASTA MARZO DEL 2021	1603
MAYO 2021 HASTA SEPTIEMBRE DEL 2021	1630
OCTUBRE DEL 2021 HASTA FEBRERO DEL 2022	1696
MAYO DEL 2022 HASTA SEPTIEMBRE DEL 2022	1622
OCTUBRE DEL 2022 HASTA FEBRERO DEL 2023	1865
MAYO DEL 2023 HASTA SEPTIEMBRE DEL 2023	2039
OCTUBRE DEL 2023 HASTA ENERO DEL 2024	2169
ABRIL DEL 2024 HASTA AGOSTO DEL 2024	2275
TOTAL	17358

3.1.7 Selección de muestra

Se selecciona como muestra intencional de la Universidad Técnica de Manabí, del Ecuador, que proporcionó el conjunto de datos necesario para el desarrollo y la validación del modelo de explicabilidad de la DEU. La selección se fundamenta en la disponibilidad, la consistencia y la cobertura temporal de la información. Para el estudio se aplican criterios de exclusión orientados a garantizar la pertinencia, la calidad y la representatividad de los datos analizados, asegurando que sean adecuados y significativos para la construcción y la evaluación del modelo propuesto en el contexto de la educación superior pública. Para asegurar la consistencia, precisión y validez del análisis, se aplicaron los siguientes criterios de exclusión:

- Estudiantes con información institucional incompleta: Se eliminaron los registros que carecían de datos esenciales de carácter académico, socioeconómico o personal, a fin de garantizar la calidad y la uniformidad del dataset.
- Estudiantes en proceso de transferencia o de cambio de carrera: Se excluyeron estos casos porque sus trayectorias académicas siguen lógicas distintas de las del abandono definitivo.
- Estudiantes de programas de posgrado o de cursos de corta duración: El estudio se enfocó exclusivamente en la población de pregrado, el grupo más representativo y relevante para el análisis de la deserción.

La selección de la muestra se realizó mediante un muestreo no probabilístico por conveniencia, aplicando los criterios previamente definidos al conjunto de datos disponible, con el propósito de contrastar las hipótesis planteadas en el estudio. Esta estrategia permitió asegurar la presencia de las variables explicativas relevantes en los distintos subconjuntos de datos, garantizando su adecuada representación a lo largo de los periodos académicos analizados. En este sentido, la tabla 2 presenta la distribución de

los registros estudiantiles por periodo académico, lo que aporta el soporte empírico necesario para el desarrollo, la evaluación y la interpretación del modelo explicativo de la DEU.

Tabla 2

Conjunto de datos de estudiantes iniciales

PERIODO	MATRICULADOS
OCTUBRE DEL 2019 HASTA: FEBRERO DEL 2020	1140
MAYO 2020 HASTA: OCTUBRE DEL 2020	1319
NOVIEMBRE DEL 2020 HASTA MARZO DEL 2021	1603
MAYO 2021 HASTA SEPTIEMBRE DEL 2021	1630
OCTUBRE DEL 2021 HASTA FEBRERO DEL 2022	1696
MAYO DEL 2022 HASTA SEPTIEMBRE DEL 2022	1622
OCTUBRE DEL 2022 HASTA FEBRERO DEL 2023	1844
MAYO DEL 2023 HASTA SEPTIEMBRE DEL 2023	1998
OCTUBRE DEL 2023 HASTA ENERO DEL 2024	2120
ABRIL DEL 2024 HASTA AGOSTO DEL 2024	2211
Total	17183

3.2. Dataset y partición.

3.2.1 Dataset

La definición del conjunto de datos utilizado en esta investigación se presenta en la tabla 3. Para desarrollar un modelo predictivo y explicable de la deserción universitaria, se utilizó un único conjunto de datos institucionales proveniente del sistema de gestión académica de la universidad. Este conjunto de datos reúne información demográfica, socioeconómica, académica, institucional y personal de los estudiantes, lo que lo convierte en una fuente sólida y representativa para analizar el fenómeno de la deserción estudiantil.

El conjunto de datos incluye variables como género, edad, nivel académico, programa de estudios, procedencia geográfica, características del hogar, situación económica, acceso

a la conectividad, antecedentes familiares en educación y en condiciones de estudio, entre otras. Su estructura permite capturar de manera integral las diferentes dimensiones que influyen en la permanencia o el abandono de los estudiantes, ofreciendo un marco empírico sólido para el entrenamiento, la validación y la explicabilidad del modelo de ML propuesto.

Tabla 3

Descripción de las variables del conjunto de datos sobre la deserción estudiantil universitaria.

Variable	Tipo	Descripción
idpersonal	int	Identificar numérico del estudiante
cedula	int	Número de identificación
apellido1	string	Apellido paterno del estudiante.
apellido2	string	Apellido materno del estudiante.
nombres	string	Nombres del estudiante.
codigo_matricula	int	Código de matrícula del estudiante.
telefono_personal_celular	int	Teléfono celular del estudiante.
correo_personal_institucional	string	Correo institucional del estudiante.
correo_personal_alternativo	string	Correo personal del estudiante.
Carrera	string	Carrera en la que el estudiante está inscrito.
nivel_estudiante	int	Nivel académico cursado.
paralelo	string	Paralelo inscrito del estudiante
genero	string	Género reportado por el estudiante.
modalidad_carrera	string	Carrera del estudiante.
nacionalidad	string	Nacionalidad del estudiante.
pais_residencia	string	País de residencia del estudiante.
pais_origen	string	País de origen del estudiante.
provincia_residencia	string	Provincia en la que reside el estudiante.
provincia_origen	string	Provincia de origen del estudiante.
canton_residencia	string	Canton donde reside el estudiante.
canton_origen	string	Cantón de origen del estudiante.

tipo_zona_procedencia	string	Tipo de zona de procedencia
fecha_nacimiento	date	Fecha de nacimiento del estudiante
edad	int	Edad del estudiante durante el periodo analizado.
tipo_sangre	string	Tipo de sangre del estudiante.
estado_civil	string	Estado civil del estudiante.
etnia	string	Etnia descrita del estudiante.
nota_colegio	float	Nota obtenida en estudios secundarios.
sostenibilidad_colegio	string	Sostenibilidad del colegio.
hijos_numero	int	Número de hijo del estudiante.
salario_ingreso_total	string	Rango de ingresos
integrantes_familiar	int	Número de integrantes de familia
internet_dispone_casa	string	Indica si tiene internet en casa
internet_dispone_celular	string	Indicasi tiene internet en el teléfono
internet_horas_uso	int	Número de horas de uso de internet en el día
red_social_horas_uso	int	Número de horas de uso de red social en el día
tipo_vivienda	string	Tipo de vivienda del estudiante
vivienda_localizacion	string	Ubicación de la vivienda
recibe_ayuda_familiar	string	Recibe ayuda familiar
jefe_ocupacion	string	Ocupación de estudiante
transporte_horas_traslado	float	Número de horas de traslado.
transporte_horas_permanencia	float	Número de horas de permanencia.
tipo_transporte	string	Tipo de transporte de traslado.
madre_instruccion	string	Nivel de instrucción de la madre.
padre_instruccion	string	Nivel de instrucción del padre.
gasto_alimentacion	string	Rango de gasto de alimentación
gasto_transporte	string	Rango de gasto de transporte
materias_matriculado	int	Número de materias matriculadas.
materias_aprobados	int	Número de materias aprobadas.
promedio_aprobada	float	Promedio de materias aprobadas.
semestre_tomados	int	Número de semestres matriculados.
semestre_anulados	int	Número de semestres anulados.
dispone_tablet	string	Indica si el estudiante tiene una tablet.

dispone_computador	string	Indica si el estudiante tiene un computador.
orientacion_sexual	string	Variable del dataset empleada en el análisis.
preferencia_religiosa	string	Variable del dataset empleada en el análisis.
periodo	string	Periodo académico del registro.
desertor	string	Variable objetivo.

3.2.2 Partición

Para la construcción del modelo de ML aplicado a la DEU, el dataset se dividió de acuerdo con las directrices del ML y con los estudios en el ámbito educativo. El 60% de los registros se asignó al conjunto de entrenamiento, donde los algoritmos ajustaron sus parámetros internos y aprendieron las relaciones entre las variables académicas, socioeconómicas y demográficas asociadas a la DEU. Esta fase permitió identificar patrones relevantes y reproducibles en el comportamiento de los estudiantes.

A continuación, se utilizó el 20% del dataset como conjunto de validación, lo que permitió analizar el comportamiento del modelo durante el entrenamiento, detectar el sobreajuste y optimizar los hiperparámetros antes de la evaluación final. Finalmente, el 20% restante se utilizó como conjunto de prueba, lo que permitió evaluar objetivamente el rendimiento del modelo en datos no observados y su capacidad real de generalización.

Este esquema permitió evaluar de manera efectiva el rendimiento de los modelos, lo que destaca la importancia de equilibrar adecuadamente las fases de entrenamiento y evaluación para obtener modelos robustos y comparables (Zhu et al., 2023).

Operacionalización de variables.

3.2.3. Variable independiente

Modelo de explicabilidad basado en el aprendizaje automático.

Esta variable corresponde al conjunto de técnicas, algoritmos y metodologías empleados para el diseño e implementación de un sistema que permita anticipar la probabilidad de deserción universitaria, explicar los factores asociados y simular escenarios alternativos.

3.2.4. Variable dependiente

Deserción universitaria.

Se refiere al abandono parcial o definitivo de los estudios por parte de los estudiantes matriculados en programas de educación superior, fenómeno que constituye el objeto de predicción y análisis del presente estudio.

3.2.5. Matriz de operacionalización de variables

La Tabla 44 presenta la matriz de operacionalización de variables, en la que se muestran tanto las variables independientes como las dependientes consideradas en este estudio. En dicha matriz se especifican la definición conceptual y operacional de cada variable, los indicadores seleccionados para su medición y las escalas empleadas. Este esquema permite una comprensión precisa del tratamiento de las variables en la investigación, garantizando la coherencia metodológica y el rigor en los procesos de recolección y análisis de datos.

Tabla 4

Operacionalización de variables

VARIABLES	TIPO DE VARIABLES	DEFINICIÓN DE VARIABLE	DIMENSIONES O FACTORES	MEDICIÓN DE INDICADOR O CUMPLIMIENTO	DEFINICIÓN DE INDICADOR
Deserción universitaria	Dependiente	Estudiantes que abandonan sus estudios universitarios antes de culminar el programa académico.	Factores demográficos, socioeconómicos, institucionales, personales, académicos.	Porcentaje de estudiantes que desertan.	Se divide el número de estudiantes que abandonan sus estudios entre el total de matriculados.
Modelo de predicción	Independiente	Algoritmos de aprendizaje automático orientados a predecir el riesgo de deserción.	RF, XGBoost, LightGBM, CatBoost, LR	Métricas de desempeño: Accuracy, F1 Score, Recall, Precision, ROC-AUC.	Nivel de precisión: capacidad del modelo para predecir la deserción.
Modelo de explicabilidad	Independiente	Técnicas para interpretar la influencia de las variables en la predicción del riesgo de deserción.	SHAP, LIME, árboles subrogados	Importancia de la variable en la predicción.	Nivel de contribución de cada variable a la deserción, explicado de manera comprensible para la toma de decisiones.

3.3. Técnicas e instrumentos de recolección de datos.

3.3.1 Técnicas.

Para la recolección de datos se empleó la base de datos institucional de la Universidad Técnica de Manabí (UTM), que es la fuente oficial de datos académicos, sociodemográficos y administrativos de los alumnos inscritos. Este conjunto de información fue obtenido mediante el Sistema de Gestión Académica (SGA-UTM), tras recibir la autorización formal de la Dirección de Tecnologías de la Información, y en cumplimiento de los protocolos institucionales para el acceso y el uso adecuados de los datos educativos.

La base de datos utilizada incluyó archivos históricos agrupados de las cohortes en activo, que abarcaron aspectos vinculados al camino académico, la permanencia, el desempeño, las condiciones socioeconómicas y los datos de identificación institucional de los alumnos. Se entregaron todos los datos sin información sensible, cumpliendo así con la legislación ecuatoriana en materia de protección de datos y de confidencialidad institucional, incluida la Ley Orgánica de Protección de Datos Personales (Asamblea Nacional del Ecuador, 2021)

Este proceso proporcionó una fuente de información confiable, normalizada y directamente relacionada con el fenómeno en cuestión, garantizando una metodología rigurosa y la relevancia contextual necesarias para desarrollar el modelo de explicabilidad de la deserción en IES públicas.

3.3.2 Procedimientos de análisis y modelado.

En esta investigación no se utilizaron instrumentos tradicionales como cuestionarios, encuestas o escalas estandarizadas, dado que el enfoque metodológico se basa en el análisis de datos institucionales y en la aplicación de técnicas propias de la ciencia de datos y del ML. En lugar de instrumentos de recolección directa, el enfoque central se centró en un conjunto de

procedimientos computacionales orientados al procesamiento, la validación y el modelado de datos tabulares provenientes del sistema de gestión académica.

Herramientas y entorno de desarrollo.

Para el desarrollo y la ejecución de los scripts de ML se empleó Google Colab Pro, que ofrece recursos avanzados, como GPU, entornos de Python y acceso a bibliotecas de ciencia de datos. El uso de esta herramienta resulta especialmente pertinente en investigaciones de carácter computacional, ya que permite ejecutar modelos de alta complejidad sin depender de infraestructura local, lo que garantiza la eficiencia en el procesamiento, la reproducibilidad del código y la flexibilidad para la experimentación iterativa. Además, su capacidad para gestionar grandes volúmenes de datos y su integración con herramientas colaborativas facilitan el desarrollo de un pipeline metodológico robusto y transparente.

Análisis exploratorio y de correlación.

En el estudio se incluyó un análisis exploratorio para comprender el comportamiento, la estructura y las características intrínsecas del conjunto de datos de los estudiantes. Este análisis ayuda a identificar patrones relevantes, irregularidades y anomalías, así como relaciones entre variables, lo que sienta las bases para el desarrollo y la validación de modelos predictivos en el futuro. Como indican (Patel & Patel, 2024) y (Mustofa et al., 2025) ,el análisis exploratorio desempeña un papel crucial en la limpieza, contextualización y comprensión semántica de los datos, asegurando que las relaciones halladas reflejen comportamientos reales del fenómeno y no meros artefactos estadísticos.

Además, se complementó con la matriz de correlaciones para identificar relaciones entre las variables y detectar posibles riesgos de multicolinealidad que pudieran afectar el rendimiento de los modelos predictivos. Esta etapa permitió comprender la estructura del conjunto de datos y orientar la selección de características para el análisis.

Explicabilidad de los resultados

Se incorporaron técnicas de Inteligencia Artificial Explicable (XAI), como SHAP y LIME, que permiten interpretar el comportamiento de las predicciones. Estas herramientas permiten analizar la contribución de cada variable a nivel global y local, facilitando una comprensión clara, transparente y accionable de los factores que influyen tanto en las predicciones individuales como en la estructura general del modelo.

Integración en el pipeline metodológico.

Todos los procedimientos mencionados constituyeron el pipeline metodológico empleado para entrenar y evaluar los algoritmos de clasificación supervisada seleccionados en esta investigación y, posteriormente, para evaluar la explicabilidad de las predicciones.

3.3.3 Validación y confiabilidad de los modelos de ML

La validación y la confiabilidad de los modelos de ML en esta investigación se garantizaron mediante la aplicación de procedimientos ampliamente aceptados en la ciencia de datos y la inteligencia artificial. Estas estrategias aseguran que los resultados obtenidos sean robustos, reproducibles y generalizables ante posibles variaciones en los datos institucionales.

Validación cruzada k-fold.

Se utilizó la técnica de validación cruzada con $k = 10$ folds. Este procedimiento divide el dataset en diez particiones de igual tamaño e intercala el entrenamiento y la evaluación, de modo que cada una sirve como conjunto de validación en una ocasión. Al finalizar, se obtuvieron métricas promedio que reflejan de manera confiable la capacidad predictiva de cada modelo y reducen el riesgo de sobreajuste.

Desbalance de clases.

Dado que la variable dependiente “desertor” presenta una distribución desigual entre las clases “deserta” y “no deserta”, se aplicaron varias técnicas de balanceo de clases para corregir el desequilibrio, como Undersampling, Oversampling, SMOTENC. Este método genera ejemplos sintéticos de la clase minoritaria, lo que favorece un conjunto de entrenamiento más equilibrado y evita que los algoritmos privilegien la clase mayoritaria, lo que mejora la sensibilidad y la estabilidad del modelo (Wongvorachan et al., 2023).

Métricas de desempeño.

Para evaluar el desempeño de los modelos, se usaron métricas muy conocidas en el ámbito de la predicción, ya sea en el ámbito educativo o en otra área de investigación, las cuales se detallan en la tabla 5.

Tabla 5

Descripción de métricas utilizadas.

Métrica	Descripción
Accuracy	Proporción de predicciones correctas.
Recall:	Mide la capacidad del modelo para identificar correctamente los casos positivos
Precision	Indica el grado de acierto entre los casos clasificados como positivos;
F1-score	Combina Recall y Precision mediante una media armónica, proporcionando una medida equilibrada del rendimiento del modelo al considerar simultáneamente la capacidad para identificar correctamente los casos positivos y la exactitud de las predicciones positivas
AUC	que evalúa la capacidad discriminativa del modelo para distinguir entre las clases; particularmente útil ante distribuciones desequilibradas;
Especificidad	Mide la proporción de casos negativos correctamente identificados, es decir, de estudiantes que no desertan.

El uso conjunto de estas métricas permite una valoración completa de la DEU, reduciendo la dependencia de indicadores sensibles al desbalance, como la exactitud, y fortaleciendo la

validez del modelo predictivo. Estudios recientes en educación superior destacan la pertinencia de los modelos de predicción del abandono, ya que permiten captar con mayor fiabilidad la complejidad del fenómeno y minimizar errores críticos de clasificación (Coussement et al., 2020; Matz et al., 2023; Wongvorachan et al., 2023).

3.3.4 Procedimientos

En el presente estudio se aplicó un conjunto estructurado de técnicas y procedimientos propios de la ciencia de datos y del ML para analizar la información. Estas metodologías se utilizaron para el procesamiento de datos tabulares y de registros longitudinales, con el propósito de asegurar la calidad de los datos, identificar variables relevantes y construir modelos predictivos con una capacidad de generalización adecuada. En este contexto, se implementó un proceso metodológico compuesto por diversas etapas: preparación de datos, análisis exploratorio, selección de variables, entrenamiento de modelos y evaluación del desempeño de los modelos. Para ello, se llevaron a cabo los siguientes procedimientos:

Análisis de los datos

- Se obtuvo un conjunto de datos institucionales del Sistema de Gestión Académica (SGA) de la UTM, que contiene variables demográficas, socioeconómicas, académicas e institucionales relevantes para el estudio de la DEU.
- Se realizó un análisis exploratorio de datos (EDA) para comprender la estructura de la DEU, evaluar la condición de deserción y establecer una relación preliminar entre los predictores. Además, se examinaron las frecuencias, las medidas de tendencia central, la dispersión, la asimetría y la curtosis, así como la proporcionalidad entre las clases.
- Asimismo, se identificaron patrones, valores atípicos y posibles sesgos estructurales mediante matrices de correlación (Pearson y de Cramér V), análisis de ANOVA y visualizaciones multivariadas. Este proceso permitió orientar la selección inicial de

características y detectar redundancias, lo que fortaleció la consistencia estadística del conjunto de datos.

Preparación de los datos

La fase de preparación incluyó procedimientos como:

- Limpieza y depuración de inconsistencias.
- Imputación diferenciada de valores faltantes según la naturaleza de la variable.
- Codificación controlada de variables categóricas.
- Estandarización y normalización de variables numéricas cuando fue necesario.
- Tratamiento de valores atípicos mediante criterios estadísticos.

Se controló la multicolinealidad eliminando variables con correlaciones elevadas o baja relevancia informativa, apoyándose en métricas como Information Value (IV/WOE), la importancia basada en RF y la selección progresiva con criterio AIC.

Dado que la variable dependiente presentaba un desequilibrio entre las clases, se aplicaron técnicas de balanceo exclusivamente en el conjunto de entrenamiento para evitar sesgo hacia la clase mayoritaria. Se evaluaron distintas estrategias, incluyendo Undersampling, Oversampling y SMOTENC. Este enfoque permitió mejorar la capacidad discriminativa del modelo sin alterar la distribución original de los conjuntos de validación y de prueba.

Evaluación y diseño experimental

Para garantizar un proceso de evaluación riguroso y reproducible, se implementó un esquema de validación cruzada estratificada de 10 iteraciones (Stratified k-Fold, $k = 10$), para preservar la proporcionalidad original de las clases en cada partición y asegurar que los modelos fueran evaluados en condiciones representativas de la distribución real de los datos.

- Este enfoque se complementó con una división del dataset: el 60 % para entrenamiento, el 20 % para validación y el 20 % para pruebas externas. Esta separación permitió

evaluar la capacidad de generalización de los modelos, evitando que las decisiones de ajuste incidieran en la evaluación final.

- Además, el balanceo de clases se aplica exclusivamente al conjunto de entrenamiento para garantizar que los conjuntos de validación y de prueba permanezcan intactos, conservando las distribuciones originales y permitiendo una evaluación imparcial del rendimiento de cada modelo.
- Posteriormente, se entrenó un amplio conjunto de algoritmos de ML y se evaluó su comportamiento en diferentes configuraciones experimentales que incluyeron la eliminación de variables redundantes, el ajuste del desequilibrio de clases y la optimización de hiperparámetros. Estas configuraciones permitieron contrastar el desempeño de los algoritmos desde perspectivas complementarias e identificar los modelos que mostraron mayor estabilidad y precisión en la predicción del riesgo de deserción universitaria.

Incorporación de Explicabilidad (XAI)

Se incorporaron técnicas de explicabilidad, tanto a nivel global como local, con el propósito de garantizar la interpretabilidad, la trazabilidad y la aplicabilidad institucional. Se emplearon:

- Para analizar la contribución de cada variable a las predicciones del modelo, se utilizó SHAP, que permite estimar su contribución al resultado del modelo. Como entrada, el método utilizó los datos procesados del conjunto de evaluación junto con el modelo predictivo previamente entrenado. A partir de esta información, SHAP calcula valores de contribución para cada variable mediante un procedimiento basado en la teoría de valores de Shapley, lo que permite descomponer la predicción del modelo en los aportes individuales de cada característica. Como resultado, se obtuvieron valores de SHAP que cuantifican el impacto de cada variable en la predicción del modelo. Estos resultados se

analizaron tanto a nivel global como a nivel individual. A nivel global, se evaluó la importancia promedio de las variables y la distribución de sus efectos en las predicciones del modelo. A nivel individual, se analizaron las contribuciones específicas de las variables para explicar la predicción de casos particulares del conjunto de datos.

- Además, para complementar el análisis de la interpretabilidad, se utilizó LIME, que permite aproximar localmente el comportamiento del modelo mediante representaciones lineales interpretables. Como entrada, el método empleó el modelo predictivo previamente entrenado y los registros individuales del conjunto de evaluación seleccionados para el análisis. A partir de estas observaciones, LIME genera perturbaciones sintéticas alrededor de cada instancia analizada para observar cómo cambian las predicciones del modelo ante pequeñas variaciones en las variables de entrada.

Posteriormente, el método ajusta un modelo lineal local que aproxima el comportamiento del modelo complejo en la vecindad de la instancia analizada. Como resultado de este procedimiento, se obtienen coeficientes que indican la contribución de cada variable a la predicción correspondiente.

A nivel global, se analizaron rankings de importancia, diagramas tipo beeswarm y la agregación de explicaciones locales. A nivel local, se estudiaron casos representativos (verdaderos positivos, falsos negativos, falsos positivos y verdaderos negativos) con el fin de comprender tanto los aciertos como los errores del modelo.

La convergencia interpretativa entre SHAP y LIME permitió validar la coherencia estructural del sistema y evidenciar que las decisiones del modelo se fundamentan en

patrones académicos consistentes, lo que refuerza su carácter explicable y no meramente predictivo.

Experimentos y validación de hipótesis específicas

- **Hipótesis específica 1: Identificación de factores que influyen en la predicción de la deserción universitaria**
 - Se realizó un análisis exploratorio de datos para examinar la relación entre las variables sociodemográficas, académicas y socioeconómicas y la DEU.
 - Se utilizaron técnicas de análisis estadístico y correlacional, incluyendo medidas de asociación como la correlación de Pearson, pruebas de chi-cuadrado y análisis de varianza (ANOVA), con el fin de identificar las variables con mayor influencia en la ocurrencia del abandono estudiantil.
 - Se emplearon métodos de selección de características, tales como la importancia de variables mediante RF, el Information Value y el ranking de variables, para determinar los factores más relevantes en la predicción de la deserción universitaria.
 - Se entrenaron modelos de ML sin incluir las variables seleccionadas, con el objetivo de evaluar su impacto en el desempeño predictivo.
- **Hipótesis específica 2: Selección de algoritmos de aprendizaje automático para la predicción de la deserción universitaria**
 - Se realizó un análisis comparativo de diferentes algoritmos de ML, tanto de modelos individuales como de modelos de ensamble, con el propósito de identificar los que presentaban el mejor desempeño predictivo.
 - Se implementaron diversos algoritmos de clasificación supervisada, tales como DT, RF, ET, GB, HGB, SVM, LR, NB y MLP.

- Se aplicaron estrategias de tratamiento del desbalance de clases, incluyendo undersampling, SMOTENC y oversampling, para evaluar su impacto en el desempeño de los modelos.
- Se utilizó un esquema de validación cruzada estratificada (k-fold) durante la fase de entrenamiento para garantizar la estabilidad y la capacidad de generalización de los modelos evaluados.
- Se compararon los modelos utilizando métricas de evaluación orientadas a problemas de clasificación desbalanceada, tales como Recall, Precision, F1-score y ROC-AUC, priorizando la capacidad de detectar a estudiantes en riesgo de abandono.
- Por último, se identificó el modelo con mejor desempeño global y mayor equilibrio entre sensibilidad y precisión, considerando además su estabilidad en los conjuntos de validación y de prueba.
- **Hipótesis específica 3: Implementación de modelos de explicabilidad para interpretar los factores asociados a la DEU**
 - Se implementaron técnicas de inteligencia artificial explicable (XAI) para analizar el comportamiento de los modelos predictivos y comprender la influencia de las variables en las predicciones.
 - Se aplicaron métodos de explicabilidad global, como SHAP, para identificar las variables con mayor impacto en la predicción del riesgo de deserción universitaria.
 - Asimismo, se utilizaron métodos de explicabilidad local, como LIME (Local Interpretable Model-Agnostic Explanations), para interpretar la contribución de cada variable a las predicciones individuales de los estudiantes.

- Se analizaron las relaciones entre las variables académicas, institucionales y socioeconómicas y el riesgo de abandono estudiantil, identificando patrones explicativos relevantes en el modelo.
- Se evaluó la consistencia entre los resultados obtenidos mediante técnicas de explicabilidad y los hallazgos del análisis predictivo.
- **Validación de la hipótesis general**
 - Se realizó un análisis comparativo integral de los modelos evaluados en las hipótesis específicas, con el objetivo de determinar si la integración de técnicas de aprendizaje automático y de métodos de explicabilidad mejora la capacidad de análisis y de comprensión de la deserción universitaria.
 - Se examinaron la robustez y la estabilidad de los modelos mediante diferentes configuraciones experimentales, que incluyeron estrategias de balanceo de datos, validación cruzada y optimización de hiperparámetros.
 - Se aplicaron técnicas de explicabilidad global y local para verificar la consistencia de los factores identificados como influyentes en la predicción del abandono estudiantil.
 - Finalmente, se analizaron las implicaciones de los resultados obtenidos para la gestión institucional, destacando el potencial del modelo explicable propuesto como herramienta para la detección temprana del riesgo de deserción universitaria y el diseño de estrategias de intervención académica.

3.4. Matriz De Consistencia

En la tabla 6 se presenta la matriz de consistencia con los problemas, objetivos e hipótesis identificados, como se describe a continuación.

Tabla 6

Matriz de consistencia

Aspectos	General	Específico 1	Específico 2	Específico 3
Problema	¿Cómo pueden los modelos de explicabilidad basados en el aprendizaje automático contribuir a predecir e interpretar los factores asociados a la deserción estudiantil universitaria?	¿Cuáles son los factores más relevantes que deben integrarse en un conjunto de datos para modelar adecuadamente la deserción estudiantil universitaria?	¿Qué modelos de aprendizaje automático presentan la mayor capacidad predictiva y la mayor robustez para identificar a estudiantes en riesgo de deserción universitaria?	¿Cómo pueden interpretar los modelos de explicabilidad la contribución y las interacciones entre los factores asociados a la deserción estudiantil universitaria?
Objetivos	Desarrollar un modelo de explicabilidad basado en el aprendizaje automático que permita predecir e interpretar los factores asociados a la deserción estudiantil universitaria.	Construir un conjunto de datos con los factores relevantes que influyan en la predicción de la deserción universitaria	Entrenar y evaluar algoritmos de aprendizaje automático para predecir la deserción estudiantil universitaria, considerando criterios de desempeño y de robustez del modelo.	Diseñar un modelo explicativo para interpretar la contribución y las relaciones entre los factores que inciden en la deserción estudiantil universitaria.
Hipótesis	El desarrollo de un modelo de explicabilidad basado en el aprendizaje automático permite interpretar de manera efectiva los factores que influyen en la deserción universitaria.	La construcción de un conjunto de datos integrado por variables relevantes permite modelar adecuadamente la deserción estudiantil universitaria.	El entrenamiento y la evaluación de algoritmos de aprendizaje automático permiten obtener modelos de predicción para identificar la deserción estudiantil universitaria.	El diseño de un modelo explicativo permite interpretar la contribución y las relaciones entre los factores asociados a la deserción estudiantil universitaria.

CAPÍTULO IV. FASE EXPERIMENTAL

La fase experimental constituye el centro del estudio y se orienta a validar el modelo de explicabilidad mediante técnicas de ML y herramientas de explicabilidad (XAI), aplicadas a datos académicos, socioeconómicos, institucionales y contextuales de estudiantes de universidades públicas.

Este proceso está compuesto por las siguientes partes:

- (1) Análisis y procesamiento del dataset.
- (2) Construcción de escenarios experimentales.
- (3) Implementación de modelos predictivos.
- (4) Aplicación de técnicas de explicabilidad.

4.1. Análisis de dataset

Como parte del análisis preliminar, se verificó la consistencia lógica de la información para asegurar su validez en las etapas posteriores del modelado. Se aplicaron técnicas de normalización para unificar las escalas de medición de las variables académicas, socioeconómicas y administrativas, lo que facilitó su integración en los algoritmos de ML. De manera complementaria, se implementaron estrategias para mitigar el impacto del ruido presente en los registros institucionales, considerando que ciertos factores externos, como reformas en las políticas estudiantiles, cambios administrativos o variaciones en los calendarios académicos, podrían introducir fluctuaciones en indicadores sensibles al riesgo de abandono. Para reducir estos efectos y mejorar la estabilidad del conjunto de datos, se utilizaron valores promedio y medidas agregadas en periodos comparables, con el fin de atenuar las variaciones atípicas y garantizar una representación más confiable del comportamiento estudiantil.

4.1.1 Ventajas del dataset seleccionado

El dataset contiene características que favorecen el análisis riguroso de la deserción universitaria, entre las que se pueden mencionar:

- **Volumen adecuado:** permite entrenar modelos de ML con una eficacia estadística adecuada.
- **Diversidad multidimensional:** Incluye factores individuales, institucionales, socioeconómicos y personales, lo que permite modelar la DEU desde un enfoque sistémico.
- **Coherencia semántica:** las variables se estandarizaron para garantizar la consistencia conceptual entre las cohortes y los programas académicos.
- **Pertinencia contextual:** Debido a que los datos provienen de una universidad pública, un entorno en el que la deserción presenta patrones específicos vinculados a la inequidad, la precariedad económica y las brechas digitales.

4.1.2 Limitaciones del dataset

Si bien el conjunto de datos de esta investigación constituye una base robusta y adecuada para el análisis de la DEU, resulta indispensable reconocer las limitaciones que pueden afectar la validez interna y externa de los resultados. La deserción es un fenómeno complejo, multidimensional y heterogéneo, por lo que la capacidad interpretativa y predictiva se ve afectada por la calidad, la exhaustividad y la naturaleza de los datos. Entre las limitaciones más notables tenemos:

- **Desbalanceo de la clase objetivo:** La prevalencia de desertores es menor que la de los estudiantes regulares, lo cual es común en el ámbito educativo. Este desequilibrio puede inducir a modelos que favorecen a la clase mayoritaria y obtienen un alto rendimiento en métricas como el AUC, pero con una sensibilidad limitada para detectar el abandono.

- **Sesgo territorial:** La información proviene exclusivamente de instituciones de una región geográfica específica, lo que limita la generalización inmediata de los hallazgos a otros contextos. Los patrones de permanencia y abandono académicos suelen verse profundamente influidos por factores culturales, económicos, institucionales y sociales propios de cada país o territorio, por lo que los resultados obtenidos reflejan principalmente las dinámicas locales.
- **Calidad de los datos:** Se observaron datos faltantes, inconsistencias y disparidades en los datos institucionales. Esta situación es frecuente en investigaciones en el ámbito educativo, dado que las instituciones emplean múltiples plataformas y procedimientos de captura de datos con distintos niveles de estandarización, lo que genera sesgos residuales que es necesario considerar al interpretar los resultados.
- **Presencia de ruido:** Asimismo, se detectó ruido en diversas variables, derivado de fluctuaciones administrativas, decisiones institucionales coyunturales o diferencias en la forma en que las unidades académicas registran la información. Esta variabilidad introduce distorsiones que pueden afectar la estabilidad de los modelos de predicción y de explicación.

Estas limitaciones no invalidan la pertinencia ni la solidez del estudio; por el contrario, permiten contextualizar el alcance del modelo y orientar futuras líneas de investigación que integren dimensiones psicosociales, contextuales y longitudinales, fortaleciendo la capacidad explicativa y predictiva de la DEU en entornos complejos.

4.1.3 Preparación y enriquecimiento del dataset

La preparación del dataset constituyó una etapa esencial en el desarrollo del modelo de explicabilidad para DEU, dado que la calidad de los modelos de predicción y la estabilidad de los algoritmos de ML dependen de la consistencia, la integridad y la estructuración de los datos

analizados (Han et al., 2022; Romero & Ventura, 2020). En esta fase se emplearon los procesos de selección de variables, diagnóstico y tratamiento de valores faltantes, codificación de variables categóricas, estandarización de características y construcción de escenarios, que pueden o no incluir el balanceo de clases (Chapman et al., 2000).

En primera instancia se consideraron variables personales y sociodemográficas, tales como género, edad, estado civil, etnia, nacionalidad, orientación sexual y preferencia religiosa, así como variables de procedencia y residencia (país, provincia, cantón, tipo de zona y localización de la vivienda). De forma complementaria, se incorporaron indicadores familiares y socioeconómicos, entre ellos el número de hijos, el tamaño del núcleo familiar, la ocupación del jefe del hogar, los niveles de instrucción de la madre y del padre, el tipo de vivienda, la recepción de ayuda familiar, los ingresos totales del hogar y los gastos asociados a la alimentación y al transporte.

Asimismo, se integraron factores relacionados con el acceso y uso de tecnologías digitales, tales como la disponibilidad de internet en el hogar y en el teléfono móvil, las horas de uso de internet y de redes sociales, así como el uso de dispositivos tecnológicos (computador y tablet), dada su relevancia en contextos de educación superior mediada por entornos virtuales. En el ámbito académico, se incluyeron variables relacionadas con el desempeño y la trayectoria estudiantil, como la facultad, la carrera, la modalidad de estudio, el nivel académico, el número de materias matriculadas, el promedio de calificaciones, los semestres tomados y anulados, la nota y la sostenibilidad del colegio de procedencia, así como el periodo académico del registro.

Posteriormente, se realizó un análisis de valores faltantes, identificando el número de registros ausentes en cada columna y su proporción respecto del total de observaciones. El tratamiento de datos incompletos resulta crucial en entornos educativos, ya que la ausencia de información suele asociarse con perfiles de mayor vulnerabilidad académica y social (Buuren, 2018) . De

este modo, se priorizó la aplicación de técnicas de imputación adecuadas para datos predominantemente categóricos, evitando la pérdida de información que implicaría la eliminación de registros.

Debido a que los algoritmos de ML requieren entradas numéricas, las variables categóricas se codificaron posteriormente mediante LabelEncoder, con un codificador independiente para cada variable. Esto permitió representar las categorías discretas sin incrementar excesivamente la dimensionalidad del dataset, un problema que suele presentarse con frecuencia cuando se utilizan métodos como el one-hot encoding en datasets extensos con múltiples variables categóricas (Kuhn & Johnson, 2019) . La variable objetivo también se codificó como binaria para asegurar su correcta interpretación durante el entrenamiento de los modelos.

A continuación, se definieron la matriz de características (X) y el vector objetivo (y) y se implementó una partición estratificada en entrenamiento, validación y prueba. El 60 % se utilizó para el entrenamiento de los modelos, mientras que el 40 % se dividió entre la validación y las pruebas, con un 20% en cada una. Este particionamiento estratificado se recomienda en escenarios de clases desbalanceadas, ya que evita distorsiones en la representación de la clase minoritaria (Branco et al., 2016).

En relación con los subconjuntos generados, se realizó una estandarización mediante StandardScaler, ajustando la media y la desviación estándar únicamente en los datos de entrenamiento para evitar fugas de información y, posteriormente, aplicando los parámetros resultantes a los conjuntos de validación y de prueba. Esta técnica es esencial en modelos sensibles a la escala, ya que mejora la estabilidad numérica y la eficiencia en las etapas de optimización (Han et al., 2022).

Finalmente, se diseñaron escenarios mediante técnicas de balanceo, como Undersampling, Oversampling, SMOTENC. Estas técnicas han demostrado ser efectivas para detectar la clase

minoritaria en contextos educativos y para evitar que los modelos prioricen a la clase mayoritaria (Branco et al., 2016). La comparación entre el escenario no balanceado y el balanceado permitió evaluar el impacto directo del desbalanceo en las métricas de rendimiento y en la capacidad de los modelos para identificar a estudiantes con alto riesgo de deserción.

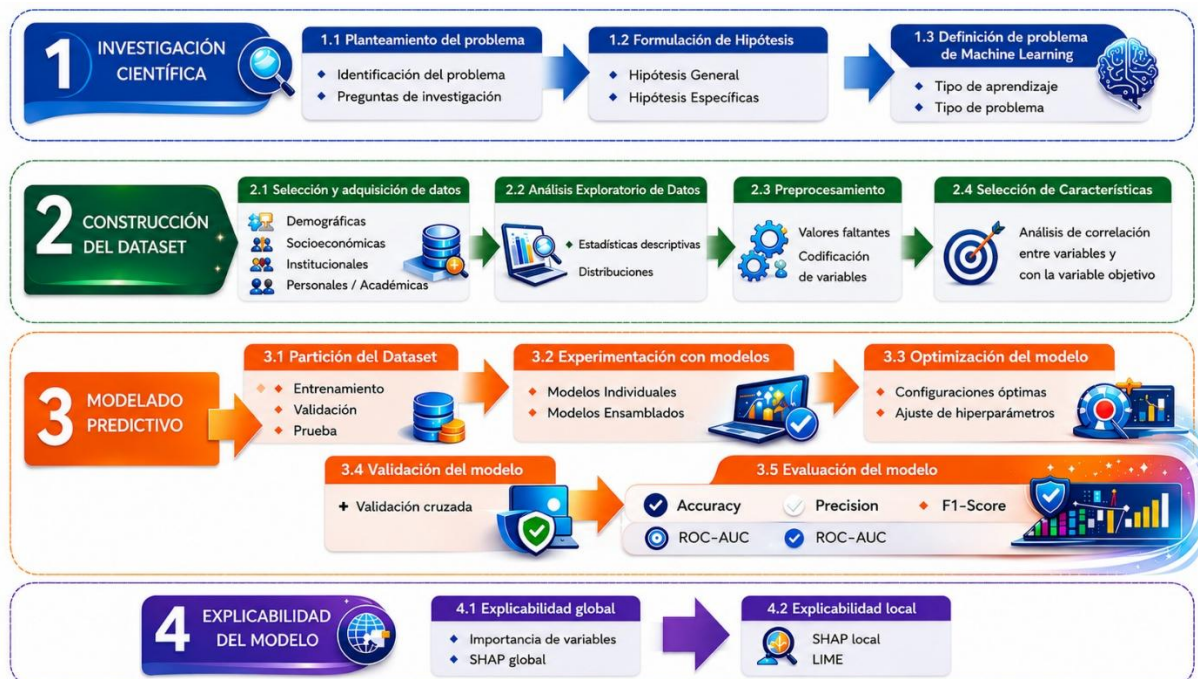
Los procedimientos aplicados, tales como la selección de variables, el tratamiento de valores faltantes, la codificación categórica, la partición estratificada, la estandarización y el balanceo de clases, constituyen un apartado fundamental en la preparación de datos robustos, reproducibles y alineados con las mejores prácticas de minería de datos educativa y de ML aplicado a la DEU (Castaño et al., 2017; Han et al., 2022; Romero & Ventura, 2020) .

4.2. Modelo Propuesto

La arquitectura metodológica de la figura 1 describe el proceso estructurado para desarrollar modelos explicables orientados a identificar el riesgo de la DEU. Este enfoque integra etapas propias de la investigación científica, la construcción del dataset, el modelado predictivo y la inteligencia artificial explicable, lo que permite abordar el fenómeno desde una perspectiva analítica y metodológicamente robusta.

Figura 1

Modelo de explicabilidad con ML para la DEU



4.2.1. Investigación científica

Permite delimitar con rigor el problema de investigación, estructurar las preguntas que orientan el análisis y formular las hipótesis que posteriormente se contrastarán mediante técnicas de ML. En esta fase inicial no se construyen modelos predictivos, sino que se establecen las bases conceptuales, metodológicas y analíticas que justifican todo el proceso posterior. En otras palabras, esta etapa cumple una función estratégica porque traduce una problemática real del contexto educativo en un problema científicamente abordable, susceptible de ser estudiado desde una perspectiva empírica, sistemática y reproducible.

Planteamiento del problema

En esta fase se formaliza el problema de investigación. Esto implica reconocer una necesidad real, relevante y respaldada por evidencia que, en el contexto de la tesis, se vincula con la deserción universitaria y con la necesidad de comprenderla, predecirla y explicarla mediante técnicas de ML. El planteamiento no debe limitarse a describir una situación problemática, sino que debe delimitar con precisión el fenómeno de estudio, su contexto de ocurrencia, su impacto

académico o institucional y la justificación de por qué requiere ser abordado científicamente. Asimismo, se formulan las preguntas de investigación, que orientan todo el desarrollo metodológico y guardan coherencia con el problema central, los objetivos de la tesis y el alcance explicativo y predictivo del estudio. En investigaciones sobre ML, las preguntas suelen centrarse en aspectos como qué variables se relacionan con el fenómeno, qué tan bien puede predecirse la ocurrencia del evento de interés y qué factores explican en mayor medida las decisiones del modelo.

Formulación de hipótesis

Una vez delimitado el problema, se procede a formular las hipótesis, diferenciando la hipótesis general de las hipótesis específicas. La hipótesis general expresa la proposición central que articula el estudio y resume la relación principal que se espera demostrar o sustentar. En una tesis con enfoque predictivo y explicativo, esta hipótesis suele vincular el uso de modelos de ML con la capacidad de identificar patrones asociados a la deserción universitaria y de explicar la influencia de variables relevantes en dicho fenómeno.

Por su parte, las hipótesis específicas desglosan esa proposición central en afirmaciones más concretas y verificables, asociadas a dimensiones particulares del problema. Estas pueden estar relacionadas con el aporte de ciertos conjuntos de variables, el desempeño diferencial de determinados algoritmos o la utilidad de técnicas de explicabilidad para interpretar los resultados obtenidos. Su formulación debe ser clara, coherente con los objetivos específicos y contrastable metodológicamente.

Definición del problema de ML

Esta fase consiste en traducir el problema científico en un problema de ML computacional. En esta parte se define, en primer lugar, el tipo de aprendizaje que se utilizará. En el caso de una tesis orientada a predecir la deserción universitaria a partir de registros históricos etiquetados,

corresponde emplear el aprendizaje supervisado, dado que se dispone de una variable objetivo previamente identificada.

En segundo lugar, se determina el tipo de problema que los modelos resolverán. Cuando la variable objetivo indica si un estudiante deserta o no, el estudio se configura como un problema de clasificación, ya que el algoritmo debe asignar cada caso a una categoría previamente definida. Esta definición no es menor, pues condiciona la selección de algoritmos, las métricas de evaluación, la estructura de los datos y los mecanismos de validación que se aplicarán en fases posteriores.

En consecuencia, constituye el punto de conexión entre la investigación científica tradicional y la modelación computacional. Aquí se establece cómo formalizar un problema educativo complejo para su tratamiento mediante herramientas de inteligencia artificial, sin perder su carácter científico ni su pertinencia contextual.

4.2.2. Construcción del dataset

La construcción del dataset constituye una fase crítica, ya que de su solidez depende, en gran medida, la validez de los resultados obtenidos. Además, esta etapa transforma el problema científico delimitado en una base de datos analítica estructurada, consistente y apta para el modelado predictivo.

Selección de variables

La primera actividad de esta fase consiste en definir las variables que conformarán el dataset. Esta selección no debe realizarse de manera arbitraria ni exclusivamente por la disponibilidad de información, sino a partir de criterios teóricos, empíricos y contextuales que sustenten su inclusión en el estudio. En ese sentido, las variables deben guardar una relación directa con el problema de investigación, con los objetivos planteados y con la literatura especializada sobre la deserción universitaria. Las variables pueden organizarse en diferentes grupos, como los

demográficos, socioeconómicos, institucionales, personales y académicos. Esta clasificación permite abordar el fenómeno desde una perspectiva multidimensional, reconociendo que la deserción no responde a una única causa, sino a la compleja interacción de diversos factores.

Comprensión de los datos mediante análisis exploratorio

Superada la auditoría inicial, se realiza un análisis exploratorio de datos (EDA) del dataset para examinar el comportamiento general de las variables, identificar patrones preliminares, detectar anomalías y comprender mejor la estructura interna de la información disponible. El EDA cumple una función diagnóstica, ya que ayuda a identificar características relevantes del conjunto de datos antes de aplicar transformaciones o entrenar modelos.

En esta etapa se analizan aspectos como la estadística descriptiva, las distribuciones de las variables y la presencia de valores atípicos. La estadística descriptiva permite resumir las principales propiedades de los datos, como la tendencia central, la dispersión y el rango. En segundo lugar, se examinan los valores faltantes, con el fin de determinar su magnitud, localización y posible impacto en la calidad del conjunto de datos, aspecto clave para definir estrategias de tratamiento adecuadas. Finalmente, se realiza el análisis de outliers, orientado a identificar observaciones atípicas que puedan influir de manera significativa en el comportamiento de los modelos o distorsionar la interpretación de los resultados.

Preprocesamiento

Luego de comprender la estructura y la calidad de los datos, se realiza el preprocesamiento del dataset, cuyo propósito es preparar la información para su uso eficaz en los algoritmos de ML. Este proceso comprende transformaciones orientadas a mejorar la compatibilidad técnica, la estabilidad analítica y el rendimiento predictivo de los modelos.

Entre las tareas principales del preprocesamiento se encuentran la limpieza de datos, el tratamiento de valores faltantes, el balanceo de clases y la codificación de variables categóricas.

La limpieza incluye la corrección o eliminación de registros erróneos, inconsistentes o irrelevantes. El tratamiento de faltantes puede implicar distintas estrategias, como la eliminación de registros, la imputación estadística o métodos más avanzados, según la magnitud del problema y la naturaleza de la variable. Por último, la codificación de variables categóricas permite convertir atributos cualitativos en valores numéricos adecuados para el entrenamiento algorítmico.

Selección de características

Esta fase tiene como objetivo identificar el subconjunto de variables más relevantes para el modelado. En consecuencia, esta etapa busca optimizar la representación del problema, conservando únicamente las variables con mayor capacidad explicativa o predictiva. Esta selección se basa en criterios como la correlación y la importancia de las variables. El análisis de correlación permite detectar relaciones fuertes entre atributos y, en algunos casos, identificar posibles redundancias. Por su parte, la estimación de la importancia permite identificar qué variables contribuyen en mayor medida al aprendizaje del modelo o a la discriminación entre clases.

4.2.3 Modelado predictivo

Constituye la fase central de la experimentación, ya que en este nivel se materializa la transición desde una base de datos previamente depurada y estructurada hasta la construcción de modelos capaces de identificar patrones, estimar comportamientos y predecir la ocurrencia del fenómeno de interés. Esta etapa tiene como propósito desarrollar, entrenar, comparar y validar distintos algoritmos de ML orientados a la predicción de la DEU, procurando no solo alcanzar un desempeño satisfactorio, sino también asegurar que los resultados obtenidos sean sólidos, reproducibles y científicamente interpretables.

Esta fase no es una mera aplicación de algoritmos a los datos, sino un proceso experimental estructurado. El modelado predictivo implica definir cómo se organizarán los datos para el entrenamiento y la evaluación, qué algoritmos se considerarán, bajo qué configuraciones se ejecutarán, cómo se validará su estabilidad y con qué métricas se comparará su rendimiento.

Partición del dataset

La partición del dataset tiene como objetivo dividir los datos en subconjuntos destinados a funciones específicas del proceso experimental. Esta división es fundamental para evitar sesgos de evaluación y garantizar que el rendimiento de los modelos se estime sobre datos no utilizados durante el entrenamiento. Durante esta etapa se permite diferenciar claramente entre la fase de ajuste del modelo y la de comprobación de su capacidad de generalización.

El dataset se dividirá en tres subconjuntos: entrenamiento, validación y prueba. El conjunto de entrenamiento se utilizará para que los algoritmos aprendan los patrones presentes en los datos históricos. El conjunto de validación servirá para ajustar configuraciones, comparar alternativas y controlar el riesgo de sobreajuste durante el proceso experimental. Finalmente, el conjunto de prueba se reservará para la evaluación final del modelo, lo que permitirá estimar su comportamiento sobre datos no vistos.

Esta partición es particularmente relevante porque fortalece la validez interna del experimento, lo que asegura que los resultados no se deban a una adaptación excesiva del modelo a los datos disponibles, sino a una capacidad real de predicción. Además, esta estrategia aporta transparencia y robustez al proceso metodológico al establecer una estructura clara para el entrenamiento, el ajuste y la evaluación final.

Experimentación con modelos

Esta fase comprende el entrenamiento y la comparación de distintos algoritmos de ML para identificar cuáles ofrecen los mejores resultados. La experimentación no se limita a un único

enfoque, sino que contempla la evaluación tanto de modelos individuales como de ensamblados, lo que permite contrastar distintas capacidades de aprendizaje frente a la DEU

Los modelos individuales pueden incluir algoritmos como la regresión logística (LR), los árboles de decisión (DT), los Random Forests (RF) u otros clasificadores supervisados. Estos modelos ofrecen la ventaja de permitir una primera aproximación al comportamiento predictivo de los datos y, en algunos casos, facilitan la interpretación de los resultados. Por su parte, los modelos ensamblados, como los métodos de boosting o las combinaciones de clasificadores, permiten aprovechar la complementariedad entre múltiples estimadores para mejorar la precisión y la robustez del sistema predictivo.

De tal forma, se realizará el entrenamiento de cada modelo seleccionado utilizando el conjunto de entrenamiento en condiciones controladas y comparables. La finalidad no es determinar únicamente qué algoritmo produce mejores métricas, sino también examinar cómo responde cada técnica ante la estructura particular del dataset, el posible desbalance de clases y la complejidad de las relaciones entre variables. En términos académicos, esta etapa constituye el núcleo experimental de la investigación, ya que permite contrastar empíricamente distintas estrategias de modelado según el objetivo predictivo planteado.

Validación del modelo

Después del entrenamiento, se validarán los modelos para verificar la estabilidad de su desempeño y reducir el riesgo de llegar a conclusiones basadas en resultados circunstanciales. La validación cumple una función esencial porque permite comprobar si el comportamiento observado en una partición específica de los datos se mantiene de forma consistente al someter el modelo a distintos subconjuntos de datos.

Esta validación se realizará mediante validación cruzada, ampliamente utilizada para estimar la capacidad de generalización de los modelos. A través de este procedimiento, el conjunto de

entrenamiento se subdivide en múltiples particiones, de modo que el algoritmo pueda entrenarse y evaluarse repetidamente con distintas combinaciones de datos. Esto permite obtener una estimación más estable y confiable del rendimiento, reduciendo la dependencia de una única división inicial.

Optimización del modelo

Una vez validados los modelos base, se optimizará el modelo, etapa en la que se ajustarán las configuraciones internas de los algoritmos para mejorar su rendimiento. En esta etapa, se buscarán configuraciones adecuadas para cada algoritmo, considerando aspectos como la profundidad de los árboles, el número de estimadores, las tasas de aprendizaje, los criterios de partición u otros parámetros específicos según la técnica utilizada. La optimización no persigue únicamente elevar los valores de desempeño de manera aislada, sino también encontrar combinaciones equilibradas que mejoren la capacidad predictiva sin comprometer la capacidad de generalización del modelo.

Evaluación del modelo

Esta fase tiene como objetivo medir objetivamente el desempeño de las distintas alternativas entrenadas y optimizadas. Esta evaluación se realizará utilizando el conjunto de prueba, que ha permanecido separado durante las etapas previas, con el fin de estimar la capacidad real de generalización del modelo sobre datos no observados durante el entrenamiento.

La evaluación se apoyará en métricas como accuracy, precision, recall, F1-score y ROC-AUC. Cada una de estas medidas aporta una perspectiva distinta sobre el rendimiento del clasificador. La exactitud permite observar la proporción global de aciertos; la precisión informa qué tan confiables son las predicciones positivas; el recall indica la capacidad del modelo para identificar correctamente los casos positivos; la puntuación F1 equilibra precisión y exhaustividad; y el área bajo la curva ROC ofrece una visión integral de la capacidad

discriminativa del modelo. La interpretación de las métricas deberá realizarse de forma integral, considerando tanto el equilibrio entre las clases como las implicaciones del problema en el contexto educativo. Esto permitirá establecer comparaciones sólidas entre los modelos evaluados y fundamentar la selección del modelo de mejor desempeño para las etapas posteriores de análisis explicativo.

4.2.4 Explicabilidad del modelo

Constituye la fase final del proceso metodológico y un componente esencial de las investigaciones con ML para abordar fenómenos complejos en el ámbito educativo. Si bien el modelado predictivo permite identificar patrones y estimar la probabilidad de ocurrencia de la DEU, dichos resultados resultan insuficientes desde una perspectiva científica si no se comprende, con un nivel razonable de profundidad, por qué el modelo toma determinadas decisiones y qué variables influyen en mayor medida en sus predicciones. En este sentido, la explicabilidad introduce una capa analítica adicional que transforma el modelo de una herramienta meramente predictiva en un instrumento interpretable y útil para la generación de conocimiento.

Explicabilidad global

Tiene como objetivo analizar el comportamiento general del modelo en el conjunto completo de datos e identificar qué variables tienen mayor influencia en la DEU. A diferencia de las métricas de evaluación, que muestran qué tan bien funciona el modelo, la explicabilidad global permite comprender cuáles son los atributos que, en términos agregados, tienen mayor peso en la toma de decisiones.

En esta etapa se emplearán mecanismos de interpretación orientados a estimar la importancia de las variables y el aporte global de cada característica, mediante técnicas como SHAP global. La importancia de las variables permite establecer una jerarquía de atributos más influyentes

para identificar qué dimensiones inciden más en el comportamiento predictivo del modelo. Sin embargo, esta aproximación por sí sola no siempre resulta suficiente, dado que suele expresar magnitudes de aporte sin describir con claridad la dirección ni la variabilidad del efecto. En consecuencia, la explicabilidad global ofrece una visión estructural del modelo, lo que permite interpretar de manera general los patrones que este ha aprendido durante el proceso experimental.

Explicabilidad local

Mientras la explicabilidad global examina el comportamiento promedio del modelo, la explicabilidad local se orienta al análisis de casos individuales. Su propósito es comprender por qué el modelo asigna una predicción específica a un estudiante en particular e identificar qué variables contribuyen positivamente o negativamente a dicha predicción. La explicabilidad local se abordará mediante técnicas como SHAP local y LIME. A través de SHAP local, será posible descomponer la predicción de un caso individual en contribuciones específicas de cada variable, mostrando qué atributos influyen en la decisión del modelo hacia una mayor probabilidad de deserción o, por el contrario, de permanencia. Por su parte, LIME complementa este análisis al generar explicaciones locales aproximadas del comportamiento del modelo en torno a una observación concreta. Su utilidad radica en que permite construir una interpretación comprensible de predicciones individuales, incluso al trabajar con modelos complejos o de difícil interpretación directa. La combinación de SHAP local y LIME fortalece el análisis, ya que permite contrastar distintas aproximaciones explicativas y obtener una visión más robusta de los factores que influyen en casos específicos.

4.3. Implementación de Modelos

La elección de los algoritmos de ML se fundamenta en la evidencia empírica de la literatura, de acuerdo con el trabajo de (Quimiz-Moreira et al., 2025) que describe de forma

amplia los modelos supervisados y no supervisados, incluyendo técnicas lineales, algoritmos basados en árboles, métodos probabilísticos, modelos de ensamblaje y arquitecturas neuronales

Entre los algoritmos más usados está DT, debido a que presenta una estructura interpretable, empleada por (Yağcı, 2022) y (Guleria & Sood, 2023) , quienes resaltan su utilidad para identificar factores críticos de riesgo académico. De forma complementaria, el algoritmo XGBoost destaca por su eficacia en entornos con datos desbalanceados y estructuras no lineales, condiciones comunes en los estudios sobre la deserción universitaria. En el trabajo de (T. Chen & Guestrin, 2016) , XGBoost ha mostrado una capacidad superior para optimizar el aprendizaje mediante el boosting de gradiente estructurado, reduciendo los errores residuales. Investigaciones más recientes refuerzan su pertinencia: (Wang et al., 2020) describen un desempeño destacado en la predicción de la DEU, mientras que (Miranda et al., 2021) reportan que supera de forma consistente a modelos clásicos y otros métodos de ensamblaje en tareas de clasificación en educación.

En esa misma línea, CatBoost y LightGBM destacan por su eficiencia computacional, su capacidad para reducir el sobreajuste y procesar grandes volúmenes de información a alta velocidad, lo cual resulta especialmente relevante en estudios educativos en los que la información proviene de fuentes heterogéneas. Como evidencian (Prokhorenkova et al., 2018) y (Ke et al., 2017) , CatBoost ofrece un tratamiento avanzado e independiente de sesgos para variables categóricas mediante técnicas de ordenamiento eficientes, mientras que LightGBM optimiza el proceso de entrenamiento mediante el crecimiento por hojas, lo que permite capturar interacciones de gran complejidad entre características a bajo costo.

En el caso de las redes neuronales, el MLP destaca por su capacidad para modelar relaciones no lineales y captar interacciones entre las variables asociadas al riesgo de deserción. Su

arquitectura de capas totalmente conectadas permite aprender representaciones profundas del comportamiento estudiantil y ofrece un rendimiento superior al de los modelos lineales en tareas de predicción en el ámbito educativo. Los estudios de (Ahuja & Khosla, 2020), (Al-Amin et al., 2021) y (Saa, 2022) evidencian que MLP alcanza altos niveles de precisión y estabilidad en contextos educativos heterogéneos, lo que justifica su inclusión entre los algoritmos evaluados en esta investigación. Asimismo, RF es uno de los algoritmos más robustos y utilizados para la predicción del riesgo de abandono universitario, por su estabilidad, su capacidad para manejar la dimensionalidad y su buen rendimiento con datos desbalanceados. Estudios como los de (Bezek Güre, 2024) y (González-Nucamendi et al., 2023) evidencian altos niveles de precisión en la clasificación de estudiantes con probabilidad de deserción, superando a modelos lineales y otros métodos de ensamblaje. Por su parte, NB continúa siendo utilizado en investigaciones que requieren rapidez computacional y eficiencia en contextos con múltiples características, tal como lo demuestran (Páez, 2022) y (A. S. Martínez et al., 2023), quienes señalan que ofrece un desempeño competitivo y una estructura interpretativa que facilita el análisis educativo.

Entre los métodos de boosting se incluyen Gradient Boosting (GB) y HistGradientBoosting (HGB), que optimizan secuencialmente los errores residuales mediante la construcción iterativa de árboles débiles. En el trabajo de (H. Chen et al., 2020) se destaca la eficacia de GB para problemas de clasificación en educación, mientras que la versión basada en histogramas mejora la eficiencia computacional y el manejo de grandes volúmenes de datos, lo cual resulta pertinente en entornos institucionales con múltiples cohortes académicas.

En el grupo de modelos lineales y de margen máximo, la regresión logística (LR) se incorpora por su alta interpretabilidad y por permitir estimar probabilidades asociadas al evento de deserción, lo que facilita su integración en análisis explicativos. Por su parte, el SVM permite

modelar fronteras de decisión no lineales mediante funciones de kernel y muestra un desempeño sólido en escenarios con interacciones complejas entre variables académicas y sociodemográficas.

Asimismo, el modelo probabilístico Naïve Bayes (NB) se incluye como referencia comparativa por su eficiencia computacional y su simplicidad estructural. Aunque asume independencia condicional entre variables, estudios como los de (Páez, 2022) y (A. S. Martínez et al., 2023) demuestran que puede ofrecer resultados competitivos en tareas de clasificación educativa. Finalmente, el Multi-Layer Perceptron (MLP) representa el enfoque basado en redes neuronales, lo que permite capturar relaciones no lineales profundas y modelar interacciones complejas entre características, tal como evidencian (Ahuja & Khosla, 2020), (Al-Amin et al., 2021) y (Saa, 2022).

Inicialmente, también se consideraron algoritmos avanzados como XGBoost, LightGBM y CatBoost, ampliamente reportados en la literatura por su alto rendimiento en problemas de clasificación estructurada (T. Chen & Guestrin, 2016; Ke et al., 2017; Prokhorenkova et al., 2018). Sin embargo, tras el análisis preliminar del comportamiento del dataset institucional y considerando los objetivos centrales de la investigación, en particular la necesidad de equilibrio entre el desempeño predictivo y la explicabilidad, se decidió no incluirlos en la fase experimental final. Esta decisión se fundamenta en tres aspectos metodológicos: (i) la similitud estructural con los métodos de boosting ya evaluados (GB e HGB), lo que podría generar redundancia experimental; (ii) el incremento de la complejidad computacional sin mejoras sustanciales en las métricas preliminares; y (iii) la priorización de modelos cuya estructura facilite una integración más transparente con técnicas de explicabilidad como SHAP y el análisis surrogado, eje central del modelo doctoral propuesto.

En conjunto, la selección definitiva de los algoritmos es: HGB, ET, DT, GB, RF, LR, SVM, MLP y NB y responde a una estrategia metodológica que combina interpretabilidad, robustez estadística, eficiencia computacional y capacidad de modelado no lineal. Esta diversidad algorítmica permite comparar distintos paradigmas bajo un mismo esquema experimental y, especialmente, fortalecer el componente de explicabilidad del modelo propuesto, elemento esencial para comprender de manera rigurosa los factores asociados a la deserción universitaria en el contexto institucional analizado.

4.4. Métricas de evaluación del modelo

Para la evaluación del desempeño de los modelos se utilizó un conjunto de métricas que consideró la naturaleza desbalanceada del fenómeno de la DEU y la necesidad de identificar con precisión a los estudiantes en riesgo. En primer lugar, se analizaron los verdaderos positivos (TP), es decir, los estudiantes que desertaron y fueron correctamente identificados por el modelo. Su análisis resulta central, dado que la finalidad del modelo de predicción es anticipar el abandono.

De manera complementaria, se incorporó la tasa de verdaderos positivos (TPR o Recall), relevante en contextos educativos, donde resulta crítico evitar que los estudiantes vulnerables pasen inadvertidos. Como señalan (Praseeda & Shivakumar, 2021), el recall es adecuado cuando la clase positiva conlleva un riesgo institucional elevado. A su vez, se evaluaron los falsos positivos (FP) y la tasa de falsos positivos (FPR), factores fundamentales para determinar la estabilidad del modelo y evitar la asignación innecesaria de intervenciones educativas a estudiantes que no presentan un riesgo real.

También se consideraron los falsos negativos (FN), errores particularmente críticos en este estudio, pues implican no identificar a estudiantes que sí desertaron. Para equilibrar la proporción entre predicciones correctas y falsas alarmas, se utilizó la métrica de precisión, que

permite evaluar la calidad de las alertas generadas por el modelo. De forma complementaria, se empleó la exactitud (Accuracy) como métrica global; no obstante, su interpretación se realizó con cautela debido a su sensibilidad a escenarios desbalanceados, en concordancia con lo advertido por (Powers, 2011), quien destaca los peligros de la exactitud como métrica única en problemas en los que una clase domina el conjunto de datos.

Dado que la deserción constituye una clase minoritaria, se incorporó el F1-score, que equilibra adecuadamente la precisión y el recall. Esta métrica resulta especialmente adecuada para problemas desbalanceados, tal como lo reportan (Powers, 2011) y (He & Garcia, 2009) en el contexto del aprendizaje con clases desproporcionadas. Su uso permitió equilibrar la capacidad del modelo para detectar desertores con la confiabilidad de las alertas emitidas.

Asimismo, se utilizaron métricas basadas en curvas. El área bajo la curva ROC (ROC-AUC) se utilizó para evaluar la capacidad discriminativa del modelo en distintos umbrales de decisión. Esta métrica es un estándar ampliamente conocido y avalado, como lo describe (Fawcett, 2006) en su análisis de las curvas ROC para clasificación binaria. Sin embargo, debido al desequilibrio del conjunto de datos, se complementó con el área bajo la curva de precisión-recall (PR-AUC), que proporciona una evaluación más robusta cuando la clase positiva es minoritaria. Como señalaron (Davis & Goadrich, 2006) y (Saito & Rehmsmeier, 2015), la curva PR-AUC es más adecuada que la ROC-AUC en escenarios en los que el interés se centra en el desempeño en la clase minoritaria.

De tal forma, las métricas indicadas permitieron una evaluación integral del modelo, abordando no solo su desempeño global, sino también su capacidad para distinguir casos relevantes, gestionar el desequilibrio de clases y reducir errores críticos. Su aplicación razonada aseguró una comparación justa entre los algoritmos y proporcionó un sólido sustento metodológico para la selección del modelo más adecuado para la predicción de la deserción universitaria.

4.5. Hiperparámetros de los modelos propuestos

Se realizó un análisis exhaustivo de los hiperparámetros de cada algoritmo para garantizar modelos adecuados, comparables y ajustados al problema de la deserción universitaria. La selección y el uso de los hiperparámetros no se limitaron a las configuraciones por defecto; por el contrario, se adoptó un enfoque experimental en el que cada hiperparámetro se modificó para evaluar su impacto en las métricas de desempeño. Este procedimiento permitió identificar configuraciones que evitaran el sobreajuste y mejoraran la capacidad de generalización en escenarios con clases desbalanceadas. En consecuencia, la tabla 7 resume los hiperparámetros críticos por algoritmo, los rangos explorados y la justificación técnica de su ajuste, ofreciendo un panorama claro y fundamentado del proceso de optimización.

Tabla 7

Hiperparámetros para algoritmos de ML

Algoritmo	Parámetro	Rango de exploración
HGB	learning_rate	0.01–0.3
	max_depth	3–15
	max_iter	100–500
	min_samples_leaf	10–50
ET	n_estimators	200–800
	max_depth	5–30
	min_samples_leaf	1–5
	max_features	sqrt, log2
DT	max_depth	3–30
	min_samples_split	2–20
	criterion	gini, entropy
GB	learning_rate	0.01–0.3
	n_estimators	100–400
	max_depth	3–7
	subsample	0.6–1.0

RF	n_estimators	100–600
	max_depth	5–30
	max_features	sqrt, log2
	min_samples_split	2–20
LR	C	0.01–10
	penalty	l1, l2
	solver	liblinear, saga
	class_weight	none, balanced
SVM	C	0.1–100
	kernel	linear, rbf, poly
	gamma	scale, auto
	class_weight	none, balanced
MLP	hidden_layer_sizes	1–3 capas; 32–256 neuronas por capa
	activation	relu, tanh
	solver	adam, sgd
	learning_rate_init	0.0005–0.01
	max_iter	200–500
NB	alpha	0.1–1.0

CAPÍTULO V. RESULTADOS Y DISCUSIÓN

5.1. Resultados de construcción de un conjunto de datos

5.1.1. Selección y adquisición de los datos

La selección y adquisición de los datos constituyeron la etapa inicial del proceso metodológico para definir y obtener el conjunto de datos necesario para el desarrollo del modelo de explicabilidad de DEU. En esta fase se recopilaron variables provenientes del sistema de gestión académica (SGA) de la UTM como se observa en la tabla 8, seleccionadas con base en criterios conceptuales y empíricos sustentados en la literatura especializada sobre deserción en educación superior, como indica el estudio de (Quimiz-Moreira et al., 2025), las cuales abarcaron dimensiones demográficas, socioeconómicas, institucionales, personales y académicas, lo que proporcionó un marco descriptivo amplio y multidimensional para el análisis del fenómeno de la DEU).

Tabla 8
Categoría de factores que influyen en la DEU.

Categoría	Factores	Referencias
Demográficas	• Género	(Al-Jallad et al., 2019; Asto-Lazaro & Cieza-
	• Edad	Mostacero, 2024; Behr et al., 2021; Berens
	• Estado civil	et al., 2019; Cai & Fleischhacker, 2024;
	• Sexo,	Coussement et al., 2020; Goran et al., 2024;
	• Lugar de residencia	Gutiérrez et al., 2024; Hassan et al., 2024;
	• Lugar de origen	Hutagaol & Suharjito, 2019; Isleib et al., 2019;
	• Nacionalidad	Martins et al., 2023; Mutrofin et al., 2019;
	• Nivel educativo de padres	Oqaidi et al., 2022b; Phan et al., 2023; Realinho et al., 2022; Sultana et al., 2017; Thammasiri
	• Tipo de colegio	et al., 2014; Vaarma & Li, 2024; Velasco et al.,
	• Etnia	2023; Won et al., 2023)

Socioeconómicas	• Becas	(Arthana et al., 2024; Coussement et al., 2020;
	• Trabajos	Díaz & Moreira, 2024; Guerra et al., 2020;
	• Ingresos familiares	Gutiérrez et al., 2024; Hinojosa et al., 2022;
	• Nivel de ingresos	Isleib et al., 2019; Kuz & Morales, 2023;
	• Puntuación de admisión	Martins et al., 2023; Mutrofin et al., 2019; Phan et al., 2023; Realinho et al., 2022; Sultana et al., 2017; Thammasiri et al., 2014; Velasco et al., 2023; Villar & de Andrade, 2024; Zapata-Medina et al., 2024)
	• Discapacidad	
	• Nivel educativo	
	• Acceso a internet	
	• Ayuda financiera	
Institucionales	• Infraestructura	(Alban & Mauricio, 2019b; Bedregal-Alpaca,
	• Servicios educativos	Cornejo-Aparicio, et al., 2020; R. Chen, 2012;
	• Equipo adecuado	Díaz et al., 2022; Kuz & Morales, 2023;
	• Tamaño de la institución	Marquez-Vera et al., 2013; Qvortrup & Lykkegaard, 2024; Villarreal-Torres et al., 2023; Won et al., 2023)
	• Área geográfica	
	• Compromiso del docente con el estudiante	
	• Clasificación de la carrera o institución	
	• Clase grupal	
	• Escala de evaluación del clima escolar	
Personales	• Año de ingreso	(Bedregal-Alpaca, Tupacyupanqui-Jaén, et al., 2020; Berens et al., 2019; Cai & Fleischhacker, 2024; Díaz et al., 2022; Guerra et al., 2020;
	• Actividades extracurriculares	
	• Compromiso	Hutagaol & Suharjito, 2019; Kim et al., 2023;
	• Participación en clase	Silva et al., 2020; Villarreal-Torres et al., 2023; Zanellati et al., 2024; Zapata-Medina et al., 2024)

	• Número de actividades voluntarias • Tiempo para estudiar • Adaptación y convivencia • Líder
Académicos	• Calificaciones (Al-Jallad et al., 2019; Arthana et al., 2024; • Promedio general Berens et al., 2019; Bouihi et al., 2024; Cai & • Nota de secundaria Fleischhacker, 2024; Coussement et al., 2020; • Asignaturas cursadas Díaz et al., 2022; Diaz & Moreira, 2024; Goran • Créditos cursados et al., 2024; Guerra et al., 2020; Gutiérrez et al., • Asistencia 2024; Hassan et al., 2024; Hinojosa et al., 2022; • Tipo de admisión Hutagaol & Suharjito, 2019; Iam-On & • Tipo de escuela Boongoen, 2017; Kim et al., 2023; Martins • Escuela et al., 2023; Mutrofin et al., 2019; Oqaidi et al., • Año académico 2022b; Quispe et al., 2024b; Realinho et al., 2022; Silva et al., 2020; Thammasiri et al., 2014; Vaarma & Li, 2024; Velasco et al., 2023; - Villar & de Andrade, 2024; Won et al., 2023; - Zapata-Medina et al., 2024)

Para garantizar la independencia analítica de esta etapa, la variable objetivo no se incorporó durante el proceso de selección y adquisición, dado que su uso se restringe a las fases posteriores de entrenamiento, validación y prueba de los modelos de ML. Esta exclusión temprana permitió evitar posibles sesgos de información y garantizó que la caracterización inicial del dataset se realizara de manera objetiva, centrada en la calidad, la coherencia y la pertinencia de las variables explicativas.

En concordancia con los criterios previamente definidos, se obtuvieron 57 variables de naturaleza categórica, numérica y temporal, que constituyen la base informativa del estudio. Durante el proceso de adquisición, se identificaron, adicionalmente, variables cuyo contenido se limitaba exclusivamente a funciones administrativas o de registro, sin aportar valor explicativo al fenómeno de la deserción universitaria. Entre estas se incluyeron identificadores personales y datos de contacto, tales como idpersonal, cédula, nombres, apellidos, código de matrícula, teléfono personal y correos electrónicos personales e institucionales, los cuales fueron excluidos del análisis con el fin de preservar la confidencialidad de los estudiantes y reducir posibles fuentes de sesgo o ruido informacional.

Como resultado de esta selección conceptual, el dataset consolidado quedó conformado por 48 variables y 17 183 observaciones, lo que constituye una base de datos robusta, representativa y estadísticamente adecuada para la realización de análisis descriptivos y exploratorios, así como para la posterior implementación de modelos de ML orientados tanto a la predicción como a la explicación del fenómeno de la deserción universitaria.

5.1.2. Comprensión y Análisis Exploratorio de los Datos (EDA)

El conjunto de datos estuvo conformado por 17183 registros, correspondientes a estudiantes matriculados en la institución durante los períodos académicos considerados. A partir del análisis exploratorio de las variables categóricas, se identificaron las categorías dominantes en cada una, lo que evidencia distintos grados de concentración y desequilibrio. En la variable género, la categoría con mayor frecuencia fue masculino, que concentró el 52,55 % de las observaciones. La modalidad de carrera presencial se constituyó claramente dominante, con el 85,05 % de los registros. Respecto a la nacionalidad, la categoría ecuatoriana concentró el 99,61 %, mientras que en pais_residencia la totalidad de los registros correspondió a Ecuador (100%). De forma similar, en pais_origen la categoría Ecuador concentró el 98,93 % de las

observaciones. En el ámbito territorial, aunque provincia_residencia presentó 26 categorías distintas y provincia_origen una cardinalidad elevada, la mayor concentración se observó en la provincia de Manabí. En canton_residencia y canton_origen, con más de 100 y aproximadamente 159 categorías, respectivamente, la mayor frecuencia se observó en el cantón Portoviejo. En tipo_zona_procedencia, la categoría cabecera cantonal concentró el 56,54 % de los registros. En estado_civil, la categoría soltera/a fue claramente la más frecuente, con el 93,27 % de las observaciones. En términos de etnia, el mestizo/a presentó la mayor frecuencia relativa. En sostenibilidad_colegio, los estudiantes provenientes de instituciones fiscales concentraron el 76,33 %. En la variable económica salario_ingreso_total, la categoría de ingreso más bajo fue la dominante, con el 62,59 %. En internet_dispone_casa, la respuesta sí concentró el 87,43 %, y en internet_dispone_celular, la categoría sí fue mayoritaria. En tipo_vivienda, la categoría casa/apartamento concentró el 66,59 %, mientras que en vivienda_localizacion la mayor frecuencia correspondió a barrio, con el 72,51 %. En recibe_ayuda_familiar, la categoría no fue dominante, con un 79,50 % de los registros. En jefe_ocupacion, a pesar de identificarse 90 categorías distintas, la categoría con mayor frecuencia individual fue la de desempleado, aunque con una concentración baja (8,88 %), lo que refleja una alta dispersión. En tipo_transporte, el transporte público se constituyó en la categoría dominante, concentrando el 86,28 %. En el ámbito educativo familiar, en madre_instruccion, la categoría bachiller completo presentó la mayor frecuencia (34,31 %), mientras que en padre_instruccion, la categoría dominante también fue bachiller completo (31,07 %). En gasto_alimentacion, el rango inferior a 9 USD concentró el 31,02 %, y en gasto_transporte, el rango más bajo concentró el 25,53 %. En cuanto a la disponibilidad de dispositivos, en dispone_tablet la categoría NO fue dominante (92,23 %), mientras que en dispone_computador el SI concentró el 74,04 %.

De igual forma, como parte de la evaluación de la calidad de los datos, se identificaron variables con porcentajes críticos de valores faltantes. La tabla 9 presenta el detalle de los valores nulos detectados en las variables institucionales, familiares y económicas. Este diagnóstico preliminar fue fundamental para anticipar la magnitud del problema de completitud y planificar las estrategias que se implementarían posteriormente en la fase de preparación de los datos.

Tabla 9

Valores nulos

N.º	Variable	Valores nulos	Total de registros	% Missing	% Completitud
1	orientacion_sexual	11 716	17 183	68,18 %	31,82 %
2	preferencia_religiosa	2 601	17 183	15,14 %	84,86 %
3	padre_instruccion	1 696	17 183	9,87 %	90,13 %
4	salario_ingreso_total	1 251	17 183	7,28 %	92,72 %
5	jefe_ocupacion	716	17 183	4,17 %	95,83 %
6	tipo_zona_procedencia	268	17 183	1,56 %	98,44 %
7	canton_residencia	179	17 183	1,04 %	98,96 %
8	dispone_tablet	113	17 183	0,66 %	99,34 %
9	gasto_alimentacion	113	17 183	0,66 %	99,34 %
10	internet_dispone_celular	113	17 183	0,66 %	99,34 %
11	gasto_transporte	113	17 183	0,66 %	99,34 %
12	dispone_computador	113	17 183	0,66 %	99,34 %
13	provincia_residencia	100	17 183	0,58 %	99,42 %
14	madre_instruccion	68	17 183	0,40 %	99,60 %
15	internet_dispone_casa	68	17 183	0,40 %	99,60 %

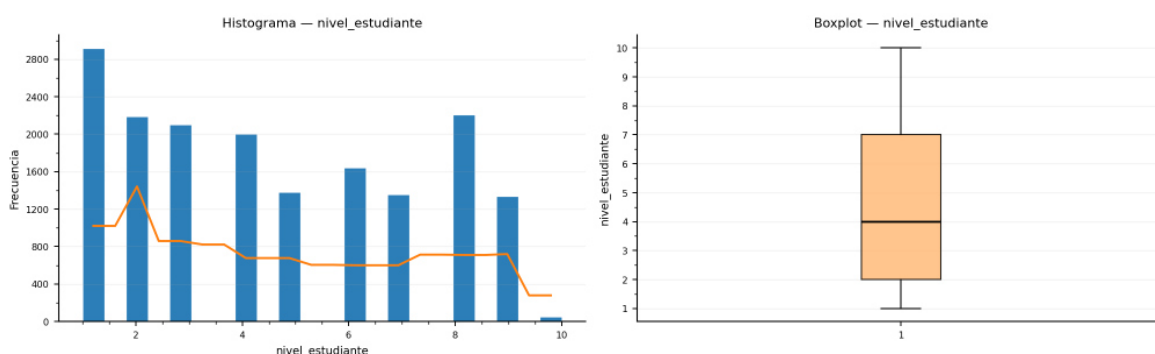
Además, el análisis descriptivo de las variables numéricas permitió identificar patrones consistentes de asimetría, de concentración central y de presencia de valores atípicos, tanto en variables sociodemográficas y de uso del tiempo como en las relacionadas con la trayectoria académica del estudiante. En conjunto, estas características estructurales del dataset resultan fundamentales para comprender la complejidad del fenómeno de la deserción universitaria y orientar adecuadamente las etapas posteriores de preprocesamiento, modelado y explicabilidad. Tal como se observa en la Figura 2, variables como `internet_horas_uso`, `red_social_horas_uso`, `hijos_numero`, `transporte_horas_permanencia`, `integrantes_familiar`, `transporte_horas_traslado` presentan distribuciones marcadamente asimétricas, con colas prolongadas hacia valores elevados. En particular, `internet_horas_uso` y `red_social_horas_uso` muestran una concentración principal entre 3 y 8 horas diarias, coexistiendo con registros extremos que superan las 20 horas, lo que evidencia patrones de uso intensivo en subgrupos específicos. De forma similar, `hijos_numero` se encuentra fuertemente sesgada hacia el valor cero, lo que confirma que la mayoría de los estudiantes no tiene hijos, aunque la presencia de casos aislados con valores elevados introduce escenarios familiares de alta carga. La variable `integrante_familiar` muestra un comportamiento similar, con una concentración en hogares pequeños y valores atípicos que reflejan estructuras familiares extensas.

En relación con las variables de movilidad, `transporte_horas_traslado` presenta una elevada concentración de valores bajos, acompañada de valores superiores a 20 horas, mientras que `transporte_horas_permanencia` muestra una distribución más centrada, aunque con casos extremos que superan las 15 horas. Estos patrones sugieren que, si bien la mayoría de los estudiantes enfrenta condiciones de desplazamiento moderadas, existen situaciones puntuales de alta exigencia logística que podrían afectar negativamente la continuidad académica.

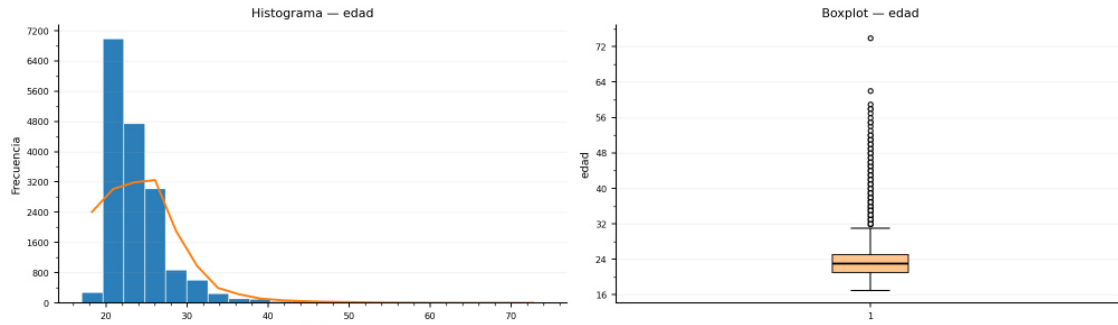
Desde la perspectiva estrictamente académica, la variable `materias_aprobados` evidencia una distribución concentrada en valores intermedios, con una mediana cercana al rango esperado de aprobación por período, lo que indica trayectorias académicas relativamente estables para una proporción significativa del estudiantado. No obstante, la presencia de valores bajos y altos, así como de algunos valores atípicos, revela la coexistencia de perfiles con rezago académico y de otros con avances acelerados. Por su parte, `promedio_aprobada` presenta una distribución más compacta y relativamente simétrica, con una fuerte concentración en rangos altos de calificación, lo que sugiere un desempeño académico general favorable; sin embargo, se identifican valores atípicos inferiores que podrían asociarse a situaciones de riesgo académico acumulado. Finalmente, la variable `semestre_tomados` muestra una distribución asimétrica a la derecha, con mayor densidad en los primeros semestres y una cola extendida hacia valores superiores, lo que refleja diferencias en la duración efectiva de la trayectoria universitaria y posibles episodios de repetición o interrupción.

Figura 2

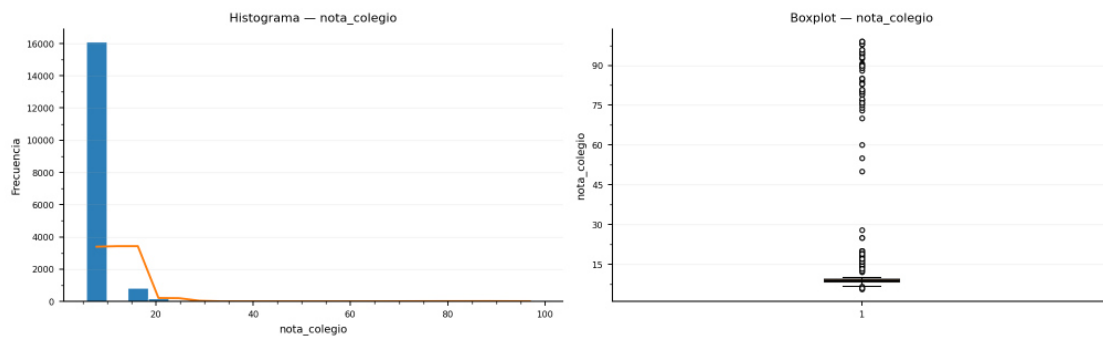
Descripción de variables numéricas



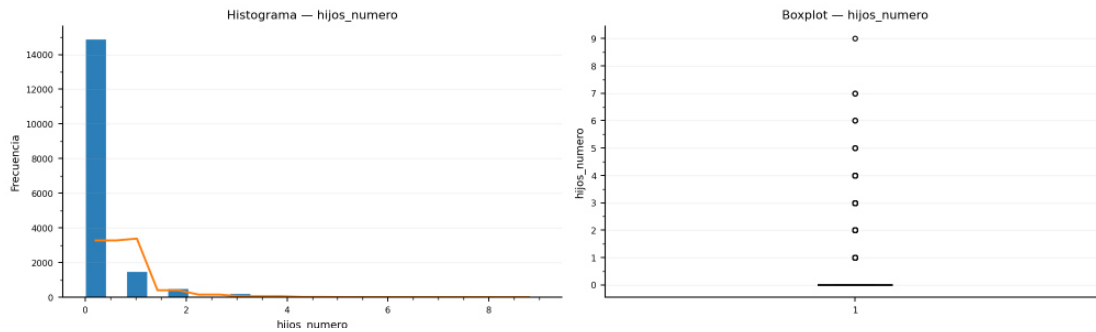
Distribución numérica — edad



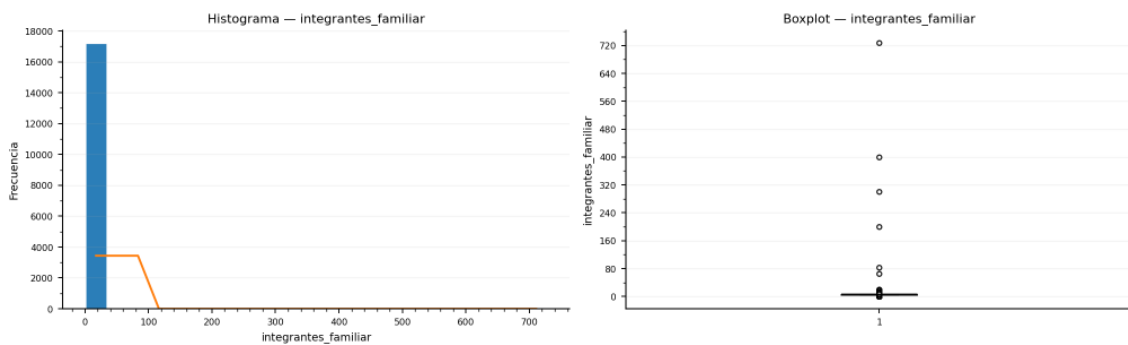
Distribución numérica — nota_colegio



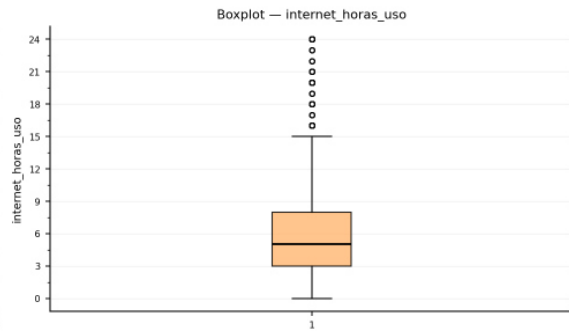
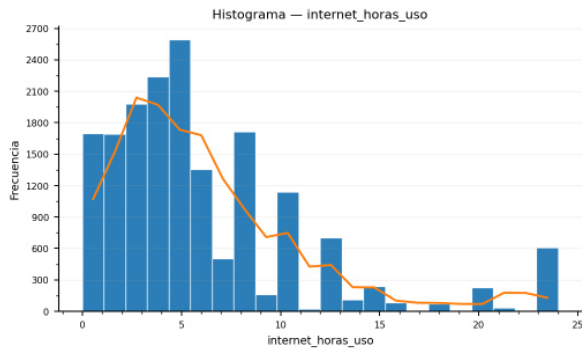
Distribución numérica — hijos_numero



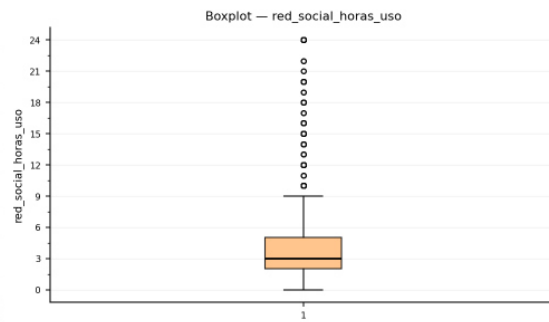
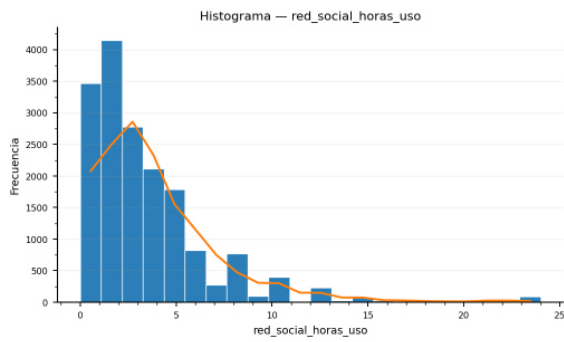
Distribución numérica — integrantes_familiar



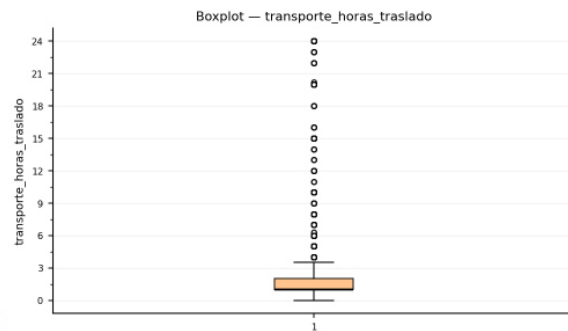
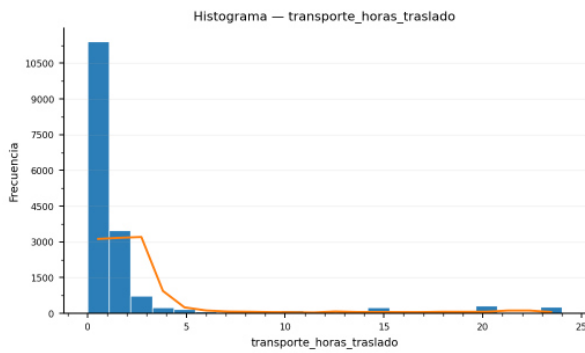
Distribución numérica — internet_horas_uso



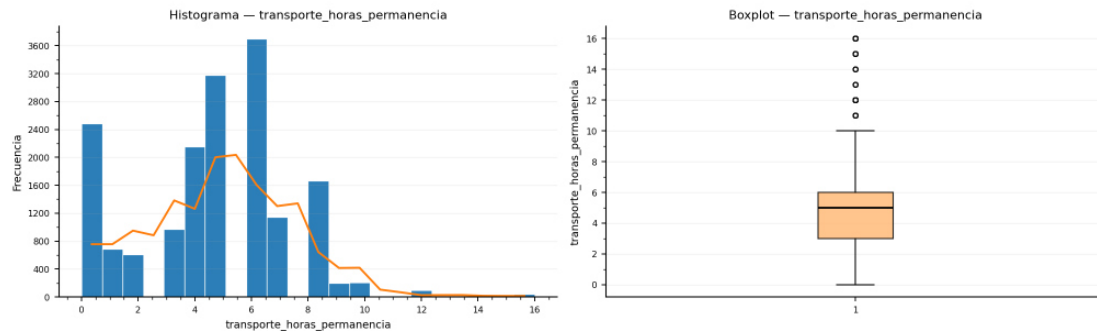
Distribución numérica — red_social_horas_uso



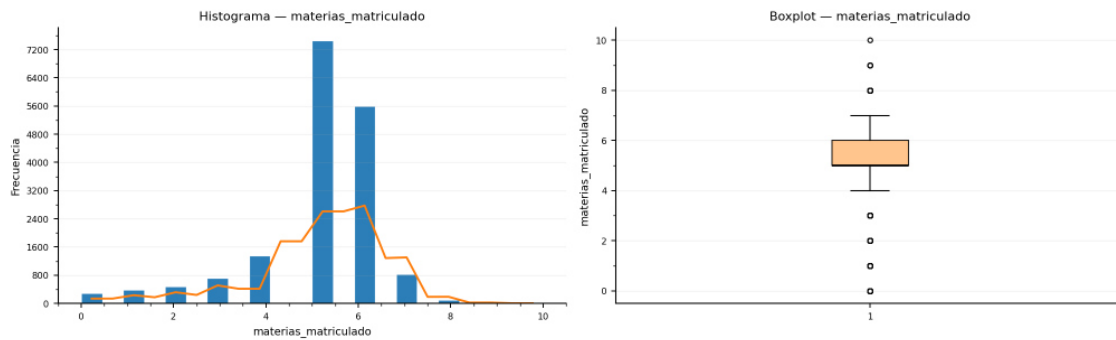
Distribución numérica — transporte_horas_traslado



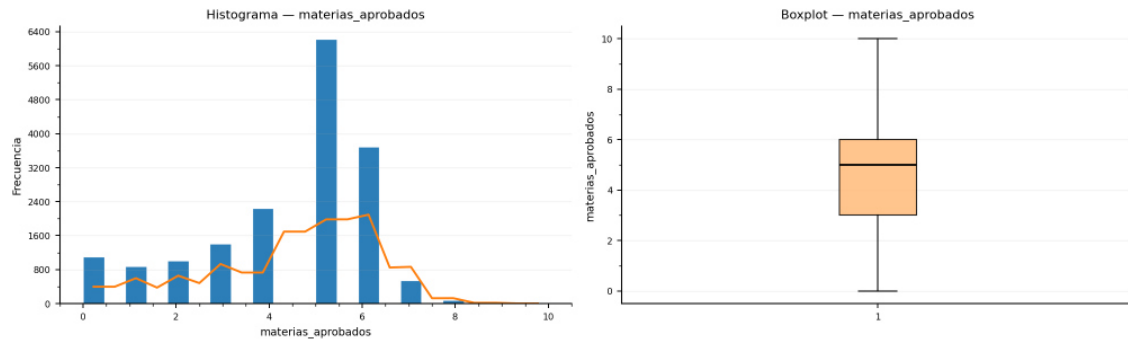
Distribución numérica — transporte_horas_permanencia



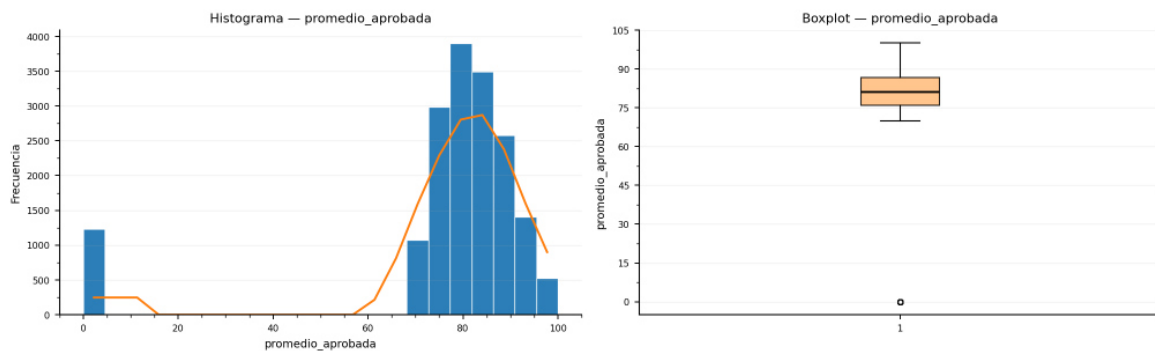
Distribución numérica — materias_matriculado

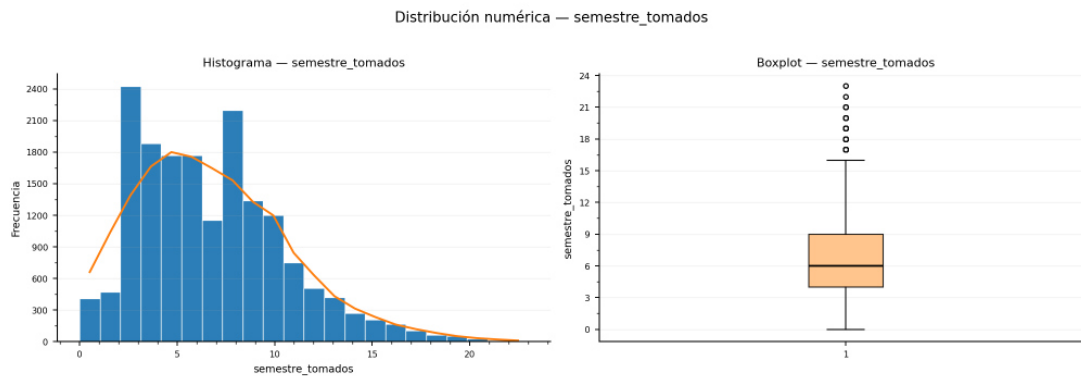


Distribución numérica — materias_aprobados



Distribución numérica — promedio_aprobada





La información de la tabla 10 permite establecer un diagnóstico integral del estado inicial del conjunto de variables y diferenciar entre aquellas que presentan anomalías, altos niveles de incompletitud o patrones atípicos y aquellas que muestran una calidad aceptable para su uso en etapas analíticas posteriores. Este diagnóstico es un elemento esencial en la fase de comprensión de los datos, ya que permite detectar de manera temprana posibles fuentes de sesgo, ruido o inestabilidad en el modelado predictivo.

Tabla 10

Selección de variables (conjunto de datos original)

Num	Variable	Data Anómala	Exceso de nulos	Valido
1	carrera			X
2	nivel_estudiante			X
3	paralelo	X		
4	genero			X
5	modalidad_carrera			X
6	nacionalidad	X		
7	pais_residencia	X		
8	pais_origen	X		
9	provincia_residencia	X		
10	provincia_origen	X		
11	canton_residencia	X		
12	canton_origen	X		

13	tipo_zona_procedencia		X
14	fecha_nacimiento	X	
15	edad		X
16	tipo_sangre	X	
17	estado_civil		X
18	etnia		X
19	nota_colegio		X
20	sostenibilidad_colegio		X
21	hijos_numero		X
22	salario_ingreso_total		X
23	integrantes_familiar		X
24	internet_dispone_casa		X
25	internet_dispone_celular		X
26	internet_horas_uso		X
27	red_social_horas_uso		X
28	tipo_vivienda		X
29	vivienda_localizacion		X
30	recibe_ayuda_familiar		X
31	jefe_ocupacion		X
32	transporte_horas_traslado		X
33	transporte_horas_permanencia		X
34	tipo_transporte		X
35	madre_instruccion		X
36	padre_instruccion		X
37	gasto_alimentacion		X
38	gasto_transporte		X
39	materias_matriculado		X
40	materias_aprobados		X
41	promedio_aprobada		X
42	semestre_tomados		X
43	semestre_anulados		X

44	dispone_tablet		X
45	dispone_computador		X
46	orientacion_sexual	X	
47	preferencia_religiosa		X
48	periodo	X	

5.1.3. Preprocesamiento

En esta fase se aplicó exclusivamente a las variables previamente identificadas como válidas en la etapa de comprensión de los datos. El dataset original estuvo conformado por 17.183 registros y 36 variables de los ámbitos sociodemográfico, económico, institucional, personal y académico. Sin embargo, el análisis de calidad desarrollado en la fase anterior permitió identificar inconsistencias, como porcentajes extremos de valores faltantes, anomalías estructurales o falta de relevancia predictiva. Una vez definido el conjunto de variables válidas, se procedió al preprocesamiento, como se observa en la figura 3, siguiendo una secuencia programada en Python orientada a garantizar la coherencia estadística del dataset. Esta fase abarcó:

Figura 3

Proceso de preprocesamiento



Detección y tratamiento de valores atípicos

En la fase de detección y tratamiento de valores atípicos, se aplicó el criterio del rango intercuartílico (IQR) con el objetivo de identificar observaciones extremas en las variables numéricas que, aun siendo plausibles desde el punto de vista empírico, podían afectar la estabilidad, la convergencia y la interpretabilidad de los modelos de ML. Como se detalla en la tabla 11, se identificaron proporciones relevantes de valores atípicos en variables asociadas a las condiciones familiares, a la trayectoria académica y al uso del tiempo.

Tabla 11

Detección y tratamiento de outliers

N.º	Variable	N.º de outliers	% de outliers	Límite inferior	Límite superior	Q1	Q3	IQR
1	integrantes_familiar	3 050	17,75 %	2,50	6,50	4,00	5,00	1,00
2	materias_matriculado	1 965	11,44 %	3,50	7,50	5,00	6,00	1,00
3	transporte_horas_traslado	1 566	9,11 %	-0,50	3,50	1,00	2,00	1,00
4	promedio_aprobada	1 234	7,18 %	60,10	102,50	70,60	86,60	16,00
5	nota_colegio	1 054	6,13 %	6,68	10,88	8,25	9,30	1,05
6	internet_horas_uso	1 046	6,09 %	-4,50	15,50	3,00	8,00	5,00
7	red_social_horas_uso	926	5,39 %	-2,50	9,50	2,00	5,00	3,00
8	edad	783	4,56 %	15,00	31,00	21,00	25,00	4,00
9	semestre_tomados	258	1,50 %	-3,50	16,50	4,00	9,00	5,00
10	transporte_horas_permanencia	192	1,12 %	-1,50	10,50	3,00	6,00	3,00
11	nivel_estudiante	0	0,00 %	-5,50	14,50	2,00	7,00	5,00
12	materias_aprobados	0	0,00 %	-1,50	10,50	3,00	6,00	3,00

En particular, la variable integrante_familiar presentó el mayor porcentaje de valores atípicos (17,75 %), seguida de materias_matriculado (11,44 %) y transporte_horas_traslado (9,11 %), lo que evidencia una alta dispersión estructural en estas variables. Asimismo, variables de desempeño académico como promedio_aprobada (7,18 %) y nota_colegio (6,13 %) presentaron observaciones extremas, asociadas a valores excepcionalmente altos o bajos. De forma similar,

internet_horas_uso y red_social_horas_uso presentaron proporciones moderadas de outliers, entre 5,38 % y 6 %, respectivamente, coherentes con patrones intensivos de uso en subgrupos específicos. En contraste, variables como nivel_estudiante y materias_aprobados no presentaron valores atípicos según el criterio del IQR, lo que sugiere una distribución más estable y acotada.

Dado que estos valores extremos reflejan situaciones reales en la población estudiantil y no errores de medición, se optó por no eliminar observaciones, a fin de preservar la variabilidad informativa del conjunto de datos. En su lugar, se aplicaron técnicas estadísticas robustas, específicamente la winsorización para limitar el efecto de los valores más extremos dentro de los límites definidos por el IQR, y el escalado robusto, con el propósito de reducir la influencia desproporcionada de observaciones alejadas de la tendencia central. La aplicación conjunta de estas transformaciones permitió atenuar el impacto de los outliers sin distorsionar la estructura subyacente de los datos, lo que mejoró la estabilidad numérica, la capacidad de generalización y la consistencia de los modelos.

Imputación de valores faltantes

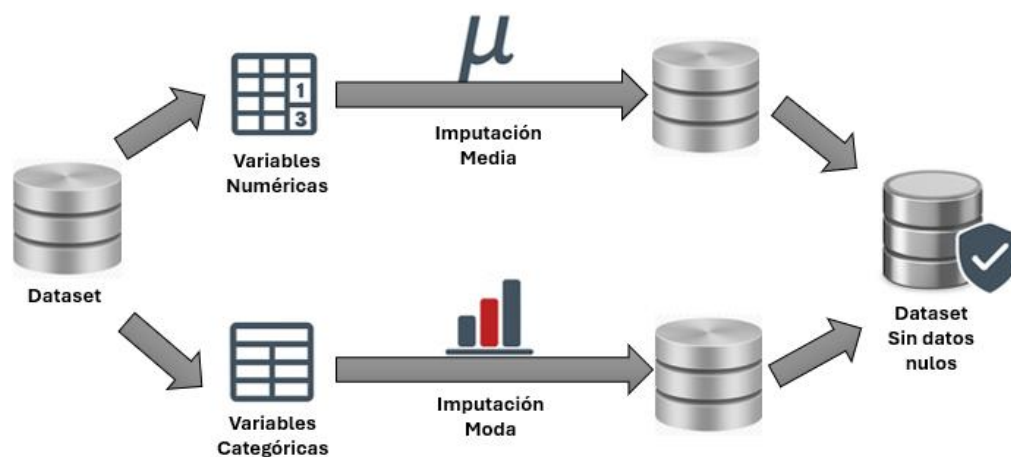
Las variables incluidas en el conjunto de datos no presentaron niveles críticos de ausencia de información; no obstante, se identificaron porcentajes moderados de valores faltantes en determinados atributos, principalmente de naturaleza categórica, que requerían un tratamiento sistemático previo a la fase de modelado. Con el fin de garantizar la consistencia estadística y semántica del dataset, se aplicó una estrategia de imputación diferenciada según el tipo de variable, siguiendo buenas prácticas ampliamente aceptadas en el preprocesamiento de datos para modelos de ML.

En el caso de las variables numéricas, se empleó la mediana como estadístico de imputación, debido a su robustez frente a valores atípicos y a su capacidad para preservar la forma general

de la distribución (Figura 4). Esta estrategia se aplicó a la variable edad, en la que se registró un único valor faltante, imputado con la mediana de 23 años, evitando así distorsiones en la tendencia central y en la dispersión previamente identificadas en el análisis exploratorio.

Figura 4

Imputación de datos



Para las variables categóricas, se utilizó la moda como criterio de imputación, ya que representa la categoría más frecuente y, semánticamente, la más consistente en cada atributo. Tal como se detalla en la tabla 12, este procedimiento se aplicó a variables como preferencia_religiosa, padre_instruccion, salario_ingreso_total, jefe_ocupacion, tipo_zona_procedencia, madre_instruccion, tipo_vivienda, tipo_transporte, así como a variables relacionadas con las condiciones de conectividad y el gasto. Cabe destacar que las mayores cantidades de imputación se registraron para preferencia_religiosa (2 601 registros) y padre_instruccion (1 696 registros), lo cual es coherente con el carácter sensible de estas variables y con los patrones de no respuesta observados en la etapa exploratoria.

Adicionalmente, cuando la ausencia de información podía reflejar condiciones contextuales específicas o patrones no completamente aleatorios, se optó por mantener la coherencia informativa mediante categorías representativas, evitando tanto la eliminación de

observaciones como la introducción de sesgos artificiales. Como resultado de este proceso, el conjunto de datos alcanzó un nivel de completitud del 100 %, lo que garantiza su idoneidad para las fases posteriores de entrenamiento, validación y explicabilidad de los modelos de ML.

Tabla 12

Proceso de imputación de valores faltantes

N.º	Variable	Tipo	N.º imputados	Método	Valor imputado
1	preferencia_religiosa	Categórica	2 601	Moda	Católico
2	padre_instruccion	Categórica	1 696	Moda	Bachiller completo
3	salario_ingreso_total	Categórica	1 251	Moda	Menos de 499 dólares
4	jefe_ocupacion	Categórica	716	Moda	Desempleado
5	tipo_zona_procedencia	Categórica	268	Moda	Cabecera cantonal
6	dispone_computador	Categórica	113	Moda	Sí
7	gasto_alimentacion	Categórica	113	Moda	Menos de \$9
8	gasto_transporte	Categórica	113	Moda	Menos de \$9
9	internet_dispone_celular	Categórica	113	Moda	Sí
10	dispone_tablet	Categórica	113	Moda	No
11	madre_instruccion	Categórica	68	Moda	Bachiller completo
12	internet_dispone_casa	Categórica	68	Moda	Sí
13	vivienda_localizacion	Categórica	68	Moda	Barrio
14	tipo_vivienda	Categórica	68	Moda	Apartamento / casa
15	tipo_transporte	Categórica	62	Moda	Tipo de Transporte
16	edad	Numérica	1	Mediana	23,0
17	etnia	Categórica	1	Moda	Mestizo
18	estado_civil	Categórica	1	Moda	Soltero/a

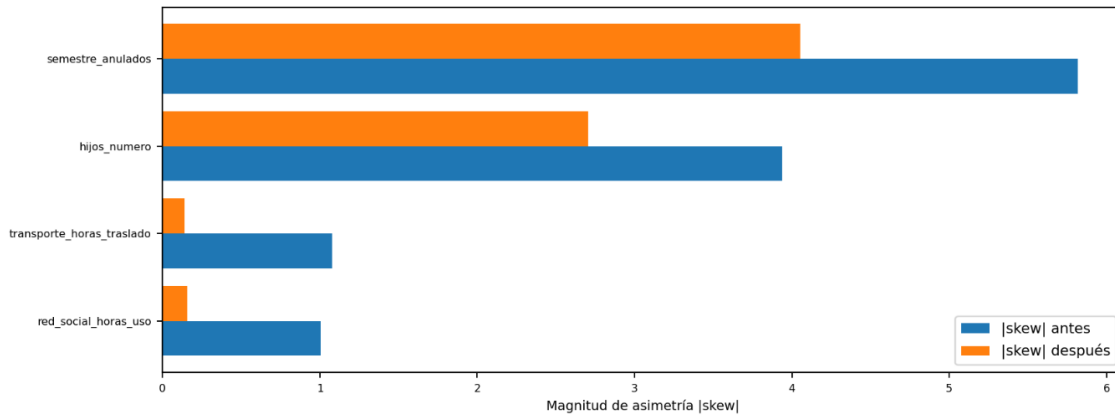
Transformaciones estadísticas

Con el propósito de estabilizar las distribuciones numéricas con asimetrías positivas pronunciadas y varianzas elevadas, se aplicaron transformaciones estadísticas específicas en la fase de preprocesamiento. En particular, se identificaron variables con altos niveles de asimetría absoluta, entre las que destacan `semestre_anulados`, `hijos_numero`, `transporte_horas_traslado` y `red_social_horas_uso`, cuyos valores iniciales de asimetría superaban los umbrales recomendados para el modelado predictivo.

Se empleó la transformación logarítmica tipo $\log_{1p}$, que permite reducir la influencia de colas largas y manejar la presencia de ceros en la distribución. Como resultado, se observó una disminución sustantiva de la asimetría, evidenciada por la reducción del coeficiente $|\text{skew}|$ en todas las variables involucradas, donde `semestre_anulados` redujo su asimetría absoluta de 5.82 a 4.05, mientras que `hijos_numero` pasó de 3.94 a 2.71. Adicionalmente, estas transformaciones se complementaron con técnicas previamente implementadas, como la winsorización y el escalado robusto, orientadas a mitigar la influencia de los valores extremos sin eliminar observaciones potencialmente informativas. La Figura 5 muestra la magnitud de la asimetría antes y después de la transformación, lo que permite verificar visualmente la mejora en la estabilidad distributiva de las variables numéricas más críticas.

Figura 5

Reducción de asimetría



Codificación de variables categóricas

El conjunto de variables válidas incluyó atributos categóricos con niveles de cardinalidad heterogéneos, lo que hizo necesario aplicar estrategias de codificación diferenciadas para garantizar representaciones numéricas adecuadas sin distorsionar la estructura semántica de los datos ni introducir expansiones dimensionales innecesarias. Como se resume en la tabla 13, la selección del método de codificación se realizó en función del número de categorías y del impacto potencial de cada variable en la complejidad del modelo.

Tabla 13

Reducción de asimetría

Variable	Cardinalidad	Método	Nuevas columnas
jefe_ocupacion	89	Frequency Encoding	1
gasto_transporte	15	One-Hot Encoding	15
gasto_alimentacion	15	One-Hot Encoding	15
salario_ingreso_total	9	One-Hot Encoding	9
madre_instruccion	9	One-Hot Encoding	9
padre_instruccion	9	One-Hot Encoding	9
etnia	8	One-Hot Encoding	8
tipo_transporte	7	One-Hot Encoding	7
estado_civil	5	One-Hot Encoding	5
vivienda_localizacion	5	One-Hot Encoding	5

tipo_vivienda	5	One-Hot Encoding	5
escuela	4	One-Hot Encoding	4
tipo_zona_procedencia	4	One-Hot Encoding	4
sostenibilidad_colegio	4	One-Hot Encoding	4
preferencia_religiosa	3	One-Hot Encoding	3
genero	2	One-Hot Encoding	2
modalidad_carrera	2	One-Hot Encoding	2
internet_dispone_casa	2	One-Hot Encoding	2
internet_dispone_celular	2	One-Hot Encoding	2
recibe_ayuda_familiar	2	One-Hot Encoding	2
dispone_tablet	2	One-Hot Encoding	2
dispone_computador	2	One-Hot Encoding	2
desertor	2	One-Hot Encoding	2

Para la mayoría de las variables categóricas de baja y media cardinalidad, se aplicó One-Hot Encoding, lo que asegura una codificación libre de jerarquías artificiales entre categorías. Este enfoque se aplicó a variables sociodemográficas, económicas, académicas y de condiciones de estudio, lo que permitió preservar la interpretabilidad de cada categoría y facilitó su análisis posterior mediante modelos explicativos. En el caso de la variable jefe_ocupacion, que presentó alta cardinalidad, se optó por Frequency Encoding para evitar una expansión excesiva del espacio de características, lo que podría comprometer la estabilidad del entrenamiento y aumentar el riesgo de sobreajuste. Esta técnica permitió capturar la información distributiva de cada categoría mediante una única representación numérica, manteniendo su capacidad informativa a un costo computacional reducido. Finalmente, la variable objetivo "desertor" fue codificada mediante One-Hot Encoding, lo que garantiza su compatibilidad con los distintos algoritmos de ML empleados en las fases posteriores de entrenamiento, validación y explicabilidad.

Como resultado de la aplicación secuencial de las fases de detección y tratamiento de valores atípicos, imputación de valores faltantes, transformaciones estadísticas y codificación de variables categóricas, se obtuvo un conjunto de datos coherente, libre de valores nulos, con distribuciones estabilizadas y dimensionalidad controlada, preservando la variabilidad informativa esencial del fenómeno analizado. El dataset final, conformado por 17183 registros y 133 variables, cumple con los criterios de calidad, robustez y reproducibilidad requeridos para avanzar de manera sólida hacia la fase de selección de características y al modelado predictivo y explicativo de la deserción universitaria.

5.1.4. Selección de características

Una vez concluidos el preprocesamiento y la depuración del conjunto de datos, se procedió a la selección de características, cuyo propósito fue identificar únicamente las variables con capacidad explicativa y predictiva de la DEU. Este proceso permitió trabajar con un conjunto de variables previamente validadas y depuradas, lo que garantizó la coherencia metodológica de la fase de modelado.

Correlación entre variables

Se realizó el análisis de correlación entre variables, considerando su naturaleza, dado que el uso de una técnica inadecuada puede conducir a interpretaciones erróneas o estadísticamente inválidas. En el caso de las variables numéricas, se empleó el coeficiente de correlación de Pearson, ya que permite cuantificar la fuerza y la dirección de la asociación lineal entre variables continuas, bajo el supuesto de que las variables son métricas y comparables. Por el contrario, las variables categóricas no satisfacen dichos supuestos, al carecer de una escala numérica inherente y presentar distribuciones no continuas; en consecuencia, para este tipo de variables se utilizó el estadístico V de Cramér, que se fundamenta en tablas de contingencia y

mide la intensidad de la asociación sin asumir linealidad ni orden numérico. Esta diferenciación metodológica garantiza la validez estadística del análisis, evita sesgos en la estimación de las dependencias y asegura que las conclusiones derivadas reflejen fielmente la estructura real de los datos.

Análisis de correlación de Pearson entre variables numéricas

La figura 6 muestra la matriz de correlación de Pearson de las variables numéricas incluidas en el análisis de selección de características. Los resultados evidencian un bajo nivel de correlación lineal entre la mayoría de los pares de variables, con coeficientes $|r|$ inferiores a 0,30, lo que indica escasa redundancia informativa y sugiere que estas variables aportan información complementaria al proceso de modelado. Variables como *nota_colegio*, *integrantes_familiar*, *internet_horas_uso*, *red_social_horas_uso* y *transporte_horas_traslado* presentan asociaciones débiles tanto entre sí como con la variable objetivo, lo que refuerza su idoneidad para ser consideradas conjuntamente en modelos predictivos sin un riesgo significativo de multicolinealidad.

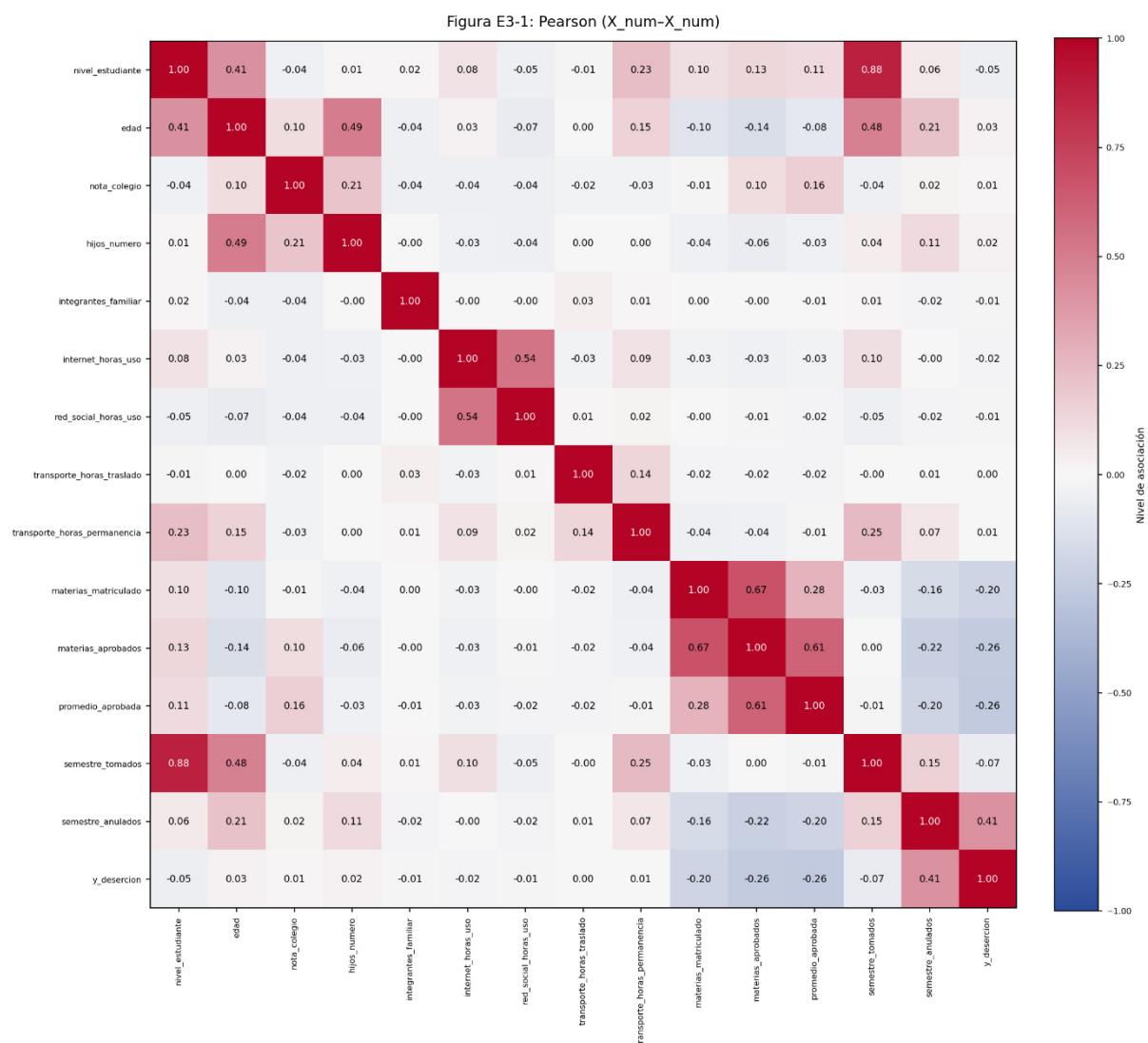
No obstante, se identifican correlaciones moderadas y fuertes en subconjuntos específicos de variables académicas y demográficas. Destaca la fuerte correlación positiva entre *nivel_estudiante* y *semestre_tomados* ($r = 0,88$), así como asociaciones moderadas entre *edad* y *hijos_numero* ($r = 0,49$) y entre *edad* y *semestre_tomados* ($r = 0,48$), lo cual resulta coherente desde una perspectiva formativa y sociodemográfica. Asimismo, se observa una correlación elevada entre *materias_matriculado* y *materias_aprobados* ($r = 0,67$), y entre *materias_aprobados* y *promedio_aprobada* ($r = 0,61$), lo que refleja la consistencia interna del desempeño académico de los estudiantes.

En relación con la variable objetivo, las asociaciones lineales son, en general, bajas a moderadas, destacándose correlaciones negativas con *materias_aprobados* ($r = -0,26$) y

promedio_aprobada ($r = -0,26$), y una correlación positiva moderada con semestre_anulados ($r = 0,41$). Estos patrones sugieren que el rendimiento académico y la trayectoria curricular son factores relevantes para explicar la DEU, aunque su relación no es estrictamente lineal.

Figura 6

Matriz de correlación de Pearson entre variables numéricas.



Análisis de correlación de Cramer's V entre variables categóricas

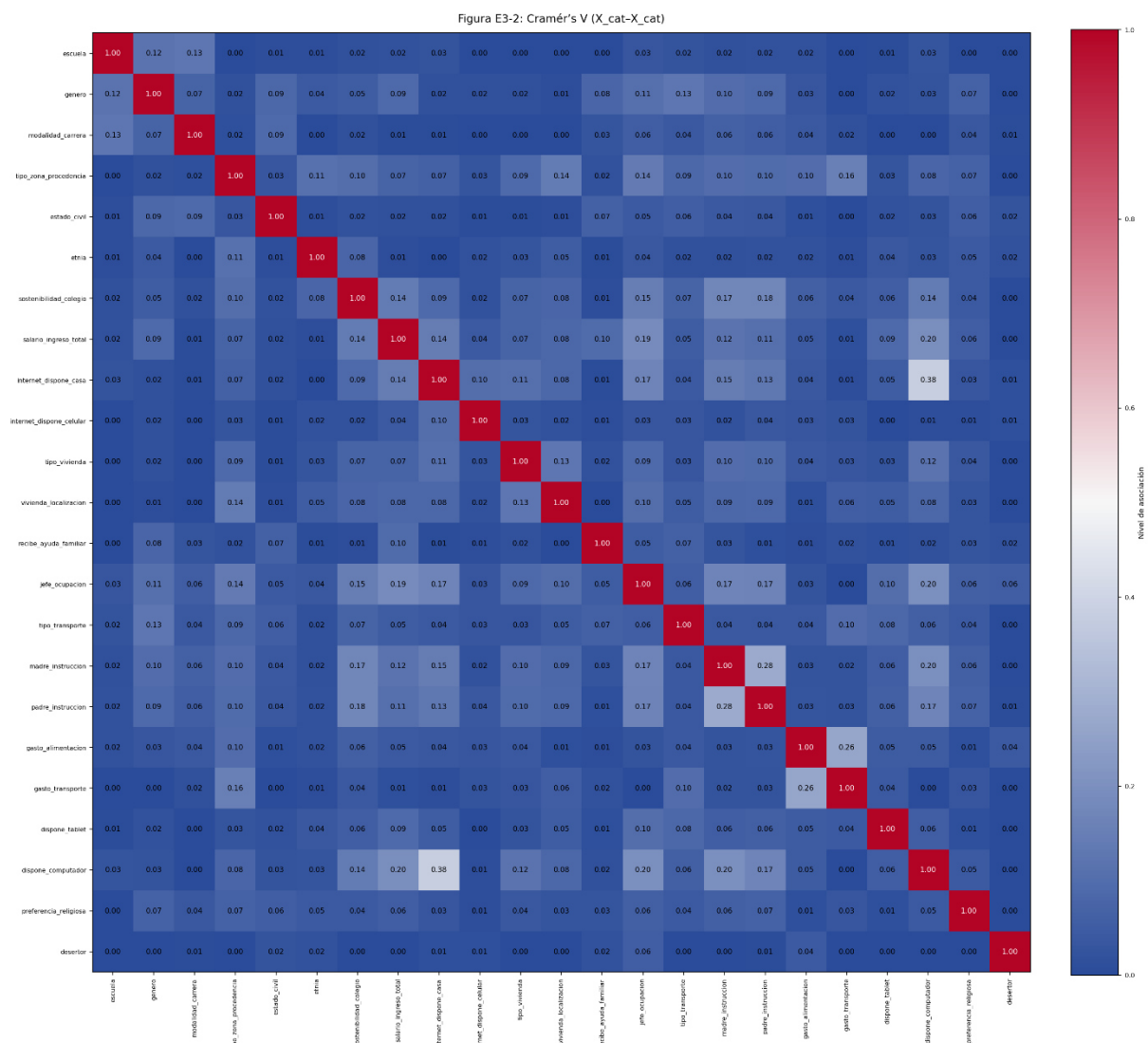
La figura 7 presenta la matriz de asociación de Cramér V aplicada a las variables categóricas consideradas en el proceso de selección de características. Los resultados evidencian que la mayoría de las asociaciones se sitúan en niveles bajos ($V < 0,30$), lo que indica una escasa dependencia entre variables categóricas y sugiere una adecuada diversidad informativa en el conjunto de datos.

No obstante, se identifican asociaciones moderadas en subconjuntos conceptualmente relacionados. Destaca la relación entre `internet_dispone_casa` y `dispone_computador` ($V = 0,38$), lo que refleja una coherencia lógica entre el acceso a la conectividad domiciliaria y la disponibilidad de recursos tecnológicos. Asimismo, se observa una asociación moderada entre `padre_instruccion` y `madre_instruccion` ($V = 0,28$), consistente con patrones de homogeneidad educativa parental, sin alcanzar niveles que impliquen redundancia crítica. De forma similar, se identifican asociaciones moderadas entre `gasto_alimentacion` y `gasto_transporte` ($V = 0,26$), lo que sugiere una relación parcial entre los patrones de gasto del hogar.

Otras variables, como género, `estado_civil`, `etnia`, `preferencia_religiosa`, `tipo_vivienda`, `vivienda_localizacion`, `recibe_ayuda_familiar` y `tipo_transporte`, presentan bajos niveles de asociación con el resto de los predictores, lo que indica que aportan información diferenciada y potencialmente complementaria. De manera consistente, la variable objetivo `desertor` muestra asociaciones muy débiles con las variables categóricas individuales, lo que sugiere que su relación con estos atributos no es estrictamente directa ni lineal, lo que refuerza la necesidad de enfoques multivariados y no lineales en el modelado posterior.

Figura 7

Matriz de correlación de Cramér-V entre variables categóricas.



A partir de los criterios definidos, se estableció un esquema sistemático para la selección de variables en el conjunto final de predictores, diferenciando explícitamente entre la variable objetivo y los atributos explicativos según su naturaleza, como se muestra en la tabla 14. En primer lugar, la variable desertor se definió como variable objetivo numérica, es decir, el resultado que se pretende predecir en los modelos de ML. Para las variables categóricas, su permanencia se fundamentó en el análisis de asociación mediante Cramér's V, lo que permitió evaluar el grado de dependencia estructural entre los predictores y evitar redundancias innecesarias. Bajo este criterio se mantuvieron variables sociodemográficas, económicas, tecnológicas y académicas, como estado_civil, etnia, género, modalidad_carrera,

madre_instruccion, padre_instruccion, jefe_ocupacion, salario_ingreso_total, gasto_alimentacion, gasto_transporte, tipo_vivienda, tipo_transporte, vivienda_localizacion, entre otras. Por su parte, las variables numéricas se conservaron en función del análisis de correlación lineal mediante el coeficiente de Pearson, priorizando aquellas que aportan información cuantitativa relevante sobre la trayectoria académica, el contexto familiar, el uso de tecnologías y la movilidad, tales como edad, hijos_numero, integrantes_familiar, internet_horas_uso, materias_aprobados, materias_matriculado, promedio_aprobada, semestre_tomados, semestre_anulados y transporte_horas_traslado. Este enfoque combinado garantiza que el conjunto final de variables conserve diversidad informativa, reduzca la multicolinealidad y mantenga la coherencia estadística y semántica, fortaleciendo así la robustez, la interpretabilidad y la capacidad predictiva de los modelos desarrollados para el análisis de la deserción universitaria.

Tabla 14

Selección de variables mediante correlación entre variables

Variable	Tipo	Técnica
desertor	Numérica	Variable objetivo
dispone_computador	Categórica	Cramér's V
dispone_tablet	Categórica	Cramér's V
escuela	Categórica	Cramér's V
estado_civil	Categórica	Cramér's V
etnia	Categórica	Cramér's V
gasto_alimentacion	Categórica	Cramér's V
gasto_transporte	Categórica	Cramér's V
genero	Categórica	Cramér's V
internet_dispone_casa	Categórica	Cramér's V
internet_dispone_celular	Categórica	Cramér's V
jefe_ocupacion	Categórica	Cramér's V

madre_instruccion	Categórica	Cramér's V
modalidad_carrera	Categórica	Cramér's V
padre_instruccion	Categórica	Cramér's V
preferencia_religiosa	Categórica	Cramér's V
recibe_ayuda_familiar	Categórica	Cramér's V
salario_ingreso_total	Categórica	Cramér's V
sostenibilidad_colegio	Categórica	Cramér's V
tipo_transporte	Categórica	Cramér's V
tipo_vivienda	Categórica	Cramér's V
tipo_zona_procedencia	Categórica	Cramér's V
vivienda_localizacion	Categórica	Cramér's V
edad	Numérica	Pearson
hijos_numero	Numérica	Pearson
integrantes_familiar	Numérica	Pearson
internet_horas_uso	Numérica	Pearson
materias_aprobados	Numérica	Pearson
materias_matriculado	Numérica	Pearson
nota_colegio	Numérica	Pearson
promedio_aprobada	Numérica	Pearson
red_social_horas_uso	Numérica	Pearson
semestre_anulados	Numérica	Pearson
semestre_tomados	Numérica	Pearson
transporte_horas_permanencia	Numérica	Pearson
transporte_horas_traslado	Numérica	Pearson

Correlación con la variable objetivo

A continuación, se evaluó de manera sistemática la relación estadística entre cada característica individual y la variable objetivo de deserción, con el propósito de identificar los atributos con mayor capacidad explicativa directa del fenómeno. Para ello, se adoptó un enfoque basado en la triangulación de cuatro técnicas complementarias para variables numéricas y de cuatro para

variables categóricas, reforzado mediante un esquema de selección progresiva (forward selection).

En el caso de las variables numéricas, se aplicaron las siguientes técnicas: el coeficiente de Pearson/Point-biserial, el ANOVA F-test, el Information Value (IV) con WOE y la importancia de las variables basada en RF. Para las variables categóricas se emplearon la prueba de Chi-cuadrado, el coeficiente de Cramér V, el Information Value (IV/WOE) y la importancia de las variables derivada del RF sobre codificaciones categóricas agregadas, lo que permitió identificar aportes no lineales relevantes.

Correlación de Pearson entre variables numéricas y la DEU

La figura 8 presenta los resultados del análisis de correlación de Pearson/Point-biserial entre las variables numéricas (X) y la variable objetivo. Los resultados evidencian que semestre_anulados es la variable con la mayor asociación lineal directa con la deserción ($r \approx 0.41$), lo que confirma su fuerte vínculo con trayectorias académicas inestables. En segundo lugar, se identifican promedio_aprobada ($|r| \approx 0.26$) y materias_aprobados ($|r| \approx 0.25$), ambas con correlaciones moderadas y de signo inverso con la deserción, lo que refleja el efecto protector del rendimiento académico sostenido. Asimismo, materias_matriculado presenta una asociación significativa ($|r| \approx 0.20$), lo que sugiere que la carga académica también influye en la probabilidad de abandono.

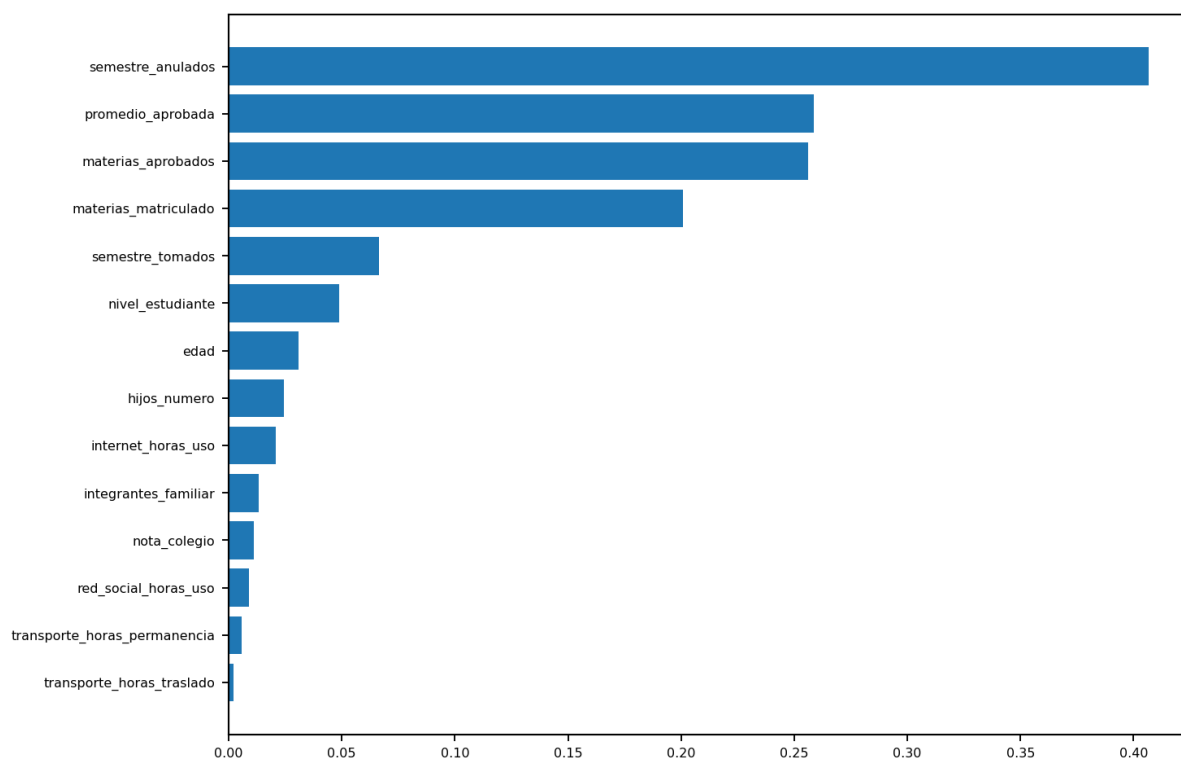
En un nivel intermedio se ubican semestre_tomados ($|r| \approx 0.07$) y nivel_estudiante ($|r| \approx 0.05$), mientras que variables demográficas y contextuales como edad, hijos_numero, internet_horas_uso e integrantes_familiar muestran asociaciones débiles ($|r| < 0.04$). Finalmente, variables como nota_colegio, red_social_horas_uso,

transporte_horas_permanencia y transporte_horas_traslado presentan correlaciones prácticamente nulas con la variable objetivo.

Figura 8

Relación entre variables numéricas (X) y la deserción universitaria (Y) mediante

Pearson/Point-biserial



ANOVA F-test aplicado a variables numéricas y la DEU

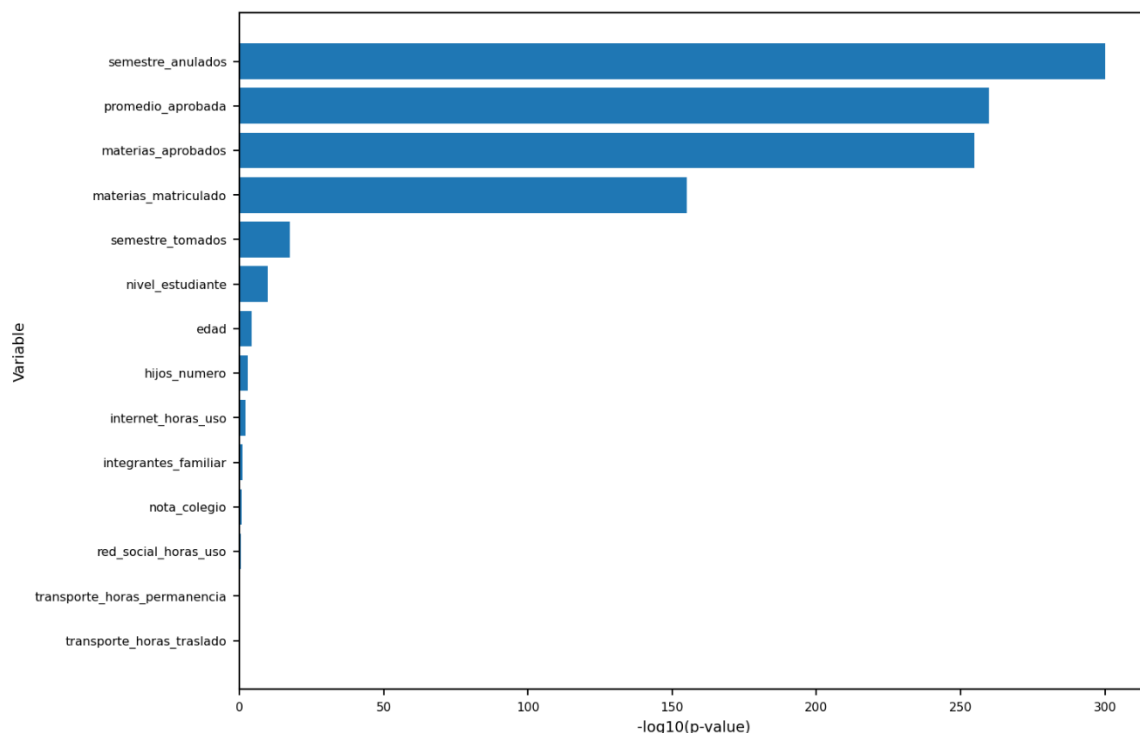
La figura 9 presenta los resultados del ANOVA F-test de las variables numéricas frente a la variable objetivo DEU, utilizando el estadístico $-\log_{10}(\text{p-value})$ como medida de significancia estadística, lo que permite identificar las variables que difieren entre los grupos de estudiantes desertores y no desertores y aporta evidencia cuantitativa sobre su capacidad de discriminación. Los resultados indican que semestre_anulados presenta el mayor nivel de significancia estadística, con un valor extremadamente elevado de $-\log_{10}(\text{p-value})$, lo que confirma que las diferencias entre ambos grupos son altamente relevantes desde una perspectiva inferencial. De

forma consistente, variables asociadas al rendimiento académico, como promedio_aprobada, materias_aprobados y materias_matriculado, también presentan valores elevados del estadístico F, lo que evidencia diferencias sustantivas en su distribución entre estudiantes que desertan y aquellos que permanecen en el sistema universitario.

En un segundo nivel de relevancia se encuentran semestre_tomados y nivel_estudiante, cuyos resultados, aunque menos intensos, siguen siendo estadísticamente significativos y reflejan la influencia de la progresión académica en la probabilidad de deserción. Por el contrario, variables de carácter demográfico y contextual, como edad, hijos_numero, internet_horas_uso, integrantes_familiar, nota_colegio, red_social_horas_uso y las variables de transporte, presentan valores bajos del estadístico, lo que sugiere diferencias marginales o no significativas entre los grupos analizados.

Figura 9

Análisis ANOVA F-test entre variables numéricas (X) y la DEU (Y)



Análisis de importancia de variables numéricas mediante Information Value (IV) y la DEU

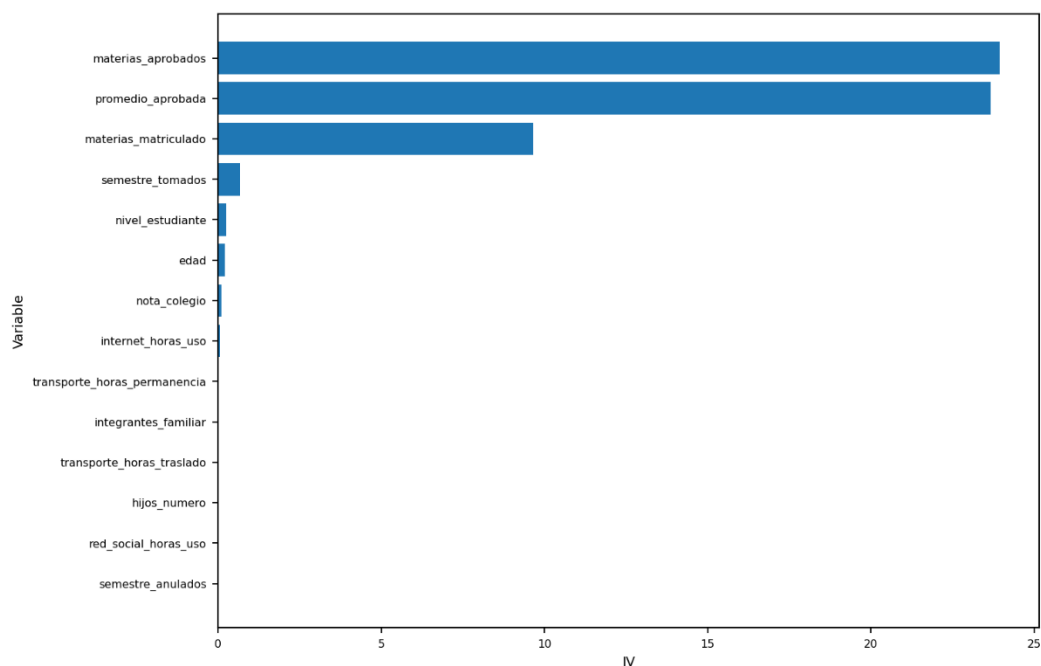
La figura 10 presenta los resultados del análisis de Information Value (IV) aplicado a las variables numéricas mediante un proceso de binning y la transformación Weight of Evidence (WOE), con el objetivo de evaluar su capacidad discriminante frente a la variable de deserción universitaria. Esta técnica permite cuantificar la capacidad de cada predictor para separar eficazmente a los estudiantes desertores de los no desertores, superando las limitaciones de los enfoques estrictamente lineales.

Los resultados muestran que las variables asociadas al desempeño académico presentan los valores de IV más altos. En particular, *materias_aprobados* y *promedio_aprobada* presentan valores de IV extremadamente altos, lo que indica un poder predictivo muy fuerte y evidencia una clara diferenciación entre los grupos de interés. De forma consistente, *materias_matriculado* también alcanza un valor elevado de IV, lo que confirma que la carga académica y el rendimiento acumulado constituyen factores críticos en la explicación del fenómeno de la deserción.

En contraste, variables como *semestre_tomados* y *nivel_estudiante* muestran valores de IV bajos a moderados, lo que sugiere una contribución explicativa limitada cuando se consideran de forma individual. Asimismo, las variables demográficas, como *edad*, *nota_colegio*, *internet_horas_uso*, *integrantes_familiar*, *transporte_horas_permanencia*, *hijos_numero* y *red_social_horas_uso*, presentan valores de IV cercanos a cero, lo que indica una capacidad discriminante débil bajo este criterio específico.

Figura 10

Importancia de variables numéricas según Information Value (IV)



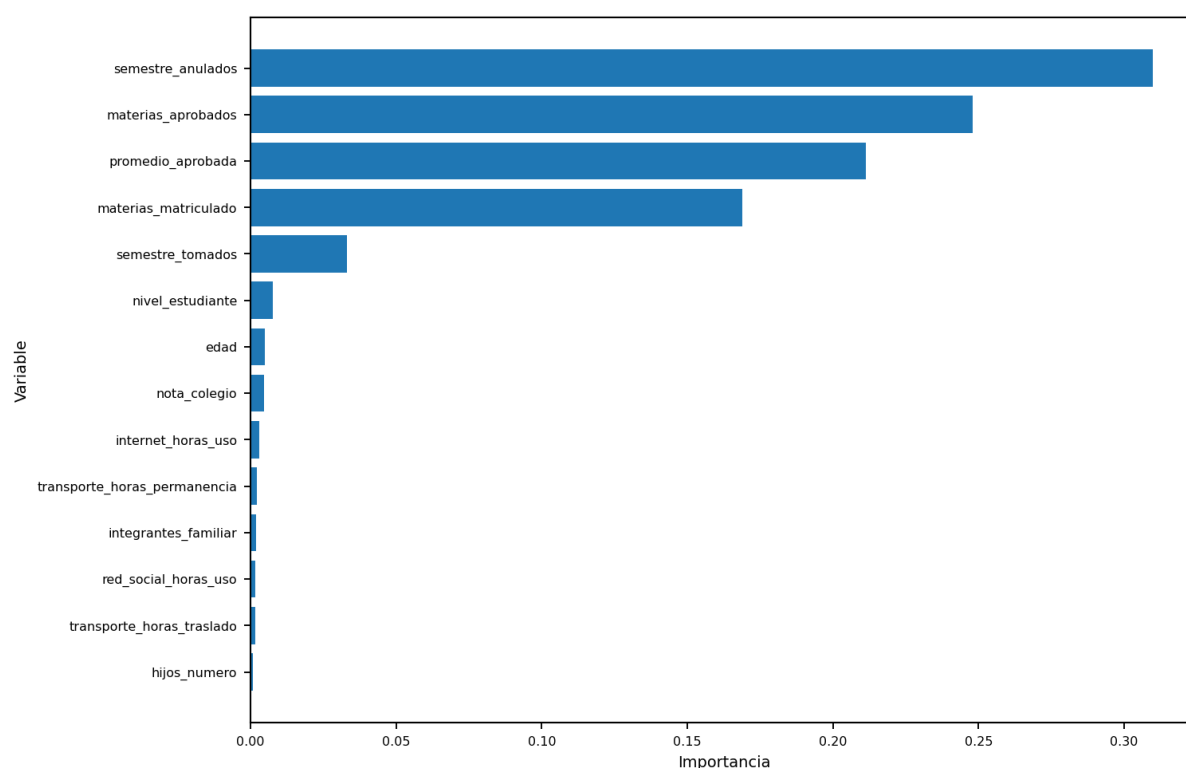
Contribución de variables numéricas en la DEU mediante Random Forest

De manera complementaria, la Figura 11 muestra los resultados del análisis de importancia de variables obtenido mediante un modelo RF, aplicado a las variables numéricas en relación con la variable objetivo de deserción universitaria. A diferencia de los enfoques estadísticos tradicionales, esta técnica permite capturar relaciones no lineales, efectos de umbral e interacciones complejas entre los predictores, lo que proporciona una perspectiva complementaria y robusta para la selección de características. Los resultados evidencian que semestre_anulados constituye el predictor numérico más influyente, concentrando la mayor proporción de la importancia relativa, lo que confirma su papel determinante en la explicación del riesgo de deserción. En un segundo nivel de relevancia se ubican materias_aprobados, promedio_aprobada y materias_matriculado, lo que consolida nuevamente el desempeño académico como eje central del fenómeno analizado. Este patrón es consistente con los hallazgos obtenidos mediante ANOVA e IV/WOE, lo que refuerza la estabilidad y la coherencia del proceso de selección. Variables como semestre_tomados y nivel_estudiante aportan una

contribución moderada, lo que sugiere que su influencia se manifiesta principalmente en combinación con otros factores académicos. En contraste, las variables de carácter demográfico y contextual, entre ellas edad, nota_colegio, internet_horas_uso, integrantes_familiar, red_social_horas_uso y las asociadas al transporte, presentan importancias marginales, lo que indica un impacto limitado al analizarse de forma aislada en un modelo no lineal.

Figura 11

Importancia de variables numéricas (X) con la DEU (Y) mediante Random Forest



Análisis de variables categóricas mediante la prueba chi-cuadrado (χ^2)

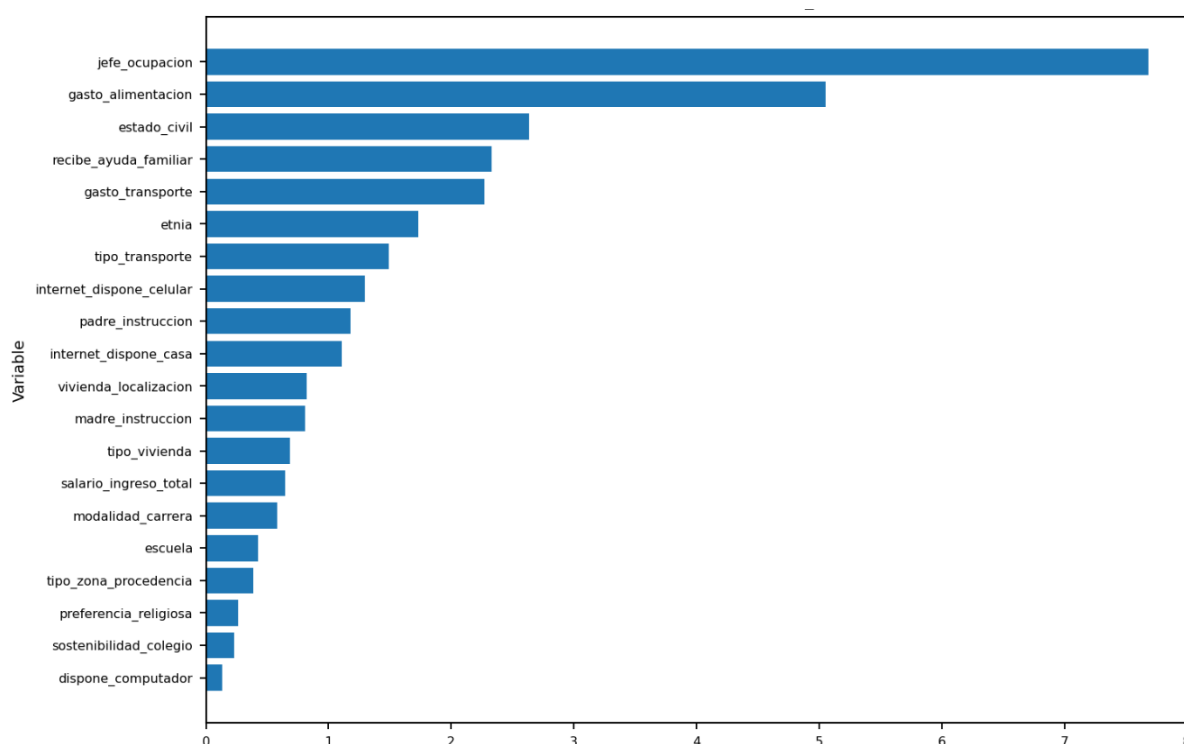
La Figura 12 presenta los resultados del análisis de asociación mediante la prueba de chi-cuadrado, aplicada a las variables categóricas transformadas mediante One-Hot Encoding agregado, con el objetivo de evaluar su relación estadística con la deserción universitaria. Este enfoque permite contrastar la hipótesis de independencia entre cada predictor categórico y la variable objetivo. Los resultados evidencian que jefe_ocupacion es la variable categórica con

el mayor nivel de asociación estadísticamente significativa con la deserción, seguida de `gasto_alimentacion`, `estado_civil`, `recibe_ayuda_familiar` y `gasto_transporte`. Estas variables reflejan dimensiones socioeconómicas y familiares que inciden de manera relevante en la permanencia estudiantil, lo cual es coherente con los enfoques teóricos que vinculan la vulnerabilidad económica y el apoyo familiar con el riesgo de abandono.

En un nivel intermedio de asociación se ubican variables como `etnia`, `tipo_transporte`, `internet_dispone_celular`, `padre_instruccion` e `internet_dispone_casa`, cuya significancia estadística sugiere una contribución contextual relevante, aunque menos determinante que la de los factores económicos directos. Por el contrario, variables como `preferencia_religiosa`, `sostenibilidad_colegio` y `dispone_computador` presentan valores de significancia baja, lo que indica una relación débil o marginal con la deserción al analizarse de forma individual.

Figura 12

Importancia de las variables categóricas respecto a la deserción universitaria mediante chi-cuadrado



Análisis de variables categóricas mediante la correlación de Cramér's V

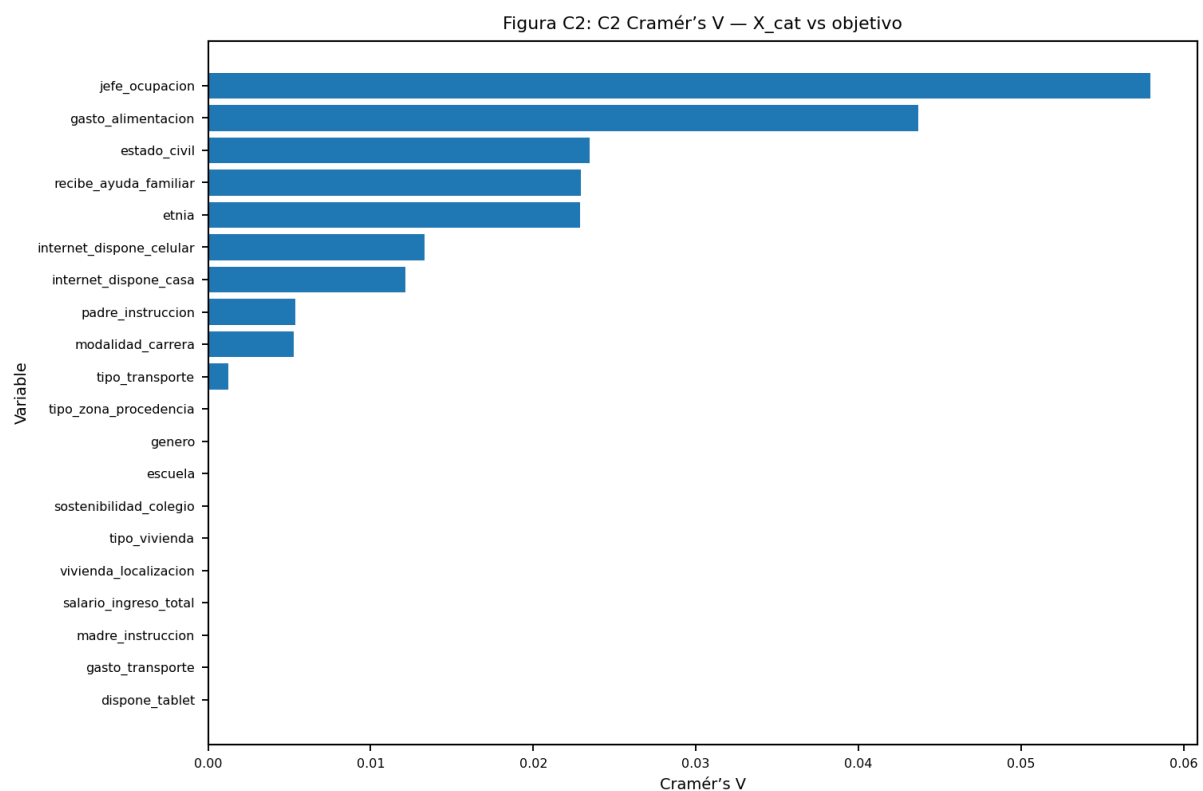
La Figura 13 muestra los resultados del análisis de asociación mediante el estadístico de Cramér's V, aplicado a las variables categóricas frente a la variable objetivo de deserción universitaria. Esta medida permite cuantificar la fuerza de asociación entre variables cualitativas, independientemente del tamaño muestral, complementando así los contrastes de significancia estadística obtenidos previamente mediante la prueba de chi-cuadrado.

Los resultados evidencian que jefe_ocupacion presenta la mayor magnitud de asociación con la deserción, seguida por gasto_alimentacion, estado_civil, recibe_ayuda_familiar y etnia, todas ellas con valores de Cramér's V superiores a los de las demás variables analizadas. Si bien los valores absolutos se sitúan en rangos bajos a moderados consistentes con fenómenos sociales complejos y multifactoriales, su relevancia radica en que reflejan relaciones estructurales no triviales entre las condiciones socioeconómicas y familiares y la permanencia estudiantil.

En un segundo nivel de asociación se ubican variables relacionadas con el acceso a tecnologías, como `internet_dispone_celular` e `internet_dispone_casa`, así como `padre_instruccion` y `modalidad_carrera`, que aportan información contextual complementaria sobre el entorno académico y familiar del estudiante. Por el contrario, variables como `tipo_zona_procedencia`, `genero`, `escuela`, `sostenibilidad_colegio`, `tipo_vivienda`, `vivienda_localizacion`, `salario_ingreso_total`, `madre_instruccion`, `gasto_transporte` y `dispone_tablet` presentan valores cercanos a cero, lo que indica una asociación débil con la variable objetivo cuando se consideran de forma aislada.

Figura 13

Asociación entre variables categóricas y la deserción universitaria mediante Cramér's V

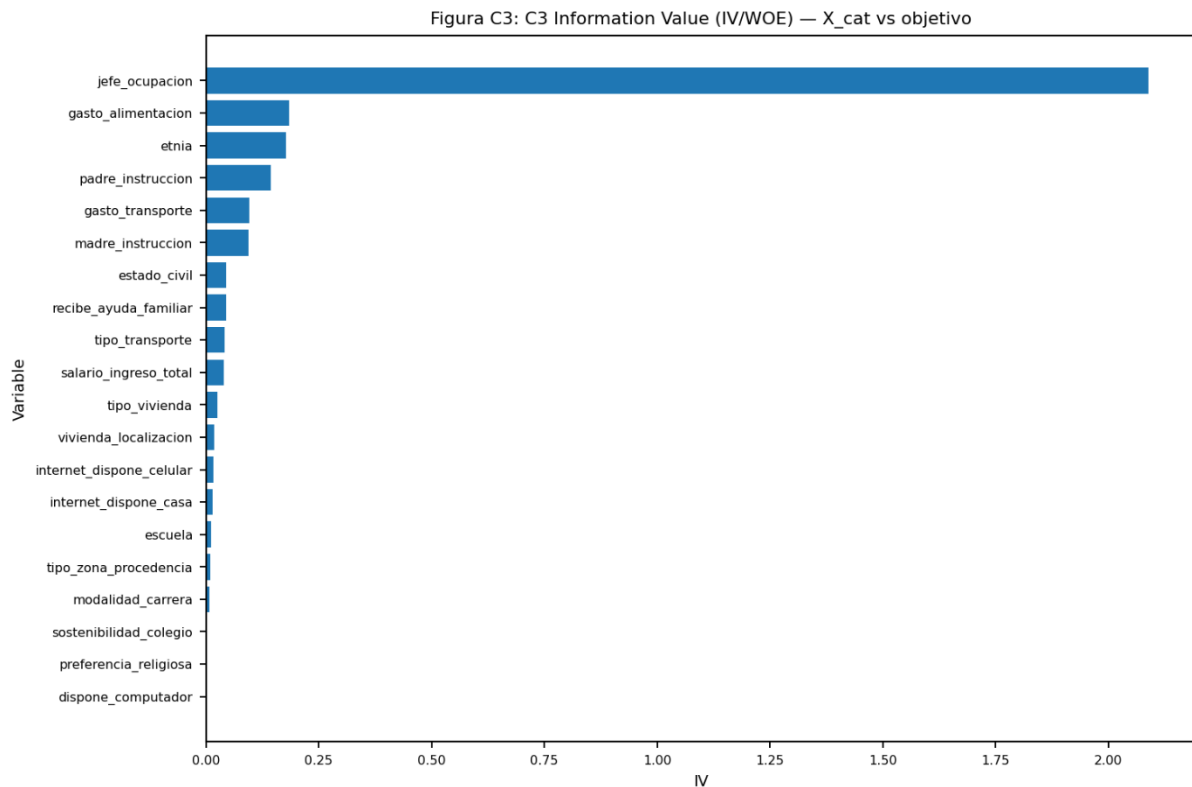


Análisis de la importancia de las variables categóricas mediante el Information Value (IV/WOE).

La figura 14 presenta el análisis del Information Value (IV/WOE) aplicado a las variables categóricas en relación con la variable objetivo de deserción universitaria, con el propósito de evaluar su poder discriminante individual. Los resultados muestran una marcada heterogeneidad en la capacidad explicativa de los predictores. En particular, jefe_ocupacion destaca por un valor de IV notablemente superior al de los demás, lo que indica un alto poder de separación entre estudiantes desertores y no desertores, lo que la consolida como una de las variables categóricas más relevantes desde una perspectiva predictiva. En un segundo nivel, variables como gasto_alimentacion, etnia, padre_instruccion y gasto_transporte presentan valores de IV bajos a moderados, lo que sugiere una contribución informativa adicional, aunque claramente inferior a la del predictor principal. El resto de las variables categóricas presenta valores de IV cercanos a cero, lo que evidencia una capacidad discriminante limitada cuando se consideran de manera aislada. En conjunto, este análisis respalda la necesidad de combinar criterios multitécnicos en la selección de características, ya que, si bien pocas variables muestran un alto poder discriminante individual, otras pueden adquirir relevancia al interactuar con predictores numéricos o en modelos no lineales, lo que justifica su retención en las fases posteriores de modelado.

Figura 14

Importancia de variables categóricas con Information Value (IV/WOE)



Análisis de importancia de variables categóricas mediante Random Forest (RF)

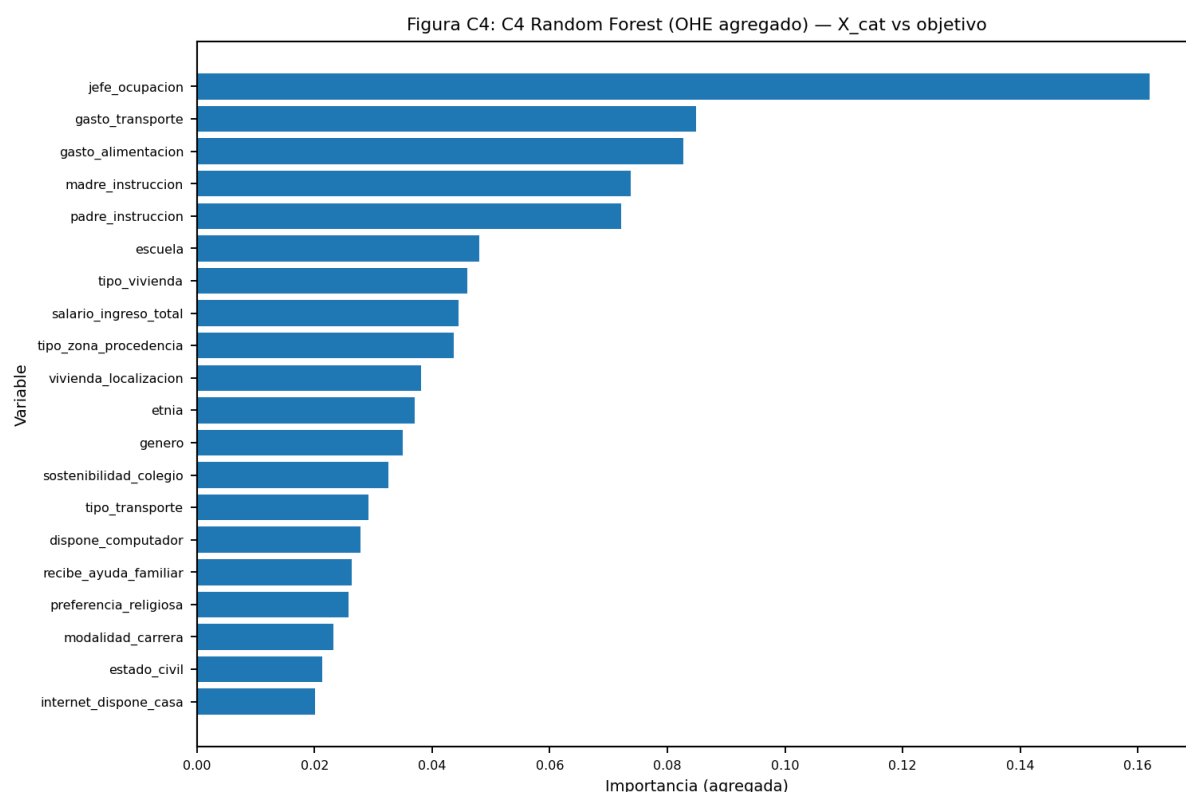
La figura 15 muestra la importancia agregada de las variables categóricas estimada mediante RF (con codificación One-Hot Encoding aplicada) frente a la variable DEU. Este enfoque permite capturar relaciones no lineales e interacciones complejas entre las categorías y el resultado, superando las limitaciones de los métodos estadísticos univariados.

Los resultados evidencian que jefe_ocupacion es, con diferencia, el predictor categórico más influyente, lo que indica que la condición laboral del jefe del hogar desempeña un papel central en la explicación del riesgo de deserción. En un segundo nivel de relevancia se ubican gasto_transporte y gasto_alimentacion, lo que refuerza la importancia de los factores socioeconómicos y de acceso en la permanencia estudiantil. Asimismo, variables asociadas al capital educativo familiar, como madre_instruccion y padre_instruccion, aportan una contribución significativa, lo cual es coherente con la literatura sobre las trayectorias académicas y el apoyo familiar.

Otras variables, como escuela, tipo_vivienda, salario_ingreso_total, tipo_zona_procedencia y vivienda_localizacion, muestran una importancia intermedia, lo que sugiere que el contexto educativo previo y las condiciones de residencia aportan información relevante cuando se consideran junto con otros predictores. Finalmente, variables como preferencia_religiosa, modalidad_carrera, estado_civil e internet_dispone_casa presentan niveles de importancia individual menores, aunque no nulos, lo que indica que su contribución puede provenir principalmente de interacciones con otros atributos.

Figura 15

Importancia de variables categóricas (X) respecto a la deserción mediante Random Forest



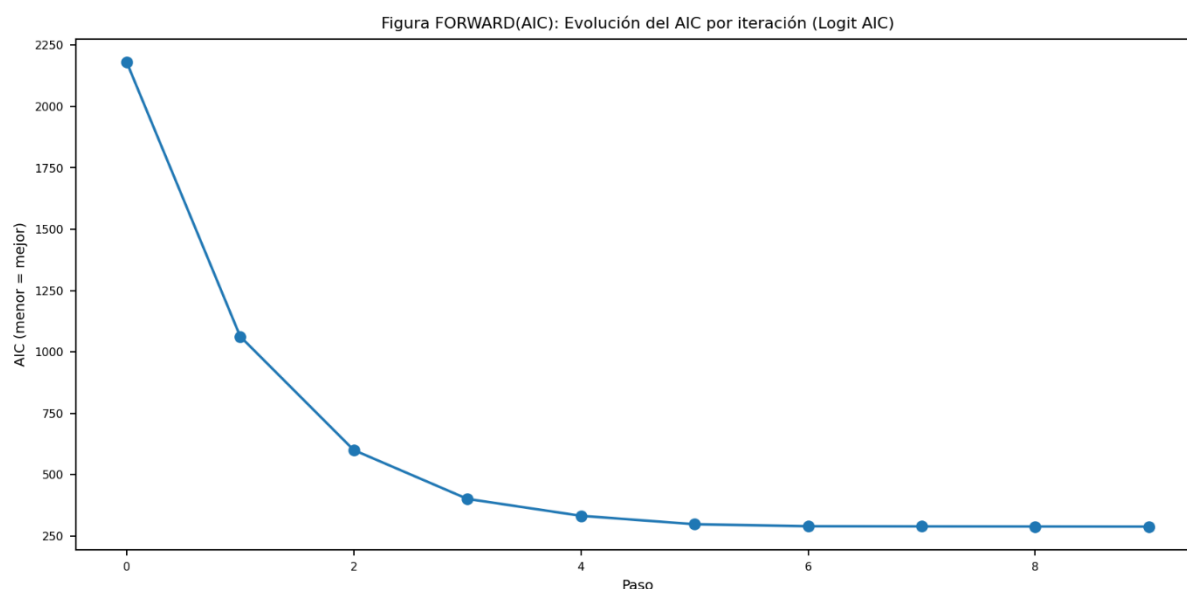
Selección progresiva de características basada en el criterio de información de Akaike

Con el propósito de optimizar el conjunto de variables predictoras y garantizar un equilibrio adecuado entre capacidad explicativa, parsimonia e interpretabilidad, se aplicó un procedimiento de selección progresiva de variables (Forward Selection) basado en el criterio

de información de Akaike (AIC). La figura 16 muestra la evolución del valor del AIC a lo largo de las iteraciones del proceso, evidenciando una disminución abrupta en las primeras etapas, seguida de una fase de estabilización progresiva. El análisis de la evolución del criterio de información de Akaike (AIC) evidencia que, tras una disminución significativa en las primeras iteraciones del proceso de selección progresiva de variables, el modelo comienza a estabilizarse aproximadamente a partir de la quinta variable. A partir de este punto, las reducciones en el valor del AIC son mínimas y las variaciones entre iteraciones consecutivas resultan marginales, lo que indica que la incorporación de nuevas variables ya no aporta mejoras sustanciales en la capacidad explicativa del modelo

Figura 16

Evolución del criterio de información de Akaike (AIC) durante el proceso de selección progresiva de variables



Finalmente, a partir del proceso de selección de características, se conformó el conjunto de datos mediante varias técnicas estadísticas y de ML, diferenciadas según la naturaleza de las variables numéricas y categóricas, tal como se presenta en la tabla 15. En particular, se

emplearon la correlación de Pearson, la información mutua (MI), la LR con regularización L1 (LASSO) y RF para variables numéricas; mientras que para las variables categóricas, se aplicaron el estadístico χ^2 (Chi-cuadrado), la información mutua, la LR penalizada y RF. De manera complementaria, se incorporó el método Stepwise Forward como técnica mixta para evaluar la contribución conjunta de variables numéricas y categóricas en un contexto multivariado.

Las variables de mayor recurrencia fueron consideradas las más estables y relevantes para la predicción de la DEU. Las variables de mayor recurrencia fueron consideradas las más estables y relevantes para la predicción de la DEU. Entre las más destacadas se encontraron materias_aprobados, promedio_aprobada, materias_matriculado, semestre_anulados, semestre_tomados, nivel_estudiante, edad, internet_horas_uso y nota_colegio. Estos resultados evidencian la influencia predominante de los factores académicos, de rendimiento estudiantil y de conectividad en la DEU.

Tabla 15

Selección de variables

Variable	Numericas				Categoricas					
	Pearson	ANOVA	IV	RF	Chi ²	Cramér's V	IV	RF	Forward (AIC)	Consenso
materias_aprobados	X	X	X	X					X	5
promedio_aprobada	X	X	X	X					X	5
materias_matriculado	X	X	X	X					X	5
semestre_anulados	X	X	X	X					X	5
semestre_tomados	X	X	X	X					X	5
nivel_estudiante	X	X	X	X					X	5
edad	X	X	X	X					X	5
internet_horas_uso	X	X	X	X					X	5
nota_colegio	X	X	X	X					X	5
jefe_ocupacion					X	X	X	X		4
gasto_alimentacion					X	X	X	X		4
etnia					X	X	X	X		4

padre_instruccion					X	X	X	X	4
gasto_transporte					X	X	X	X	4
estado_civil					X	X	X	X	4
recibe_ayuda_familiar					X	X	X	X	4
madre_instruccion					X	X	X	X	4
integrantes_familiar	X	X	X	X					4
tipo_transporte					X	X	X	X	4
hijos_numero	X	X	X	X					4
tipo_vivienda					X	X	X	X	4
transporte_horas_permanencia	X	X	X	X					4
salario_ingreso_total					X	X	X	X	4
escuela					X	X	X	X	4
red_social_horas_uso	X	X	X	X					4
tipo_zona_procedencia					X	X	X	X	4
internet_dispone_casa					X	X	X	X	4
vivienda_localizacion					X	X	X	X	4
transporte_horas_traslado	X	X	X	X					4
modalidad_carrera					X	X	X	X	4
sostenibilidad_colegio					X	X	X	X	4
internet_dispone_celular					X	X	X		3
dispone_computador					X		X	X	3
preferencia_religiosa					X		X	X	3
genero						X		X	2
dispone_tablet						X			1

5.1.5. Hipótesis Específica 1

La hipótesis planteada establece que la construcción de un conjunto de datos integrado por variables relevantes permite modelar adecuadamente la deserción estudiantil universitaria. Los resultados obtenidos respaldan esta afirmación con base en la evidencia empírica derivada del análisis y del procesamiento del conjunto de datos.

En primer lugar, la fase de preprocesamiento y depuración de los datos permitió consolidar un conjunto de datos consistente y estructuralmente robusto. Mediante la detección y el tratamiento de valores atípicos, la imputación de valores faltantes, la transformación estadística de variables

y la codificación de atributos categóricos, se obtuvo un conjunto de datos libre de inconsistencias y apto para el entrenamiento de algoritmos de ML. Como resultado de este proceso, se conformó un dataset final compuesto por 17 183 registros y 133 variables, que cumple con los criterios de calidad, estabilidad estadística y reproducibilidad necesarios para el análisis predictivo de la deserción universitaria.

A continuación, el proceso de selección de características permitió identificar las variables con mayor capacidad explicativa mediante un enfoque de triangulación metodológica que integró múltiples técnicas estadísticas y de ML, entre ellas la correlación de Pearson, ANOVA, Information Value (IV), Random Forest, pruebas χ^2 , Cramér's V y modelos de regresión penalizada. Asimismo, se incorporó un procedimiento de selección progresiva basado en el criterio de información de Akaike (AIC), lo que permitió optimizar el conjunto de predictores, manteniendo un equilibrio entre la capacidad explicativa, la parsimonia y la interpretabilidad del modelo. Los resultados de este análisis evidenciaron que diversas variables presentan una relación significativa con la deserción estudiantil. En particular, se identificó que semestre_anulados presenta la mayor asociación con la variable objetivo ($|r| \approx 0.41$), seguida de promedio_aprobada ($|r| \approx 0.26$) y materias_aprobados ($|r| \approx 0.25$), lo que confirma la relevancia de los factores asociados a la DEU. Asimismo, variables como materias_matriculado, edad, internet_horas_uso y nota_colegio también contribuyen al poder explicativo del modelo, aunque con niveles de asociación moderados o bajos.

De manera complementaria, el análisis de la importancia de las variables y el consenso entre diferentes técnicas de selección confirmaron la estabilidad de un conjunto de predictores clave que integran dimensiones académicas, socioeconómicas, familiares y sociodemográficas, lo que evidencia que la DEU es un fenómeno multidimensional cuya explicación requiere considerar múltiples factores interrelacionados.

Finalmente, el desempeño de los modelos predictivos entrenados con este conjunto de datos confirma su capacidad para representar adecuadamente el fenómeno analizado. En particular, varios modelos alcanzaron niveles de precisión superiores al 0,99, con valores elevados de recall, F1-score y ROC-AUC, lo que evidencia una alta capacidad discriminativa para identificar estudiantes en riesgo de deserción.

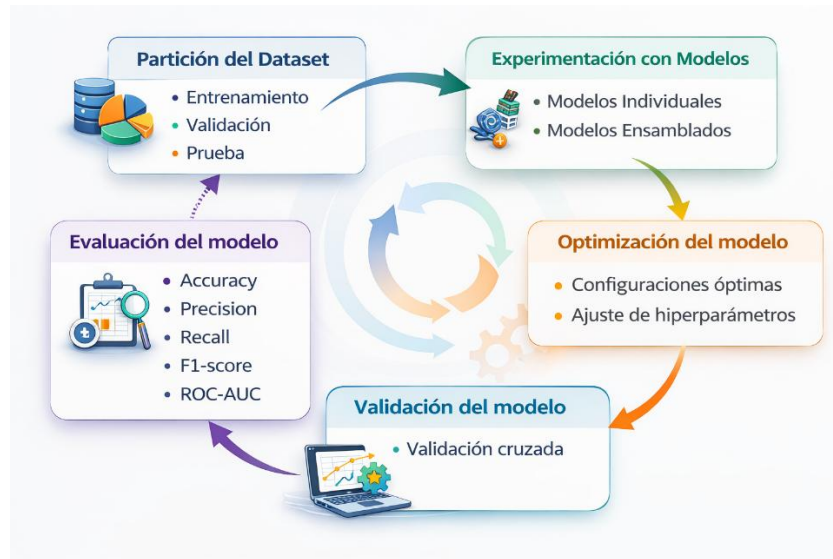
En conjunto, estos resultados demuestran que el dataset construido no solo cumple con criterios de calidad y consistencia estadística, sino que además contiene variables con capacidad explicativa suficiente para sustentar el modelado predictivo del fenómeno estudiado. Por tanto, la evidencia empírica obtenida permite aceptar la hipótesis planteada, al confirmarse que la construcción de un conjunto de datos integrado por variables relevantes constituye un elemento fundamental para modelar adecuadamente la deserción estudiantil universitaria.

5.2. Resultados del entrenar y evaluar algoritmos de predicción

En esta fase se examina el comportamiento de los modelos de ML desarrollados para predecir la DEU. Como se observa en la figura 17, la finalidad es presentar de manera ordenada los resultados obtenidos a partir del proceso de partición del conjunto de datos, del entrenamiento de modelos individuales y ensamblados, del ajuste de hiperparámetros, de la validación cruzada y de la evaluación mediante métricas de desempeño. Esta fase no solo permite identificar qué algoritmos alcanzan los mejores niveles de precisión predictiva, sino también valorar su estabilidad, robustez y capacidad de generalización a datos no observados.

Figura 17

Proceso de modelado predictivo



5.2.1 División del conjunto de datos

Para el desarrollo experimental, el conjunto de datos se segmentó en tres subconjuntos independientes: entrenamiento (60%), validación (20%) y prueba (20%), siguiendo un esquema de partición estratificada según la variable objetivo (DEU). Esta estrategia permitió preservar la proporción original entre estudiantes desertores y no desertores en cada subconjunto, garantizando la representatividad de la distribución de clases a lo largo de todo el proceso de modelado.

Como se presenta en la tabla 16, el conjunto de datos utilizado para el modelado de la DEU se conformó por un total de 17183 registros, de los cuales 193 corresponden a estudiantes desertores y 16990 a estudiantes no desertores, lo que evidencia un marcado desbalance de clases, característico de los problemas de predicción. El subconjunto de entrenamiento para la validación cruzada incluyó 10337 registros, compuestos por 116 desertores y 10221 no desertores, y sirvió como base principal para entrenar los modelos de clasificación. Por su parte, el subconjunto de validación constó de 3438 registros (39 desertores y 3399 no desertores) y se utilizó para ajustar hiperparámetros y seleccionar configuraciones óptimas. Finalmente, el subconjunto de prueba, compuesto por 3398 observaciones (38 desertores y 3360 no

desertores), está destinado a la evaluación final del desempeño de los modelos, garantizando una estimación objetiva, robusta y no sesgada de su capacidad de generalización a datos no observados.

Tabla 16

Distribución del conjunto de datos DEU para entrenamiento, validación y prueba

Conjunto de datos	Cantidad para entrenamiento (validación cruzada)				Para prueba y evaluación		Total
	Para entrenamiento		Para validación		Desertor	No desertor	
	Desertor	No desertor	Desertor	No desertor	Desertor	No desertor	
Modelo explicabilidad con conjunto de datos de la DEU	116	10221	39	3399	38	3360	17183

5.2.2 Ejecución de algoritmos de predicción

Con el conjunto de datos final, se aplicaron varios algoritmos de ML supervisado para evaluar su rendimiento inicial en la identificación de estudiantes en riesgo de deserción. La selección de los algoritmos se fundamentó en lo reportado en la literatura científica, en concordancia con la revisión sistemática presentada por (Quimiz-Moreira et al., 2025), donde se documenta de manera exhaustiva el uso de modelos lineales, algoritmos basados en árboles, métodos probabilísticos, técnicas de ensamblaje y arquitecturas neuronales en el estudio de la DEU. En coherencia con ello, se seleccionaron algoritmos representativos de las principales familias de modelado empleadas en problemas de predicción educativa.

En particular, se incorporó el Árbol de Decisión (DT) por su elevada interpretabilidad y su capacidad para identificar explícitamente los factores críticos asociados al riesgo académico, una característica ampliamente documentada en la literatura sobre deserción universitaria (Guleria & Sood, 2023; Yağcı, 2022). Asimismo, se incluyeron RF y ET como modelos de

ensamble basados en árboles, dada su robustez frente a la alta dimensionalidad, su estabilidad ante el ruido y su buen desempeño en contextos caracterizados por un marcado desbalance de clases, condiciones recurrentes en los conjuntos de datos asociados a la DEU (Bezek Güre, 2024; González-Nucamendi et al., 2023).

De igual forma, se consideraron algoritmos del tipo boosting, como GB y HGB, que han demostrado una elevada capacidad para modelar relaciones no lineales complejas y mejorar el rendimiento predictivo mediante la combinación secuencial de clasificadores débiles. En particular, HGB destaca por su eficiencia computacional y escalabilidad en conjuntos de datos de tamaño medio a grande, lo cual resulta especialmente pertinente en entornos educativos con múltiples variables heterogéneas.

Adicionalmente, se incorporó el Multi-Layer Perceptron (MLP) entre los modelos neuronales, por su capacidad para captar interacciones complejas y patrones no lineales entre variables académicas, sociodemográficas y contextuales asociadas al riesgo de deserción. Estudios previos han evidenciado que el MLP alcanza niveles competitivos de desempeño y estabilidad en diversos escenarios educativos (Ahuja & Khosla, 2020; Al-Amin et al., 2021; Saa, 2022), lo que justifica su inclusión en el análisis comparativo.

Finalmente, se incluyeron modelos base ampliamente utilizados en la literatura, como LR, NB y SVM, que sirven como referentes metodológicos por su simplicidad conceptual, eficiencia computacional y valor interpretativo. Estos modelos permiten establecer líneas de base sólidas para contrastar el desempeño de los enfoques más complejos y evaluar objetivamente las ganancias obtenidas mediante técnicas de ensamble y de aprendizaje no lineal (Liang et al., 2020; Martinez, 2021; Páez, 2022).

Experimento 1 sin balanceo

El experimento se constituyó en un escenario base para la predicción de la deserción universitaria, en el que se emplearon las variables seleccionadas sin aplicar técnicas de balanceo de clases. El análisis se estructuró en dos fases claramente diferenciadas: evaluación del entrenamiento, orientada a verificar la estabilidad y la capacidad de generalización de los modelos, y evaluación en el conjunto de prueba, destinada a estimar el desempeño real sobre datos no observados.

Con el fin de evaluar la capacidad de generalización de los modelos predictivos y evitar conclusiones sesgadas derivadas del ajuste al conjunto de entrenamiento, se incorporó una fase de validación independiente (X_{val}) en el Experimento 1, correspondiente al escenario de predicción sin técnicas de balanceo de clases. Esta etapa resulta crítica en contextos de deserción universitaria, donde el desempeño aparente en el entrenamiento puede no reflejar adecuadamente el comportamiento del modelo ante datos no observados.

En esta fase, se compararon de manera sistemática modelos individuales y ensamblados, considerando métricas complementarias orientadas tanto a la capacidad discriminativa como al equilibrio entre clases, específicamente accuracy, recall, precision, F1-score, ROC-AUC y especificidad. El análisis de estos indicadores permite identificar no solo el modelo con mayor desempeño promedio, sino también aquel que mantiene un comportamiento estable y robusto frente a la variabilidad inherente de los datos.

Los resultados obtenidos en la validación, presentados en la tabla 17, constituyen la base técnica para la selección del modelo candidato a evaluación final en el conjunto de prueba (X_{test}), así como para las etapas posteriores de análisis de explicabilidad, en las que la estabilidad del modelo y la consistencia de sus decisiones son requisitos fundamentales.

Tabla 17

Resultados de validación de modelos predictivos (X_val) – Experimento 1 (sin balanceo)

Grupo	Modelo	Accuracy	Recall	Precision	F1	ROC-AUC	Especificidad
Ensamblado	HGB	0.996727	0.96875	0.794872	0.873239	0.999069	0.997057
Ensamblado	ET	0.996000	0.96875	0.756098	0.849315	0.998816	0.996321
Individual	DT	0.996000	0.93750	0.769231	0.845070	0.967094	0.996689
Ensamblado	GB	0.995636	0.96875	0.738095	0.837838	0.999028	0.995953
Ensamblado	RF	0.995636	0.93750	0.750000	0.833333	0.998799	0.996321
Individual	LR	0.995273	1.00000	0.711111	0.831169	0.998873	0.995217
Individual	SVM	0.994909	1.00000	0.695652	0.820513	0.997896	0.994849
Individual	MLP	0.995273	0.87500	0.756757	0.811594	0.998574	0.996689
Individual	NB	0.984000	1.00000	0.421053	0.592593	0.998436	0.983812

En términos generales, los modelos de aprendizaje ensamblado superan de forma consistente a los modelos individuales, especialmente cuando se analizan métricas robustas frente al desbalance de clases, como el F1-score y el ROC-AUC. El modelo HGB presenta el mejor desempeño global en validación, alcanzando un F1-score de 0.873 y un AUC-ROC de 0.999, lo que indica una excelente capacidad de discriminación entre estudiantes desertores y no desertores. Este comportamiento sugiere que los métodos basados en boosting con histogramas capturan de manera eficiente relaciones no lineales y efectos de interacción entre variables académicas y sociodemográficas, y además mantienen un tiempo de entrenamiento reducido.

De forma similar, los modelos Extra Trees, GB y RF muestran resultados altamente competitivos, con valores de F1-score superiores a 0.83 y de ROC-AUC cercanos a 1.0, lo que confirma la estabilidad del enfoque basado en árboles. No obstante, se observa que estos modelos presentan ligeras diferencias en la relación precisión–recall, lo que refleja distintos compromisos entre la detección de desertores y la minimización de falsos positivos.

Entre los modelos individuales, la LR y el SVM alcanzan un recall máximo de 1,000, lo que indica una alta capacidad para identificar correctamente a los estudiantes en riesgo de deserción. Sin embargo, este desempeño se acompaña de valores de precisión más bajos, lo que se traduce en un mayor número de falsos positivos y, en consecuencia, en puntuaciones F1 inferiores a las de los modelos ensamblados. Este comportamiento es consistente con modelos lineales que priorizan la sensibilidad por encima de la especificidad.

El modelo NB, aunque presenta un recall perfecto, muestra un desempeño más limitado en precisión y F1-score, lo que evidencia su menor capacidad para modelar dependencias complejas entre variables en este contexto. Por su parte, el MLP alcanza resultados aceptables, aunque no supera a los métodos basados en ensamblajes y presenta un mayor costo computacional relativo. Para complementar estos hallazgos, los resultados obtenidos en el conjunto de prueba (X_{test}), presentados en la tabla 18, confirman la consistencia y la capacidad de generalización de los modelos evaluados.

Tabla 18

Resultados en el conjunto de prueba (X_{test})– Experimento 1 (sin balanceo)

Grupo	Modelo	Accuracy	Recall	Precision	F1	ROC-AUC	Especificidad
Ensamblado	HGB	0.997672	0.925	0.880952	0.902439	0.999584	0.998528
Ensamblado	GB	0.997381	0.900	0.878049	0.888889	0.999669	0.998528
Individual	DT	0.997381	0.900	0.878049	0.888889	0.949264	0.998528
Ensamblado	RF	0.997381	0.900	0.878049	0.888889	0.999551	0.998528
Ensamblado	ET	0.997381	0.900	0.878049	0.888889	0.999481	0.998528
Individual	MLP	0.996800	0.850	0.871795	0.860759	0.999183	0.998528
Individual	SVM	0.995345	0.925	0.740000	0.822222	0.998719	0.996173
Individual	LR	0.994763	1.000	0.689655	0.816327	0.998948	0.994701
Individual	NB	0.977888	1.000	0.344828	0.512821	0.998800	0.977627

En esta etapa, HGB se consolida como el modelo con mejor desempeño integral, alcanzando un F1-score superior a 0.90 y un ROC-AUC cercano a 0.9996, lo que evidencia una capacidad de discriminación prácticamente perfecta entre estudiantes desertores y no desertores en datos no vistos. De manera coherente con los resultados de validación, GB, RF y Extra Trees mantienen valores de F1 alrededor de 0.89 y ROC-AUC superiores a 0.999, lo que confirma la estabilidad del enfoque basado en modelos ensamblados y su baja susceptibilidad al sobreajuste. En contraste, los modelos individuales, tal como se observa en la Tabla de Prueba, reproducen patrones similares a los identificados previamente: la LR y la NB alcanzan un recall de 1.000, pero con niveles de precisión menores, lo que afecta negativamente su F1-score; mientras que SVM y MLP presentan un equilibrio moderado entre sensibilidad y precisión, sin superar el rendimiento de los métodos ensamblados. En conjunto, la evidencia resumida en la Tabla de Prueba refuerza la idea de que, en un escenario sin balanceo de clases, los modelos de aprendizaje ensamblado no solo dominan en la validación, sino que mantienen su superioridad en la fase de prueba, consolidándose como la alternativa más robusta para la predicción de la deserción universitaria en este primer experimento.

A partir del diagnóstico automático de Train–Validación (tabla 19), se observa que el análisis de la brecha F1 (Train–Val) resulta clave para evaluar la capacidad de generalización de los modelos. Los resultados evidencian que varios modelos ensamblados basados en árboles (ET, DT, GB y RF) presentan un probable sobreajuste, reflejado en brechas elevadas entre entrenamiento y validación, pese a su alto desempeño aparente. En contraste, LR y SVM muestran brechas reducidas y un comportamiento estable, siendo estos los modelos con la mejor capacidad de generalización en este escenario sin balanceo. El modelo HGB destaca por su alto desempeño en la validación, aunque requiere un ajuste intermedio que justifica realizar

evaluaciones complementarias. Finalmente, NB evidencia un subajuste, lo que confirma su menor idoneidad para el problema de predicción de la deserción universitaria.

Tabla 19

Diagnóstico automático de generalización (Train vs Validación) basado en F1-score

Modelo	F1 Train	F1 Validación	GAP (F1)	Diagnóstico
HGB	0.965251	0.873239	0.092012	Ajuste intermedio
ET	1.000000	0.849315	0.150685	Probable sobreajuste
DT	1.000000	0.845070	0.154930	Probable sobreajuste
GB	0.957854	0.837838	0.120017	Probable sobreajuste
RF	1.000000	0.833333	0.166667	Probable sobreajuste
LR	0.810127	0.831169	-0.021042	Buen ajuste
SVM	0.853333	0.820513	0.032821	Buen ajuste
MLP	0.972973	0.811594	0.161379	Probable sobreajuste
NB	0.517312	0.592593	-0.075281	Subajuste

Experimento 2: predicción + UNDERSAMPLE

Para analizar cómo afecta el desequilibrio de clases a la capacidad predictiva y estudiar la estabilidad de los algoritmos con una distribución más balanceada, se aplicaron modelos de predicción mediante Undersampling, utilizando un método de submuestreo aleatorio de la clase mayoritaria antes del entrenamiento. La tabla 20 permite analizar no solo el rendimiento absoluto en métricas como F1-score y ROC-AUC, sino también la variación de la sensibilidad, la precisión y la brecha de generalización (Train vs Validation), aportando evidencia empírica sobre qué arquitecturas son más resilientes ante la reducción de la información de la clase mayoritaria.

Tabla 20

Resultados de validación de modelos predictivos (X_val) – Experimento 2 (undersampling)

Grupo	Modelo	Accuracy	Recall	Precision	F1	ROC-AUC	Especificidad
Ensamblado	RF	0.989091	1.0000	0.516129	0.680851	0.998229	0.988962

Individual	NB	0.984000	1.0000	0.421053	0.592593	0.997792	0.983812
Ensamblado	ET	0.982182	1.0000	0.395062	0.566372	0.998356	0.981972
Ensamblado	GB	0.978182	1.0000	0.347826	0.516129	0.998962	0.977925
Individual	DT	0.978182	1.0000	0.347826	0.516129	0.989862	0.979725
Ensamblado	HGB	0.977091	1.0000	0.336842	0.503937	0.998350	0.976821
Individual	LR	0.972727	1.0000	0.299065	0.460432	0.995125	0.972406
Individual	MLP	0.971636	1.0000	0.290909	0.450704	0.995332	0.971302
Individual	SVM	0.970909	1.0000	0.285714	0.444444	0.995792	0.970567

En primer lugar, se observa que todos los modelos alcanzan un recall de 1, lo que indica que el conjunto de alumnos desertores fue reconocido con precisión en el conjunto de validación. Este resultado corrobora que el submuestreo de la clase mayoritaria mejora directamente la sensibilidad del modelo al priorizar la identificación de casos positivos (deserción) y al reducir al máximo los falsos negativos. Esto es especialmente importante en contextos institucionales donde se busca detectar a tiempo a los alumnos en riesgo.

Sin embargo, este aumento constante del recall se traduce en una disminución significativa de la precisión, lo que indica un incremento de los falsos positivos. En este contexto, el modelo RF (ensamblado) tiene la mejor precisión general, con un F1-score de 0.680851 y una exactitud de 0.516129, lo que le permite superar a todos los demás algoritmos, tanto ensamblados como individuales. También mantiene un ROC-AUC de 0.998229, lo que indica una capacidad discriminativa sobresaliente, incluso con una distribución artificialmente equilibrada. Esta conducta indica que los procedimientos basados en el bagging son más robustos frente a la pérdida de información asociada al undersampling.

Por otro lado, Extra Trees y GB también presentan valores de ROC-AUC muy altos (más allá de 0.998), pero con F1-scores más bajos (0.566372 y 0.516129, respectivamente), lo que indica una relación deficiente entre la precisión y el recall. En el caso de los modelos individuales, NB

tiene un rendimiento relativamente competitivo ($F1 = 0.592593$), pero la LR, la MLP y la SVM muestran una disminución continua del F1-score, que oscila entre 0.44 y 0.46; esto indica que son estructuralmente menos estables cuando la clase mayoritaria se reduce.

Tabla 21

Resultados en el conjunto de prueba (X_{test})– Experimento 2 (undersampling)

Modelo	Accuracy	Recall	Precision	F1	ROC-AUC	Especificidad
RF	0.992726	1.0000	0.615385	0.761905	0.998867	0.992641
HGB	0.983416	1.0000	0.412371	0.583942	0.999110	0.983220
DT	0.980506	1.0000	0.373832	0.544218	0.990138	0.980277
GB	0.980506	1.0000	0.373832	0.544218	0.990138	0.980277
NB	0.977888	1.0000	0.344828	0.512821	0.998528	0.977627
ET	0.976433	1.0000	0.330579	0.496894	0.998727	0.976155
LR	0.974687	1.0000	0.314961	0.479042	0.995143	0.974389
MLP	0.972360	1.0000	0.296296	0.457143	0.996136	0.972034
SVM	0.971778	1.0000	0.291971	0.451977	0.996048	0.971445

A continuación, se presentan los resultados obtenidos en el conjunto de prueba (X_{test}) bajo la estrategia de Undersampling, para evaluar la capacidad real de generalización de los modelos sobre datos no utilizados durante el entrenamiento ni la validación. Los resultados se observan en la tabla 21, donde todos los modelos presentan un recall de 1.000, lo que confirma que la identificación total de los estudiantes desertores se asegura mediante el balanceo por submuestreo, tal como sucede en la validación. Este descubrimiento es estructural y consistente, lo que indica que el modelo otorga sistemáticamente prioridad a la sensibilidad. No obstante, este beneficio sigue asociado a una disminución de la precisión, lo que indica un aumento de los falsos positivos.

En comparación con otros modelos, el más balanceado es RF, que obtiene un F1-score de 0.761905, el más alto entre todos los resultados del conjunto de prueba. Además, presenta un ROC-AUC de 0.998867 y una precisión de 0.615385. Esta conducta demuestra una mayor estabilidad ante la disminución de los datos de la clase mayoritaria y una combinación adecuada de sensibilidad y control de errores.

HGB, con un rendimiento competitivo ($F1 = 0.583942$; $ROC-AUC = 0.999110$), presenta una precisión inferior a la de RF. Los modelos basados en GB y en árboles individuales presentan valores intermedios de F1 (≈ 0.54), mientras que los modelos lineales individuales y de margen (SVM, MLP y LR) presentan puntuaciones de F1 más bajas (entre 0.45 y 0.48). Esto indica una menor resistencia estructural frente al undersampling. Los resultados de la prueba corroboran el patrón detectado en la validación: aunque el undersampling optimiza la sensibilidad del sistema predictivo, disminuye la precisión. Los modelos ensamblados, sobre todo los de RF, son los que, en general, alcanzan el equilibrio óptimo.

También, para analizar el comportamiento de generalización de los modelos evaluados, se realizó un diagnóstico automático comparando el desempeño obtenido en los conjuntos de entrenamiento y de validación, utilizando como métrica de referencia el F1-score, como se observa en la tabla 22.

Tabla 22

Diagnóstico de generalización (train vs validación) con F1-score – Experimento 2 (undersampling)

MODELO	F1 TRAIN	F1 VAL	GAP(F1)	DIAGNÓSTICO
RF	0.713092	0.680851	0.032241	Ajuste intermedio (revisar la curva y las métricas).
NB	0.516260	0.592593	-0.076332	Probable UNDERFITTING (bajo desempeño general).
ET	0.497087	0.566372	-0.069284	Probable UNDERFITTING (bajo desempeño general).

GB	0.526749	0.516129	0.010620	Probable UNDERFITTING (bajo desempeño general).
DT	0.526749	0.516129	0.010620	Probable UNDERFITTING (bajo desempeño general).
HGB	0.542373	0.503937	0.038436	Probable UNDERFITTING (bajo desempeño general).
LR	0.474954	0.460432	0.014522	Probable UNDERFITTING (bajo desempeño general).
MLP	0.441379	0.450704	-0.009325	Probable UNDERFITTING (bajo desempeño general).
SVM	0.440620	0.444444	-0.003825	Probable UNDERFITTING (bajo desempeño general).

Los resultados evidencian que la mayoría de los modelos presentan indicios de subajuste (underfitting), lo que sugiere un desempeño moderado en ambas fases de evaluación. En particular, el modelo Random Forest mostró el mejor comportamiento relativo, alcanzando un F1-score de 0.713 en entrenamiento y de 0.681 en validación, con un GAP de 0.032, lo que indica un ajuste intermedio y una capacidad de generalización razonable. Por su parte, modelos como NB y ET presentaron diferencias negativas entre entrenamiento y validación (-0.076 y -0.069 , respectivamente), lo que refleja un rendimiento limitado y un posible subajuste de los modelos. De manera similar, algoritmos como GB, DT, HGB, LR, MLP y SVM mostraron valores de F1 relativamente bajos y diferencias reducidas entre entrenamiento y validación, lo que refuerza la interpretación de un underfitting y sugiere que estos modelos no logran capturar completamente la complejidad de los patrones presentes en los datos.

Experimento 3: predicción + OVERSAMPLING

De igual forma, se analizaron modelos de predicción mediante Oversampling con el propósito de evaluar su solidez ante un aumento sintético de la clase minoritaria y de analizar las consecuencias de una estrategia alternativa para balancear clases. En este contexto, se empleó una técnica de sobremuestreo para aumentar la representación de los alumnos desertores en el conjunto de entrenamiento. Esto se hizo para reducir el sesgo hacia la clase mayoritaria y mejorar la capacidad del modelo para detectar situaciones de riesgo. Esta estrategia, a diferencia del Undersampling, preserva la información original de la clase mayoritaria. Esto permite

analizar cómo el aumento de la clase minoritaria afecta métricas clave como Recall, Precision, F1-score y ROC-AUC.

Los resultados obtenidos se muestran en la tabla 23, la cual evidencia que la aplicación de la estrategia de Oversampling permitió mejorar sustancialmente la capacidad de los modelos para identificar correctamente la clase minoritaria (desertores), manteniendo simultáneamente niveles elevados de desempeño global. En términos generales, los modelos ensamblados continúan mostrando un comportamiento superior al de los modelos individuales, lo que confirma la robustez de los métodos basados en la agregación de múltiples estimadores en escenarios de desbalance.

Tabla 23

Resultados de validación de modelos predictivos (X_val) – Experimento 3 (oversampling)

Grupo	Modelo	Accuracy	Recall	Precision	F1	ROC-AUC	Especificidad
Ensamblado	GB	0.996000	0.96875	0.756098	0.849315	0.998977	0.996321
Ensamblado	ET	0.996000	0.96875	0.756098	0.849315	0.998885	0.996321
Ensamblado	HGB	0.995636	0.96875	0.738095	0.837838	0.999034	0.995953
Ensamblado	RF	0.995636	0.96875	0.738095	0.837838	0.998724	0.995953
Individual	DT	0.995636	0.93750	0.750000	0.833333	0.996910	0.996321
Individual	LR	0.995273	1.00000	0.711111	0.831169	0.998885	0.995217
Individual	SVM	0.995273	1.00000	0.711111	0.831169	0.997769	0.995217
Individual	MLP	0.995636	0.90625	0.763158	0.828571	0.997229	0.996689
Individual	NB	0.984000	1.00000	0.421053	0.592593	0.998655	0.983812

Específicamente, Extra Trees y GB logran el mejor desempeño global, con valores de exactitud de 0.9960 y de recall cercanos a 0.969, lo que evidencia una alta capacidad para identificar a alumnos en riesgo de abandono escolar. El F1-score (0.8493) muestra una buena proporción entre precisión y recall, lo que indica que el incremento sintético de la clase minoritaria mejoró la sensibilidad del modelo sin perjudicar significativamente la precisión. Asimismo, si el valor de ROC-AUC es mayor que 0.998, puede observarse una gran capacidad discriminativa entre ambas clases.

A pesar de que el modelo HGB mantiene un rendimiento competitivo, ha experimentado una ligera disminución de la precisión (0.7381), lo que indica que la clase minoritaria podría estar sujeta a un ligero incremento de falsos positivos debido a un aumento artificial. No obstante, el alto Recall (0.9687) confirma que el modelo prioriza detectar a los alumnos en riesgo de desertar, lo cual concuerda con la meta estratégica de reducir los falsos negativos en un asunto relacionado con la DEU.

Por su parte, RF presenta un comportamiento estable, con métricas similares a las de los modelos de boosting, lo que refuerza la evidencia de que los métodos de ensamble son menos sensibles a la variabilidad introducida por el sobremuestreo sintético. En contraste, los modelos individuales muestran una mayor variabilidad. El DT mejora considerablemente su Recall (0.9375), lo que sugiere que el Oversampling beneficia especialmente a los modelos con menor capacidad estructural. Sin embargo, su F1-score (0.8333) sigue siendo inferior al de los modelos ensamblados, lo que confirma su menor estabilidad predictiva.

En cuanto a RF, muestra un comportamiento constante y presenta métricas similares a las de los modelos de boosting, lo que corrobora que los métodos de ensamble son menos susceptibles a la variabilidad generada por el sobremuestreo sintético. Los modelos individuales, en cambio, presentan una mayor variabilidad. El DT mejora notablemente su Recall (0.9375), lo que indica que el oversampling favorece especialmente a los modelos con menor capacidad estructural. No obstante, su puntuación F1 (0.8333) sigue por debajo de la de los modelos ensamblados, lo que confirma una menor estabilidad predictiva. El SVM y la LR alcanzan un Recall de 1, lo que significa que, en la validación, clasifican correctamente todos los casos positivos. Sin embargo, esta gran sensibilidad se acompaña de una disminución de la precisión (0.7111), lo que conlleva un aumento de los falsos positivos. Esta conducta indica que el sobremuestreo puede generar una frontera de decisión más amplia para la clase minoritaria en modelos lineales,

lo que puede afectar el equilibrio entre especificidad y sensibilidad. Por último, NB tiene el rendimiento más bajo, con una precisión (0.4211) y un F1-score (0.5926) notablemente menores, lo que demuestra que este modelo probabilístico no se ajusta adecuadamente a la distribución sintética generada por el Oversampling.

En síntesis, los resultados confirman que la estrategia de sobremuestreo fortalece la capacidad de detección de la clase minoritaria sin comprometer la capacidad discriminativa global del sistema. Sin embargo, también se observa que el impacto del Oversampling depende del tipo de algoritmo: los modelos de ensamble presentan mayor estabilidad y un mejor equilibrio entre métricas, mientras que los modelos lineales tienden a maximizar el Recall a costa de la Precision. Este hallazgo es relevante para la construcción del modelo explicable final, ya que evidencia la necesidad de seleccionar un algoritmo que no solo maximice el desempeño predictivo, sino que también mantenga la estabilidad estructural frente a técnicas de balanceo sintético.

También se realizó la prueba en el conjunto de datos X_{test} , utilizando métricas de evaluación para analizar el rendimiento. Tal como se muestra en la tabla 24, se compararon modelos individuales y ensamblados, considerando métricas como accuracy, recall, precision, F1-score, ROC-AUC y especificidad, así como el tiempo de entrenamiento de cada algoritmo

Tabla 24

Resultados en el conjunto de prueba (X_{test})– Experimento 3 (oversampling)

MODELO	ACCURACY	RECALL	PRECISION	F1	ROC_AUC	ESPECIFICIDAD
HGB	0.997963	0.975	0.866667	0.917647	0.999764	0.998234
DT	0.997672	0.875	0.921053	0.897436	0.937058	0.999117
ET	0.997381	0.900	0.878049	0.888889	0.999459	0.998528
MLP	0.997381	0.875	0.897436	0.886076	0.999029	0.998822
RF	0.997090	0.925	0.840909	0.880952	0.999647	0.997939
GB	0.996509	0.900	0.818182	0.857143	0.996254	0.997645

LR	0.995636	1.000	0.727273	0.842105	0.999058	0.995584
SVM	0.995636	0.975	0.735849	0.838710	0.986363	0.995879
NB	0.977888	1.000	0.344828	0.512821	0.998900	0.977627

Los resultados evidencian que los modelos presentan un desempeño elevado en términos generales, destacándose especialmente los métodos basados en ensambles de árboles. En particular, el modelo HGB obtuvo el mejor rendimiento global, con una precisión de 0.997963, un recall de 0.975, un F1-score de 0.917647 y un AUC-ROC de 0.999764, lo que indica una excelente capacidad de discriminación entre clases. De manera similar, RF mostró un desempeño competitivo, con un F1-score de 0.880952 y un recall de 0.925, lo que evidencia una alta capacidad para identificar correctamente los casos positivos. Por su parte, modelos como ET, GB y MLP también alcanzaron valores elevados de accuracy y ROC-AUC, aunque con ligeras variaciones en la precisión y el recall. En contraste, el modelo NB presentó el desempeño más bajo en términos de F1-score (0.512821) debido a su baja precisión, lo que sugiere limitaciones para manejar la distribución de las clases en el conjunto de datos. En conjunto, estos resultados indican que los modelos de ensamble basados en árboles ofrecen el mejor equilibrio entre precisión y capacidad predictiva para el problema estudiado.

A continuación, en la tabla 25 se muestra el diagnóstico (Train vs Validation) utilizando la métrica F1. La finalidad de esto es analizar el poder de generalización de los modelos y detectar posibles sobreajustes o subajustes en el marco del experimento de Oversampling.

Tabla 25

Diagnóstico de generalización (train vs validación) con F1-score – Experimento 3 (oversampling)

Modelo	F1_Train	F1_Val	GAP(F1)	Diagnóstico
ET	1.000000	0.849315	0.150685	Probable OVERFITTING (brecha Train-Val alta)
GB	0.951673	0.849315	0.102358	Ajuste intermedio (revisar curva y métricas)

HGB	0.984615	0.837838	0.146778	Probable OVERFITTING (brecha Train–Val alta)
RF	1.000000	0.837838	0.162162	Probable OVERFITTING (brecha Train–Val alta)
DT	1.000000	0.833333	0.166667	Probable OVERFITTING (brecha Train–Val alta)
LR	0.820513	0.831169	-0.010656	Buen ajuste (generaliza bien)
SVM	0.823151	0.831169	-0.008018	Buen ajuste (generaliza bien)
MLP	0.988417	0.828571	0.159846	Probable OVERFITTING (brecha Train–Val alta)
SVM	0.517312	0.592593	-0.075281	Probable UNDERFITTING (bajo desempeño general)

El análisis comparativo del rendimiento en entrenamiento (F1_train) y en validación (F1_val) evidencia diferencias significativas en la capacidad de generalización de los algoritmos analizados. Primero, los modelos MLP, DT, RF y ET tienen valores de F1_train cercanos o iguales a 1.000, pero difieren notablemente respecto al F1_val (en varios casos, las brechas superan 0.15). Esta conducta indica un probable sobreajuste (overfitting), en el que el modelo capta con exactitud las particularidades del conjunto de entrenamiento, incluidos patrones que podrían haber sido introducidos por el sobremuestreo sintético, pero su capacidad de generalización a datos no observados se ve disminuida. El aumento artificial de la clase minoritaria puede intensificar la varianza en modelos basados en árboles profundos o en estructuras muy flexibles, como lo evidencia la magnitud de la brecha entre Train y Val.

En el caso de GB, a pesar de que se observa una diferencia positiva entre la validación y el entrenamiento ($GAP \approx 0.10$), esta es leve en comparación con otros ensambles, lo que indica un ajuste medio. Esta conducta evidencia un equilibrio más marcado entre la capacidad de predicción y la regularización implícita del modelo, lo cual concuerda con el carácter secuencial y regularizado del boosting. Por otro lado, los modelos lineales (como SVM y LR) exhiben brechas negativas mínimas (GAP cercanas a 0), lo que indica una consistencia significativa entre la validación y el entrenamiento. Este patrón muestra una adecuación y una generalización apropiadas, dado que el desempeño se mantiene constante en ambos conjuntos. La

regularización propia de estos modelos ayuda a prevenir el aprendizaje excesivo de patrones sintéticos generados por el sobremuestreo.

Por el contrario, NB presenta un F1_train significativamente menor que el F1_val (con un GAP negativo significativo) y, además, un rendimiento general bajo. Este comportamiento indica un probable subajuste (underfitting), lo que sugiere que el modelo carece de la capacidad estructural necesaria para abordar la complejidad del problema de la deserción universitaria, incluso con una estrategia de balanceo sintético.

Experimento 4: predicción + SMOTENC

Por último, se realizó el experimento incorporando la técnica SMOTENC para abordar el desbalance de clases, dado el carácter mixto del conjunto de datos (variables numéricas y categóricas). Los resultados se describen en la tabla 26. A diferencia del sobremuestreo tradicional, esta técnica permite generar instancias sintéticas que preservan la estructura de las variables nominales, evitando distorsiones en atributos institucionales relevantes para el análisis de la DEU. El propósito de esta fase es evaluar si un balanceo más coherente estructuralmente mejora la capacidad de detección de estudiantes en riesgo (Recall y F1-score) sin incrementar el sobreajuste observado en experimentos previos.

Tabla 26

Resultados de validación de modelos predictivos (X_val) – Experimento 4 (SMOTENC)

Grupo	Modelo	Accuracy	Recall	Precision	F1	ROC-AUC	Especificidad
Ensamblado	HGB	0.996727	0.96875	0.794872	0.873239	0.999069	0.997057
Ensamblado	ET	0.996000	0.96875	0.756098	0.849315	0.998816	0.996321
Individual	DT	0.996000	0.93750	0.769231	0.845070	0.967094	0.996689
Ensamblado	GB	0.995636	0.96875	0.738095	0.837838	0.999028	0.995953
Ensamblado	RF	0.995636	0.93750	0.750000	0.833333	0.998799	0.996321
Individual	LR	0.995273	1.00000	0.711111	0.831169	0.998873	0.995217
Individual	SVM	0.994909	1.00000	0.695652	0.820513	0.997896	0.994849

Individual	MLP	0.995273	0.87500	0.756757	0.811594	0.998574	0.996689
Individual	NB	0.984000	1.00000	0.421053	0.592593	0.998436	0.983812

En la mayoría de los modelos evaluados, los resultados obtenidos mediante el balanceo con SMOTENC muestran un rendimiento muy competitivo, con una precisión superior al 0,99 en la mayoría de los casos. Este comportamiento confirma que la técnica de sobremuestreo condicionada por el tipo de variable permite mantener la estabilidad global del sistema sin comprometer la capacidad discriminativa entre clases. El modelo HGB sobresale por su rendimiento general, superior al de los demás algoritmos estudiados, con un F1-score de 0.873239 y una precisión (Accuracy) de 0.996727. La capacidad para detectar con precisión a los alumnos en riesgo de deserción es alta, como lo demuestran su Recall de 0.96875 y su Precision de 0.794872, que evidencian un equilibrio positivo entre la sensibilidad y la capacidad de control de falsos positivos. Además, el ROC-AUC de 0.999069 indica una capacidad discriminativa casi perfecta, lo que lo convierte en el modelo más sólido en esta configuración.

Los modelos de ensamble ET, GB y RF mantienen un comportamiento sólido, con valores de Recall cercanos a 0.97 en los métodos de boosting y ligeramente menores en el de RF (0.93750). Pese a que sus valores de precisión son ligeramente inferiores a los de HGB, los F1-scores superiores a 0.83 demuestran un desempeño equilibrado y constante. Estos hallazgos indican que los métodos basados en la agregación siguen mostrando mayor estabilidad frente al desequilibrio estructural del problema.

El DT, que forma parte del conjunto de modelos individuales, muestra un rendimiento competitivo (F1-score: 0.845070 y precisión: 0.769231). Esto sugiere que la técnica SMOTENC beneficia incluso a los modelos menos complejos. Por otro lado, SVM y LR alcanzan un recall de 1, lo que significa que todos los casos positivos en la validación son detectados. Sin embargo, esta sensibilidad máxima se acompaña de valores de precisión

inferiores (0.711111 y 0.695652, respectivamente), lo que evidencia un incremento de falsos positivos y, por tanto, un equilibrio menos favorable en términos de F1.

El modelo MLP presenta un comportamiento estable, aunque con un Recall ligeramente inferior (0.87500), lo que afecta su F1-score (0.811594). En última instancia, NB presenta el peor rendimiento: un F1-score de 0.592593 y una precisión bastante baja (0.421053), lo que evidencia su escasa capacidad para modelar adecuadamente la complejidad estructural del fenómeno de deserción universitaria.

Con el propósito de evaluar la capacidad real de generalización del modelo bajo la estrategia de balanceo mediante SMOTENC, a continuación, se presentan los resultados obtenidos en el conjunto de prueba en la tabla 27. Esta fase es fundamental dentro del protocolo metodológico, ya que permite analizar el desempeño predictivo en datos completamente no utilizados durante el entrenamiento ni la validación, garantizando así una estimación objetiva del comportamiento del sistema en escenarios reales

Tabla 27

Resultados en el conjunto de prueba (X_test)– Experimento 4 (SMOTENC)

Grupo	Modelo	Accuracy	Recall	Precision	F1	ROC-AUC	Especificidad
Ensamblado	HGB	0.997672	0.925	0.880952	0.902439	0.999584	0.998528
Ensamblado	GB	0.997381	0.900	0.878049	0.888889	0.999669	0.998528
Individual	DT	0.997381	0.900	0.878049	0.888889	0.949264	0.998528
Ensamblado	RF	0.997381	0.900	0.878049	0.888889	0.999551	0.998528
Ensamblado	ET	0.997381	0.900	0.878049	0.888889	0.999481	0.998528
Individual	MLP	0.996800	0.850	0.871795	0.860759	0.999183	0.998528
Individual	SVM	0.995345	0.925	0.740000	0.822222	0.998719	0.996173
Individual	LR	0.994763	1.000	0.689655	0.816327	0.998948	0.994701
Individual	NB	0.977888	1.000	0.344828	0.512821	0.998800	0.977627

Con excepción de NB, todos los modelos mantienen valores de Accuracy superiores a 0.99, lo que evidencia una alta estabilidad predictiva fuera de los conjuntos de entrenamiento y validación. El modelo HGB confirma su liderazgo observado en la validación, alcanzando un Accuracy de 0.997672 y el mayor F1-score (0.902439) entre todos los algoritmos evaluados. Su Recall de 0.925 y su Precision de 0.880952 reflejan un equilibrio sólido entre la detección de estudiantes en riesgo y el control de los falsos positivos. Asimismo, el ROC-AUC de 0.999584 evidencia una capacidad discriminativa prácticamente óptima, lo que consolida su robustez en condiciones reales de prueba.

Los modelos de ensamble GB, RF y Extra Trees presentan un comportamiento muy similar entre sí, con valores de F1-score de 0.888889 y de recall de 0.900. Esta consistencia sugiere que la estrategia SMOTENC permitió mantener la estabilidad estructural en los métodos basados en árboles, reduciendo las brechas observadas en fases anteriores y favoreciendo una generalización homogénea.

En el grupo de modelos individuales, el DT alcanza métricas comparables a las de los ensambles ($F1 = 0.888889$), lo que indica que el balanceo condicionado por el tipo de variable también beneficia a los modelos menos complejos. El MLP mantiene un desempeño competitivo ($F1 = 0.860759$), aunque con un recall ligeramente menor (0.850), lo que sugiere una leve reducción de la sensibilidad frente a los métodos de boosting. Por su parte, SVM y LR presentan comportamientos distintos. El SVM mantiene un equilibrio razonable ($F1 = 0.822222$), mientras que la LR alcanza un Recall perfecto (1.000), aunque con una precisión inferior (0.689655), lo que indica un mayor número de falsos positivos. Este patrón confirma que los modelos lineales tienden a priorizar la sensibilidad en escenarios sintéticamente balanceados. Finalmente, NB presenta el desempeño más bajo ($F1 = 0.512821$), con una precisión

considerablemente reducida (0.344828), lo que evidencia limitaciones estructurales para modelar la complejidad del fenómeno DEU, incluso con técnicas avanzadas de balanceo de clases.

Finalmente, se realizó un diagnóstico automático comparando el desempeño obtenido en los conjuntos de entrenamiento y de validación, utilizando como métrica de referencia el F1-score. Este análisis permite identificar posibles problemas de sobreajuste (overfitting) o de subajuste (underfitting) mediante la comparación del rendimiento del modelo durante el entrenamiento con el observado en la fase de validación. Como se observa en la tabla 28, se calculó el GAP(F1), definido como la diferencia entre el F1-score del conjunto de entrenamiento y el del conjunto de validación.

Tabla 28

Diagnóstico de generalización (train vs validación) con F1-score – Experimento 4 (SMOTENC)

,	F1_TRAIN	F1_VAL	GAP(F1)	DIAGNÓSTICO
HGB	0.965251	0.873239	0.092012	Ajuste intermedio (revisar la curva y las métricas).
ET	1.000000	0.849315	0.150685	Probable OVERFITTING (brecha Train-Val alta).
DT	1.000000	0.845070	0.154930	Probable OVERFITTING (brecha Train-Val alta).
GB	0.957854	0.837838	0.120017	Probable OVERFITTING (brecha Train-Val alta).
RF	1.000000	0.833333	0.166667	Probable OVERFITTING (brecha Train-Val alta).
LR	0.810127	0.831169	-0.021042	Buen ajuste (generaliza bien).
SVM	0.853333	0.820513	0.032821	Buen ajuste (generaliza bien).
MLP	0.972973	0.811594	0.161379	Probable OVERFITTING (brecha Train-Val alta).
NB	0.517312	0.592593	-0.075281	Probable UNDERFITTING (bajo desempeño general).

Los resultados evidencian que varios modelos presentan indicios de sobreajuste, en particular los basados en árboles de decisión y en métodos de ensamble. En este sentido, algoritmos como ET, DT, GB y RF alcanzan valores de F1-score cercanos o iguales a 1.0 en el conjunto de entrenamiento, mientras que en la validación disminuyen, lo que genera brechas GAP(F1) superiores a 0.12. Esta diferencia sugiere que los modelos logran un ajuste muy elevado a los

datos de entrenamiento, pero presentan una menor capacidad de generalización al aplicarse a datos no observados. De manera similar, el modelo MLP también muestra signos de sobreajuste, con un F1-score de 0.973 en entrenamiento y de 0.812 en validación, lo que genera un GAP(F1) de 0.161, uno de los valores más altos entre los algoritmos evaluados.

Por otra parte, algunos modelos presentan un comportamiento más equilibrado. HGB muestra un GAP (F1) de 0.092, lo que sugiere un ajuste intermedio, mientras que LR y SVM evidencian una mayor estabilidad entre entrenamiento y validación, lo que indica una mejor capacidad de generalización. En contraste, NB presenta el menor rendimiento global y una diferencia negativa entre entrenamiento y validación, lo que sugiere un posible subajuste.

Consolidado de los experimentos

Con el propósito de comparar el desempeño de los diferentes algoritmos evaluados en los cuatro escenarios, se elaboró una tabla consolidada del F1-score, presentada en la tabla 29, en la que se sintetizan los resultados obtenidos para cada modelo en los experimentos realizados: sin balanceo, con undersampling, con oversampling y con SMOTENC. Esta métrica fue seleccionada por su capacidad para integrar la precisión y el recall, lo cual resulta especialmente relevante en problemas de clasificación con clases desbalanceadas, como ocurre en el análisis de la DEU.

Tabla 29

Consolidado de modelos predictivos con F1-score

MODELO	sin balanceo	Undersampling	Oversampling	SMOTENC
HGB	0.902439	0.583942	0.917647	0.902439
DT	0.888889	0.544218	0.897436	0.888889
GB	0.888889	0.544218	0.857143	0.888889
ET	0.888889	0.496894	0.888889	0.888889
RF	0.888889	0.761905	0.880952	0.888889
MLP	0.860759	0.457143	0.886076	0.860759

LR	0.816327	0.479042	0.842105	0.816327
SVM	0.822222	0.451977	0.838710	0.822222
NB	0.512821	0.512821	0.512821	0.512821

Los resultados muestran que el mejor desempeño global se obtuvo con el modelo HGB con oversampling, que alcanzó un F1-score de 0.917647, lo que indica una elevada capacidad de clasificación equilibrada. De manera similar, DT también presenta un rendimiento destacado en el mismo escenario experimental, con un F1-score de 0.897436, lo que evidencia la capacidad de los modelos basados en árboles para capturar patrones relevantes en el conjunto de datos cuando se aplican técnicas de balanceo adecuadas. Asimismo, modelos de ensamble como RF, GB y ET muestran resultados competitivos, cercanos a 0.89 en el F1-score, lo que confirma la robustez de los métodos basados en múltiples árboles para este tipo de problemas.

En contraste, algoritmos como MLP, SVM y LR presentan desempeños moderados, con valores de F1-score inferiores a los obtenidos por los modelos de árboles, aunque mantienen una estabilidad aceptable entre los distintos experimentos. Por otro lado, el modelo NB presenta el rendimiento más bajo en todos los escenarios evaluados, con un F1-score de 0.512821, lo que sugiere limitaciones para captar la complejidad de las relaciones presentes en los datos.

Los resultados indican que el oversampling permitió obtener los mejores resultados globales, lo que sugiere que el balanceo mediante la generación de muestras adicionales contribuye a mejorar el equilibrio entre precisión y recall, optimizando el desempeño del modelo en términos de F1-score. Estos hallazgos refuerzan la importancia de aplicar estrategias adecuadas para tratar el desbalance de clases y mejorar la capacidad predictiva de los modelos de ML.

5.2.3 Optimización del modelo predictivo con oversampling

Como se observa en la tabla 30, la incorporación de la fase de optimización mediante el ajuste del umbral permitió mejorar de forma consistente el equilibrio predictivo de los modelos

previamente entrenados con oversampling. Este hallazgo es particularmente relevante, ya que en experimentos anteriores el sobremuestreo había mostrado un incremento significativo en la capacidad de detección de la clase minoritaria, aunque con ciertos indicios de sobreajuste en modelos de alta complejidad.

Tabla 30

Comparación F1_base vs F1_opt (Validación)

Grupo	Modelo	F1_base	F1_opt	THR_opt_val	Accuracy	Recall	Precision	Roc-Auc	Especificidad
Individual	SVM	0.837838	0.853333	0.30	0.996000	1.00000	0.744186	0.997769	0.995953
Ensamblado	HGBM	0.837838	0.853333	0.10	0.996000	1.00000	0.744186	0.999069	0.995953
Individual	LR	0.831169	0.849315	0.95	0.996000	0.96875	0.756098	0.998885	0.996321
Ensamblado	RF	0.842105	0.842105	0.50	0.995636	1.00000	0.727273	0.998793	0.995585
Ensamblado	ET	0.831169	0.842105	0.55	0.995636	1.00000	0.727273	0.998919	0.995585
Ensamblado	GBM	0.826667	0.842105	0.25	0.995636	1.00000	0.727273	0.999000	0.995585
Individual	MLP	0.805556	0.837838	0.10	0.995636	0.96875	0.738095	0.998609	0.995953
Individual	DT	0.831169	0.831169	0.05	0.995273	1.00000	0.711111	0.997609	0.995217
Individual	NB	0.592593	0.592593	0.05	0.984000	1.00000	0.421053	0.998655	0.983812

La comparación entre F1_base y F1_opt evidencia que ajustar el punto de corte probabilístico constituye un mecanismo eficaz para mejorar el desempeño del modelo sin alterar su arquitectura. En particular, SVM e HGB incrementan su F1-score de 0.837838 a 0.853333, lo que consolida una mejora cuantitativa relevante en el equilibrio entre Recall y Precision. Este comportamiento confirma que, cuando el Oversampling se combina con una optimización posterior del umbral, se logra mitigar parcialmente la pérdida de precisión derivada del aumento sintético de la clase minoritaria.

Asimismo, la LR muestra una mejora sustancial de 0.831169 a 0.849315, lo que demuestra que los modelos lineales también se benefician del ajuste fino del umbral en escenarios desbalanceados. En los modelos de ensamble, si bien las mejoras son moderadas, se observa

estabilidad estructural, particularmente en Extra Trees y GB, cuyos valores optimizados alcanzan 0.842105.

Los resultados confirman que la estrategia de Oversampling, complementada con una fase explícita de optimización del umbral, alcanza un desempeño superior al de la configuración base. No solo se incrementa el F1-score en la mayoría de los modelos, sino que también se mantiene la capacidad discriminativa ($\text{ROC-AUC} > 0.99$ en varios casos), lo que evidencia que el proceso de balanceo no deteriora la capacidad de separación entre clases.

En el contexto de la predicción de la DEU, la evaluación del desempeño debe basarse en el conjunto de prueba (X_{test}), ya que es el único conjunto completamente independiente de los procesos de entrenamiento y validación. A continuación, en la tabla 31 se contrastaron los resultados obtenidos en la validación del conjunto de prueba (X_{test}) tras aplicar el umbral óptimo aprendido.

Tabla 31

Resultados en Prueba (X_{test}) – F1_base vs F1_opt

Modelo	F1_base	F1_opt	THR_opt_val	Accuracy	Recall	Precision	ROC-AUC	Especificidad
ET	0.888889	0.888889	0.55	0.997090	1.000	0.800000	0.999492	0.997056
RF	0.888889	0.888889	0.50	0.997090	1.000	0.800000	0.999514	0.997056
LR	0.842105	0.886364	0.95	0.997090	0.975	0.812500	0.999058	0.997351
GB	0.886364	0.886364	0.25	0.997090	0.975	0.812500	0.999139	0.997351
MLP	0.878049	0.880952	0.10	0.997090	0.925	0.840909	0.999367	0.997939
HGB	0.873563	0.873563	0.10	0.996800	0.950	0.808511	0.999588	0.997351
DT	0.869565	0.869565	0.05	0.996509	1.000	0.769231	0.998234	0.996467
SVM	0.847059	0.847059	0.30	0.996218	0.900	0.800000	0.986363	0.997351
NB	0.512821	0.512821	0.05	0.977888	1.000	0.344828	0.998900	0.977627

Se observa que los modelos de ensamble continúan mostrando un desempeño superior y consistente, con valores de F1_opt cercanos o iguales a 0.888889 en los casos de ET y de RF.

Ambos modelos alcanzan un Recall de 1.000 y una Precision de 0.800000, lo que indica una detección completa de la clase minoritaria sin un deterioro significativo en la tasa de falsos positivos. Este comportamiento refleja una adecuada capacidad de generalización y confirma que el ajuste del umbral no introdujo inestabilidad en los datos no observados.

Por su parte, LR y GB presentan un F1_opt de 0.886364, junto con un Recall de 0.975 y una Precision de 0.812500. Estos resultados evidencian un equilibrio favorable entre sensibilidad y precisión, lo que confirma la efectividad del ajuste probabilístico aplicado durante la validación. En el caso de la LR, el umbral elevado (0.95) sugiere una configuración conservadora para la clasificación positiva, lo que contribuye a mantener bajo control los falsos positivos.

El modelo MLP alcanza un F1_opt de 0.880952 y una precisión de 0.840909, lo que constituye el valor más alto entre los modelos individuales. No obstante, su Recall de 0.925 indica una ligera reducción en la detección total de la clase minoritaria en los ensambles principales. Aun así, mantiene una capacidad discriminativa muy elevada ($\text{ROC-AUC} \approx 0.999367$), lo que confirma su estabilidad estructural.

En el caso de HGB, el F1_opt de 0.873563 refleja un desempeño competitivo, con valores robustos de ROC-AUC (0.999588) y de especificidad (0.997351). Aunque su F1 es ligeramente inferior a la de ET y RF, mantiene una sólida capacidad de discriminación global, lo que respalda su solidez como modelo candidato.

El DT conserva la estabilidad ($\text{F1_opt} = 0.869565$), mientras que el SVM presenta un desempeño ligeramente inferior ($\text{F1_opt} = 0.847059$) y una reducción moderada de la sensibilidad ($\text{Recall} = 0.900$). Finalmente, NB continúa mostrando limitaciones estructurales ($\text{F1_opt} = 0.512821$), lo que confirma que el ajuste del umbral no compensa la baja capacidad de modelado inherente a este enfoque probabilístico frente a la complejidad del fenómeno DEU.

Complementariamente, la evaluación del diagnóstico automático, basada en la métrica F1_opt, permite analizar con mayor precisión la capacidad de generalización estructural de los modelos tras la fase de optimización. Como se muestra en la tabla 32, la mayoría de los algoritmos presentan brechas reducidas entre entrenamiento y validación (GAP F1_opt), lo que evidencia una mejora sustancial respecto a los escenarios previos con umbral base.

Tabla 32

Análisis de Estabilidad Train–Validación con F1 Optimizado

Modelo	F1_opt_Train	F1_opt_Val	THR_opt_val	GAP (F1_opt)	Diagnóstico
HGB	0.895105	0.853333	0.10	0.041772	Buen ajuste (generaliza bien)
SVM	0.895105	0.853333	0.30	0.041772	Buen ajuste (generaliza bien)
LR	0.876325	0.849315	0.95	0.027010	Buen ajuste (generaliza bien)
GB	0.882759	0.842105	0.25	0.040653	Buen ajuste (generaliza bien)
ET	0.888889	0.842105	0.55	0.046784	Buen ajuste (generaliza bien)
RF	0.879725	0.842105	0.50	0.037620	Buen ajuste (generaliza bien)
MLP	0.901408	0.837838	0.10	0.063571	Ajuste intermedio (revisar regularización)
DT	0.853333	0.831169	0.05	0.022165	Buen ajuste (generaliza bien)
NB	0.517312	0.592593	0.05	0.075281	Underfitting probable

En particular, HGB y SVM muestran una brecha de apenas 0.041772, lo que los clasifica como modelos con un buen ajuste y una adecuada capacidad de generalización. Este comportamiento confirma que el ajuste del umbral contribuyó a estabilizar el desempeño sin inducir un sobreajuste significativo, lo que consolida la coherencia entre el entrenamiento y la validación. La LR presenta una de las brechas más pequeñas (0.027010), lo que refuerza su estabilidad estructural y su comportamiento consistente en la optimización probabilística. De manera similar, los modelos de ensamble GB, ET y RF mantienen brechas inferiores a 0.05, lo que evidencia que la combinación de oversampling y ajuste de umbral logró controlar la varianza observada en fases anteriores. El MLP, aunque clasificado como de ajuste intermedio (GAP = 0.063571), muestra una brecha considerablemente menor que la registrada antes de la

optimización, lo que sugiere una mejora en la estabilidad general, aunque aún con margen para regularización adicional. El DT presenta la brecha más baja entre los modelos basados en árboles (0.022165), lo que confirma un comportamiento equilibrado tras la optimización. Por el contrario, NB presenta un patrón de underfitting ($GAP = 0.075281$), con valores de F1 bajos tanto en entrenamiento como en validación, lo que evidencia limitaciones estructurales independientes del ajuste del umbral.

En conjunto, este diagnóstico confirma que la fase de optimización no solo mejoró el F1 tanto en validación como en prueba, sino que también redujo sustancialmente las brechas Train–Val, lo que fortaleció la estabilidad del pipeline propuesto. Desde una perspectiva metodológica, la evidencia sugiere que los modelos de ensamble, particularmente ET y RF en prueba, junto con HGB por su consistencia global, presentan el mejor equilibrio entre desempeño predictivo y capacidad de generalización y constituyen candidatos sólidos para la fase posterior de explicabilidad (XAI).

5.2.4 Optimización mediante hiperparámetros

Con el objetivo de fortalecer la capacidad predictiva y la estabilidad estructural de los modelos, como se observa en la tabla 33, se incorpora en esta fase un proceso de optimización de hiperparámetros. A diferencia de los parámetros internos aprendidos durante el entrenamiento, los hiperparámetros controlan la complejidad, la regularización y el comportamiento estructural del algoritmo, influyendo directamente en el equilibrio entre sesgo y varianza. Los hiperparámetros de la tabla 30 corresponden a la configuración óptima obtenida mediante validación cruzada estratificada, centrada en los 3 modelos con el mejor F1-score. Esta estrategia permitió ajustar la complejidad estructural de cada modelo, reducir el riesgo de sobreajuste observado en las fases preliminares y mejorar la capacidad de generalización en el conjunto de prueba.

Tabla 33

Hiperparámetros Óptimos de los Modelos con Mejor Desempeño Predictivo

Modelo	Hiperparámetro	Valor
ET	n_estimators	1000
	max_depth	10
	min_samples_split	2
	min_samples_leaf	2
	max_features	$\sqrt{p}$
	bootstrap	No
RF	n_estimators	800
	max_depth	10
	min_samples_split	2
	min_samples_leaf	2
	max_features	$\sqrt{p}$
	bootstrap	Sí
	max_samples	0.85
LR	C	0.5
	solver	lbfgs
	penalty	L2
	max_iter	8000

Los resultados presentados en la Tabla 34 evidencian el impacto positivo de la optimización de hiperparámetros, combinada con el ajuste de umbral, sobre el equilibrio predictivo de los modelos seleccionados. En términos generales, se observa una mejora consistente del F1-score respecto a la configuración base, lo que confirma que la calibración estructural de los modelos contribuyó a reducir el sesgo–varianza y a estabilizar el desempeño en la validación.

Tabla 34

Resultados en Validación tras Optimización de Hiperparámetros

Modelo	F1_base	F1_opt	THR_opt_val	Accuracy	Recall	Precision	ROC-AUC	Especificidad
LR	0.831169	0.864865	0.91	0.996364	1.000000	0.761905	0.998885	0.996321
ET	0.837838	0.857143	0.91	0.996727	0.843750	0.870968	0.998931	0.998528
RF	0.837838	0.842105	0.39	0.995636	1.000000	0.727273	0.998793	0.995585

En el caso de la LR, el F1-score aumenta de 0.831169 a 0.864865 tras la optimización, lo que representa el mejor desempeño global en este grupo comparativo. Este resultado se explica por la combinación de una regularización moderada ($C = 0.5$, penalización L2) y un umbral óptimo elevado (0.91), lo que permitió mantener un Recall perfecto (1.000000) y mejorar simultáneamente la Precision (0.761905). La elevada Accuracy (0.996364) y el ROC-AUC cercano a 0.999 corroboran una alta capacidad discriminativa sin evidencia de deterioro estructural. Para ET, el F1-score aumenta de 0.837838 a 0.857143, lo que evidencia una mejora significativa tras ajustar los hiperparámetros ($n_estimators = 1000$, $max_depth = 10$) y aplicar un umbral óptimo de 0.91. A diferencia de la LR, este modelo presenta un equilibrio distinto entre sensibilidad y precisión (Recall = 0.84375 y Precision = 0.870968), lo que evidencia una estrategia más conservadora en la clasificación positiva. El alto valor de ROC-AUC (0.998931) confirma su estabilidad discriminativa. En el caso de RF, aunque el F1-score solo muestra una ligera variación ($0.837838 \rightarrow 0.842105$), el modelo mantiene un Recall de 1.000000 y una Precision de 0.727273 con un umbral óptimo de 0.39. Este comportamiento indica que el modelo prioriza por completo la detección de la clase minoritaria, aunque con un compromiso moderado con la precisión. No obstante, la estabilidad en ROC-AUC (0.998793) y en especificidad (0.995585) confirma que el ajuste estructural logró controlar la varianza sin afectar la capacidad global de discriminación.

Como se evidenció previamente en la fase de validación, el ajuste del umbral óptimo permitió maximizar el F1-score sin comprometer la estabilidad estructural de los modelos. Esta mejora se traslada de manera consistente al conjunto de prueba, lo que confirma que el umbral aprendido no responde a una adaptación específica al conjunto de validación, sino a una calibración robusta y generalizable.

Tabla 35

Resultados en el conjunto de prueba (X_test) con optimización de hiperparámetros

Modelo	F1_base	F1_opt	THR_opt_val	Accuracy	Recall	Precision	ROC-AUC	Especificidad
ET	0.883721	0.891892	0.91	0.997672	1.000	0.970588	0.999551	0.999706
RF	0.883721	0.886364	0.39	0.997090	0.975	0.812500	0.999647	0.997351
LR	0.816327	0.869565	0.91	0.996509	1.000	0.769231	0.998955	0.996467

Los resultados de la tabla 35 muestran que los tres algoritmos evaluados mejoran el F1-score tras el proceso de optimización de hiperparámetros y el ajuste del umbral de decisión. Entre ellos, el modelo que alcanza el mejor desempeño global es ET, cuyo F1-score aumenta de 0.883721 a 0.891892, lo que constituye el valor más alto entre los modelos analizados. Además, este modelo presenta un recall de 1, lo que indica que identifica correctamente la totalidad de los casos positivos, junto con una precisión de 0,970588, un accuracy de 0,997672, un ROC-AUC de 0,999551 y una especificidad de 0,999706, lo que evidencia una elevada capacidad discriminativa y un equilibrio adecuado entre las métricas de clasificación.

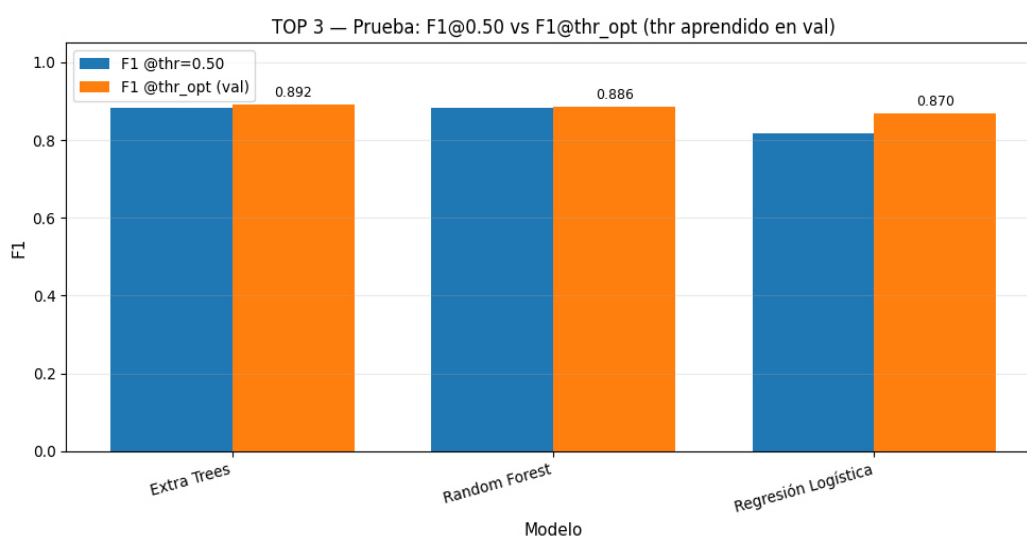
Por su parte, el RF también muestra una mejora tras la optimización del umbral, alcanzando un F1-score de 0.886364, un recall de 0.975 y un AUC-ROC de 0.999647, lo que refleja un desempeño competitivo, aunque ligeramente inferior al de ET. En el caso de la Regresión Logística (LR), se observa una mejora significativa del F1-score de 0.816327 a 0.869565, manteniendo un recall de 1.000, un accuracy de 0.996509 y un ROC-AUC de 0.998955, lo que confirma su capacidad para detectar correctamente los casos positivos. Los resultados

evidencian que la optimización del umbral de clasificación y de los hiperparámetros contribuye a mejorar el equilibrio entre precisión y recall, siendo Extra Trees el modelo que presenta el mejor desempeño global en el conjunto de prueba, por lo que se selecciona como modelo final para la predicción de la DEU.

Con el fin de facilitar la comparación visual del impacto de la optimización del umbral de decisión en el desempeño de los modelos, la Figura 19 presenta una representación gráfica que contrasta el F1-score calculado con el umbral estándar de 0.50 con el obtenido con el umbral óptimo aprendido en el conjunto de validación.

Figura 18

Comparación del F1-score en Prueba con Umbral Base y Umbral Optimizado



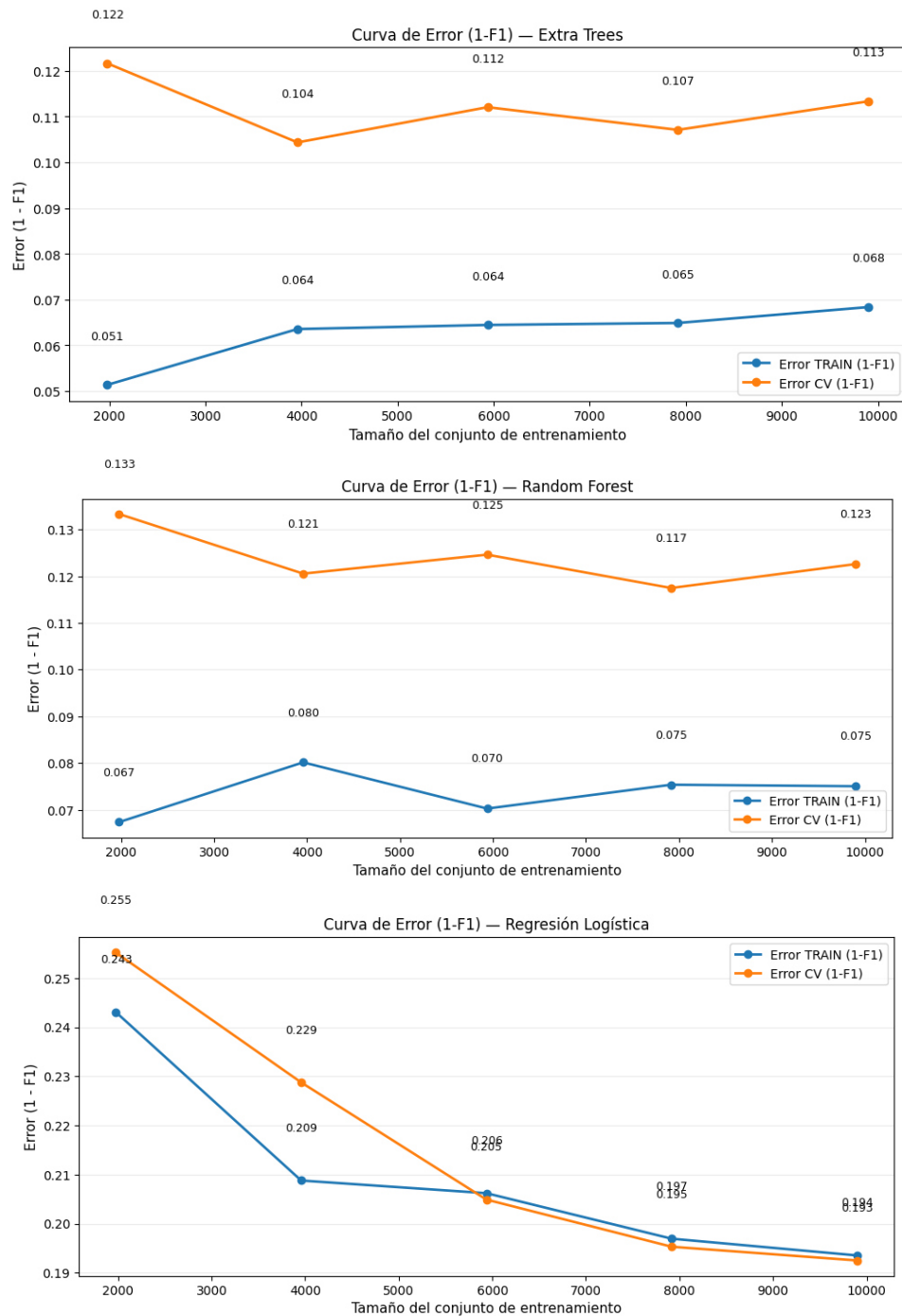
En general, ET muestra el mejor desempeño global, alcanzando el mayor F1-score en el conjunto de prueba (0.892) tras optimizar el umbral de decisión. Aunque la mejora respecto al umbral estándar es moderada, esta variación resulta consistente y estable, lo que sugiere que la estructura del modelo ya se encontraba cerca del punto de decisión óptimo. Este comportamiento confirma la robustez predictiva de ET, así como su capacidad para mantener un equilibrio adecuado entre precisión y recall incluso con datos no observados. Por su parte,

la LR muestra una mejora más pronunciada al pasar del umbral estándar (0.50) al optimizado, lo que incrementa su F1-score en el conjunto de prueba. Este resultado valida empíricamente la estrategia de calibrar el punto de corte probabilístico durante la fase de validación, ya que el beneficio observado se mantiene en el conjunto de prueba. En el caso de RF, la variación es mínima, lo que indica que el modelo ya operaba cerca de un punto de equilibrio adecuado bajo el umbral base y que la optimización del umbral no introduce inestabilidad ni degrada su desempeño.

Para complementar el análisis cuantitativo, se emplean las curvas de aprendizaje basadas en el error ($1-F1$) para los modelos ET, RF y LR. La figura 20 permite examinar la evolución del error de entrenamiento y de la validación cruzada a medida que aumenta el tamaño del conjunto de datos, aportando evidencia estructural sobre el comportamiento del sesgo y de la varianza. En los modelos de ensamble (ET y RF) se observa una brecha controlada y relativamente estable entre ambas curvas, lo que confirma un adecuado equilibrio entre ajuste y generalización, coherente con los altos valores de F1 obtenidos en validación y en prueba. Por su parte, la LR muestra una convergencia progresiva entre las curvas a medida que aumenta el tamaño de la muestra, lo que indica una reducción del error y una mayor estabilidad del modelo. Las gráficas refuerzan la interpretación de los resultados numéricos y consolidan la evidencia de que la optimización implementada produce modelos robustos, sin indicios significativos de sobreajuste y con una capacidad de generalización consistente.

Figura 19

Curvas de Aprendizaje (Error $1-F1$)



Con base en los resultados obtenidos en el conjunto de prueba y considerando el equilibrio entre Recall, Precision y F1-score, el ET se consolida como la alternativa más sólida. Si bien la LR presenta una mejora significativa tras la optimización del umbral, ET alcanza el F1 más alto, 0.892, manteniendo un Recall perfecto y una precisión adecuada, además de una elevada

capacidad discriminativa (ROC-AUC cercano a 1). Esta combinación de sensibilidad total, estabilidad estructural y equilibrio predictivo lo posiciona como el modelo con el mejor desempeño integral en la predicción de la DEU. A partir de esta selección, las técnicas de XAI, en particular SHAP y LIME, se aplican al modelo Extra Trees para interpretar la contribución de las variables a la predicción del abandono estudiantil. De esta manera, el análisis de explicabilidad se centra en el modelo con mejor desempeño predictivo, lo que permite comprender cómo las características del estudiante influyen en la probabilidad de deserción y facilita la generación de estrategias institucionales de intervención basadas en evidencia.

5.2.5. Hipótesis Específica 2

La segunda hipótesis indica que el entrenamiento y la evaluación de algoritmos de ML permiten obtener modelos de predicción capaces de identificar la DEU. Los resultados obtenidos en la fase de modelado predictivo respaldan empíricamente esta afirmación a partir del desempeño de diversos algoritmos de ML entrenados sobre el conjunto de datos previamente construido y depurado.

En primer lugar, se evaluó un amplio conjunto de modelos predictivos, incluidos algoritmos individuales y métodos ensamblados, con el objetivo de analizar su capacidad para identificar a estudiantes en riesgo de abandono. Este proceso permitió comparar distintos enfoques de modelado y seleccionar aquellos con mayor capacidad de discriminación respecto de la variable objetivo de deserción. Asimismo, se aplicaron estrategias de balanceo de clases, como UNDERSAMPLING, OVERSAMPLING y SMOTENC, con el propósito de abordar el desbalance entre estudiantes desertores y no desertores, preservando la naturaleza mixta de las variables numéricas y categóricas del dataset.

Los resultados obtenidos indicaron que los algoritmos ensamblados alcanzaron los niveles de rendimiento más altos. En las fases iniciales de experimentación, el modelo ET destacó con una exactitud de 0.9960 y un F1-score de 0.8493, mientras que el modelo GB alcanzó valores comparables, con una exactitud de 0.9956 y un ROC-AUC cercano a 0.999, lo que evidencia una elevada capacidad para discriminar entre estudiantes desertores y no desertores. Adicionalmente, el análisis comparativo entre modelos permitió identificar diferencias en su capacidad de generalización. Mientras que los métodos ensamblados demostraron un desempeño superior en términos de estabilidad y precisión, algunos modelos más simples, como NB, mostraron niveles de rendimiento considerablemente menores, lo que sugiere un posible fenómeno de subajuste (underfitting) ante la complejidad del problema analizado.

Posteriormente, los modelos con mejor desempeño fueron sometidos a un proceso de optimización de hiperparámetros y evaluados en el conjunto de prueba para validar su capacidad de generalización. Los resultados finales evidencian que los modelos seleccionados mantienen un desempeño altamente competitivo. En particular, el ET alcanzó un F1-score optimizado de 0.8919 (89.19 %), junto con una exactitud de 0.9977, un recall de 1.000, una precisión de 0.9706 y un ROC-AUC de 0.9996, lo que indica una elevada capacidad para identificar correctamente a los estudiantes desertores. De forma similar, el modelo RF obtuvo un F1-score de 0.8864 (88.64 %) y una exactitud de 0.9971, mientras que el modelo LR alcanzó un F1-score de 0.8696 (86.96 %) y una exactitud de 0.9965.

La consistencia de los valores de F1-score, superiores al 86 % en todos los modelos evaluados, evidencia un equilibrio adecuado entre precisión y recall, lo cual resulta relevante en contextos de datos desbalanceados como el análisis de la DEU. En este sentido, estos resultados confirman que los modelos desarrollados presentan un desempeño estable y robusto en la identificación

de estudiantes en riesgo de abandono, lo que garantiza una detección confiable de la minoría de estudiantes desertores.

Los resultados indican que el entrenamiento sistemático y la evaluación comparativa de algoritmos de ML permiten desarrollar modelos predictivos capaces de identificar con alta precisión a los estudiantes en riesgo de deserción. En consecuencia, la evidencia empírica derivada del proceso de modelado corrobora la hipótesis planteada, al demostrar que el uso de técnicas de ML constituye una herramienta efectiva para la predicción de la deserción estudiantil universitaria.

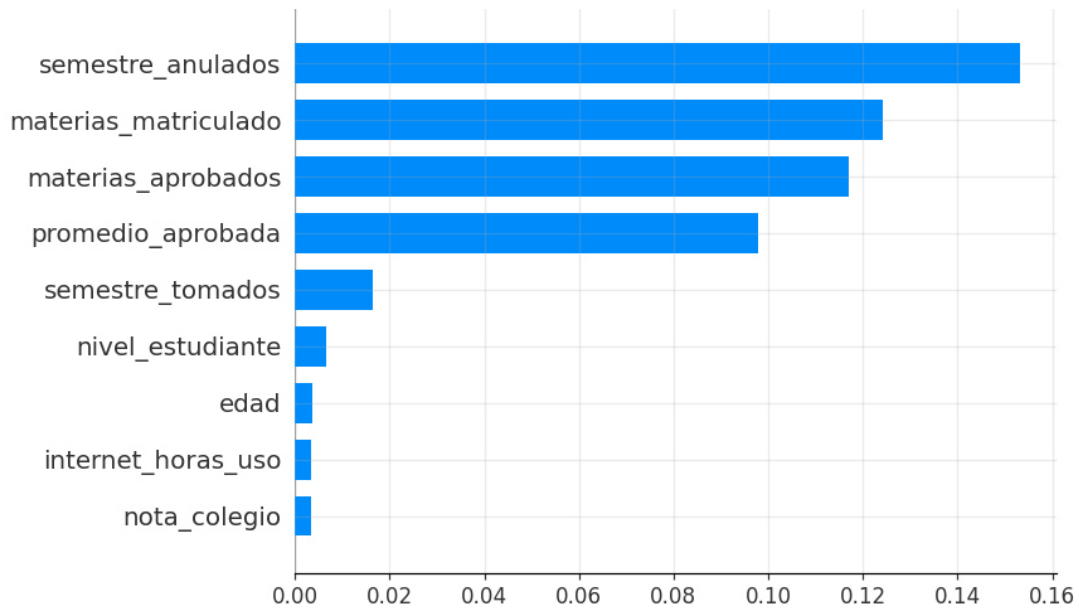
5.3. Resultados de diseño de modelo de explicabilidad

5.3.1 Explicabilidad Global

Como resultado del análisis de explicabilidad con SHAP y LIME, se generaron gráficos de barras que ilustran las características que más contribuyen a explicar el modelo predictivo de la DEU. La Figura 21 muestra la importancia de las variables del modelo Extra Trees mediante valores SHAP, lo que permite identificar las características que más contribuyen a la predicción de la DEU. Los resultados evidencian que semestre_anulados es la variable con mayor peso explicativo, lo que sugiere que el historial de asignaturas anuladas constituye un indicador crítico del perfil de riesgo. En segundo y tercer lugar se ubican materias_matriculado y materias_aprobados, lo que refleja que la carga académica y el desempeño acumulado influyen de manera significativa en la probabilidad de permanencia o de abandono. Asimismo, la variable promedio_aprobada es de relevancia destacada, lo que confirma que el rendimiento académico global mantiene una asociación consistente con el fenómeno analizado. En contraste, variables como semestre_tomados y nivel_estudiante muestran una contribución moderada y aportan información complementaria sobre la trayectoria académica.

Figura 20

Resultados de explicabilidad global con SHAP

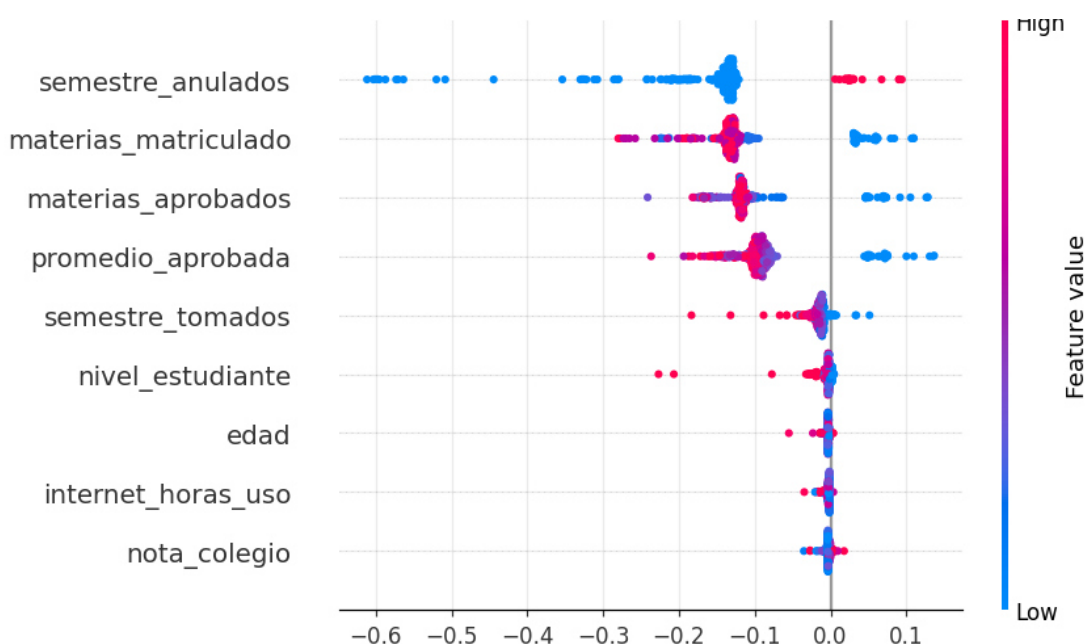


En continuidad con la importancia global estimada mediante SHAP, el diagrama permitió profundizar no solo en la jerarquía de relevancia de las variables, sino también en la dirección y la magnitud de su influencia en la predicción del modelo. Los resultados se presentan en la figura 22, donde los valores a la derecha del eje central indican que la variable incrementa la probabilidad de deserción, mientras que los valores a la izquierda indican que la reduce. Además, el color indica si el valor de la variable es alto (rojo) o bajo (azul). Por tanto, la variable que más contribuye al riesgo de deserción es semestre_anulados. Cuando un estudiante tiene un número elevado de asignaturas anuladas (puntos rojos a la derecha), el modelo aumenta significativamente la probabilidad estimada de abandono. Por el contrario, cuando este valor es bajo (en azul a la izquierda), el riesgo disminuye. Esto indica que la acumulación de anulaciones es un fuerte indicador de vulnerabilidad académica. Por el contrario, materias_aprobados y promedio_aprobada presentan un comportamiento inverso: valores altos de promedio y una mayor cantidad de materias aprobadas tienden a ubicarse a la izquierda, lo que reduce la probabilidad de deserción. Por el contrario, niveles de aprobación o promedios bajos desplazan

la predicción hacia la derecha, lo que incrementa el riesgo. Esto confirma que el rendimiento académico sostenido actúa como factor protector. En el caso de `materias_matriculado`, una carga académica elevada puede tener efectos mixtos según el desempeño asociado. No obstante, cuando se combina con un bajo rendimiento, puede incrementar el riesgo. Variables como `semestre_tomados` y `nivel_estudiante` aportan de manera moderada y sirven como indicadores del avance en la trayectoria académica. Su efecto es menor que el del rendimiento acumulado, pero ayuda a complementar la evaluación del perfil del estudiante. Por último, `edad`, `internet_horas_uso` y `nota_colegio` tienen una influencia menor que la de las variables académicas universitarias. Aunque pueden aportar información contextual, su impacto en la predicción del modelo es secundario.

Figura 21

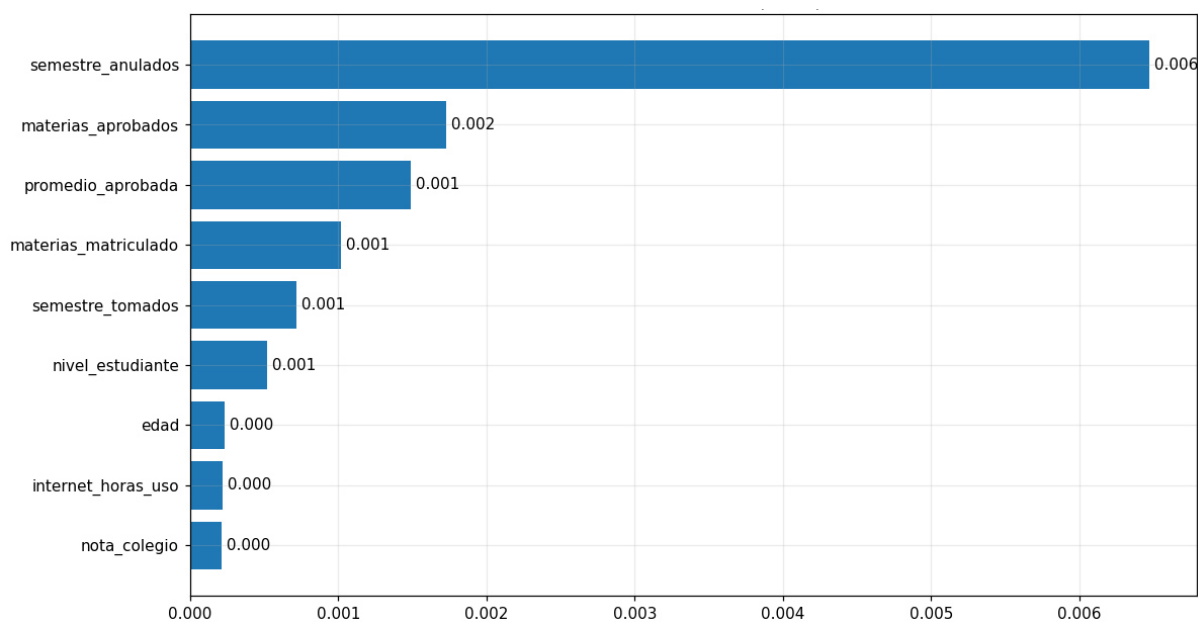
Resultados de explicabilidad global con SHAP



A continuación, también se analizó la importancia global mediante LIME, calculada a partir de 120 explicaciones locales agregadas. Los resultados obtenidos se muestran en la figura 23.

Figura 22

Resultados de explicabilidad global con LIME



Se confirma que semestre_anulados es la variable con mayor peso explicativo en el modelo, con una contribución sustancialmente superior a la de los demás predictores. Este hallazgo refuerza la evidencia obtenida mediante SHAP y sugiere que el historial de asignaturas anuladas constituye un factor determinante para identificar el riesgo de deserción. En un segundo nivel de relevancia se ubican materias_aprobados y promedio_aprobada, lo que indica que el desempeño académico acumulado continúa siendo un componente estructural clave en la predicción. De manera consistente, materias_matriculado y semestre_tomados aportan información complementaria sobre la trayectoria académica del estudiante, lo que consolida el peso explicativo de las variables estrictamente académicas en el modelo.

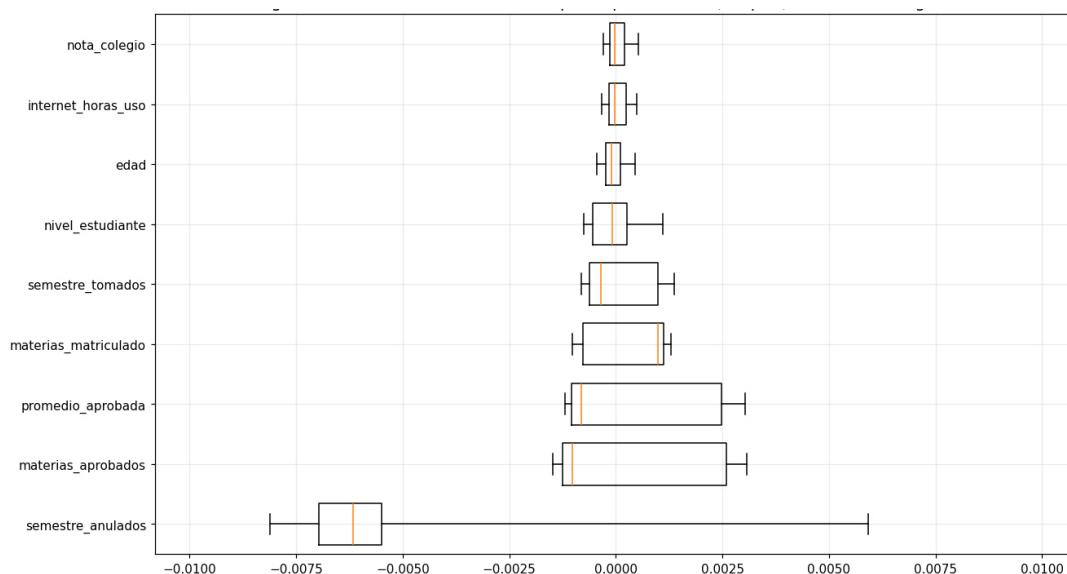
Por otro lado, variables como nivel_estudiante, edad, internet_horas_uso y nota_colegio aportan una contribución promedio menor, lo que sugiere que, aunque proporcionan

información contextual relevante, su influencia es secundaria frente a los indicadores de rendimiento y de progresión académica.

De igual forma, la figura 24 muestra el análisis de la influencia de cada variable en la predicción del modelo mediante LIME. Cada caja representa la distribución de los pesos locales de una variable en distintas observaciones. Cuando los valores se ubican hacia la derecha, la variable tiende a aumentar la probabilidad de deserción; cuando se ubican hacia la izquierda, contribuye a disminuirla. Además, la amplitud de la caja indica qué tan variable es su impacto entre distintos estudiantes.

Figura 23

Distribución de pesos por variable con LIME



Se observa que semestre_anulados es la variable con el mayor efecto en la predicción. Su distribución es más extensa, lo que indica una influencia fuerte y consistente y sugiere que un mayor número de materias anuladas tiende a incrementar el riesgo estimado de abandono. Las variables materias_aprobados y promedio_aprobada también muestran una influencia significativa. En general, cuando el estudiante tiene más materias aprobadas o un promedio más alto, el peso de estas variables tiende a reducir el riesgo. Por el contrario, los bajos niveles de

rendimiento académico tienden a aumentar la probabilidad de deserción. Por su parte, `materias_matriculado` y `semestre_tomados` presentan un impacto moderado. Su influencia depende más del contexto académico del estudiante y de su interacción con otras variables. En cambio, factores como `edad`, `internet_horas_uso` y `nota_colegio` muestran una distribución más concentrada en torno a cero, lo que indica que su efecto sobre la predicción es menor que el de las variables directamente relacionadas con el desempeño universitario.

5.3.2 Explicabilidad Local

En esta investigación se implementó la explicabilidad local mediante SHAP y LIME, enfocada en cuatro ejemplos clave que permitieron analizar en detalle el comportamiento del modelo ExtraTrees ante distintas combinaciones de resultados esperados y reales y aportaron evidencia empírica sobre su capacidad de generalización y precisión. Los casos analizados se detallan en la tabla 36, que se describe a continuación:

Tabla 36

Escenarios de correspondencia entre la realidad observada y la predicción

Resultado	Predicción
Deserta	Deserta
Deserta	No deserta
No deserta	Deserta
No deserta	No Deserta

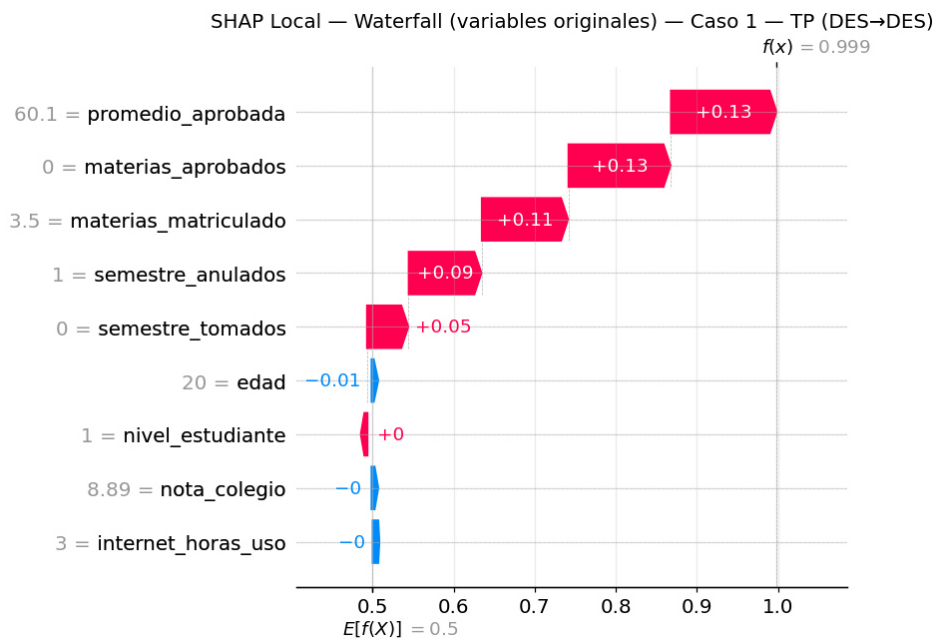
Explicabilidad Local Caso 1: Verdadero Positivo (DES→DES)

La Figura 25 muestra la explicación local del modelo ET para el caso individual con $\text{idx}=3215$, clasificado como deserción (verdadero positivo: DES→DES). Este tipo de visualización permite descomponer la predicción final del modelo en las contribuciones aditivas de cada variable, de acuerdo con la formulación teórica de los valores de Shapley. En este contexto, el modelo parte de un valor esperado base $E[f(X)] = 0.5$, correspondiente a la probabilidad promedio de deserción en el conjunto de entrenamiento, y cada variable desplaza esta probabilidad hacia arriba o hacia abajo hasta alcanzar la predicción específica para el individuo analizado, en este caso $f(x) = 0.999$.

Las variables académicas presentan las mayores contribuciones positivas al incremento del riesgo. En particular, `promedio_aprobada` y `materias_aprobados` aportan cada una aproximadamente $+0.13$ al valor base, seguidas por `materias_matriculado` ($+0.11$), `semestre_anulados` ($+0.09$) y `semestre_tomados` ($+0.05$). Estas contribuciones indican que, dadas las características concretas del estudiante, el patrón aprendido por el modelo asocia dicha configuración académica con una alta probabilidad de deserción. Por el contrario, la edad, `nota_colegio`, `internet_horas_uso` y `nivel_estudiante` presentan contribuciones cercanas a cero, lo que indica que su influencia en este caso particular es limitada y no compensa el efecto acumulado de los indicadores académicos. La magnitud relativamente pequeña de estos efectos confirma que, para este individuo, los factores contextuales desempeñan un papel secundario frente a las variables asociadas al rendimiento universitario.

Figura 24

Explicabilidad local SHAP (Deserta-Deserta)



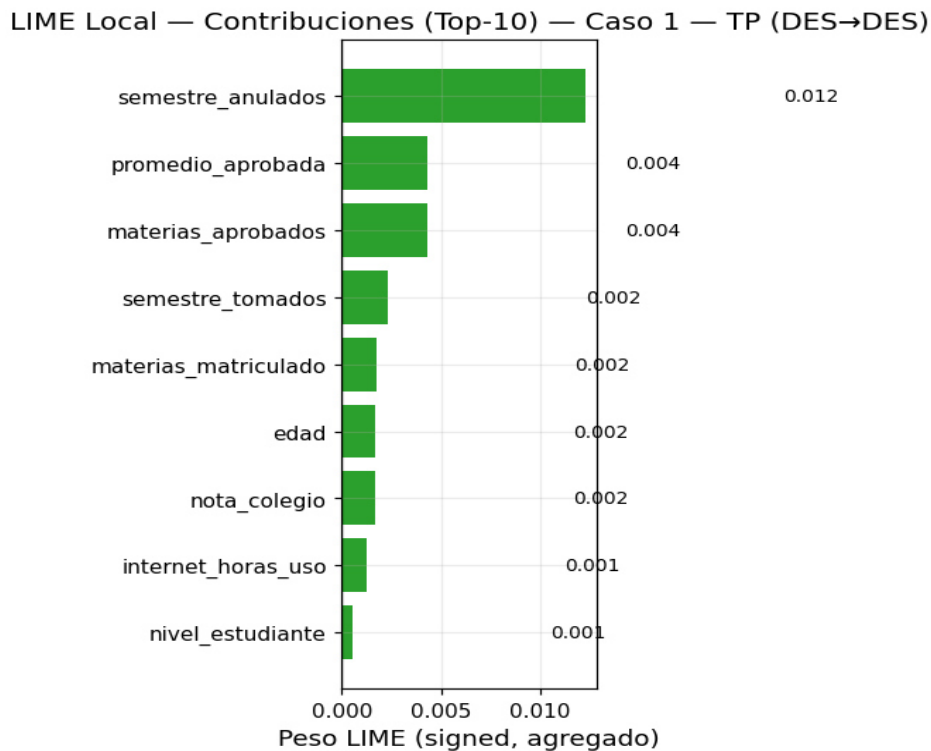
De forma complementaria, la Figura 26 muestra las contribuciones locales estimadas mediante LIME para el mismo caso en el que el estudiante fue clasificado como desertor. A diferencia del método SHAP, que se fundamenta en la teoría de valores de Shapley y garantiza propiedades aditivas globales, LIME aproxima localmente el comportamiento de un modelo complejo mediante un modelo lineal interpretable en torno a la instancia analizada.

Los resultados evidencian que semestre_anulados es el que aporta la mayor contribución positiva (0.012), lo que lo consolida como el factor más determinante en la explicación local. Este hallazgo es coherente con el análisis previo basado en SHAP, lo que refuerza la consistencia entre los métodos de explicabilidad. En el segundo nivel de relevancia se ubican promedio_aprobada y materias_aprobados (0.004 cada una), lo que confirma que el rendimiento académico acumulado desempeña un papel estructural en la decisión del modelo.

Variables como semestre_tomados, materias_matriculado, edad y nota_colegio presentan contribuciones intermedias (0.002), lo que indica que, aunque influyen en la predicción, su peso es mucho menor que el de los indicadores principales de desempeño universitario.

Figura 25

Explicabilidad local LIME (Deserta-Deserta)

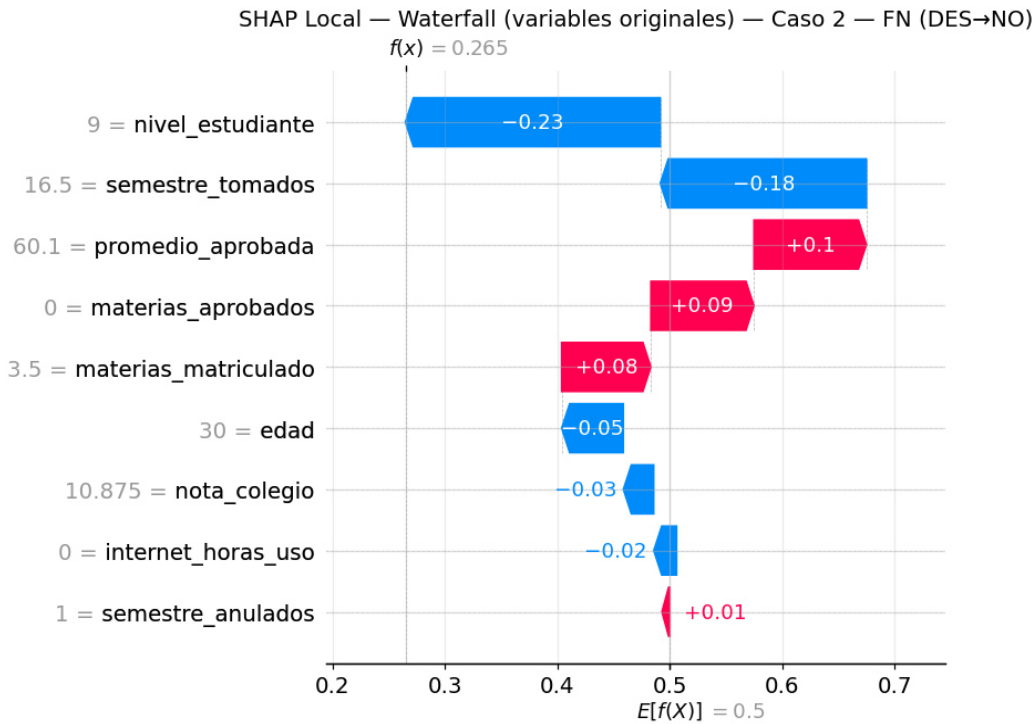


Explicabilidad Local — Caso 2: Verdadero Negativo (DES→NO)

A continuación, se presenta el análisis académico del Caso 2, correspondiente a la instancia con $\text{idx} = 266$, de falso negativo (DES→NO), en la que el estudiante realmente pertenecía a la clase “Deserción”, pero el modelo lo clasificó como “No deserción”. La figura 27 muestra la explicación local mediante SHAP (gráfico de cascada), lo que permite descomponer la predicción individual en contribuciones aditivas.

Figura 26

Explicabilidad local SHAP (Deserta- No Deserta)



En este caso, la probabilidad final estimada por el modelo es $f(x) = 0.265$, un valor inferior al umbral de decisión, lo que explica su clasificación como “NO deserción”. Sin embargo, el análisis detallado revela una interacción compleja entre variables que inciden en la predicción en direcciones opuestas. Las contribuciones negativas más relevantes, es decir, aquellas que reducen la probabilidad de deserción, corresponden a `nivel_estudiante` (-0.23) y `semestre_tomados` (-0.18). Estas variables indican que el estudiante se encontraba en un nivel académico avanzado y con un número considerable de semestres cursados, factores que el modelo interpreta como señales de permanencia y estabilidad académica. Asimismo, variables como `edad` (-0.05), `nota_colegio` (-0.03) e `internet_horas_uso` (-0.02) contribuyen marginalmente a reducir el riesgo estimado.

En contraste, variables como `promedio_aprobada` ($+0.10$), `materias_aprobados` ($+0.09$) y `materias_matriculado` ($+0.08$) empujan la predicción hacia la clase “Deserción”. Sin embargo,

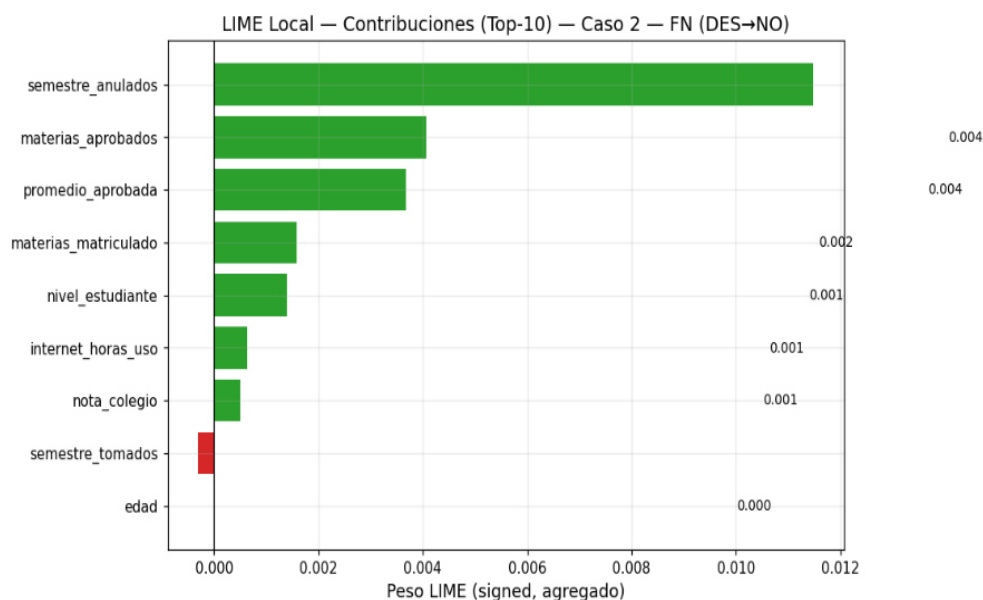
su efecto acumulado no logra compensar el fuerte impacto protector asociado al nivel académico y al tiempo de permanencia en la universidad.

Desde una perspectiva interpretativa, este caso ilustra un fenómeno típico en modelos predictivos complejos: la predominancia de variables estructurales (como la trayectoria académica consolidada) puede neutralizar las señales tempranas de riesgo académico. En términos sustantivos, el modelo parece priorizar la madurez académica y la permanencia prolongada como indicadores de continuidad, lo que conduce a una subestimación del riesgo real para el estudiante.

De forma complementaria, en la Figura 28 se presenta la interpretación correspondiente al Caso 2 — Falso Negativo (DES→NO) utilizando LIME, cuyo objetivo es aproximar localmente la decisión del modelo mediante un modelo lineal interpretable.

Figura 27

Explicabilidad local LIME (Deserta- No Deserta)



La figura muestra las contribuciones locales agregadas (Top-10) de este estudiante en particular. A diferencia de SHAP, que descompone la predicción en términos de

desplazamiento respecto al valor base, LIME aproxima la frontera de decisión en el entorno inmediato del individuo, asignando pesos locales que indican qué variables empujan la predicción hacia una u otra clase. En este caso, la variable con mayor peso positivo es `semestre_anulados`, lo que indica que la presencia de anulaciones académicas constituye la señal más fuerte que impulsa la predicción hacia la clase “Deserción”. Le siguen `materias_aprobados`, `promedio_aprobada` y `materias_matriculado`, todas con contribuciones positivas, lo que sugiere que el patrón de desempeño y la carga académica del estudiante presentan señales compatibles con un riesgo.

Sin embargo, también se observan contribuciones negativas, aunque de menor magnitud, en particular en `semestre_tomados`, lo que indica que el tiempo acumulado en la institución actúa como factor protector en la decisión local. Variables como `edad`, `nota_colegio` e `internet_horas_uso` presentan valores cercanos a cero, lo que evidencia una influencia marginal en esta predicción específica.

La explicación obtenida mediante SHAP y LIME permite comprender con precisión las causas del error de predicción. Ambas herramientas coinciden en señalar que, en este caso, la edad y las condiciones socioculturales influyeron significativamente en la clasificación, lo que resalta la importancia de ajustar el modelo para mejorar su sensibilidad ante perfiles con bajo riesgo aparente pero alta probabilidad real de abandono. Esta interpretación contribuye a fortalecer la validez explicativa del modelo al permitir identificar no solo los aciertos, sino también las razones de sus errores predictivos.

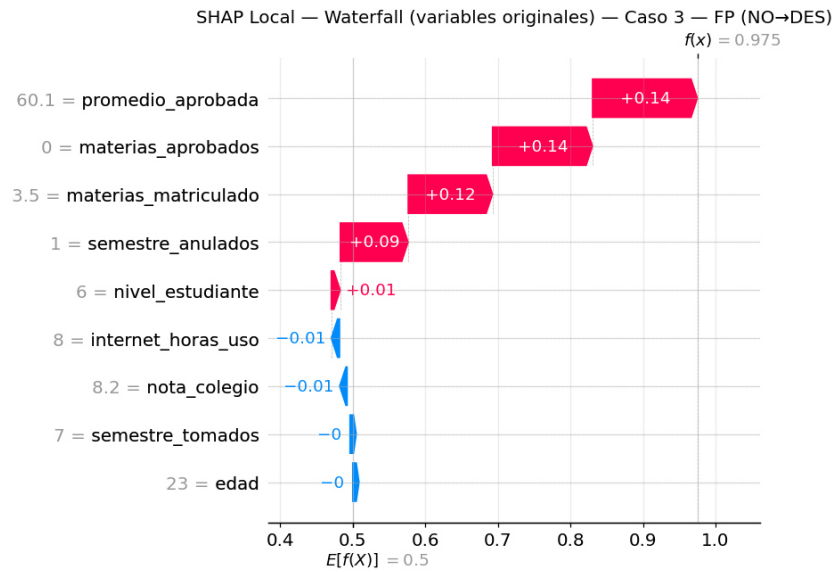
Explicabilidad Local — Caso 3: Negativo Verdadero (NO→DES)

En el Caso 3 FP (NO→DES), correspondiente a la instancia con `idx = 5177`, se observa un falso positivo. En la explicación local mediante el gráfico SHAP de la figura 29, la predicción parte

de un valor base $E[f(X)] = 0.5$ y se desplaza progresivamente hasta $f(x) = 0.975$ debido a la acumulación de contribuciones positivas. Las variables con mayor impacto en la sobreestimación del riesgo fueron principalmente académicas: promedio_aprobada y materias_aprobados aportaron aproximadamente +0.14 cada una, seguidas por materias_matriculado (+0.12) y semestre_anulados (+0.09). Estas variables, en conjunto, explican el desplazamiento significativo hacia la clase de deserción. La variable nivel_estudiante mostró una contribución marginal positiva de 0.01. En contraste, las variables con efecto atenuador fueron leves: internet_horas_uso y nota_colegio presentaron aportes negativos cercanos a -0.01, mientras que semestre_tomados y edad tuvieron un impacto prácticamente neutro en esta instancia específica. Este caso evidencia que el modelo activó patrones internos asociados a trayectorias académicas potencialmente vulnerables, en particular al desempeño y a los antecedentes de anulación de semestre, lo que generó una probabilidad extremadamente alta (0.975) que superó el umbral de clasificación (0.910). Sin embargo, el estudiante no abandonó, lo que sugiere que el modelo puede estar captando señales de riesgo académico estructural que no necesariamente se traducen en una deserción efectiva.

Figura 28

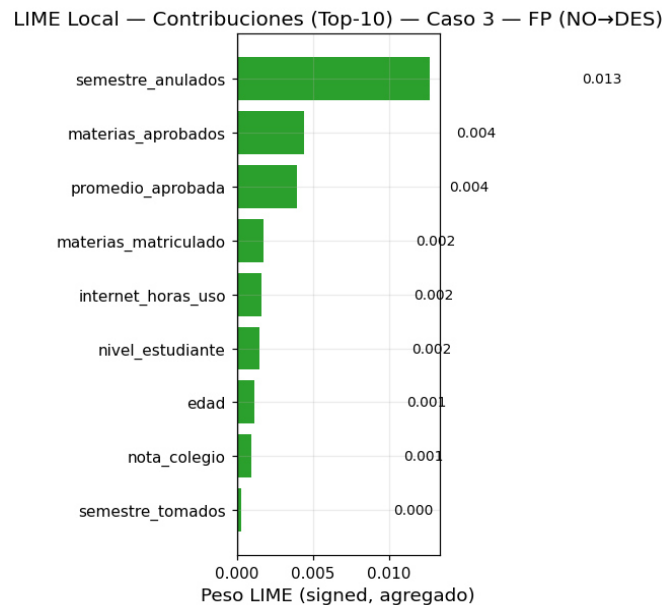
Explicabilidad local SHAP (No deserta- Deserta)



Complementando la explicabilidad local obtenida con SHAP, el análisis con LIME para $idx = 5177$ confirma la estructura explicativa observada previamente. Aunque desde una aproximación lineal local, los resultados se evidencian en la figura 30.

Figura 29

Explicabilidad local LIME (No deserta- Deserta)



Se observa que todas las variables mostradas presentan pesos positivos, lo que indica que en el entorno local de esta instancia cada una contribuye a incrementar la probabilidad de

clasificación como deserción. La variable con mayor peso explicativo es `semestre_anulados` (0.013), lo que la convierte en el factor más influyente en la predicción del modelo. Le siguen `materias_aprobados` (0.004) y `promedio_aprobada` (0.004), lo que refuerza la relevancia de los indicadores académicos en la construcción del riesgo estimado. En un segundo nivel de impacto aparecen `materias_matriculado` (0.002), `internet_horas_uso` (0.002) y `nivel_estudiante` (0.002), mientras que `edad` (0.001) y `nota_colegio` (0.001) presentan contribuciones marginales. La variable `semestre_tomados` muestra un efecto prácticamente nulo.

Por lo tanto, LIME confirma que la predicción de deserción no se explica por un único atributo, sino por la acumulación de múltiples señales académicas que, en conjunto, activan la clasificación positiva. Sin embargo, a diferencia de SHAP, que cuantifica la contribución exacta respecto del valor base del modelo, LIME aproxima el comportamiento del modelo mediante una representación local lineal, lo que explica la menor magnitud de los pesos observados.

En síntesis, el análisis de esta instancia demuestra el valor de la explicabilidad local para identificar los factores que originan errores predictivos, lo que permite detectar sesgos y comprender de forma transparente cómo el modelo toma sus decisiones. La concordancia entre las explicaciones de SHAP y LIME refuerza la validez interpretativa y la trazabilidad del modelo, y aporta una base sólida para futuras optimizaciones orientadas a minimizar los falsos positivos en la detección del riesgo de deserción universitaria.

La convergencia entre SHAP y LIME fortalece la robustez explicativa del modelo: ambos métodos identifican a `semestre_anulados`, `materias_aprobados` y `promedio_aprobada` como variables centrales en la toma de decisiones. No obstante, el hecho de que el estudiante no haya desertado sugiere que el modelo está interpretando correctamente un patrón estructural de riesgo académico, pero no necesariamente un abandono efectivo. Este hallazgo es relevante

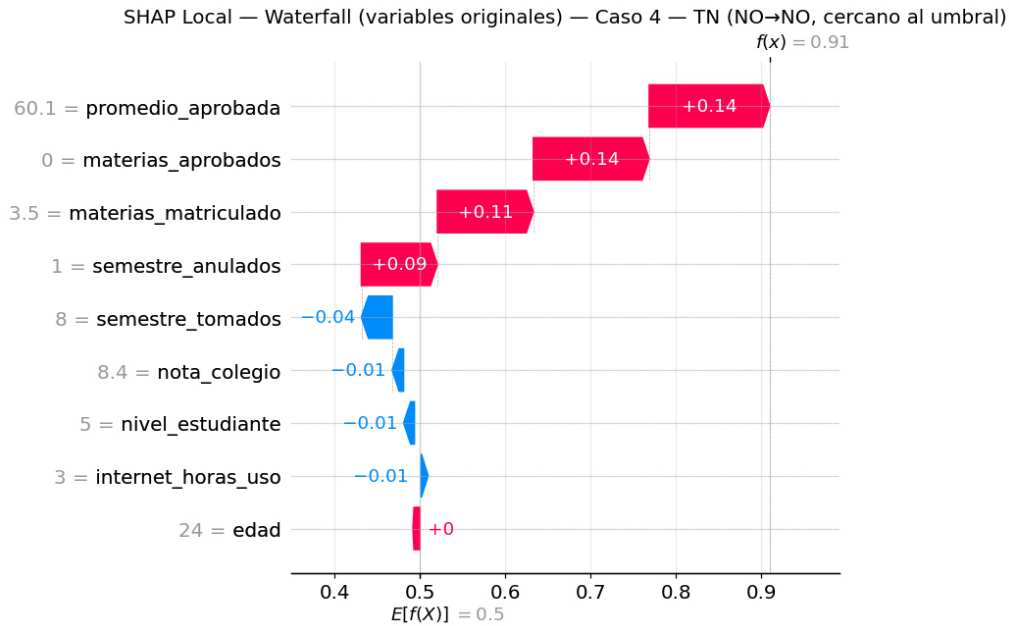
para la gestión institucional, ya que indica que el sistema de alerta temprana podría estar captando a estudiantes en situación de vulnerabilidad académica que finalmente logran persistir, lo que refuerza la necesidad de utilizar la predicción como herramienta preventiva y no como un diagnóstico determinista.

Explicabilidad Local — Caso 4: Negativo Negativo (NO→NO)

El Caso 4 (idx = 9417) corresponde a un verdadero negativo (TN), como se observa en la figura 31, ya que el estudiante no desertó y el modelo mantuvo esa clasificación. Sin embargo, la probabilidad estimada de deserción (0.910) evidencia que la instancia presenta una alta intensidad de señales asociadas al riesgo, aunque estas no fueron suficientes para modificar la decisión final del modelo. El análisis mediante el gráfico SHAP indica que hay factores que incrementan la probabilidad estimada de deserción; destacan principalmente variables de naturaleza académica: promedio_aprobada (+0.14), materias_aprobados (+0.14), materias_matriculado (+0.11) y semestre_anulados (+0.09). Estas contribuciones confirman el patrón estructural identificado en el análisis global, en el que el desempeño y la trayectoria académica constituyen los predictores con mayor capacidad explicativa del modelo. No obstante, también se observan efectos compensatorios relevantes. En particular, semestre_tomados (−0.04) emerge como el principal factor reductor del riesgo estimado, acompañado de contribuciones negativas de menor magnitud asociadas a nota_colegio (−0.01), nivel_estudiante (−0.01) e internet_horas_uso (−0.01). La variable edad tiene un impacto prácticamente neutro. La interacción entre estos factores positivos y negativos genera un equilibrio explicativo que sustenta la clasificación de no desertor.

Figura 30

Explicabilidad local SHAP (No deserta- No deserta)



Complementando el análisis con SHAP, la explicación local obtenida mediante LIME para el Caso 4 (idx = 9417) confirma la estructura explicativa observada, aunque desde una aproximación lineal local del comportamiento del modelo.

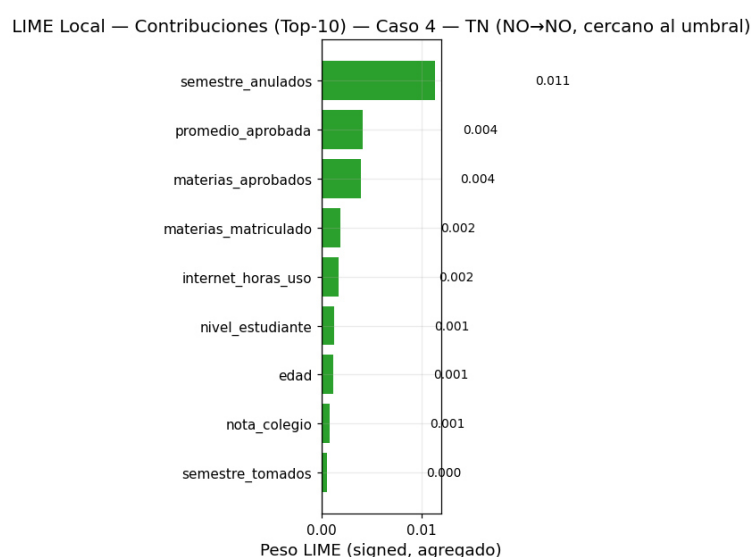
En la figura 32, se aprecia que la variable con mayor peso explicativo es semestre_anulados (0.011), lo que indica que, en el entorno local de esta instancia, este atributo constituye el principal impulsor del riesgo estimado. Le siguen promedio_aprobada (0.004) y materias_aprobados (0.004), lo que consolida nuevamente el predominio de los indicadores académicos en la construcción de la probabilidad de deserción. En un segundo nivel de relevancia aparecen materias_matriculado (0.002) e internet_horas_uso (0.002), mientras que nivel_estudiante (0.001), edad (0.001) y nota_colegio (0.001) presentan contribuciones marginales. La variable semestre_tomados muestra un efecto prácticamente nulo en la aproximación lineal local. Desde una perspectiva interpretativa, LIME confirma que la predicción no se explica por un único atributo determinante, sino por la agregación de múltiples señales académicas que, en conjunto, incrementan el riesgo estimado. A diferencia de SHAP —que descompone la predicción respecto del valor base del modelo—, LIME aproxima el

comportamiento del clasificador en la vecindad de la instancia, lo que explica la menor magnitud absoluta de los pesos observados.

La convergencia entre ambos métodos de explicabilidad refuerza la consistencia interna del modelo propuesto: tanto SHAP como LIME identifican a semestre_anulados, promedio_aprobada y materias_aprobados como variables centrales en la estructura decisional. No obstante, el hecho de que el estudiante haya mantenido su permanencia académica evidencia que la presencia de indicadores de riesgo no implica necesariamente un abandono efectivo, lo que confirma el carácter probabilístico y multifactorial del fenómeno de la deserción universitaria.

Figura 31

Explicabilidad local LIME (No deserta- No deserta)



5.3.3. Hipótesis Específica 3

La tercera hipótesis de investigación plantea que el diseño de un modelo explicativo permite interpretar la contribución y las relaciones entre los factores asociados a la DEU. Esta hipótesis se fundamenta en la necesidad de superar el enfoque puramente predictivo de los modelos de ML y avanzar hacia la comprensión de los factores que inciden en la deserción.

Una vez identificado el modelo predictivo con mayor desempeño, se procedió a la fase de explicabilidad, cuyo propósito fue analizar de manera sistemática la contribución de las variables y la forma en que interactúan para influir en la predicción de la deserción estudiantil. Para ello, se emplearon técnicas de inteligencia artificial explicable (XAI), que permiten descomponer las decisiones del modelo y cuantificar la influencia de cada atributo en el resultado de la predicción.

El análisis de explicabilidad permitió identificar que los factores que más contribuyen al comportamiento del modelo se asocian principalmente con la dimensión académica del estudiante. Variables como semestre_anulados, materias_aprobados, promedio_aprobada y materias_matriculado evidencian una influencia significativa en la probabilidad de deserción, lo que sugiere que la trayectoria y el desempeño académicos constituyen determinantes centrales del abandono universitario. Este resultado es consistente con los análisis previos de correlación y selección de características, en los que dichas variables también mostraron una elevada capacidad explicativa.

De manera complementaria, el modelo explicativo permitió identificar la contribución de variables pertenecientes a otras dimensiones relevantes del fenómeno, tales como factores socioeconómicos, familiares y sociodemográficos, entre los que destacan los atributos del contexto familiar, el acceso a recursos tecnológicos y las condiciones socioeconómicas del estudiante. Estos resultados evidencian que la deserción universitaria no responde a un único factor, sino que se configura como un fenómeno multidimensional en el que interactúan variables académicas, sociales y económicas.

Asimismo, el análisis de explicabilidad permitió visualizar las relaciones entre variables y su impacto relativo en la predicción del modelo, lo que facilita comprender cómo diferentes combinaciones de factores pueden incrementar o reducir la probabilidad de deserción. Este tipo

de interpretación aporta un valor adicional frente a los modelos predictivos tradicionales, ya que no solo permite identificar a estudiantes en riesgo, sino también comprender los factores que lo explican.

En consecuencia, el modelo explicativo desarrollado proporciona una representación interpretable del proceso de predicción, lo que permite analizar tanto la contribución individual de las variables como las relaciones entre los distintos factores asociados al abandono universitario. Por lo tanto, la evidencia obtenida respalda la hipótesis planteada, al demostrar que el diseño de un modelo explicativo basado en técnicas de inteligencia artificial explicable permite interpretar de manera sistemática los factores que inciden en la DEU.

5.4. Discusión

La presente sección examina los resultados de la literatura especializada con el propósito de evaluar la solidez metodológica y el aporte científico del modelo de explicabilidad propuesto para la DEU. Para ello, se organiza en tres ejes complementarios: la pertinencia del conjunto de datos construido, el desempeño comparativo de los modelos predictivos y la capacidad explicativa. A diferencia de una lectura descriptiva, este análisis busca establecer coincidencias, diferencias y contrastes con estudios previos, así como identificar las posibles razones que expliquen dichas convergencias o discrepancias.

En primer lugar, los resultados obtenidos confirman que la DEU no puede comprenderse a partir de una sola dimensión, sino mediante la integración de variables académicas, sociodemográficas y socioeconómicas. Este hallazgo es consistente con (Lakkaraju et al., 2015) quienes sostienen que la incorporación de múltiples dimensiones del contexto estudiantil mejora la capacidad de predicción de los modelos aplicados. Sin embargo, la contribución de esta investigación no solo reafirma esa idea general, sino que muestra con mayor precisión que

la naturaleza multidimensional del fenómeno no implica una participación equivalente de todas las variables. En el modelo desarrollado, las variables académicas concentraron el mayor peso explicativo, mientras que las sociodemográficas y socioeconómicas tuvieron una participación complementaria, aunque menos decisiva.

Este predominio de la dimensión académica coincide con (Aulck et al., 2016), quienes identificaron que el progreso académico y el rendimiento del estudiante son predictores más robustos de la deserción. También guarda relación con (Gray et al., 2014), quienes señalaron que el desempeño en los primeros semestres es uno de los factores más determinantes de la permanencia universitaria. En el mismo sentido, (Costa et al., 2017) indicaron que el rendimiento académico presenta una mayor capacidad explicativa que otras variables contextuales en modelos de ML aplicados a la educación superior. No obstante, en la presente investigación resultó ser más crucial, lo que permite sostener que la trayectoria académica temprana no actúa únicamente como predictor importante, sino también como eje estructurador del riesgo de abandono.

En contraste, las variables sociodemográficas mostraron una asociación más moderada con la DEU. Factores como la edad o ciertas características del entorno social conservaron valor explicativo, aunque en menor medida que las variables académicas. Este patrón se aproxima a lo planteado por (Schuchart & Schimke, 2022) , quienes señalan que el efecto de este tipo de variables suele estar mediado en gran medida por el rendimiento académico. De forma similar, (Sarra et al., 2019) sostienen que los factores socioeconómicos amplían la comprensión de la DEU, aunque normalmente no superan el peso de los indicadores de desempeño. Por tanto, los resultados obtenidos no niegan la importancia del contexto social del estudiante, pero sí sugieren que su influencia opera con mayor fuerza cuando se articula con la dimensión académica.

No obstante, la literatura no es uniforme al respecto. (Quispe et al., 2024a) reportaron una mayor participación de los factores socioeconómicos en la predicción del abandono estudiantil. En la presente investigación, en cambio, esas variables tuvieron una menor contribución que la de los factores académicos. Esta discrepancia no debe entenderse como una refutación directa, sino como una señal de que la estructura explicativa de la DEU no es homogénea entre contextos. Mientras que en ciertos escenarios las condiciones socioeconómicas pueden constituir el núcleo del riesgo de abandono, en otros recae en el historial académico, el rezago, la acumulación de dificultades formativas o el desempeño temprano.

Una situación similar se observa en las variables conductuales provenientes de plataformas digitales. En el trabajo de (Cheng et al., 2025), este tipo de trazas puede influir significativamente en la predicción del abandono. En el presente estudio no se incorporaron variables relacionadas con entornos virtuales, ya que el conjunto de datos se construyó a partir de información académica, sociodemográfica y socioeconómica estructurada. Aun así, los modelos alcanzaron desempeños muy altos, lo que introduce un contraste relevante con esa línea de investigación. Desde ese punto de vista, el presente estudio muestra que es posible obtener predicciones robustas incluso cuando no se dispone de analítica conductual digital. Esto representa una ventaja práctica para instituciones que aún no cuentan con ecosistemas de seguimiento virtual lo suficientemente desarrollados.

En segundo lugar, los resultados del modelado predictivo evidencian que los algoritmos de ML permiten identificar con alta capacidad discriminativa a los estudiantes en riesgo de deserción. Los modelos ET, RF, LR alcanzaron un rendimiento sobresaliente, con valores de accuracy cercanos a 0.997 o superiores. Sin embargo, una lectura metodológicamente rigurosa exige ir más allá de la exactitud global, especialmente en un problema como la DEU, en el que el

desequilibrio entre desertores y no desertores puede distorsionar la interpretación del rendimiento.

Desde esa perspectiva, ET fue el modelo con el mejor equilibrio global, con un F1-score de 0.892, superior al de RF (0.886) y al de LR (0.870). Aunque la diferencia entre ET y RF es pequeña en términos absolutos, sigue siendo relevante porque el F1-score sintetiza el equilibrio entre la sensibilidad y la precisión. En ese equilibrio, ET mostró una ventaja real. Además, ET mantuvo un recall perfecto (1.000), al igual que LR, pero lo hizo con una precisión claramente mayor (0.971 frente a 0.769). Esta diferencia es crítica, ya que indica que ambos modelos detectaron todos los casos de deserción del conjunto de prueba, pero LR generó un número mucho mayor de falsos positivos. La comparación entre ET y RF también aporta elementos significativos. Aunque RF alcanzó un ROC-AUC ligeramente superior en el valor no redondeado, con 0.999647 frente a 0.999551 de ET, esa ventaja fue marginal y no se tradujo en un mejor desempeño operativo global. De hecho, RF mostró un recall menor (0.975 frente a 1.000) y una precisión considerablemente menor (0.813 frente a 0.971). Por tanto, aunque RF fue excelente en términos de discriminación general entre clases, ET fue más consistente en la clasificación final de los estudiantes en riesgo.

En este escenario, una lectura centrada exclusivamente en la exactitud puede conducir a interpretaciones engañosas. De hecho, estudios recientes sobre la predicción de la deserción estudiantil han reportado valores de accuracy de hasta 97.6 %, como el trabajo de (Albugami et al., 2026) . Sin embargo, en contextos con un desbalance notable, este indicador debe interpretarse con cautela, ya que un modelo podría alcanzar una exactitud muy elevada simplemente clasificando la mayoría de los casos como no desertores, sin identificar correctamente a los estudiantes en riesgo. Frente a esta limitación, resulta más apropiado recurrir a métricas como precisión, recall y F1-score, ya que ofrecen una valoración más

rigurosa del equilibrio entre la identificación efectiva de los desertores y la reducción de falsos positivos, tal como plantean (Saito & Rehmsmeier, 2015).

La ventaja observada en los modelos de ensamble coincide con la reportada por (Sarra et al., 2019), quienes señalaron que este tipo de algoritmos presenta una mayor capacidad de generalización en contextos educativos. También se relaciona con (Mahyudi et al., 2025), quienes mostraron que el RF capta eficazmente relaciones no lineales complejas entre variables académicas. En una línea similar, (Bettahi et al., 2025) reportaron altas métricas de AUC y F1-score en modelos como RF y XGBoost aplicados a la predicción de la DEU. La presente investigación coincide con esta evidencia en cuanto a la superioridad general de los métodos de ensamble, pero aporta un nivel adicional de precisión al mostrar que, aun dentro de ese grupo de algoritmos, ET fue superior a RF en el equilibrio conjunto de F1-score, recall, precision y especificidad.

Asimismo, los resultados respaldan lo planteado por (Khatun et al., 2025) , quienes sostienen que la combinación entre ML y la optimización de hiperparámetros puede mejorar notablemente el rendimiento predictivo. En el presente estudio, la calidad del desempeño alcanzado no parece depender únicamente del algoritmo seleccionado, sino también del proceso metodológico empleado, que incluyó el tratamiento del desbalance, la selección de variables, la validación y la optimización. Por ello, una parte de la mejora observada respecto de estudios anteriores puede atribuirse no solo a la arquitectura del modelo, sino también al rigor del procedimiento de modelado.

Al comparar estos resultados con los antecedentes internacionales, se observan diferencias llamativas. En el trabajo de (J. Martínez & Castillo, 2024) se señaló que los modelos basados en árboles potenciados pueden superar el 95 % de exactitud en la predicción del abandono. En la presente investigación, los tres algoritmos superaron ese umbral, con una precisión entre 99,7

% y 99,8 %, lo que sugiere una mejora clara en la exactitud global. Sin embargo, la contribución más relevante no es solo esa diferencia, sino que el rendimiento también se mantuvo en métricas más rigurosas, como un F1-score de hasta 89,2 % y un recall de 100 % en ET. De manera similar, (Vaarma & Li, 2024) reportaron niveles de exactitud cercanos al 86 % en la educación superior finlandesa. En comparación, los resultados obtenidos aquí fueron claramente superiores. Esa brecha de aproximadamente en accuracy constituye una ventaja visible, aunque debe interpretarse con cautela, dado que las diferencias en el contexto institucional, la composición de los datos, las variables disponibles y las estrategias de validación pueden influir considerablemente en el rendimiento.

En tercer lugar, respecto a la explicabilidad, los hallazgos muestran que fue posible interpretar de manera consistente la contribución de las variables al comportamiento del modelo predictivo. Este componente adquiere especial relevancia en el ámbito universitario, donde la adopción de sistemas inteligentes no debería justificarse únicamente por su precisión, sino también por su capacidad de ofrecer explicaciones comprensibles y útiles para la toma de decisiones. En este sentido, los resultados se alinean con (S. M. Lundberg & Lee, 2017), quienes demostraron que SHAP permite estimar la contribución individual de cada variable en modelos complejos.

De forma complementaria, (Khosravi et al., 2022; Ribeiro et al., 2016) mostraron que las técnicas de XAI permiten identificar variables clave asociadas al rendimiento académico y al riesgo de abandono. Del mismo modo, (Albreiki, 2022) utilizó enfoques explicativos para analizar factores asociados al abandono en entornos educativos digitales, mientras que (Subhiksha, 2026) también plantea que los modelos explicativos pueden fortalecer los sistemas de alerta temprana orientados a prevenir la DEU. La presente investigación coincide con estos aportes en cuanto al valor de la explicabilidad, pero introduce un matiz importante: la

explicabilidad no fue tratada como una capa secundaria añadida al final del proceso, sino como una dimensión central del modelo propuesto. Esto significa que el objetivo no fue solo alcanzar predicciones precisas, sino también transformar el comportamiento interno del modelo en evidencia interpretable y útil para la gestión institucional.

Desde esta perspectiva, la principal fortaleza del estudio no radica únicamente en identificar qué estudiantes tienen mayor probabilidad de abandonar sus estudios, sino en esclarecer por qué el modelo asigna ese nivel de riesgo y qué variables explican con mayor intensidad dicha decisión. Este aspecto resulta esencial en contextos universitarios, ya que permite diseñar intervenciones más focalizadas, reducir la opacidad de los modelos de caja negra y respaldar las decisiones institucionales con criterios más transparentes.

Además, los resultados permiten discutir críticamente una de las tensiones más frecuentes en la literatura sobre inteligencia artificial explicable: la aparente oposición entre la precisión predictiva y la interpretabilidad. En el trabajo de (Doshi-Velez & Kim, 2017) se advierte que el aumento de la interpretabilidad puede implicar una reducción del rendimiento cuando se sustituyen modelos complejos por otros intrínsecamente interpretables, pero más simples. En el presente estudio, esa tensión no se manifestó de forma directa. En lugar de renunciar a modelos de alto rendimiento, se aplicaron técnicas de explicabilidad posentrenamiento, lo que permitió mantener un desempeño sobresaliente y, al mismo tiempo, obtener interpretaciones detalladas del papel de las variables. El caso de ET resulta ilustrativo: fue posible mantener una precisión de 0.998, un F1-score de 0.892, un recall de 1.000 y una precisión de 0.971, mientras se generaban explicaciones consistentes sobre la contribución de los factores asociados a la DEU. Esto muestra que, al menos en este estudio, la explicabilidad no redujo la capacidad predictiva del modelo, sino que la amplió en su utilidad analítica y práctica.

En síntesis, la discusión desarrollada permite afirmar que el aporte principal de esta investigación no se limita a confirmar que el ML puede predecir la deserción estudiantil, sino a demostrar que dicha predicción puede construirse con rigor metodológico, evaluarse mediante métricas adecuadas, contrastarse críticamente con la literatura y explicarse de forma consistente para apoyar la gestión institucional. Los resultados evidencian que las variables académicas mantienen una primacía explicativa sobre las sociodemográficas y socioeconómicas; que los modelos de ensamble basados en árboles muestran una ventaja clara frente a los modelos lineales; y que, entre ellos, ET emergió como el algoritmo más sólido al ofrecer el mejor equilibrio entre F1-score (0.892), recall (1.000), precision (0.971), ROC-AUC (1.000) y especificidad (1.000). En consecuencia, el estudio no solo aporta un modelo predictivo eficaz, sino también una propuesta analítica más robusta, transparente y aplicable para comprender e intervenir en el fenómeno de la DEU.

CONCLUSIONES

La presente investigación permitió diseñar un modelo de explicabilidad de la DEU, basado en técnicas de ML y en un enfoque estructurado de XAI

En relación con el OE1/HE1, se consolidó un dataset robusto compuesto por 17 183 registros provenientes del SGA-UTM, el cual fue sometido a un proceso riguroso de preparación de datos que incluyó preprocesamiento, análisis exploratorio, imputación diferenciada, transformación estadística y codificación controlada de variables. La combinación de métodos estadísticos y algoritmos de selección de características, entre los que destacan la correlación de Pearson, Cramér's V, ANOVA, Information Value/Weight of Evidence (IV/WOE), RF y AIC, permitió depurar las características, reducir la redundancia y fortalecer la coherencia del conjunto de datos. Este proceso no solo optimizó la calidad del dataset, sino que también garantizó su idoneidad para el modelado predictivo y minimizó posibles sesgos derivados de la redundancia o de la colinealidad.

Respecto al OE2/HE2, los resultados confirmaron que la selección algorítmica influye de manera determinante en la efectividad predictiva del modelo. Los métodos de ensamble basados en árboles superaron de forma consistente a los modelos individuales, especialmente en escenarios con clases desbalanceadas. La combinación de técnicas de balanceo (Oversampling), optimización de hiperparámetros y calibración del umbral probabilístico permitió alcanzar el mejor rendimiento global del sistema. En el conjunto de prueba independiente, el algoritmo ET obtuvo el desempeño más destacado, alcanzando un F1-score de aproximadamente 0.892, junto con un recall de 1.000, una precisión equilibrada y un AUC-ROC cercano a 1.0, lo que evidencia una elevada capacidad discriminativa del modelo para identificar a estudiantes en riesgo de deserción. Adicionalmente, la reducción de las brechas entre los conjuntos de entrenamiento y validación, junto con la convergencia estable observada en las curvas de

aprendizaje, evidenció un adecuado proceso de generalización y un control efectivo del sobreajuste.

En cuanto al OE3/HE3, la incorporación de algoritmos de explicabilidad global y local, como SHAP y LIME, permitió no solo realizar predicciones precisas, sino también interpretar los factores asociados al riesgo de deserción. A nivel global, variables académicas como `semestre_anulados`, `materias_aprobados`, `promedio_aprobada` y `materias_matriculado` destacaron como los principales predictores explicativos, mostrando un patrón consistente entre ambas técnicas de interpretación. A nivel local, el análisis de distintos escenarios de predicción, incluidos los aciertos y los errores del modelo, permitió identificar las interacciones entre variables que influyen en cada decisión individual. La concordancia observada entre las interpretaciones generadas por SHAP y LIME sugiere que el modelo se fundamenta en patrones académicos estructurales consistentes, lo que permite considerarlo una herramienta probabilística de apoyo preventivo

En términos generales, los resultados demuestran que el diseño de un modelo explicativo basado en ML permite comprender los factores asociados a la DEU. La integración de predicción robusta, validación rigurosa y explicabilidad estructurada no solo mejora la detección temprana del riesgo de abandono, sino que también aporta evidencia interpretable que puede apoyar la toma de decisiones institucionales basadas en datos.

RECOMENDACIONES

Se recomienda orientar futuras investigaciones hacia la incorporación de transformers en el análisis de la DEU, debido a su potencial para modelar relaciones complejas entre variables y mejorar el desempeño predictivo en escenarios educativos de alta heterogeneidad. Esta línea de trabajo debería complementarse con técnicas de explicabilidad, a fin de garantizar que los resultados no solo sean precisos, sino también interpretables y útiles para la toma de decisiones institucionales. Asimismo, se sugiere ampliar el conjunto de variables analizadas para fortalecer progresivamente la capacidad predictiva y explicativa del modelo en distintos contextos de educación superior

Un aspecto fundamental es mantener la privacidad de la información estudiantil y garantizar el uso ético de algoritmos de predicción y de explicabilidad, considerando los principios de privacidad, confidencialidad y consentimiento informado. En esa línea, se debe asegurar que la información se utilice únicamente con fines institucionales y que no promueva la exclusión ni la estigmatización en el ámbito de la educación superior.

Se recomienda también el uso de modelos de simulación para recrear nuevos escenarios que permitan analizar cómo los cambios en los factores académicos, socioeconómicos o institucionales afectan a la DEU. Además, la integración de la predicción, la explicabilidad y la simulación aumentará la capacidad analítica, lo que favorecerá una gestión universitaria más eficiente.

REFERENCIAS BIBLIOGRÁFICAS

- Alban, G., & Mauricio, G. (2019). La deserción universitaria en América Latina: una revisión de causas y consecuencias. *Revista Iberoamericana de Educación Superior*, 10(28), 105–122. <https://doi.org/10.22201/iisue.20072872e.2019.28.582>
- Al-Barrak, M. A., & Al-Razgan, M. S. (2015). Predicting students final GPA using decision trees: A case study. *International Journal of Information and Education Technology*, 6(7), 528–533. <https://doi.org/10.7763/IJiet.2016.V6.745>
- Albreiki, B. (2022). Explainable AI in education: Interpreting predictions for student performance. *International Journal of Artificial Intelligence in Education*, 32(1), 101–117. <https://doi.org/10.1007/s40593-021-00241-9>
- Al-Jallad, M., Hamdan, A., & Bader, A. (2019). Predicting student dropout using machine learning techniques: A case study of a Jordanian university. *International Journal of Advanced Computer Science and Applications*, 10(12), 437–445. <https://doi.org/10.14569/IJACSA.2019.0101256>
- Améstica-Rivas, B., Aranda, C., & González, V. (2020). Análisis de factores asociados a la deserción universitaria en los primeros años de carrera. *Revista Electrónica de Investigación Educativa*, 22(1), 1–15. <https://doi.org/10.24320/redie.2020.22.e01.2303>
- Barallobres, G. (2016). La deserción universitaria: causas, consecuencias y propuestas de intervención. *Revista de Educación Superior*, 45(3), 55–70.
- Bedregal-Alpaca, L., Cornejo-Aparicio, F., & Díaz-Galindo, H. (2020). Modelo predictivo de deserción universitaria basado en aprendizaje automático. *Revista de Tecnología e Innovación*, 5(1), 32–45. <https://doi.org/10.23854/rtiv5n1p32>
- Berka, P., & Marek, L. (2021). Exploring the causes and consequences of student dropout in higher education: A multidimensional perspective. *Education Sciences*, 11(8), 379. <https://doi.org/10.3390/educsci11080379>
- Blundo, C., Biondi, A., & Serrano, J. (2022). Evaluación de modelos predictivos en deserción universitaria: una revisión crítica. *Revista Iberoamericana de Inteligencia Artificial*, 25(72), 33–46. <https://doi.org/10.4116/ai.2022.72.03>
- Bowers, A. J., Sprott, R., & Taff, S. A. (2013). Do we know who will drop out? A review of the predictors of dropping out of high school: Precision, sensitivity, and specificity. *The High School Journal*, 96(2), 77–100. <https://doi.org/10.1353/hsj.2013.0000>
- Caballero-González, Y. A., Romero-Frías, E., & Fernández-Sanz, L. (2022). Towards interpretable machine learning in higher education dropout prediction. *IEEE Access*, 10, 31724–31735. <https://doi.org/10.1109/ACCESS.2022.3160279>
- Cabrera, A., Ávila, H., & González, J. (2006). La deserción universitaria en España: magnitud y variables asociadas. Madrid: Ministerio de Educación y Ciencia.

- Cardona, S., Bernal, R., & Peña, X. (2020). Inversión pública en educación superior y retorno social: evidencias para América Latina. *Economía y Desarrollo*, 17(2), 45-62.
- Chung, J., & Lee, J. (2019). Dropout prediction in MOOCs using learning activity features. *EDM 2019: Proceedings of the 12th International Conference on Educational Data Mining*, 114–123.
- Consejo de Educación Superior [CES]. (2023). Estadísticas de matrícula en instituciones de educación superior del Ecuador 2022–2023. <https://www.ces.gob.ec>
- Coussement, K., Lessmann, S., & Verstraeten, G. (2020). Budget allocation and student retention: The hidden impact of performance-based funding. *Journal of Higher Education Policy and Management*, 42(6), 649-666. <https://doi.org/10.1080/1360080X.2020.1731948>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Dass, L., Suárez, D., & Chen, Y. (2021). Understanding dropout patterns in Latin American universities: A data mining approach. *Journal of Educational Data Science*, 2(1), 45–61.
- Deist, T. M., Dankers, F. J. W. M., Ojha, P., Scott Marshall, M., Janssen, T., Faivre-Finn, C., ... & Lambin, P. (2019). Simulation modelling in health care: Applications and lessons learned from other industries. *British Journal of Radiology*, 92(1100), 20180785. <https://doi.org/10.1259/bjr.20180785>
- Deist, T. M., Patti, F., Wang, Y., Krane, D., & Dekker, A. (2019). Simulation in machine learning: Applications and perspectives. *Computer Methods and Programs in Biomedicine*, 173, 41–49. <https://doi.org/10.1016/j.cmpb.2019.02.012>
- Díaz, G., Romero, C., & Ventura, S. (2020). A review of early warning systems in educational data mining: The case of university dropout prediction. *IEEE Access*, 8, 147419–147431. <https://doi.org/10.1109/ACCESS.2020.3014614>
- dos Santos de Assis, J., Vieira, F., & Silva, T. (2022). Higher education dropout in Latin America: Systematic review and future research agenda. *Journal of Educational Research and Practice*, 12(3), 87–103. <https://doi.org/10.5590/JERAP.2022.12.3.07>
- Eegdeman, I., Bakker, A., & Steenbeek, H. (2022). Early warning systems for reducing student dropout in higher education: A systematic review. *Educational Research Review*, 36, 100455. <https://doi.org/10.1016/j.edurev.2022.100455>
- Espinosa Castro, Y., & Mariño Castro, J. (2020). Causas de la deserción universitaria y su impacto en el rendimiento académico: Estudio de caso. *Revista Educación y Desarrollo Social*, 14(2), 29–42. <https://doi.org/10.18359/reds.4264>

- Felderer, M., Körner, M., & Winkler, D. (2022). Challenges and future directions for machine learning in dropout prediction. *Educational Technology Research and Development*, 70, 1421–1440. <https://doi.org/10.1007/s11423-021-10028-w>
- Freitas, R., Lacerda, A., & Correia, P. (2020). Dropout analysis at the Federal Institute of Ceará: A case study. *Revista Brasileira de Educação Profissional e Tecnológica*, 22(3), 1-15.
- González-Nieto, N. A., & Rodríguez-Hernández, C. F. (2023). Educación superior y retención estudiantil: Retos de la universidad contemporánea. *Education in the Knowledge Society*, 24, e31018. <https://doi.org/10.14201/eks.31018>
- Hernández-Orallo, J., Martínez-Plumed, F., & Schmid, U. (2019). Simulation as a scientific tool: The role of digital twins in education and research. *AI & Society*, 35(3), 669–680. <https://doi.org/10.1007/s00146-019-00883-4>
- Hernández-Sampieri, R., & Collado, C. (2018). *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta* (6.ª ed.). McGraw-Hill Education.
- Hernández-Sampieri, R., & Fernández-Collado, C. (2017). *Fundamentos de metodología de la investigación* (4.ª ed.). McGraw-Hill Education.
- Hernández-Sampieri, R., & Mendoza Torres, C. P. (2021). *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta* (7.a ed.). McGraw-Hill Education
- Heuillet, A., Cakir, A., & Le Grand, B. (2021). Explainability in deep learning for time-series data: A survey. *Digital Signal Processing*, 115, 103056. <https://doi.org/10.1016/j.dsp.2021.103056>
- Huo, Y., Wang, X., & Chen, L. (2020). Interpretable machine-learning approaches for early prediction of student dropout in higher education. *IEEE Access*, 8, 137 853-137 862.
- Kelleher, J. D., & Tierney, B. (2022). *Ciencia de datos* (English UC Language Center, Trad.; reimpresión noviembre 2022). Ediciones Universidad Católica de Chile. (Trabajo original publicado en 2018).
- Kuh, G. D., Kinzie, J., & Buckley, J. A. (2020). What matters to student success: A review of the literature. National Postsecondary Education Cooperative. Recuperado de https://nces.ed.gov/ipeds/data/ipedsreports/2020/ipeds2020_whatmatters.pdf
- Kumar, R., Alhashmi, S., & Tripathi, P. (2017). A study on academic risk factors leading to university dropout. *International Journal of Educational Research*, 82, 34–41. <https://doi.org/10.1016/j.ijer.2017.01.001>
- Leal, J., Teixeira, C. A., & da Silva, J. F. (2022). Explainable machine learning models in education: Understanding student dropout risk. *Education and Information Technologies*, 27(4), 4781–4797. <https://doi.org/10.1007/s10639-021-10749-4>

- Lee, S., Han, J., & Whang, K.-Y. (2021). Machine learning in educational data mining: Recent applications and developments. *Computers & Education: Artificial Intelligence*, 2, 100028. <https://doi.org/10.1016/j.caeai.2021.100028>
- Lee, S., Park, H., & Lee, H. (2021). Deep learning-based dropout prediction model in higher education. *IEEE Access*, 9, 52666–52675. <https://doi.org/10.1109/ACCESS.2021.3069992>
- Li, J., & Carroll, C. (2020). Understanding student dropout and persistence in higher education: A comprehensive review. *Journal of College Student Retention: Research, Theory & Practice*, 22(1), 1–24. <https://doi.org/10.1177/1521025119873932>
- Li, Y., & Carroll, M. (2020). Factors influencing student dropout in higher education: A predictive modeling approach. *International Journal of Educational Technology in Higher Education*, 17(1), 1–22. <https://doi.org/10.1186/s41239-020-00189-7>
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 4765–4774). https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf
- Lykourantzou, I., Giatrakos, N., & Gunopulos, D. (2009). Dropout prediction in higher education: A multidimensional approach. *Journal of Educational Data Mining*, 1(3), 1–15.
- Márquez, J. (2020). Aplicaciones del aprendizaje automático en la predicción educativa: Avances y desafíos. *Revista Iberoamericana de Tecnología Educativa*, 16(1), 33–47. <https://doi.org/10.21556/edutec.2020.16.1.1991>
- Mena, M., Miranda, L., & Bravo, C. (2021). Abandono universitario en Europa y América Latina: Un análisis comparativo. *Revista Iberoamericana de Educación Superior*, 12(34), 65–82. <https://doi.org/10.22201/iisue.20072872e.2021.34.585>
- Mnyawami, B., Kaijage, S., & Mtey, E. (2022a). Investigating causes of university student dropout using educational data mining: A case of Tanzanian universities. *International Journal of Education and Development using Information and Communication Technology (IJEDICT)*, 18(1), 101–119.
- Mnyawami, B., Kaijage, S., & Mtey, E. (2022b). Application of machine learning algorithms in predicting student dropout: A case study in higher education. *International Journal of Computer Applications*, 184(31), 1–8. <https://doi.org/10.5120/ijca2022921962>
- Mnyawami, D. A., Oluoch, D., Kimwele, M., & Omwenga, E. (2022). Machine learning techniques for predicting student dropout in higher learning institutions. *Education and Information Technologies*, 27, 763–787. <https://doi.org/10.1007/s10639-021-10674-3>

- Moreira da Silva, M., Ribeiro, J. C., & Gomes, A. (2022). Retention strategies in higher education: The Portuguese case. *International Journal of Educational Management*, 36(2), 402–417. <https://doi.org/10.1108/IJEM-08-2020-0405>
- Mutrofin, H., Hastuti, D., & Raharjo, S. T. (2019). Early prediction of student dropout using machine learning and academic data. *Journal of Theoretical and Applied Information Technology*, 97(15), 4187–4199.
- Núñez-Naranjo, C. (2022). Estrategias institucionales para la permanencia estudiantil en educación superior. *Revista Colombiana de Educación Superior*, 21(1), 112–130. <https://doi.org/10.23854/rces.v21i1.3456>
- OCDE. (2019). *Education at a Glance 2019: OECD indicators*. París: Organisation for Economic Co-operation and Development.
- Olaya, D., Vieira, M., & Cantillo, V. (2020). Causas de la deserción universitaria en América Latina: un enfoque institucional. *Revista Latinoamericana de Ciencias Sociales*, 18(3), 275–294.
- Opazo, H., Martínez, J., & Tello, M. (2021). Panorama de la educación superior en América Latina: Retos de acceso y permanencia. *Revista de Políticas Públicas*, 15(2), 22–39.
- Palacios-Pacheco, X., López-Morales, M., & Romero, C. (2018). Educational data mining to predict dropout and failure in undergraduate university studies. *Interactive Learning Environments*, 28(5), 596–610. <https://doi.org/10.1080/10494820.2018.1548483>
- Paramo, P., & Maya, M. F. (2012). Factores asociados a la deserción universitaria: Una propuesta de clasificación. *Revista Colombiana de Educación*, 62, 105–123. <https://doi.org/10.17227/01203916.62rce105.123>
- Paula Seminara, M., & Teresita Aparicio, M. (2018). Retención y abandono en la educación superior: Diagnóstico regional en América Latina. *Revista de Educación Superior en América Latina*, 10(1), 99–116.
- Praseeda, C. K., & Shivakumar, B. L. (2021). Fuzzy particle swarm optimization based feature selection and hybrid kernel distance based possibilistic fuzzy local information C-means clustering for churn prediction in telecom industry. *SN Applied Sciences*, 3, 613. <https://doi.org/10.1007/s42452-021-04576-7>.
- Radovanović, M., Nanopoulos, A., & Ivanović, M. (2021). A survey of reasoning in machine learning: From traditional logic to knowledge graphs. *Artificial Intelligence Review*, 54(3), 1759–1798. <https://doi.org/10.1007/s10462-020-09834-7>
- Raji, I. D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J., & Denton, E. (2024). Explainable machine learning in practice: Pitfalls, challenges, and adoption in education. *Computers & Education: Artificial Intelligence*, 5(1), 100087. <https://doi.org/10.1016/j.caeai.2023.100087>

- Rivero Pérez, L., Medina, J. M., & Aguilera, A. (2014). Técnicas de aprendizaje automático aplicadas a la minería de datos educativos. *Revista de Educación a Distancia*, (41), 1–17. <https://doi.org/10.6018/red/41/3>
- Rodríguez Rodríguez, M. D., Muñoz, R. I., & Torres, R. C. (2021). Fundamentos epistemológicos de la investigación científica: Una revisión crítica. *Revista Educación y Humanismo*, 23(40), 17–31. <https://doi.org/10.17081/eduhum.23.40.4281>
- Russo, G. (2019). Tackling dropout in higher education: Policy insights from international practices. *Higher Education Studies*, 9(3), 12–20. <https://doi.org/10.5539/hes.v9n3p12>
- Sampieri, R. H., Collado, C. F., & Lucio, M. P. B. (2022). *Metodología de la investigación* (7.^a ed.). McGraw-Hill Education.
- Silva, M., Pinto, A., & Abarca, R. (2020). Factores asociados a la deserción universitaria: una revisión sistemática. *Educación y Sociedad*, 41(150), 1–18. <https://doi.org/10.1590/es.150.2020>
- Soares, J., Cruz, A. M., & Carvalho, M. (2020). The socioeconomic impact of student dropout in public universities: Challenges and policy perspectives. *Education Economics*, 28(5), 459–476. <https://doi.org/10.1080/09645292.2020.1800922>
- Suárez-Montes, N., & Díaz-Subieta, R. (2015). Deserción en la educación superior colombiana: Análisis y propuestas de intervención. *Revista Colombiana de Educación*, 68, 231–254. <https://doi.org/10.17227/01203916.68rce231>
- Tinto, V. (2017). Through the eyes of students. *Journal of College Student Retention: Research, Theory & Practice*, 19(3), 254–269. <https://doi.org/10.1177/1521025115621917>
- Tinto, V., & Pusser, B. (2006). Moving from theory to action: Building a model of institutional action for student success. National Postsecondary Education Cooperative Report. U.S. Department of Education. https://nces.ed.gov/npec/pdf/Tinto_Pusser_Report.pdf
- Trochez, J., Guerrero, M., & Bastidas, J. (2020). Análisis de los factores que inciden en la deserción estudiantil universitaria: un enfoque desde la adaptación académica. *Revista Colombiana de Educación Superior*, 22(1), 75–92.
- Tsai, C.-F., Shen, H.-Y., & Luo, J. (2020). Predicting university dropout using machine learning: A case study. *IEEE Access*, 8, 209825–209834. <https://doi.org/10.1109/ACCESS.2020.3039530>
- UNESCO. (2015). *Reformas educativas y expansión de la educación superior: Tendencias mundiales. Informe Mundial de Educación Superior 2015*. París: Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura.

- Vaarma, M., & Li, S. (2024). Predicting student dropout with interpretable machine learning models. *International Journal of Educational Technology in Higher Education*, 21(1), 1–20. <https://doi.org/10.1186/s41239-024-00478-z>
- van Konijnenburg, L., Oudshoorn, N., & Krahmer, E. (2018). Investing in higher education: The economic return of preventing student dropout. *International Journal of Educational Development*, 62, 1–10. <https://doi.org/10.1016/j.ijedudev.2018.01.002>
- Zambrano, D., Cano, J., & Páez, L. (2018). Calidad educativa y abandono escolar: Causas estructurales de la deserción en la educación superior. *Revista Latinoamericana de Estudios Educativos*, 14(2), 55–70.
- Zambrano, D., Gómez, J., & Rodríguez, E. (2018). Modelado predictivo de la deserción estudiantil utilizando técnicas de minería de datos. *Revista Educación en Ingeniería*, 13(25), 62–71.