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

Dirección General de Estudios de Posgrado
Facultad de Química e Ingeniería Química
Escuela Profesional de Ingeniería Química

Análisis de la variabilidad y trazabilidad en los tiempos de retardo de detonadores No Eléctricos de una empresa de explosivos: Una propuesta basada en Python

TESIS

Para optar el Título Profesional de Ingeniero Químico

AUTOR

Giovany Paolo MICUNCO ROMERO

ASESOR

Juana SANDIVAR ROSAS

Lima - Perú

2025



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

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

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

Referencia bibliográfica

Micunco G. Análisis de la variabilidad y trazabilidad en los tiempos de retardo de detonadores No Eléctricos de una empresa de explosivos: Una propuesta basada en Python [Tesis de pregrado]. Lima: Universidad Nacional Mayor de San Marcos, Facultad de Química e Ingeniería Química, Escuela Profesional de Ingeniería Química; 2025.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Giovany Paolo Micunco Romero
Tipo de documento de identidad	DNI
Número de documento de identidad	74097311
URL de ORCID	
Datos de asesor	
Nombres y apellidos	Juana Sandivar Rosas
Tipo de documento de identidad	DNI
Número de documento de identidad	08667298
URL de ORCID	https://orcid.org/0000-0001-8494-2940
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Víctor Manuel Fernández Guzmán
Tipo de documento	DNI
Número de documento de identidad	06690674
Miembro del jurado 1	
Nombres y apellidos	José Manuel García Pantigozo
Tipo de documento	DNI
Número de documento de identidad	06078851
Datos de investigación	
Línea de investigación	No aplica

Grupo de investigación	No aplica
Agencia de financiamiento	Sin financiamiento
Ubicación geográfica de la investigación	Edificio: Universidad Nacional Mayor de San Marcos País: Perú Departamento: Lima Provincia: Lima Distrito: Cercado de Lima Calle: Av. Carlos Germán Amezaga #375 Latitud: -12.056042374026323 Longitud: -77.08323214426966
Año o rango de años en que se realizó la investigación	Abril 2025 – agosto 2025
URL de disciplinas OCDE	Ingeniería química https://purl.org/pe-repo/ocde/ford#2.04.01 Ingeniería de procesos https://purl.org/pe-repo/ocde/ford#2.04.02 Sistemas de automatización, Sistemas de control https://purl.org/pe-repo/ocde/ford#2.02.03

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
(Universidad del Perú, DECANA DE AMÉRICA)
FACULTAD DE QUÍMICA E INGENIERÍA QUÍMICA
Central Telefónica: 619-7000

ESCUELA PROFESIONAL DE INGENIERÍA QUÍMICA

ACTA DE SUSTENTACIÓN DE TESIS

Los suscritos miembros del Jurado Calificador nombrados por el señor Presidente del Comité Ejecutivo del "IV Programa Ciclo Taller de Titulación Profesional por la Modalidad de Tesis y Trabajo de Suficiencia Profesional para la Facultad de Química e Ingeniería Química 2025-I", bajo la Presidencia del **Dr. Víctor Manuel Fernández Guzmán; Ing. José Manuel García Pantigozo (Miembro)** y la **Dra. Juana Sandivar Rosas (Asesora);** habiendo presentado para tal efecto la **TESIS,** titulada **Análisis de la variabilidad y trazabilidad en los tiempos de retardo de detonadores No Eléctricos de una empresa de explosivos: Una propuesta basada en Python,** después de **SUSTENTADA Y APROBADA** la tesis elaborada por el bachiller en Ingeniería Química: **MICUNCO ROMERO, GIOVANY PAOLO;** para optar el **TÍTULO PROFESIONAL** de **INGENIERO QUÍMICO,** acordaron calificarlo con la **NOTA** de:

DIECINUEVE 19

(LETRAS) (NÚMEROS)

Ciudad Universitaria, 16 de agosto del 2025

Dr. VÍCTOR MANUEL FERNÁNDEZ GUZMÁN
Presidente

Ing. JOSÉ MANUEL GARCÍA PANTIGOZO
Miembro

Dra. JUANA SANDIVAR ROSAS
Asesora

Dr. EDER CLIDIO VICUÑA GALINDO
Director de la E.P. de Ingeniería Química

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

CERTIFICADO DE SIMILITUD

Yo, **Juana Sandivar Rosas** en mi condición de asesora acreditado con el Dictamen N°000001-2025-EPIQ-FQIQ/UNMSM de la tesis, cuyo título es: **Análisis de la variabilidad y trazabilidad en los tiempos de retardo de detonadores No Eléctricos de una empresa de explosivos: Una propuesta basada en Python,** presentado por el bachiller: **Micunco Romero, Giovany Paolo,** para optar el título Profesional de Ingeniería Química, 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 5% 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 título correspondiente.

Lima, 26 de agosto del 2025

Firma del Asesora _____

Nombres y Apellidos de la asesora Juana Sandivar Rosas

DNI: 8667298

AGRADECIMIENTOS

A Dios, por darme la bendición de una familia maravillosa que ha sido mi mayor fortaleza en cada etapa de este camino.

A mi universidad, por brindarme no solo los conocimientos técnicos, sino también la capacidad de analizar y resolver problemas con responsabilidad y criterio.

A mi entorno cercano, por sus valiosos consejos y apoyo constante, que han reforzado y complementado mi bienestar personal y profesional.

DEDICATORIA

Dedico este trabajo, con todo mi cariño y gratitud, a mis padres, quienes con su amor, esfuerzo y ejemplo han hecho posible cada uno de mis logros.

ÍNDICE

RESUMEN 4
ABSTRACT 5
I. INTRODUCCIÓN 6
I.1. Introducción 6
I.2. Planteamiento del problema: 9
- Determinación del problema 9
- Formulación del problema 15
I.3. Objetivos 16
I.4. Importancia y alcance de la investigación 17
I.5. Limitaciones de la investigación 19
II. REVISIÓN DE LITERATURA 20
II.1. Marco teórico 20
II.2. Antecedentes del estudio 26
II.3. Bases teóricas 30
II.4. Definición de términos 37
III. HIPÓTESIS Y VARIABLES 44
III.1. Hipótesis 44
III.2. Variables 45
III.3. Operacionalización de las variables 46
IV. MATERIALES Y MÉTODOS 49
IV.1. Área de estudio 49
IV.2. Diseño de investigación 49
IV.3. Población y muestra 49
IV.4. Procedimientos, Técnicas e instrumentos de recolección de información 50
IV.5. Análisis estadístico 57
V. RESULTADOS 59
V.1. Presentación y análisis de los resultados 59
VI. DISCUSIÓN 72
VII. CONCLUSIONES Y RECOMENDACIONES 76
VIII. BIBLIOGRAFÍAS Y ANEXOS 79

RESUMEN

Esta investigación tiene como objetivo analizar la variabilidad y trazabilidad en los tiempos de retardo de detonadores no eléctricos (también llamados detonadores pirotécnicos) fabricados en Planta de Mixtos (fabricados como retardos), en la Planta dosificadora de cargas explosivas (fabricados como detonadores) y la Planta de ensamble (fabricados como producto terminado) de una planta de fabricación de explosivos mediante modelos de aprendizaje automático (Random Forest, Gradient Booster y Redes Neuronales).

Se emplearon datos reales de fabricación (Tabla 7), que incluyen diversos tiempos de retardo (17-1000 ms), longitudes de cilindro (9-42 mm) y tipos de mixtos pirotécnicos (AG, BG, CG, DG).

Por medio de un sistema integrado de análisis de datos basado en Python, se analizó la variabilidad en los tiempos de retardos empleando modelos de aprendizaje automático y redes neuronales donde se obtuvo como resultado lo siguiente: Random Forest – MSE(Error Cuadrático Promedio): 39.2752, MAE (Error Absoluto Medio): 3.6297, R^2: 0.9996; Gradient Boosting - MSE: 47.9773, MAE: 4.2971, R^2: 0.9995 y Red Neuronal - MSE = 55.9307, MAE = 4.4314 ms, R^2 = 0.9994, lo que se puede inferir que Random Forest es el que mejor predice: tiene el menor error promedio (MAE), menor error cuadrático (MSE) y mayor R^2 . GDB y la Red Neuronal que también funciona bien, pero con errores un poco mayor a comparación del Random Forest. Otra conclusión adicional es que nuestra variable objetivo (Tiempo de retardo) se predice con alta precisión.

Palabras clave: detonadores no eléctricos, variabilidad, trazabilidad, aprendizaje automático, Python, control de calidad.

ABSTRACT

This research aims to analyze the variability and traceability of the delay times of non-electric detonators (also known as pyrotechnic detonators) manufactured in the Mixed Compounds Plant (produced as delays), the Explosive Charges Dosing Plant (produced as detonators), and the Assembly Plant (produced as finished products) of an explosives manufacturing facility using machine learning models (Random Forest, Gradient Boosting, and Neural Networks). Real manufacturing data (Table 7) were used, including various delay times (17–1000 ms), cylinder lengths (9–42 mm), and types of pyrotechnic compounds (AG, BG, CG, DG).

Through an integrated data analysis system based on Python, the variability of delay times was analyzed using machine learning models and neural networks, yielding the following results: Random Forest — MSE (Mean Squared Error): 39.2752, MAE (Mean Absolute Error): 3.6297, R^2: 0.9996; Gradient Boosting — MSE: 47.9773, MAE: 4.2971, R^2: 0.9995; and Neural Network — MSE = 55.9307, MAE = 4.4314 ms, R^2 = 0.9994. It can be inferred that Random Forest provides the best prediction performance, with the lowest average error (MAE), lowest squared error (MSE), and highest R^2. Gradient Boosting and the Neural Network also performed well, although with slightly higher errors compared to Random Forest. An additional conclusion is that the target variable (delay time) can be predicted with high accuracy.

Keywords: non-electric detonators, variability, traceability, machine learning, Python, quality control.

I. INTRODUCCIÓN

I.1. Introducción

En la industria minera, los detonadores no eléctricos (pirotécnicos) son componentes esenciales para las operaciones de voladura, donde la precisión en los tiempos de retardo (intervalo en milisegundos entre la iniciación y la detonación) es crucial para optimizar la fragmentación del macizo rocoso, minimizar vibraciones indeseadas y garantizar la seguridad operativa (Agrawal & Mishra, 2018). En minas a tajo abierto, como las operadas en Perú, una sincronización precisa de los detonadores asegura una voladura eficiente que maximiza la extracción de mineral, reduce los costos asociados a la trituración secundaria y minimiza el impacto ambiental, como la generación de polvo y ruido (Romero Paúcar, 2016). Una variabilidad en los tiempos de retardo superior al 5%, según registros de control de calidad de 2024, puede provocar fallos en la secuencia de detonación, incrementando riesgos estructurales y generando pérdidas económicas por rechazos de lotes. Más allá de la minería, los detonadores no eléctricos son fundamentales en la demolición controlada, donde la precisión en los tiempos de retardo evita daños a estructuras aledañas, y en operaciones militares, donde la sincronización es vital para la coordinación táctica en maniobras explosivas (Brunhoferová et al., 2025).

Esta investigación propone un sistema integrado de análisis de datos basado en Python, que combina herramientas estadísticas y algoritmos de aprendizaje automático (Random Forest, Gradient Boosting y Redes Neuronales), para

caracterizar, predecir y reducir la variabilidad en los tiempos de retardo de los detonadores no eléctricos, considerando factores como la humedad del mixto (<0.1%), los días transcurridos, la longitud del cilindro de retardo (9.0-42 mm), el tipo de mixto (AG, BG, CG, DG) y la etapa de producción (plantas de Mixtos, Dosificadora de Cargas Explosivas y Ensamble). Los algoritmos seleccionados fueron elegidos por su capacidad para modelar relaciones no lineales y procesar datos históricos complejos (Raven et al., 2024). Random Forest, al combinar múltiples árboles de decisión, ofrece alta precisión (MAE: 3.6297 ms, R^2: 0.9996) y permite identificar la importancia de variables como el Tiempo Nominal (0.985658) y la Longitud del Cilindro (0.013550), facilitando la detección de factores que generan desviaciones y mejorando la trazabilidad al asociar predicciones con lotes específicos en cada etapa de producción. Gradient Boosting, mediante la corrección iterativa de errores, optimiza las predicciones en entornos dinámicos (MAE: 4.2971 ms), permitiendo ajustes en tiempo real que reducen la variabilidad. Las Redes Neuronales capturan interacciones complejas entre variables como la humedad y el tipo de mixto (MAE: 4.8292 ms), apoyando la identificación de lotes no conformes. Este sistema, implementado con librerías como Scikit-learn, Keras y Pandas, exporta resultados a archivos CSV para rastrear cada lote desde la fabricación hasta el ensamble, fortaleciendo la trazabilidad y permitiendo intervenciones correctivas que reducen rechazos y optimizan la calidad del producto, especialmente en aplicaciones mineras donde la sincronización precisa es crítica.

Por ello, surge la pregunta general:

- ¿Cómo puede un sistema integrado de análisis de datos basado en Python, que combina herramientas estadísticas y algoritmos de aprendizaje automático, analizar la variabilidad y trazabilidad en los tiempos de retardo de los detonadores y así poder implementar ajustes en los parámetros de producción mejorando el control de calidad en las etapas de producción?

Y surgen las preguntas específicas:

- ¿Cuáles son los factores principales que influyen en la variabilidad de los tiempos de retardo de los detonadores en las etapas del proceso de fabricación?
- ¿Qué algoritmos de aprendizaje automático implementados en Python son más adecuados para predecir desviaciones en los tiempos de retardo durante el proceso de fabricación, y así facilitar intervenciones correctivas que mejoren la trazabilidad y el control de calidad?

El objetivo general es desarrollar un sistema integrado de análisis de datos en Python, utilizando herramientas estadísticas y algoritmos de aprendizaje automático, para analizar la variabilidad y trazabilidad en los tiempos de retardo de los detonadores no eléctricos, mejorando el control de calidad en plantas de Mixtos, Dosificadoras de cargas explosivas y ensamble.

Los objetivos específicos son:

- Identificar los factores que generan variabilidad en los tiempos de retardo, utilizando herramientas estadísticas y algoritmos de aprendizaje automático.
- Implementar algoritmos de aprendizaje automático en Python para predecir desviaciones en los tiempos de retardo, permitiendo intervenciones correctivas que fortalezcan la trazabilidad y el control de calidad.

Esta investigación es relevante por su enfoque innovador al aplicar ciencia de datos y aprendizaje automático en la industria de explosivos, un sector tradicionalmente dominado por métodos convencionales (Soori et al., 2024). Al optimizar la calidad de los detonadores, el sistema reducirá costos operativos de producción, ahorro de materias primas y materiales, reducción de lotes no conformes, disminución de horas-hombre, entre otras mejoras inherentes al proceso global de fabricación de detonadores no eléctricos. El alcance abarca las etapas de producción en una planta de explosivos, utilizando datos históricos de tiempos de retardo (enero 2024 - marzo 2025).

I.2. Planteamiento del problema:

- Determinación del problema

  En la industria minera, específicamente en las operaciones de voladuras, el empleo de detonadores no eléctricos es común debido a su bajo costo, simplicidad operativa y resistencia a interferencias electromagnéticas

(Romero, 2016). Sin embargo, un factor crítico de estos sistemas es la dispersión del tiempo de retardo, es decir, la variación entre el tiempo nominal y el tiempo real de detonación en campo.

Estudios como el realizado por Agrawal y Mishra (2018) muestra que la dispersión en el tiempo de retardo afecta significativamente los efectos de la voladura reduciendo la eficiencia de fragmentación del macizo rocoso e incrementando los riesgos operacionales y estructurales del entorno. El estudio concluye que la dispersión temporal de los detonadores pirotécnicos ($\pm$10% para NONEL) tiene un impacto significativo en los efectos de voladura, aumentando la carga efectiva por retardo debido al solapamiento de detonaciones. Esto resulta en vibraciones inducidas en el suelo más altas de lo predicho, lo que puede comprometer la seguridad y la eficiencia en operaciones mineras e incrementando el riesgo de los operadores del proceso y comunidades cercanas provocando malestar entre las partes interesadas.

En el contexto de la minería a tajo abierto, las voladuras generan vibraciones que pueden afectar la seguridad estructural, la estabilidad de taludes y la calidad de vida de comunidades cercanas, generando conflictos sociales y desafíos para la licencia social de las operaciones mineras. Estudios realizados por Acosta et al. (2022) en la mina Cerro Corona (Perú), las comunidades de Hualgayoc (1500 m) y Pilancones (700 m) están expuestas a estas vibraciones, lo que requiere estrategias para minimizar su impacto. Se realizó estudios basados en el análisis de onda

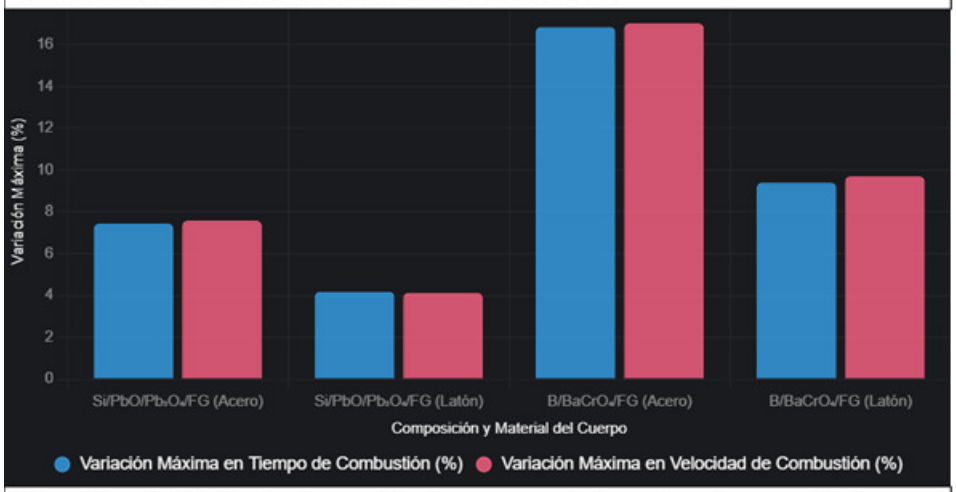
elemental para optimizar los tiempos de retardo entre taladros, reduciendo la Velocidad Pico de Partícula (PPV) en un rango de 3.6% a 27% en puntos críticos. Utilizando equipos de monitoreo de vibraciones y software como Blastware, Cycast y Scilab, se analizaron vibraciones de taladros aislados, determinando retardos óptimos (por ejemplo 25 ms) que promueven interacciones destructivas de ondas, minimizando el PPV sin comprometer significativamente la fragmentación (variación de 3-5% en P80). Estos resultados destacan la importancia de técnicas avanzadas de monitoreo y simulación para garantizar la seguridad y sostenibilidad en operaciones mineras, reduciendo riesgos estructurales y molestias en comunidades aledañas.

Según el estudio realizado por Khan et al. (2023) el tipo de composición pirotécnica tiene un impacto significativo en la variabilidad de los tiempos de combustión. Se obtuvo la conclusión que la mezcla Si/PbO/Pb_3O_4/FG muestra menor dispersión (4.17-7.43%) debido a su estabilidad química, mayor densidad y partículas más finas, lo que la hace más adecuada para aplicaciones que requieren alta precisión. Por otro lado, B/$BaCrO_4$/FG muestra mayor variabilidad (9.39-16.83%) debido a su reacción más rápida y menos controlada, lo que puede limitar su uso en escenarios donde la consistencia es crítica. La combinación de estas composiciones con un material de cuerpo como el latón, que reduce la dispersión al mejorar la disipación de calor, resalta aún más la importancia de

seleccionar la composición adecuada para optimizar el rendimiento y la seguridad en dispositivos pirotécnicos de retardo (Ver Figura 1).

Figura 1

Comparación de Variabilidad en Composiciones Pirotécnicas



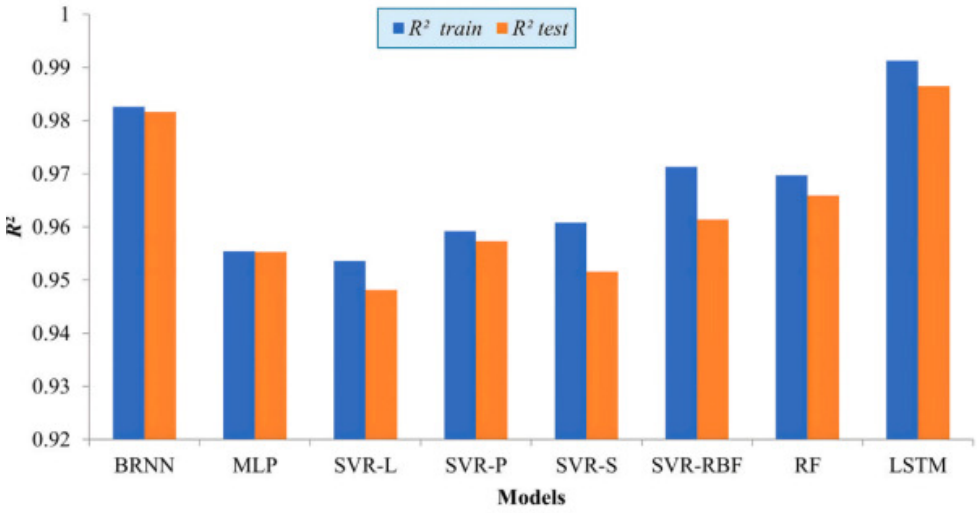
Nota: Adaptado de "Effect of body material and temperature variation on the performance of the time delay pyrotechnic compositions", por Khan, A., Malik, A. Q., Lodhi, Z. H., & Ul Abdin, Z. (2018). Defence Technology, 14(3), 264.

En la etapa de voladura en operaciones mineras, tiene como consecuencia fragmentos de roca (flyrock) que son expulsados de manera incontrolada durante una voladura en una mina o cantera. Estos fragmentos pueden ser proyectados a grandes distancias debido a la energía de la explosión, representando un peligro significativo para la seguridad de los trabajadores, equipos y estructuras cercanas, así como para comunidades aledañas. Ante este punto negativo de seguridad, Taiwo, B. O. et al. (2024) se centra en la aplicación de técnicas de Machine Learning (ML) para predecir el flyrock inducidos por voladuras en operaciones mineras. Se emplearon los algoritmos como Redes Neuronales Artificiales (ANN), Sistemas de Inferencia Neuro-Fuzzy Adaptativos (ANFIS), Perceptrón Multicapa (MLP), Máquinas de Soporte Vectorial (SVM) y Bosque

Aleatorio (RF). El conjunto de datos se dividió (80% entrenamiento, 20% prueba), y los modelos se evaluaron con métricas como el coeficiente de determinación (R^2), Raiz del Error Cuadrático Medio (RMSE), Error Cuadrático Medio (MSE), Error Absoluto Medio (MAE) y Varianza Explicada (VAF). El modelo LSTM logró el mayor rendimiento (R^2 = 0.99, MSE = 0.0001, MAE = 0.0001), con el SVM de kernel RBF Gaussiano superando a otras variantes de SVM (Ver Figura 2). Con este tipo de investigaciones buscan integrar las técnicas de Machine Learning para mejorar la predicción y gestión del flyrock inducido por voladuras, mejorando la seguridad en operaciones mineras a pequeña escala.

Figura 2

Coeficiente de correlación de los 8 modelos desarrollados (entrenados y de prueba)



Nota: Extraído de “Machine learning based prediction of flyrock distance in rock blasting: A safe and sustainable mining approach", por Taiwo, B. O., et al., (2024). Green and Smart Mining Engineering Journal.

Problema Central:

En la fabricación de detonadores no eléctricos, el control de calidad se realiza en tres etapas del proceso: Planta de Mixtos, Planta Dosificadora y Planta de Ensamble. En cada etapa, se mide el tiempo de retardo de los detonadores para verificar que esté dentro del rango nominal especificado. Un lote conforme en una etapa avanza a la siguiente, y un lote no conforme es rechazado. Sin embargo, se presentan inconsistencias: lotes aprobados en la Planta de Mixtos pueden resultar no conformes en la Planta Dosificadora, y lotes conformes en las dos primeras etapas pueden fallar en la Planta de Ensamble. Esto sugiere una variabilidad no controlada en el proceso de fabricación que no se detecta en los controles de calidad actuales.

Estas inconsistencias generan pérdidas económicas por lotes rechazados, retrasos en la producción y potenciales riesgos en la calidad del producto final, lo que podría afectar su desempeño en aplicaciones críticas como la minería.

No existe información documentada sobre casos similares en la industria de explosivos durante la fabricación, solo en aplicaciones en mina, lo que dificulta la comparación con estándares externos. Además, se requiere recopilar y estructurar datos históricos de las tres plantas para alimentar los modelos predictivos.

- Formulación del problema
    - Problema General:

      ¿Cómo puede un sistema integrado de análisis de datos basado en Python, que combina herramientas estadísticas y algoritmos de aprendizaje automático, analizar la variabilidad y trazabilidad en los tiempos de retardo de los detonadores y así poder implementar ajustes en los parámetros de producción mejorando el control de calidad en las etapas de producción?
    - Problemas específicos:
        - ¿Cuáles son los factores principales que influyen en la variabilidad de los tiempos de retardo de los detonadores en las etapas del proceso de fabricación?
        - ¿Qué algoritmos de aprendizaje automático implementados en Python son más adecuados para predecir desviaciones en los tiempos de retardo durante el proceso de fabricación, y así facilitar intervenciones correctivas que mejoren la trazabilidad y el control de calidad?

I.3. Objetivos

- Objetivo General:
  - Desarrollar un sistema integrado de análisis de datos en Python, utilizando herramientas estadísticas y algoritmos de aprendizaje automático (Random Forest, Gradient Boosting y Redes Neuronales), para analizar la variabilidad y trazabilidad (basado en archivos CSV) en los tiempos de retardo de los detonadores no eléctricos, y así poder implementar métodos de intervención como ajustes en tiempo real de parámetros de producción (humedad, longitud del cilindro, Máquina empleada y tipo de mixto) mejorando el control de calidad en las plantas de Mixtos, Dosificadora de Cargas Explosivas y Ensamble.
- Objetivos específicos:
  - Identificar los factores que generan variabilidad en los tiempos de retardo, como la humedad del mixto (<0.1%), los días transcurridos, la longitud del cilindro de retardo (9.0-42 mm), el tipo de mixto (AG, BG, CG, DG) y la máquina empleada, utilizando herramientas estadísticas (ANOVA) y algoritmos de aprendizaje automático (Importancia de características en Random Forest) en Python en base a los datos históricos de producción (enero 2024 - marzo 2025).

- Implementar algoritmos de aprendizaje automático (Random Forest, Gradient Boosting y Redes Neuronales) en Python para predecir desviaciones en los tiempos de retardo con alta precisión (MAE < 5 ms), permitiendo intervenciones correctivas específicas (Tipo de mixto, longitud del cilindro, humedad del mixto,etc.) y facilitando la trazabilidad basado en Python con Pandas que genere archivos CSV para registrar y rastrear los tiempos de retardo predichos y reales por lote, fortaleciendo el control de calidad en cada etapa de producción.

I.4. Importancia y alcance de la investigación

- Importancia de la investigación: La presente investigación busca ampliar los conocimientos sobre la variabilidad de los tiempos de retardo en detonadores no eléctricos mediante una propuesta innovadora en Python para poder analizar por medio de algoritmos Machine Learning, los tiempos de retardo identificando patrones complejos y poder caracterizar, predecir y reducir la variabilidad en los tiempos de retardo de detonadores no eléctricos.
- Justificación académica: Este trabajo desarrolla un sistema integrado de análisis de datos en Python que combina técnicas estadísticas y machine learning para evaluar la variabilidad en los tiempos de retardo de los detonadores no eléctricos, aportando una metodología novedosa que puede servir como base para investigaciones futuras en procesos industriales. A

pesar de que no existe información documentada sobre casos similares en la industria de explosivos durante la fabricación, solo en aplicaciones en mina, este estudio abre la brecha del estudio del control de calidad, llenando un vacío en la literatura.

- Contribución: Tras la implementación de los hallazgos, se espera una reducción en la variabilidad de los tiempos de retardo, lo que podría disminuir los lotes bloqueados en el polvorín, optimizar la eficiencia operativa y mejorar la seguridad en el despacho de productos conformes. Estas mejoras potenciales refuerzan la aplicabilidad práctica del estudio en la industria.
- Alcance de la Investigación

Acerca del Proceso: Esta investigación describe las tres etapas de producción en cadena de la empresa de explosivos. Las plantas en mención son la Planta de Mixtos, Dosificadora y Ensamble.

Datos Utilizados: un volumen de 10,000 registros aproximadamente, las variables clave (tiempos de retardo, longitud del cilindro, etc.).

Herramientas: Se emplea Python, donde se ejecuta herramientas estadísticas (ANOVA, CV) y algoritmos de ML.

Límites Temporales y Geográficos: Se utilizan datos de tiempos del período enero 2024 - marzo 2025. La ubicación de estos datos fueron recolectados en la ciudad de Lima, Perú.

Exclusiones: Se aclara que no se analizan detonadores electrónicos ni impactos económicos y/o ambientales, dejando estas áreas como posibles investigaciones futuras.

I.5. Limitaciones de la investigación

La investigación cuenta limitaciones debido a la confidencialidad de datos sensibles y a la delicadeza del rubro, como la composición química de los mixtos pirotécnicos, lo que justifica el uso de datos de longitud de cilindros de retardo (9.0-42 mm), tipo de mixto (nombres de mixtos inventados), datos simulados de humedad (<0.1%), y etapa de producción (Planta).

La base de datos aplicada para el análisis no será compartida debido a la información sensible que conlleva. No obstante, la metodología es aplicable a datos de tiempos reales disponibles, asegurando su validez. La calidad de los resultados depende de la precisión de los datos históricos, requiriendo un proceso riguroso de limpieza y organización.

La medición de los tiempos de retardo de los detonadores no eléctricos se da en condiciones de humedad relativa mayores a 60%, debido a las bases de seguridad donde indica que mediciones de tiempos menores a la humedad relativa indicada, aumenta las probabilidades de una detonación inesperada por estática.

II. REVISIÓN DE LITERATURA

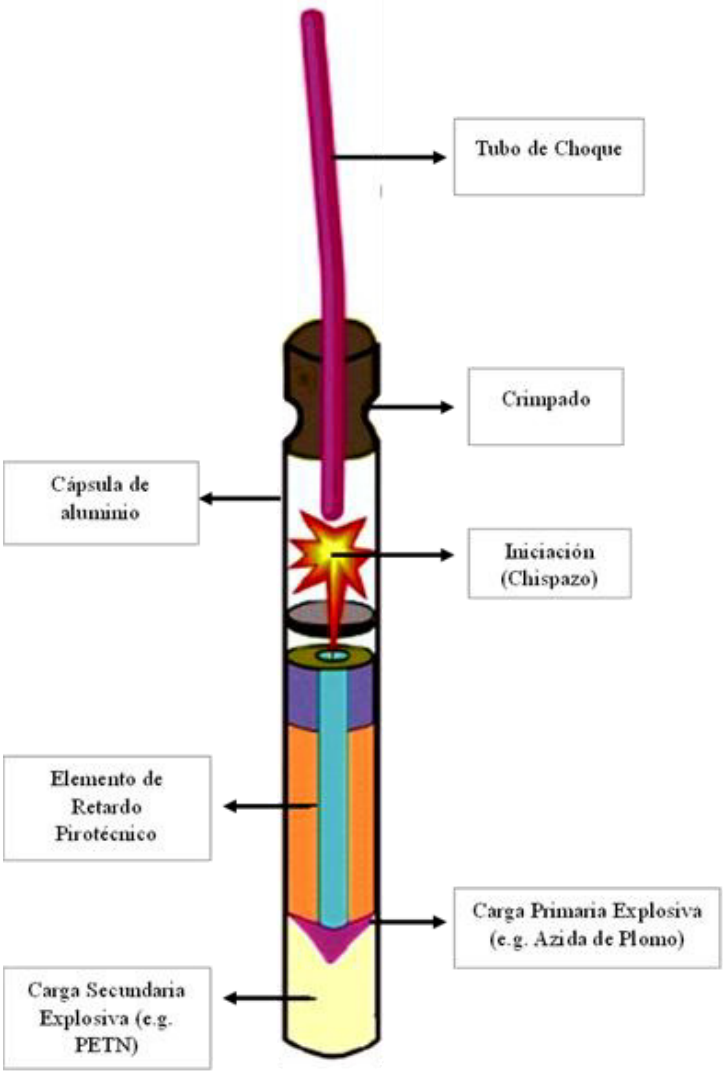
II.1. Marco teórico

Detonadores Pirotécnicos o Detonadores No Eléctricos

Los detonadores no eléctricos (pirotécnicos) son dispositivos esenciales en la industria minera, demolición y operaciones militares, donde la precisión en los tiempos de retardo (intervalo entre ignición y detonación) optimiza la fragmentación del macizo rocoso y minimiza riesgos operativos (Agrawal & Mishra, 2018). Estos detonadores (Ver Figura 4), compuestos por una mezcla iniciadora y retardante alojadas en un cilindro de aluminio, utilizan una onda de choque (tecnología de tubo de choque) para activar explosivos primarios (azida de plomo) y secundarios (PETN, TNT), intensificando la detonación (Brunhoferová et al., 2025). La variabilidad en los tiempos de retardo, influenciada por factores químicos, físicos y procesales, afecta la seguridad y eficiencia, motivando el desarrollo de sistemas de control de calidad basados en Machine Learning (ML).

Figura 4

Partes de un Detonador No Eléctrico (Pirotécnico)



Nota: Adaptado de "*Reliability analysis for mine blast performance based on delay type and firing time* ", por Silva, J., et al., (2018). International Journal of Mining Science and Technology.

Composición Química y Cinética de Combustión

Las mezclas pirotécnicas, como $Zn/KMnO_4$, $Si/PbO/Pb_3O_4/FG$ y $B/BaCrO_4/FG$, determinan la cinética de combustión y los tiempos de retardo. En $Si/PbO/Pb_3O_4/FG$, la descomposición de Pb_3O_4 ($Pb_3O_4 \rightarrow 3PbO + ½O_2$) libera oxígeno que oxida el silicio ($Si + O_2 \rightarrow SiO_2$), generando una combustión sin gas con tasas de 28.01-34.38 mm/s, influenciadas por la temperatura (Khan et al., 2023). A mayor temperatura, la energía de activación disminuye, reduciendo el tiempo de combustión (0.320 s vs. 0.393 s). En $B/BaCrO_4/FG$, la oxidación del boro ($4B + BaCrO_4 \rightarrow 4BO + Ba + Cr$) produce tasas más altas (51.61-52.95 mm/s) y tiempos más cortos (0.208-0.213 s), pero con mayor variabilidad (9.39-16.83%) debido a su reactividad (Khan et al., 2023). $Zn/KMnO_4$ ofrece una alternativa segura, con combustión rápida y óxidos de manganeso a 450°C sin gases (Polis et al., 2023).

Procesos de Fabricación y Variabilidad

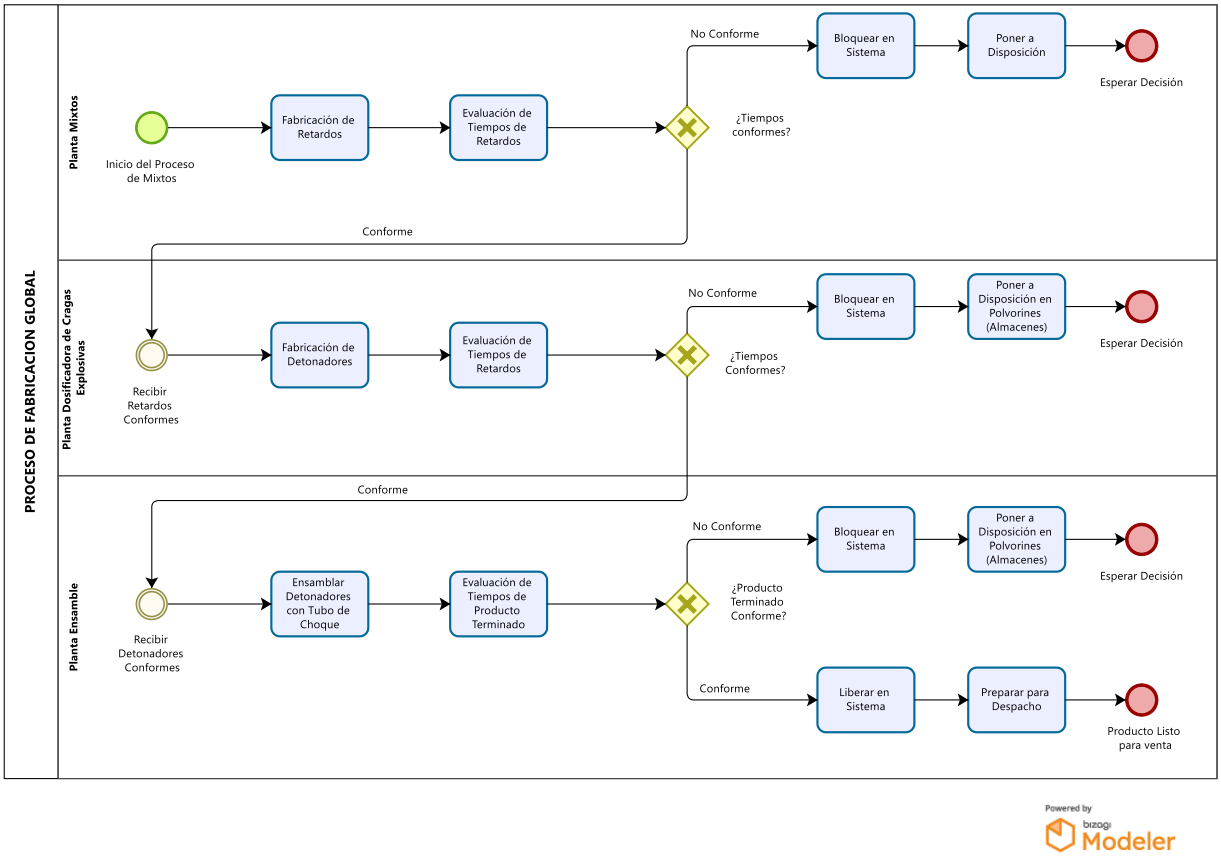
La fabricación de detonadores no eléctricos se estructura en tres etapas interconectadas, ilustradas en Ver Figura 3, un diagrama de bloques que resume el Proceso de Fabricación Global. En la Planta de Mixtos, la primera etapa, el personal de producción arma los elementos de retardo en una planta piloto, ajustando manualmente la altura del mixto en cilindros de aluminio para alcanzar el tiempo nominal. Este paso es crítico, ya que errores en altura o presión pueden propagarse a etapas posteriores, como destaca Khan et al. (2018). Tras evaluar los tiempos, si los resultados son conformes según los rangos especificados, el lote avanza; de lo contrario, se bloquea y se dirige a un Polvorín para decisión final.

La segunda etapa, la Planta Dosificadora de Cargas Explosivas, dosifica las cargas secundarias (PETN) y primarias (azida de plomo) junto al elemento de retardo en una cápsula de aluminio. Se realiza una evaluación intermedia de tiempos para verificar estabilidad, y solo los lotes conformes pasan adelante; los no conformes se bloquean nuevamente.

Finalmente, la Planta de Ensamble completa el proceso al integrar el tubo de choque, realizando un control final de tiempos de retardo. Si el producto cumple las especificaciones, se libera para despacho; de lo contrario, se almacena en el Polvorín. Este flujo asegura trazabilidad y calidad, reflejando la importancia de cada etapa en la reducción de variabilidad.

Figura 3

Diagrama de bloques del Proceso de fabricación Global de Detonadores No Eléctricos



Aprendizaje Automático y Control de Calidad

El aprendizaje automático (ML) desempeña un rol esencial en el control de calidad de los detonadores, modelando relaciones complejas entre variables como Tiempo Nominal, Longitud de Cilindro, y %Humedad para predecir tiempos de retardo con alta precisión. Random Forest, con un MAE de 3.6297 ms y R^2 de 0.9996, destaca por su capacidad para identificar patrones no lineales, asignando un 98.52% de importancia al Tiempo Nominal (Ver Tabla 6). Gradient Boosting (MAE: 4.2971 ms, R^2: 0.9995) y Redes Neuronales (MAE: 4.4314 ms, R^2: 0.9994) complementan este

enfoque, ajustando predicciones a datos históricos (2024-2025) y reduciendo errores en lotes críticos (Raven et al., 2024).

Estos modelos, implementados con Scikit-learn, procesan variables clave para detectar desviaciones y optimizar ajustes en las etapas de fabricación. En este estudio, el ML no solo predice tiempos dentro de los rangos de conformidad, sino que también identifica lotes no conformes, como un tiempo de 31 ms para un tiempo nominal de 42 ms (fuera de 32-45 ms), antes de la Planta Dosificadora, minimizando desperdicios. Taiwo et al. (2024) demuestran que ML, aplicado a voladuras, reduce la variabilidad en tiempos, mejorando la seguridad y la trazabilidad en la Industria 4.0.

Evaluación de Modelos en Contextos Industriales

La evaluación de modelos de ML en procesos industriales requiere métricas como MSE, MAE y (R^2). Raven et al. (2024) destacan que los modelos basados en RL y redes neuronales deben ser validados rigurosamente contra datos operativos para garantizar su robustez, especialmente en aplicaciones de optimización en tiempo real. En la fabricación de detonadores, estas métricas son esenciales para evaluar la precisión de las predicciones de tiempos de retardo, asegurando que los modelos sean confiables en entornos dinámicos.

Desafíos y Consideraciones

Calidad de los Datos: Raven et al. (2024) señalan que los modelos de ML requieren datos históricos de alta calidad, y cambios en las condiciones operativas pueden exigir reentrenamiento. En la fabricación de detonadores, la variabilidad en los datos de

Planta Mixtos o Planta Dosificadora debe ser abordada mediante técnicas de preprocesamiento (Limpieza de datos).

II.2. Antecedentes del estudio

En los procesos industriales se busca constantemente la optimización por medio de diversas rutas. Una de esas rutas es la conocida Industria 4.0 que integra sistemas ciberfísicos (CPS), gemelos digitales (DT), y aprendizaje automático (ML) mediante IoT (Internet of Things), computación en la nube, y Big Data (Schmitt et al., 2020).

En el rubro de los explosivos focalizado a los detonadores, se conoce que los detonadores electrónicos tienen menor dispersión que los detonadores pirotécnicos. Según Romero Paucar (2016) refleja una dispersión temporal de 0.24 ms en detonadores electrónicos frente a 21.88 ms de los pirotécnicos. Estudios similares como el de Silva et al. (2017) llegan a la misma conclusión, indicando que los detonadores pirotécnicos tienen una mayor dispersión (σ=6.195 ms para 700 ms) que los electrónicos (σ=0.342 ms para 700 ms) incluyendo un nivel de confiabilidad para el sistema pirotécnico casi nula (5.08×10^{-22}) debido a la alta variabilidad.

Uno de los principales factores causante de la variabilidad en los tiempos de retardos (pirotécnicos) es la composición química. Por ello se busca mejoras en el proceso o las materias primas utilizadas para la fabricación del pirotécnico. Según Tichapondwa et al. (2021) menciona que la precisión del tiempo de retardo está ligada a la química y el rendimiento de la composición química del pirotécnico.

El empleo de Machine Learning, como redes neuronales y bosques aleatorios, predice la calidad del producto en manufactura basado en medidas históricas, aunque enfrenta retos cuando se tiene datos insuficientes. Uno de los principales objetivos, con el empleo de modelos de aprendizaje automático, es reducir los desperdicios (merma, no conformes) y mejorar la estabilidad del proceso mediante predicciones de calidad. Con ello se optimizaría los tiempos de respuesta en la toma de decisiones (Msakni et al., 2023). Factores como humedad, presión, y material del cuerpo influyen, con el latón reduciendo variaciones en un 44% (Khan et al., 2018).

Modelos de Machine Learning (ML) mejoran la detección de defectos en un 20-29% (Schmitt et al., 2020; Msakni et al., 2023). En Planta de Mixtos, un sistema en Python con Pandas, alineado Machine Learing y Redes neuronales, analiza datos de humedad (<0.1%), Longitud del cilindro de retardo (9.0-42 mm), y tipo de mixto (AG-DG), apoyando la trazabilidad de lotes del retardo con sus respectivas evaluaciones de tiempos.

Raven et al. (2024) destacan la aplicación de algoritmos de ML, como redes neuronales y aprendizaje por refuerzo (RL), en la optimización de procesos industriales complejos, como la unidad de hidrofracturación (HCU). En un caso de estudio, un optimizador basado en RL ajustó las temperaturas promedio ponderadas de lecho (WABTs) en tiempo real, logrando mejoras en los rendimientos de productos de alto valor (+0.6% en queroseno, +0.4% en diésel) y reduciendo productos de menor valor (-0.4% en petróleo pesado) (Raven et al., 2024). Este enfoque es relevante para la fabricación de detonadores, ya que los

modelos de ML pueden predecir y reducir la variabilidad en los tiempos de retardo al analizar datos históricos de las etapas de fabricación (Planta Mixtos, Planta Dosificadora, ensamble final). Sin embargo, Raven et al. (2024) señalan que los modelos de ML requieren datos de alta calidad y reentrenamiento ante cambios en las condiciones operativas, un desafío aplicable al control de procesos en la fabricación de detonadores.

En minería, Agrawal y Mishra (2018) estudiaron la predicción de la Velocidad Pico de Partícula (PPV) inducida por voladuras, utilizando modelos probabilísticos que podrían beneficiarse de técnicas de ML. Sus resultados muestran un RMSE de 1.361 para cuerdas detonantes y 0.8155 para sistemas NONEL, lo que indica la importancia de métricas como MSE y MAE para evaluar la precisión de las predicciones en contextos con alta variabilidad (Agrawal y Mishra, 2018). Este enfoque puede adaptarse para modelar la variabilidad en los tiempos de retardo de detonadores, utilizando algoritmos de ML para correlacionar parámetros de fabricación con el rendimiento del producto.

Msakni et al. (2023) aplicaron redes neuronales y bosques aleatorios para predecir la ubicación de orificios en vigas de parachoques, logrando alta precisión en el control de calidad automotriz. Este enfoque es transferible a la fabricación de detonadores, donde redes neuronales podrían identificar patrones en los datos de las plantas para predecir desviaciones en los tiempos de retardo. Schmitt et al. (2020) utilizaron Gradient Boosted Trees para inspeccionar la calidad de placas de circuitos impresos (PCBs), alcanzando una exactitud del 92.6% y reduciendo el volumen de inspección por rayos X en un 29%. Este precedente sugiere que los

algoritmos de ML pueden optimizar el control de calidad en la fabricación de detonadores, reduciendo la variabilidad y mejorando la trazabilidad.

Tao et al. (2019) destacan el uso de gemelos digitales en la Industria 4.0 para modelar procesos en tiempo real, lo que podría aplicarse a la simulación de las etapas de fabricación de detonadores (Planta Mixtos, Planta Dosificadora, ensamble final) para monitorear y predecir tiempos de retardo. La integración de ML con tecnologías como IoT y gemelos digitales, como se describe en Soori et al. (2024), ofrece un marco para desarrollar sistemas integrados de análisis de datos en Python, alineados con el objetivo de esta tesis.

Estos antecedentes muestran que los enfoques de ML, combinados con métricas estadísticas como MSE, MAE y R^2, son prometedores para abordar la variabilidad en procesos industriales. Sin embargo, desafíos como la calidad de los datos, la aplicabilidad de los modelos y la necesidad de reentrenamiento persisten, especialmente en aplicaciones críticas como la fabricación de detonadores no eléctricos.

Factores de Variabilidad en los Tiempos de Retardo

La literatura científica ha identificado múltiples factores que contribuyen a la variabilidad en los tiempos de retardo. Factores como la composición química, % Humedad, longitud del cilindro y presión empleada juegan un papel fundamental dentro de la variabilidad de los tiempos según Tichapondwa et al. (2021) y Khan et al. (2018). El tamaño de las partículas, la heterogeneidad de la mezcla, la temperatura de calentamiento, la oxidación superficial y las proporciones de la mezcla influyen en esta variabilidad de los tiempos de retardos debido a que estos

factores están fuertemente ligados a la cinética de reacción (Ji et al., 2018). La longitud del cilindro (9-42 mm) y el tipo de mezcla (AG, BG, CG, DG) tienen un impacto menor, según los modelos Random Forest como se visualiza en el Tabla 6 (Ver anexo)

II.3. Bases teóricas

Las bases teóricas de este estudio se centran en la aplicación de aprendizaje automático y herramientas estadísticas para optimizar el control de calidad en la fabricación de detonadores no eléctricos, con un enfoque en la reducción de la variabilidad en los tiempos de retardo. Los fundamentos teóricos clave son:

Aprendizaje Automático en Control de Procesos: El aprendizaje automático permite modelar relaciones no lineales en procesos industriales complejos. Raven et al. (2024) destacan el uso de redes neuronales y RL para optimizar procesos como la HCU, donde los modelos no lineales predicen variables clave (WABTs) basadas en datos históricos. En la fabricación de detonadores, las redes neuronales pueden correlacionar parámetros de las etapas de fabricación (propiedades químicas, condiciones ambientales) con los tiempos de retardo, mejorando la trazabilidad y el control de calidad.

Métricas de Evaluación: Las métricas MSE, MAE y R^2 son esenciales para evaluar la precisión y robustez de los modelos de ML. El MSE penaliza errores grandes, el MAE es robusto frente a valores atípicos, y el R^2 mide la proporción de varianza explicada. Estas métricas son críticas para validar modelos que

predicen tiempos de retardo, asegurando que las predicciones sean fiables en entornos industriales dinámicos (Raven et al., 2024; Agrawal y Mishra, 2018).

Error Cuadrático Medio (MSE)

El Error Cuadrático Medio (Mean Squared Error, MSE) es una métrica comúnmente utilizada para evaluar el rendimiento de modelos de regresión. Representa el promedio de los cuadrados de las diferencias entre los valores predichos por el modelo ($\widehat{y_i}$) y los valores reales (y_i). Su fórmula matemática es:

$$\frac{1}{n}\sum_{i-1}^{n}(y_i - \widehat{y_i})^2$$

donde *n* es el número total de observaciones. El MSE penaliza de manera más significativa los errores grandes debido al término cuadrático, lo que lo hace sensible a valores atípicos. En Python, esta métrica puede calcularse utilizando la biblioteca Scikit-learn mediante la función mean_squared_error del módulo sklearn.metrics.

Error Absoluto Medio (MAE)

El Error Absoluto Medio (Mean Absolute Error, MAE) mide el promedio de las diferencias absolutas entre los valores predichos ($\widehat{y_i}$) y los valores reales (y_i). Su fórmula es:

$$\frac{1}{n}\sum_{i-1}^{n}|y_i - \widehat{y_i}|$$

A diferencia del MSE, el MAE no eleva al cuadrado los errores, por lo que es menos sensible a valores atípicos y proporciona una interpretación más directa del error promedio. En Python, el MAE se calcula con la función mean_absolute_error de Scikit-learn.

Coeficiente de Determinación (R^2)

El coeficiente de determinación (R^2) es una métrica que indica la proporción de la varianza en la variable dependiente que es explicada por el modelo. Su valor oscila entre 0 y 1, donde un $R^2 = 1$ indica un modelo que explica perfectamente la varianza de los datos, y un $R^2 = 0$ indica que el modelo no explica ninguna varianza. La fórmula es:

$$R^2 = 1 - \frac{\sum_{i-1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i-1}^{n}(y_i - \widehat{y})^2}$$

donde $\widehat{y}$ es la media de los valores reales. En Python, el R^2 puede obtenerse utilizando la función r2_score del módulo sklearn.metrics. Es importante destacar que un R^2 negativo puede ocurrir en modelos mal ajustados.

Machine Learning

El aprendizaje automático (machine learning, ML) es una rama de la inteligencia artificial que permite a los sistemas aprender patrones a partir de datos sin programación explícita. En el contexto industrial, el ML es fundamental para modelar relaciones no lineales y optimizar procesos complejos, como la fabricación de detonadores no eléctricos. Según Raven et al. (2024), los algoritmos de ML, como redes neuronales y aprendizaje por refuerzo (RL), son

ideales para aplicaciones en tiempo real. En la fabricación de detonadores, el ML puede utilizarse para predecir tiempos de retardo basándose en variables como la composición química, las condiciones ambientales de las plantas (Planta Mixtos, Planta Dosificadora) y los parámetros de ensamble final. Por ejemplo, modelos supervisados, como redes neuronales, pueden correlacionar datos históricos de fabricación con tiempos de retardo, mientras que el aprendizaje por refuerzo puede optimizar parámetros en tiempo real para reducir la variabilidad (Raven et al., 2024).

En minería, Agrawal y Mishra (2018) destacan la importancia de modelar fenómenos no lineales, como la Velocidad Pico de Partícula (PPV) en voladuras, lo que sugiere que algoritmos de ML pueden mejorar la predicción de tiempos de retardo al integrar datos de sensores en tiempo real. Los enfoques de ML, como los descritos por Soori et al. (2024), también se benefician de la integración con tecnologías de la Industria 4.0, como IoT y gemelos digitales, para monitorear y optimizar procesos en tiempo real, mejorando la trazabilidad y el control de calidad en la fabricación de detonadores.

Redes Neuronales

Las redes neuronales son modelos de aprendizaje automático que procesan datos a través de capas de nodos interconectados, ideales para modelar relaciones no lineales en procesos industriales. En Raven et al. (2024), las redes neuronales se utilizan como base para modelos no lineales en la optimización de procesos de hidrofracturación, donde se entrenan con datos históricos para predecir variables clave como WABTs y optimizar rendimientos de productos. Estas redes son parte

de un enfoque de aprendizaje por refuerzo (RL) que ajusta dinámicamente las condiciones operativas en tiempo real, logrando mejoras significativas en la rentabilidad (aumento en rendimientos de productos de mayor valor y reducción en productos de menor valor). Este enfoque se alinea con aplicaciones en la Industria 4.0, como el mantenimiento predictivo y el control de calidad, descritas en Soori et al. (2024).

En Python, las redes neuronales se implementan con TensorFlow o Keras, como se muestra en el ejemplo:

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense
model = Sequential([
    Dense(64, activation='relu', input_shape=(n_features,)),
    Dense(32, activation='relu'),
    Dense(1)
])
model.compile(optimizer='adam', loss='mse')
```

Código ejemplar del empleo de redes neuronales en Python

Librerías de Python para Análisis de Datos y Machine Learning

El desarrollo de un sistema integrado de análisis de datos en Python para caracterizar y predecir tiempos de retardo en detonadores no eléctricos se basa en varias librerías clave de Python, ampliamente utilizadas en aplicaciones industriales y de investigación:

- Scikit-learn: Esta librería proporciona herramientas para implementar algoritmos de ML y métricas de evaluación. Por ejemplo, las funciones mean_squared_error, mean_absolute_error y r2_score permiten calcular MSE, MAE y R^2, respectivamente, para evaluar modelos que predicen

tiempos de retardo. Además, Scikit-learn ofrece algoritmos como bosques aleatorios y regresión lineal, útiles para análisis iniciales de datos de fabricación (Msakni et al., 2023).

- TensorFlow y Keras: TensorFlow es una plataforma de código abierto para construir modelos de ML, mientras que Keras, integrada en TensorFlow, simplifica la creación de redes neuronales. Estas librerías son ideales para modelar relaciones no lineales en los tiempos de retardo, como se describe en Raven et al. (2024) para la optimización de procesos industriales. Un ejemplo de implementación para predecir tiempos de retardo es:

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense
model = Sequential([
    Dense(64, activation='relu', input_shape=(n_features,)),
    Dense(32, activation='relu'),
    Dense(1)
])
model.compile(optimizer='adam', loss='mse')
```

Código ejemplar para predecir tiempos de retardo en Python

- Pandas: Esta librería es esencial para la manipulación y análisis de datos, permitiendo procesar datos históricos de las etapas de fabricación. Pandas facilita la limpieza, filtrado y agregación de datos, como tiempos de retardo y parámetros de proceso, antes de entrenar modelos de ML (Soori et al., 2024).
- NumPy: Proporciona soporte para cálculos numéricos y manejo de arreglos multidimensionales, fundamental para procesar datos de entrada

en modelos de ML. Por ejemplo, NumPy se utiliza para normalizar datos de tiempos de retardo antes de entrenar redes neuronales.

- Matplotlib y Seaborn: Estas librerías permiten visualizar datos y resultados, como gráficos de dispersión de tiempos de retardo predichos versus reales, o histogramas de variabilidad en las etapas de fabricación. La visualización es clave para interpretar los resultados de los modelos y comunicar hallazgos a los ingenieros de procesos (Raven et al., 2024).
- GBD (Gradient Boosting Decision Trees): Es un algoritmo de aprendizaje automático que construye arboles de decisión de forma secuencial donde cada rabil corrige los errores del anterior optimizando una función de perdida mediante descenso de gradiente.

Industria 4.0 y Trazabilidad: La Industria 4.0 integra tecnologías como IoT, gemelos digitales y ML para optimizar procesos industriales. Tao et al. (2019) describen cómo los gemelos digitales modelan sistemas en tiempo real, un enfoque aplicable a la simulación de las etapas de fabricación de detonadores para monitorear la variabilidad. Soori et al. (2024) enfatizan que las redes neuronales, combinadas con IoT, mejoran el control de calidad en tiempo real, lo que es crucial para garantizar la consistencia de los detonadores.

Optimización en Tiempo Real y Control de Calidad: Raven et al. (2024) destacan que los enfoques de ML, como RL, permiten optimizar procesos en tiempo real, ajustando variables dinámicamente. En la fabricación de detonadores, los algoritmos de ML pueden optimizar parámetros en Planta Mixtos, Planta

Dosificadora y el ensamble final para minimizar la variabilidad en los tiempos de retardo, mejorando la calidad del producto.

Análisis Estadístico en Python: Python ofrece bibliotecas como Scikit-learn, TensorFlow y Pandas para implementar modelos de ML y análisis estadístico. Por ejemplo, las métricas MSE, MAE y R^2 se calculan con mean_squared_error, mean_absolute_error y r2_score de Scikit-learn, mientras que las redes neuronales se implementan con Keras (Raven et al., 2024). Estas herramientas son ideales para desarrollar el sistema integrado de análisis de datos propuesto en el objetivo general.

Estas bases teóricas proporcionan un marco robusto para desarrollar un sistema de análisis de datos que caracterice, prediga y reduzca la variabilidad en los tiempos de retardo de detonadores, alineándose con los principios de la Industria 4.0 y el control de calidad.

II.4. Definición de términos

A continuación, se presentan las definiciones de los términos clave utilizados en la investigación sobre la variabilidad en los tiempos de retardo de detonadores no eléctricos (pirotécnicos), para mayor claridad y precisión en el entendimiento del estudio.

Variabilidad en procesos industriales

La variabilidad es una característica inherente a cualquier proceso de manufactura y puede originarse por diversas fuentes, incluyendo variaciones en las materias

primas, cambios en las condiciones ambientales, diferencias en el funcionamiento de la maquinaria y variaciones en los procedimientos operativos. Esta variabilidad puede afectar la calidad del producto final y debe ser cuidadosamente monitoreada y controlada para cumplir con los requisitos de calidad establecidos (Montgomery, 2020).

Tiempo de Retardo

El tiempo de retardo se refiere al intervalo de tiempo, medido en milisegundos (ms), entre la ignición del mixto pirotécnico en un detonador no eléctrico y la detonación efectiva del explosivo principal (Silva.et. al, 2017). Este parámetro es esencial para sincronizar secuencias de voladura en minería, asegurando una fragmentación óptima y minimizando riesgos operativos como vibraciones excesivas o detonaciones prematuras.

En esta investigación, La variabilidad en los tiempos de retardo de los detonadores es la variable dependiente, analizada mediante correlaciones, gráficos de dispersión, y modelos de aprendizaje automático como Random Forest.

Técnicas de análisis de variabilidad

El análisis estadístico es una herramienta esencial para identificar, cuantificar y reducir la variabilidad en procesos industriales. Técnicas como los gráficos de dispersión, los histogramas, y el control estadístico de procesos (SPC) permiten visualizar la variabilidad, identificar tendencias y tomar decisiones basadas en datos para mejorar la calidad del producto.

En la industria de explosivos, específicamente en la fabricación de detonadores, la aplicación de estas técnicas es esencial con el fin de monitorear los tiempos de

retardo. El análisis estadístico detecta desviaciones en los tiempos establecidos y esto permite garantizar que se mantenga dentro de los límites de control (Evans & Lindsay, 2020).

Trazabilidad

Según la Norma Internacional ISO 9001:2015, la trazabilidad es esencial en el control de calidad industrial, debido a que nos permite mapear el historial, la identificación, la aplicación o ubicación de un producto a lo largo de la cadena de producción.

Aplicando a los detonadores, la trazabilidad asegura que cada unidad pueda ser relacionada con su lote de fabricación de componentes, condiciones de producción y resultados de calidad.

Integrar la trazabilidad desde etapas tempranas, permite una mejor visión de la variabilidad de los tiempos de retardo y reduce los riesgos a fallas del producto terminado.

Herramientas basadas en Python para Análisis de Datos

El análisis de datos es una disciplina clave para la mejora continua en procesos industriales, permitiendo identificar patrones, variabilidad y oportunidades de optimización (McKinney, 2017). Python es una de las principales herramientas para el análisis de datos debido a su versatilidad, flexibilidad y amplia disponibilidad de bibliotecas especializadas.

La aplicación de esta herramienta al análisis de tiempos de retardo en los detonadores permite detectar patrones, evaluar el cumplimiento de las

especificaciones, identificar correlaciones entre lotes y automatizar toda la información en reportes de control de calidad.

Control de Calidad en Procesos Industriales

El control de calidad es un conjunto de técnicas y actividades operacionales que son utilizadas para el cumplimiento de los requisitos de la calidad para un producto o servicio específico. En esta etapa, como control de calidad, permite verificar la estabilidad del proceso, identificar la variabilidad no deseada y tomar decisiones basadas en datos objetivos (Montgomery, 2020).

Aprendizaje Automático (Machine Learning)

El aprendizaje automático (ML) es un campo de la inteligencia artificial que emplea algoritmos para modelar relaciones complejas en datos y realizar predicciones. En esta investigación, se utiliza el modelo de bosques aleatorios (Random Forest), que combina múltiples árboles de decisión para predecir la variabilidad en los tiempos de retardo y evaluar la importancia de variables como humedad, altura de mixto, presión de prensado, y días transcurridos (Msakni et al., 2023). Raven et al. (2024) destacan que modelos como Random Forest y aprendizaje por refuerzo (RL) son efectivos para optimizar procesos industriales complejos, como los petroquímicos, al manejar no linealidades y mejorar la eficiencia en tiempo real. En Planta de Mixtos, Random Forest se emplea en los gráficos de importancia de variables y tendencias temporales, alineándose con enfoques de Industria 4.0 para control de calidad.

Porcentaje de Humedad en los mixtos Pirotécnicos

La humedad es el contenido de agua presente en los mixtos pirotécnicos del detonador, expresado como un porcentaje (%). En planta de Mixtos se controla rigurosamente por debajo del 0.1%., ya que niveles superiores pueden alterar la estabilidad química de los mixtos, incrementando los tiempos de retardo y afectando la fiabilidad del detonador. *Dato basado en procedimientos internos de control de calidad de una empresa del sector de explosivos, no publicados por motivos de confidencialidad.*

Periodo de Evaluación de los Tiempos de Retardo

Los días transcurridos o periodo de evaluación representan el tiempo transcurrido desde la fabricación del detonador hasta su análisis o uso, medido en "días". Este factor puede influir en los tiempos de retardo debido a la degradación química de los mixtos pirotécnicos, especialmente en presencia de humedad o condiciones ambientales adversas (Khan et al., 2018). En esta investigación, los días transcurridos son una variable independiente, analizada mediante gráficos de dispersión y predicciones de Random Forest para identificar tendencias temporales por tipo de mixto (AG-DG).

Altura de Mixto de Retardo

La altura de mixto es la dimensión vertical del material pirotécnico compactado dentro del detonador, medida en milímetros (mm). Esta variable influye en la densidad y la homogeneidad del mixto, afectando directamente la velocidad de combustión y, por ende, el tiempo de retardo. Ji et al. (2018) señalan que variaciones en la altura pueden generar inconsistencias en la propagación de la

llama, lo que incrementa la variabilidad temporal. En este estudio, la altura de mixto se considera como la longitud del cilindro (9.00 – 42 mm) evaluada mediante análisis estadísticos y modelos predictivos para determinar su impacto en los tiempos de retardo.

Mixtos Pirotécnicos (AG-DG)

Los mixtos pirotécnicos son composiciones químicas diseñadas para iniciar la reacción de detonación en los detonadores no eléctricos. En esta investigación, se identifican como AG, BG, CG y DG por motivos de confidencialidad, representando diferentes formulaciones (Al/Zr/$KClO_4$, Zn/$KMnO_4$) con propiedades específicas como velocidad de combustión y sensibilidad ambiental (Polis et al., 2023; Ji et al., 2018). En Planta de Mixtos, los mixtos se aplican en las etapas de producción y su tipo es una variable categórica que influye en los tiempos de retardo, analizada mediante diagramas de caja (boxplots) y tendencias temporales con predicciones de Random Forest.

Error Cuadrático Medio (MSE): Métrica que calcula el promedio de los errores al cuadrado entre valores predichos y reales, utilizada para evaluar la precisión de modelos de regresión en la predicción de tiempos de retardo (Raven et al., 2024).

Error Absoluto Medio (MAE): Métrica que mide el promedio de las diferencias absolutas entre valores predichos y reales, robusta frente a valores atípicos en datos de fabricación (Msakni et al., 2023).

Coeficiente de Determinación (R^2): Indicador que mide la proporción de varianza explicada por un modelo, crucial para validar predicciones en procesos industriales (Schmitt et al., 2020).

Redes Neuronales: Modelos de ML que procesan datos a través de capas de nodos interconectados, ideales para modelar relaciones no lineales en la fabricación de detonadores (Raven et al., 2024).

Aprendizaje por Refuerzo (RL): Técnica de ML donde un agente aprende a optimizar decisiones basadas en una función de recompensa, aplicada en la optimización de procesos industriales (Raven et al., 2024).

Industria 4.0: Paradigma que integra tecnologías como IoT y gemelos digitales para optimizar procesos industriales, mejorando la trazabilidad y el control de calidad (Soori et al., 2024).

Detonadores No Eléctricos: Dispositivos utilizados en minería para iniciar voladuras, cuya variabilidad en los tiempos de retardo afecta la seguridad y eficiencia (Agrawal y Mishra, 2018).

Trazabilidad: Capacidad de rastrear y documentar el historial de un producto a través de las etapas de fabricación, esencial para el control de calidad de detonadores (Soori et al., 2024).

Control de Calidad: Proceso para asegurar que los productos cumplan con especificaciones, utilizando métricas de ML para reducir la variabilidad en tiempos de retardo (Raven et al., 2024).

Días Transcurridos: es la diferencia entre la fecha de evaluación de los tiempos de retado con la fecha de fabricación del Mixto empleado en los retardos.

Crimpar: Acto de unir el tubo de choque y el detonador por medio de la deformación y compresión de ambos materiales.

III. HIPÓTESIS Y VARIABLES

III.1. Hipótesis

Hipótesis General

El desarrollo de un sistema integrado de análisis de datos basado en Python, que combina herramientas estadísticas y algoritmos de aprendizaje automático, permite analizar la variabilidad y trazabilidad de los tiempos de retardo de detonadores no eléctricos mejorando el control de calidad en las plantas de Mixtos, Dosificadora de Cargas Explosivas y Ensamble.

Hipótesis Específicas

- La identificación de los factores que generan variabilidad en los tiempos de retardo, como la humedad del mixto (<0.1%), la longitud del cilindro de retardo (9.0-42 mm), el tipo de mixto (AG, BG, CG, DG) y la etapa de producción, mediante herramientas estadísticas (ANOVA) y algoritmos de aprendizaje automático (Importancia de características en Random Forest), permitirá determinar su impacto en la variabilidad (medido por el coeficiente de variación, CV, objetivo: <5%).
- La implementación de algoritmos de aprendizaje automático (Random Forest, Gradient Boosting y Redes Neuronales) en Python para predecir la desviaciones en los tiempos de retardo con alta precisión (MAE < 5 ms) permitirá reducir la variabilidad (CV < 5%) mediante intervenciones correctivas específicas, como ajustes en la composición del mixto en la Planta de Mixtos; mientras que un sistema de trazabilidad basado en Python con Pandas registrará los tiempos predichos y reales por lote en archivos CSV para todos los lotes

mediante la automatización del registro de datos fortaleciendo el control de calidad en cada etapa de producción.

III.2. Variables

En este apartado se describe los elementos clave que serán medidos y analizados para cumplir con el objetivo general de la tesis: Desarrollar un sistema integrado de análisis de datos en Python, utilizando herramientas estadísticas y algoritmos de aprendizaje automático para analizar la variabilidad y en los tiempos de retardo de los detonadores no eléctricos, y así poder implementar métodos de intervención como ajustes en tiempo real de parámetros de producción mejorando el control de calidad en las plantas de Mixtos, Dosificadora de Cargas Explosivas y Ensamble. Las variables se definen en función de su rol en el estudio y su relación con los algoritmos de aprendizaje automático implementados en el sistema.

a) Variable Independiente

Sistema Integrado de Análisis de Datos en Python: Este sistema consiste en un conjunto de herramientas estadísticas y algoritmos de aprendizaje automático implementados en Python, diseñados para procesar datos de las etapas de fabricación y predecir la variabilidad en los tiempos de retardo de detonadores no eléctricos. Incluye librerías como Scikit-learn para métricas de evaluación (MSE, MAE, R^2), TensorFlow y Keras para modelos de redes neuronales, y Pandas y NumPy para el manejo y preprocesamiento de datos (Raven et al., 2024). La variable independiente se operacionaliza a través de la aplicación del sistema, que

procesa parámetros de entrada (composición química, temperatura, presión) y genera predicciones de tiempos de retardo. El impacto del sistema se mide por su capacidad para reducir la variabilidad, evaluada mediante métricas de precisión como MSE y MAE, y la mejora en la trazabilidad del proceso de fabricación.

b) Variable Dependiente

Variabilidad en Tiempos de Retardo (milisegundos): Representa la dispersión en los tiempos de activación de los detonadores no eléctricos, medida en milisegundos (ms) como una variable numérica continua. Esta variabilidad se calcula como la desviación estándar o el rango de los tiempos de retardo registrados en las etapas de fabricación. La reducción de esta variabilidad es crucial para garantizar la sincronización de voladuras en minería, mejorando la seguridad y eficiencia (Agrawal y Mishra, 2018). El sistema integrado de análisis de datos utilizará algoritmos de ML, como redes neuronales, para predecir y minimizar esta variabilidad, evaluando su desempeño con métricas como MSE, MAE y R^2 (Raven et al., 2024).

III.3. Operacionalización de las variables

El Sistema Integrado de Análisis de Datos en Python procesa las variables de entrada (Días Transcurridos, %Humedad, Tiempo Nominal, Máquina, Mezcla Retardante o Mixto, Longitud Cilindro y datos de tiempos) para predecir los tiempos de retardo y reducir su variabilidad. Los datos, recolectados de las plantas de mixtos, dosificadora y ensamble, se preprocesan utilizando Pandas para limpieza y StandardScaler para normalización. Modelos como Random Forest, Gradient Boosting y redes neuronales (implementados con Scikit-learn y Keras)

generan predicciones, evaluadas mediante métricas como MSE, MAE y R^2 (Raven et al., 2024). La variabilidad en los tiempos de retardo se cuantifica como la desviación estándar de las mediciones reales (Tiempo Retardo), comparada con las predicciones para evaluar la efectividad del sistema. La importancia de las variables de entrada se analiza mediante la métrica de importancia de características de Random Forest, y se generan gráficos de control (X-bar y R) para monitorear la estabilidad del proceso en cada planta.

A continuación, se describe por medio de la matriz de consistencia la operacionalización de las variables (Tabla 1).

Tabla 1

Matriz de consistencia

Problema	Objetivo	Hipótesis	Variables	Metodología	Instrumentos
General: ¿Cómo puede un sistema integrado de análisis de datos basado en Python, analizar la variabilidad y trazabilidad en los tiempos de retardo de los detonadores?	**General**: Desarrollar un sistema integrado de análisis de datos en Python para analizar la variabilidad y trazabilidad en los tiempos de retardo de los detonadores no eléctricos.	**General**: El desarrollo de un sistema integrado de análisis de datos basado en Python permite analizar la variabilidad y trazabilidad de los tiempos de retardo de detonadores no eléctricos.	**Independiente**: Sistema integrado de análisis de datos en Python. **Dependiente**: Variabilidad en tiempos de retardo (milisegundos).	- Estudio retrospectivo, cuantitativo no experimental. - Análisis estadístico y modelado predictivo. - Datos históricos (enero 2024 - marzo 2025).	- Python (pandas, NumPy, scikit-learn). - Gráficos- - CSV para trazabilidad. - Procedimiento operativo
Específico 1: ¿Cuáles son los factores principales que influyen en la variabilidad	**Específico 1**: Identificar los factores que generan variabilidad en los tiempos de	**Específica 1**: La identificación de los factores que generan variabilidad		- Desviación estándar, correlación. - Random Forest para importancia de características.	- Python (pandas, seaborn, scikit-learn). - Gráficos de correlación. - Métricas:

Problema	Objetivo	Hipótesis	Variables	Metodología	Instrumentos
de los tiempos de retardo de los detonadores en las etapas del proceso de fabricación?	retardo utilizando herramientas estadísticas y algoritmos de aprendizaje automático en Python en base a datos históricos de producción.	en los tiempos de retardo mediante herramientas estadísticas y algoritmos de aprendizaje automático, permitirá determinar su impacto en la variabilidad		- ANOVA	Desviación estándar. - Procedimiento operativo
Específico 2: ¿Qué algoritmos de aprendizaje automático implementados en Python son más adecuados para predecir desviaciones en los tiempos de retardo que faciliten la trazabilidad durante el proceso de fabricación?	**Específico 2**: Implementar algoritmos de aprendizaje automático en Python para predecir desviaciones en los tiempos de retardo con alta precisión (MAE < 5 ms) facilitando la trazabilidad durante el proceso de fabricación.	**Específica 2**: La implementación de algoritmos de aprendizaje automático en Python para predecir las desviaciones en los tiempos de retardo con alta precisión (MAE < 5 ms) permitirá reducir la variabilidad (CV < 5%) y fortalecer la trazabilidad durante el proceso de fabricación.		- Modelado: Random Forest, regresión. - Evaluación: MSE, MAE R^2.	- Python (scikit-learn). - Métricas: MSE, MAE R^2. - Procedimiento operativo

IV. MATERIALES Y MÉTODOS

IV.1. Área de estudio

Este estudio se realizó en la recolección de datos de tiempos en las plantas de mixtos, dosificadora y ensamble de una empresa del rubro de explosivos ubicada en Lima, Perú. La data proporcionada de las evaluaciones de tiempos corresponde al periodo enero 2024 a marzo 2025.

IV.2. Diseño de investigación

El estudio adopta un diseño observacional, cuantitativo no experimental, de tipo correlacional y retrospectivo, orientado a analizar la variabilidad en los tiempos de retardo de detonadores en las diferentes etapas de producción, utilizando un sistema desarrollado en Python con Pandas alineado con ML (Raven et al., 2024), y detonadores no eléctricos (Romero Paucar, 2015-2016; Msakni et al., 2023; Tao et al., 2019). Este diseño permite identificar relaciones entre la variable independiente (El sistema integrado de análisis de datos basado en Python) y la variable dependiente (Variabilidad en tiempos de retardo) basándose en datos simulados y reales recolectados de procedimientos operativos internos. Se utilizan técnicas estadísticas (correlaciones, desviaciones estándar) y aprendizaje automático (bosques aleatorios) para modelar la variabilidad, apoyando la trazabilidad de los lotes de estudio.

IV.3. Población y muestra

La población del estudio comprende los datos de producción de detonadores en Planta de Mixtos, Planta Dosificadora y Planta de Ensamble, generados por las máquinas 1, 2 y 3 (Planta de mixtos), abarcando lotes comprometidos. La muestra

incluye datos simulados y reales de tiempos de retardo (17-1000 ms), humedad de mixto (0-0.1%), longitud del cilindro del retardo (9.0-42 mm), tipo de mixto (AG, BG, CG, DG) y etapa de producción, recolectados de procedimientos operativos IV.4. Procedimientos, Técnicas e instrumentos de recolección de información

V.4.1. Metodología

La metodología de esta investigación se basa en un diseño retrospectivo, cuantitativo no experimental, utilizando datos históricos de producción (enero 2024 - marzo 2025) para caracterizar, predecir y reducir la variabilidad en los tiempos de retardo de los detonadores no eléctricos, mejorando la trazabilidad y el control de calidad en las plantas de Mixtos, Dosificadora de Cargas Explosivas y Ensamble. El sistema integrado de análisis de datos se implementa en Python, utilizando herramientas estadísticas y algoritmos de aprendizaje automático (Random Forest, Gradient Boosting, Redes Neuronales).

V.4.1.1. Recolección de Datos

Procedimientos Operativos

Los datos se recopilan en las tres etapas de producción de los detonadores no eléctricos: Planta de Mixtos, Dosificadora de Cargas Explosivas y Ensamble. Los procedimientos operativos son los siguientes:

- Planta de Mixtos: Se prepara la mezcla retardante (AG, BG, CG, DG) bajo condiciones controladas (Humedad relativa<40%). La humedad del mixto (<0.1%) se mide con higrómetros calibrados, longitud del cilindro (altura del mixto) que se consigue con un indicador llamado tiempo de combustión que mide el tiempo expresado en milisegundos

que toma para combustionar en cada milímetro de mixto y la composición química se registra para cada lote. Los datos de tiempos de retardo son evaluados en un equipo llamado "cronógrafo" (cronómetros de alta precisión) se registran a un archivo Excel.

- Dosificadora de Cargas Explosivas: El mixto se prensa en cilindros de retardo, controlando la presión de prensado (80-150 MPa) con sensores digitales y la dosificación de cargas explosivas primarias y secundarias. Los datos de tiempos de retardo son evaluados en un equipo llamado "cronógrafo" (cronómetros de alta precisión) se registran a un archivo Excel.
- Ensamble: Los detonadores se ensamblan, verificando y midiendo los tiempos de retardo reales en pruebas de detonación controlada. Cada lote se identifica con un código único, y los datos de tiempos de retardo son evaluados en un equipo llamado "cronógrafo" (cronómetros de alta precisión) se registran a un archivo Excel.

La medición de estos tiempos se da en condiciones de humedad relativa mayores a 60%, debido a las bases de seguridad donde indica que mediciones de tiempos menores a la humedad relativa indicada, aumenta las probabilidades de una detonación inesperada.

Al tener tres archivos de Excel distintos por planta, se procede limpiar la base de datos y a continuación, se integra mediante el Sistema integrado de análisis de datos en Python,

Datos empleados

Los datos históricos (enero 2024 - marzo 2025) incluyen aproximadamente 10,000 registros de lotes, combinando datos reales y simulados debido a restricciones de confidencialidad (Véase I.5). Cada registro contiene:

- Variable dependiente: La variabilidad en los tiempos de retardo de los detonadores
- Variables independientes: Sistema integrado de análisis de datos en Python que involucra los siguientes factores:
  - ❖ Tiempo Nominal (ms)
  - ❖ Humedad del mixto (categorizada: <0.1%)
  - ❖ Máquina (Categórica)
  - ❖ Longitud del cilindro (mm)
  - ❖ Mezcla Retardante o tipo de Mixto (AG, BG, CG, DG)
  - ❖ Días transcurridos desde la producción del lote.
  - ❖ Etapa de producción (Mixtos, Dosificadora, Ensamble)

Los datos se almacenan en un archivo Excel, procesado con Pandas para generar archivos CSV que registran los tiempos de retardo predichos y reales por lote, asegurando trazabilidad.

V.4.1.2. Análisis Estadístico y Modelado

Análisis Estadístico

Se utilizan herramientas estadísticas en Python (Pandas, NumPy, SciPy) para caracterizar la variabilidad de los tiempos de retardo:

- Dispersión: Cálculo del coeficiente de variación (CV) para cada mezcla retardante (objetivo: CV < 5%), basado en la desviación estándar y la media de los tiempos de retardo (Tabla 2). Se calcularon la media, la desviación estándar, el número de mediciones de los tiempos de retardo. Los resultados indican que la mezcla AG (17–25 ms) tiene una desviación estándar menor en comparación con la mezcla DG (700–1000 ms), lo que refleja mayor estabilidad en tiempos de retardo cortos. El coeficiente de variación calculada como: CV=(Desviación Estándar)/Media×100

  Estos hallazgos son consistentes con estudios previos que destacan la influencia de la composición química en la consistencia de los detonadores (Agrawal y Mishra, 2018).

  Tabla 2

  Coeficiente de Variación por Mezcla Retardante

Mezcla Retardante	Rango Tiempo_Retardo_ms	Desviación Estándar (ms)	Nº Mediciones	Coef. Variación (%)
AG	17 - 25 ms	3.08	1620	15.35
BG	42 - 125 ms	30.31	2243	41
CG	150 - 600 ms	134.35	4832	40.85
DG	700 - 1000 ms	125.17	2476	15.2

- Análisis de Varianza (ANOVA): Aplicación de un ANOVA multifactorial con un nivel de significancia del 5% para comparar las medias de los tiempos de retardo entre los niveles de los factores Tiempo Nominal, Máquina (1, 2, 3), Días Transcurridos (categorizada), Longitud del Cilindro y Humedad (categorizada). La variable Mezcla Retardante no se incluyó debido a la

multicolinealidad con Tiempo Nominal, ya que cada mezcla está diseñada para un rango específico de tiempos nominales, y su efecto es absorbido por Tiempo Nominal en el modelo estadístico. Los resultados muestran:

- Tiempo Nominal ($F = 881265.2556$, $p < 0.0001$, significativo).
- Máquina ($F = 71.3680$, $p < 0.0001$, significativo).
- Humedad ($F = 17.7080$, $p < 0.0001$, significativo)
- Días Transcurridos ($F = 70.7105$, $p < 0.0001$, significativo)
- Longitud del Cilindro ($F = 70.6994$, $p < 0.0001$, significativo).
- Interacción Tiempo Nominal × Máquina ($F = 13.2911$, $p < 0.0001$, significativo)
- Interacción Tiempo Nominal × Humedad ($F = 9.6382$, $p < 0.0001$, significativo)
- Interacción Tiempo Nominal × Días Transcurridos ($F = 16.9519$, $p < 0.0001$, significativo)
- Interacción Tiempo Nominal × Longitud del Cilindro ($F = 0.4389$, $p = 1.0$, no significativo).

Estos resultados indican que Tiempo Nominal, Máquina, Humedad, Días Transcurridos y Longitud del Cilindro, junto con las interacciones Tiempo Nominal × Máquina, Tiempo Nominal × Humedad, y Tiempo Nominal × Días Transcurridos, son factores clave que afectan la variabilidad de los tiempos de retardo.

- Análisis de importancia de características: Se empleó Random Forest para identificar la contribución de los factores (Tiempo Nominal: 0.985658, Longitud del Cilindro: 0.013550, entre otras) a la variabilidad de los tiempos de retardo, complementando el ANOVA al priorizar factores clave para intervenciones correctivas en cada etapa de producción.

Modelado con Machine Learning

Los algoritmos de aprendizaje automático (Random Forest, Gradient Boosting, Redes Neuronales) se implementan con Scikit-learn y Keras:

- Random Forest: Configurado con 100 árboles y criterio de entropía, identifica la importancia de variables (Tiempo Nominal: 0.985658).
- Gradient Boosting: Utiliza 100 estimadores con una tasa de aprendizaje de 0.1, optimizando predicciones iterativamente.
- Redes Neuronales: Arquitectura con 3 capas ocultas (64, 32, 16 neuronas), función de activación ReLU y optimizador Adam.

V.4.1.3. Validación de los Modelos

La evaluación de los modelos se realiza con las siguientes métricas:

- Error Absoluto Medio (MAE): Promedio de los errores absolutos entre los tiempos de retardo predichos y reales, calculado como:

$$\frac{1}{n}\sum_{i-1}^{n}|y_i - \widehat{y_i}|$$

Objetivo: MAE < 5 ms.

- Error Cuadrático Medio (MSE): Promedio de los errores al cuadrado, calculado como:

$$\frac{1}{n}\sum_{i-1}^{n}(y_i - \hat{y}_i)^2$$

Objetivo: Minimizar MSE.

- Coeficiente de Determinación (R^2): Proporción de la varianza explicada, calculado como:

$$R^2 = 1 - \frac{\sum_{i-1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i-1}^{n}(y_i - \widehat{y^{-}})^2}$$

Objetivo: $R^2 > 0.95$.

Los modelos se comparan según MAE, MSE y R^2 utilizando el conjunto completo de datos históricos. Los resultados muestran: Random Forest (MSE: 39.2752, MAE: 3.6297 ms, R^2: 0.9996), Gradient Boosting (MSE: 47.9773, MAE: 4.2971 ms, R^2: 0.9995), y Redes Neuronales MSE = 55.9307, MAE = 4.4314 ms, R^2 = 0.9994). El modelo Random Forest, con el menor MAE (3.6297 ms) y el mayor R^2 (0.9996), se selecciona para las predicciones operativas, integrándose en el sistema de trazabilidad que genera archivos CSV para registrar los tiempos predichos y reales por lote.

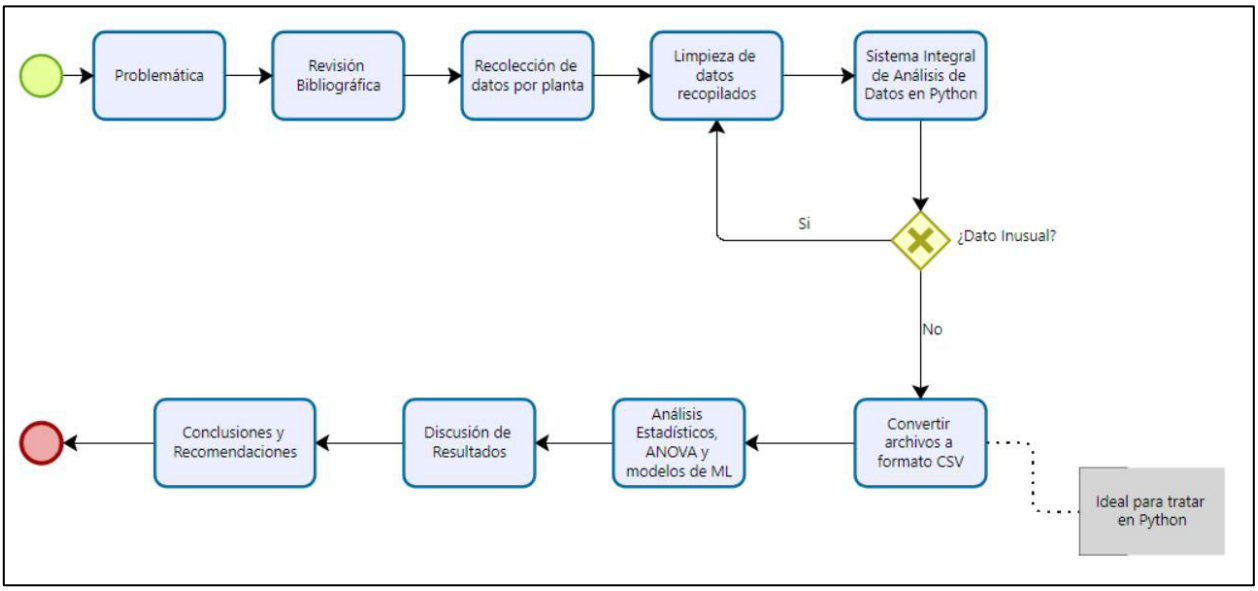
V.4.1.4. Trazabilidad

El sistema de trazabilidad, implementado con Pandas, genera archivos CSV que registran los tiempos de retardo predichos y reales, junto con los factores involucrados, para cada lote. Esto permite rastrear desviaciones desde su

fabricación, reduciendo la variabilidad (CV < 5%), con un índice de trazabilidad del 100% si el sistema se implementa sin problemas técnicos.

Figura 5

Procedimiento empleado para el tratamiento y análisis de tiempos de retardo



Nota: *Elaboración propia.*

IV.5. Análisis estadístico

Para garantizar la validez del análisis y minimizar el impacto de factores externos, se controlarán las siguientes variables:

Modelo de Detonador: Se utilizarán exclusivamente detonadores no eléctricos de un modelo específico, asegurando consistencia en el diseño y materiales.

Condiciones Ambientales Estándar: Las mediciones de tiempos de retardo se realizarán bajo condiciones controladas (temperatura de 25°C aprox., humedad

relativa mayor a 60%) en las plantas de Lima, Perú, para reducir la influencia de factores ambientales externos.

Altura de carga de pirotécnico: Por motivos de confidencialidad, se va a considerar la longitud del cilindro del retardo como la altura de carga pirotécnica empleada.

Lote de Producción: Los datos se analizarán por lotes de retardo para controlar variaciones debidas a diferencias en los lotes de fabricación.

V. RESULTADOS

V.1. Presentación y análisis de los resultados

Los resultados del Sistema Integrado de Análisis de Datos en Python, desarrollado para predecir los tiempos de retardo de detonadores no eléctricos, se derivan de un conjunto de datos de aproximadamente 10,000 registros, abarcando las mezclas retardantes AG, BG, CG, y DG con rangos de Tiempo Nominal de 17 a 1000 ms. Las métricas de evaluación de los modelos, calculadas sobre el conjunto de prueba (20% de los datos), se presentan a continuación:

- Random Forest: MSE = 39.2752, MAE = 3.6297 ms, $R^2 = 0.9996$
- Gradient Boosting: MSE = 47.9773, MAE = 4.2971 ms, $R^2 = 0.9995$
- Red Neuronal: MSE = 55.9307, MAE = 4.4314 ms, $R^2 = 0.9994$

Estas métricas cumplen el objetivo de MAE < 5 ms, con Random Forest mostrando el mejor desempeño. La Tabla 3, con 22 registros representativos, ilustra las predicciones por modelo y sirve como ejemplo:

Tabla 3

Predicciones de Tiempos de Retardo por Modelo

Mezcla Retardante	Tiempo Nominal (ms)	Tiempo Promedio Real (ms)	Desviación Estándar (ms)	Predicción Random Forest (ms)	Predicción Gradient Boosting (ms)	Predicción Red Neuronal (ms)
AG	17	16.97	1.68	17.76	17.54	17.75
AG	25	22.03	1.92	22.30	22.16	22.44
BG	42	38.17	1.78	38.32	38.33	38.49
BG	50	41.05	2.2	41.76	41.15	41.54
BG	75	70.42	2.63	71.44	70.32	70.88
BG	100	94.76	2.55	94.44	94.41	94.21

Mezcla Retardante	Tiempo Nominal (ms)	Tiempo Promedio Real (ms)	Desviación Estándar (ms)	Predicción Random Forest (ms)	Predicción Gradient Boosting (ms)	Predicción Red Neuronal (ms)
BG	125	118.45	3.77	117.45	118.47	116.82
CG	150	146.95	4.42	149.53	146.75	148.93
CG	175	171.86	2.62	172.19	171.71	172.73
CG	200	198.70	4.93	199.73	198.48	198.84
CG	250	243.90	6.71	245.83	245.70	247.39
CG	300	293.40	4.25	294.17	293.70	293.33
CG	350	349.16	5.9	351.00	351.69	348.96
CG	400	392.53	6.74	392.19	392.96	392.48
CG	450	449.44	4.37	449.94	449.76	449.27
CG	500	501.58	4.01	502.46	502.46	500.19
CG	600	594.10	9.88	594.71	593.03	594.43
DG	600	585.24	7.14	588.92	587.73	592.32
DG	700	706.12	11.59	715.17	714.24	715.20
DG	800	790.69	12.17	797.74	793.68	796.50
DG	900	889.51	18.24	898.65	899.18	897.05
DG	1000	979.41	20.91	980.70	981.07	983.09

Nota: *Generado por el Sistema Integral de Análisis de Datos en Python, Micunco, G. (2025)*

El ANOVA multifactorial analizó la influencia de %Humedad, Días Transcurridos, y Longitud Cilindro como factores, con los siguientes resultados (Tabla 4):

Tabla 4

ANOVA Multifactorial

Factor	df	Suma sq	Media sq	F	PR(>F)
Tiempo Nominal	20.0	1.030679e+09	5.153395e+07	881265.2556	0.0
Máquina	2.0	8.346800e+03	4.173400e+03	71.3680	0.0
Humedad	2.0	2.071026e+03	1.035513e+03	17.7080	0.0
Días Transcurridos	2.0	8.269905e+03	4.134953e+03	70.7105	0.0
Longitud del Cilindro	8.0	3.307443e+04	4.134303e+03	70.6994	0.0
Tiempo Nominal: Máquina	40.0	3.108915e+04	7.772287e+02	13.2911	0.0
Tiempo Nominal: Humedad	40.0	2.254462e+04	5.636155e+02	9.6382	0.0
Tiempo Nominal: días Transcurridos	40.0	3.965192e+04	9.912981e+02	16.9519	0.0
Tiempo Nominal: Longitud Cilindro	160.0	4.106863e+03	2.566790e+01	0.4389	1.0
Residual	11083.0	6.481031e+05	5.847720e+01	NaN	NaN

Nota: *Generado por el Sistema Integral de Análisis de Datos en Python, Micunco, G. (2025)*

Interpretación ($\alpha = 0.05$):

- Rechazar H0 para Tiempo Nominal, Máquina, Humedad, Días Transcurridos, Longitud Cilindro, y sus interacciones con Tiempo Nominal ($p = 0.0$), indicando efectos significativos.
- No rechazar H0 para C (Tiempo Nominal):C(Longitud Cilindro) ($p = 1.0$), sugiriendo que esta interacción no afecta significativamente el Tiempo Retardo.

Análisis de Errores por Modelo

Los errores de predicción, calculados sobre 10,000 registros, reflejan el desempeño de cada modelo. Random Forest exhibe un MSE de 39.2752 y un MAE de 3.6297 ms, con R^2 de 0.9996, mostrando alta precisión, aunque errores mayores se observan en rangos largos de DG (Diferencia entre tiempo promedio Real y Predicción Random Forest = 9.14 ms para DG – 900-

Ver Tabla 3). Gradient Boosting presenta un MSE de 47.9773 y un MAE de 4.2971 ms, con R^2 de 0.9995, indicando mayor variabilidad, posiblemente por ajustes menos robustos en interacciones. Redes Neuronales, con MSE de 55.9307 y MAE de 4.4314 ms (R^2: 0.9994), muestran el mayor error, sugiriendo sensibilidad a datos insuficientes para optimización. Todos cumplen el objetivo de MAE < 5 ms, con Random Forest como el más consistente.

Interpretación de Parámetros y su Modelado

La variabilidad en los tiempos de retardo depende de parámetros identificados en el ANOVA (Ver Tabla 4). Tiempo Nominal (F = 881265.2556, p = 0.0) domina, con una importancia de 0.985658 en Random Forest, reflejada en la cercanía a los tiempos reales (979.41 ms vs. 1000 ms para DG). %Humedad (F = 17.7080, p = 0.0) afecta la estabilidad química, aumentando desviaciones en CG y DG (585.24 ms vs. 600 ms), modelada con codificación categórica en Pandas. Longitud Cilindro (F = 70.6994, p = 0.0) modula la presión, influyendo en tiempos largos (DG - 1000), mientras que Días Transcurridos (F = 70.7105, p = 0.0) refleja envejecimiento, impactando DG.

Mezcla Retardante no se incluyó en el ANOVA debido a que su suma de cuadrados resultó negativa, atribuible a la multicolinealidad con Tiempo Nominal. Esto indica que no aporta información adicional significativa en el modelo estadístico, ya que su efecto está casi completamente absorbido por Tiempo Nominal, que captura la mayor parte de la variabilidad en Tiempo Retardo (importancia: 0.985658). Sin embargo, esto no implica que Mezcla Retardante sea irrelevante en el contexto físico o químico de los detonadores no eléctricos, donde sus diferencias (AG, BG, CG, DG) influyen en la detonación, siendo modelada como variable categórica en Random Forest para reflejar estas propiedades.

Estos parámetros se modelan en Random Forest con árboles de decisión que capturan interacciones (Tiempo Nominal × Dias Transcurridos, F = 16.9519, p = 0.0), Gradient Boosting ajusta iterativamente, y Redes Neuronales las aprenden en capas ocultas. La interacción no significativa Tiempo Nominal × Longitud Cilindro (p = 1.0) indica efectos independientes, manejados por Random Forest.

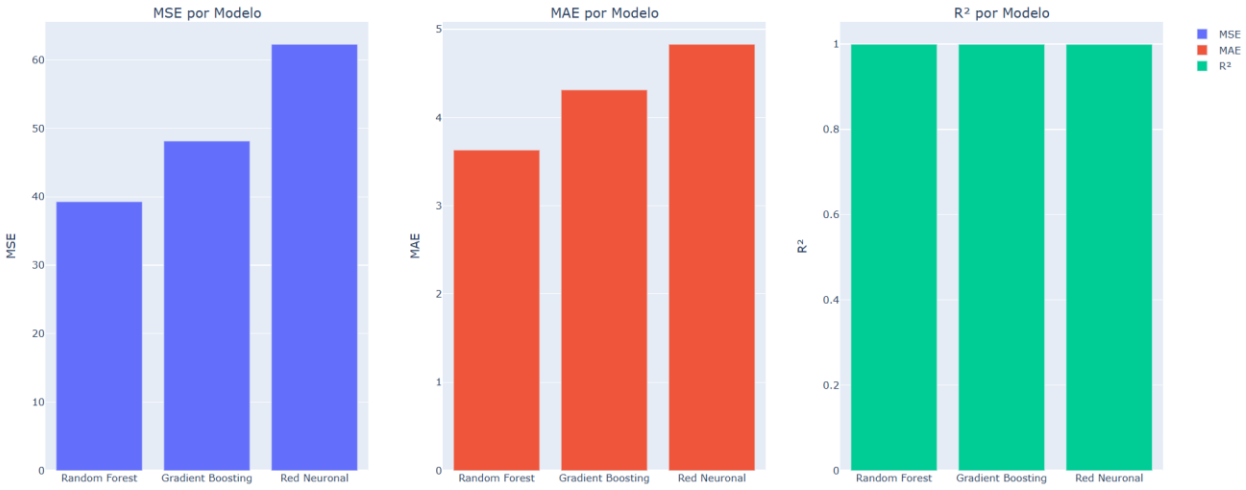
Comparación con Técnicas de Modelado y Estadísticas

- Comparación con Otras Técnicas de Modelado

  Random Forest (MAE: 3.6297 ms, R^2: 0.9996) supera a Gradient Boosting (MAE: 4.2971 ms) y Redes Neuronales (MAE: 4.4314 ms) en precisión y simplicidad, siendo ideal para trazabilidad. Gradient Boosting ofrece mejor ajuste en subgrupos (BG), mientras que Redes Neuronales manejan patrones complejos, pero su MSE más alto (55.9307) indica mayor sensibilidad a datos limitados (Ver Figura 6). La ventaja de Random Forest radica en su capacidad para integrar variables categóricas (Mezcla Retardante) y no como el ANOVA.

Figura 6

Gráfico de barras de Errores por cada Modelo



- Comparación con Técnicas Estadísticas

  Se realizó un modelo de regresión lineal múltiple tomando las primeras 50 filas de la base de datos. La ecuación es la siguiente:

  $$Tiempo\ Promedio\ Real = X + Y \times Tiempo\ Nominal + Z \times \%Humedad + m \times Longitud\ Cilindro + n \times Planta + p \times Maquina + L \times Mezcla\ Retardante + q$$

  Donde:

  - X: intercepto
  - Y: Coef. Tiempo Nominal
  - Z: Coef. %Humedad
  - m: Coef. Longitud Cilindro
  - n: Coef. Planta
  - p: Coef. Maquina
  - L: Coef. Mezcla retardante

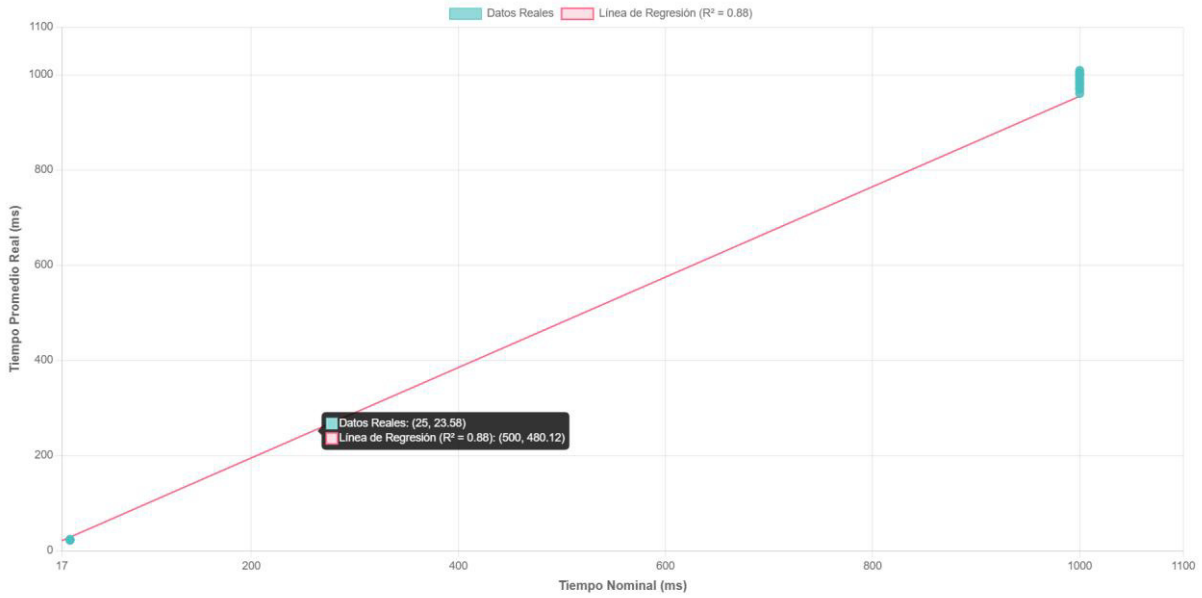
- q: Coef. Residuo.

Este modelo rindió un (R^2) aproximado de 0.88 y un MAE de 8.7 ms, indicando una explicación moderada de la variabilidad en los tiempos de retardo (Ver Figura 7).

Por otro lado, los modelos de machine learning implementados (Random Forest, Gradient Boosting y redes neuronales) superaron significativamente el rendimiento de la regresión lineal. Random Forest obtuvo un (R^2) de 0.9996 y un MAE de 3.6297 ms, Gradient Boosting un (R^2) de 0.9995 y un MAE de 4.2971 ms, y las redes neuronales un (R^2) de 0.9994 y un MAE de 4.4314 ms, basados en aproximadamente 10,000 registros. Estos resultados demuestran la superioridad de los modelos de machine learning al capturar patrones no lineales y complejas interacciones entre variables, que la regresión lineal no puede modelar eficazmente.

Figura 7

Regresión Lineal de datos reales



Evaluación del Sistema Integrado de Análisis de Datos en Python para Trazabilidad

El Sistema Integrado en Python, implementado con Scikit-learn y Pandas, mejora la trazabilidad al predecir tiempos de retardo con un MAE de 3.6297 ms (Random Forest), generando archivos CSV con Tiempo Nominal, Real, y Predicho por lote. Sus características incluyen:

- Detección de anomalías: Identifica datos inusuales (desviaciones > 9 ms como en DG - 900), asegurando una limpieza efectiva de la base de datos y un análisis de datos preciso.
- Trazabilidad por planta: Facilita el seguimiento por planta al visualizar tendencias de tiempos de retardo en cada evaluación, mejorando la monitorización del proceso productivo.
- Recolección de registros: Acumula más datos para enriquecer el aprendizaje de Machine Learning, incrementando la robustez de las predicciones a medida que crece la base de datos.
- Gestión de archivos: Genera archivos CSV nativos, optimizados para su integración con Python, y además procesa y convierte archivos Excel a CSV, garantizando un manejo flexible y eficiente de diversos formatos de datos.

Por ejemplo, para DG - 900 (Real: 889.51 ms, Predicción: 898.65 ms), detecta una desviación de 9.14 ms, alertando sobre posibles variaciones en Longitud Cilindro o Días Transcurridos. Comparado con métodos manuales, reduce el tiempo de análisis y asegura un R^2 de 0.9996. Integrar datos en tiempo real de %Humedad y Longitud Cilindro podría reducir el MAE (< 3 ms), alineándose con CV < 5%.

Para una mejor trazabilidad de las tres plantas, se vinculó computacionalmente los números de lotes, para un correcto seguimiento del comportamiento de la variabilidad de los tiempos de retardos en cada etapa de producción. En el siguiente cuadro muestra como el sistema integrado de análisis de datos vincula los números de lotes en las diferentes plantas de evaluación indicando su información relevante y datos estadísticos. También, muestra toda la producción con el lote de mixto involucrado en la producción del detonador, con el fin de identificar algún patrón en la posible desviación. Además, da una alerta en caso de que el coeficiente de variación sea superior al 5% (se resalta en rojo) para que se pueda dar una corrección oportuna durante el proceso de fabricación de los detonadores y así evitar la producción de un volumen de gran escala. Se toma en cuenta el lote ejemplo C0027 de la producción en la planta de Ensamble.

Tabla de Trazabilidad: Lote C0027

Planta	Fecha de Producción	Fecha de Evaluación	Días Transcurridos Promedio	Lote Ensamble	Lote Dosificadora	Lote Mixto	Número Mediciones	% Humedad	Longitud Cilindro (mm)	Mezcla Retardante	Tiempo Nominal (ms)	Promedio Tiempo Real (ms)	Desviación Estándar (ms)	Coeficiente de Variación (%)
Planta Mixtos	16/04/2024	16/04/2024	0			24041609-11	80	0.08	9	AG	25	21.1017	0.6538	3.1
Planta Dosificadora	16/04/2024	16/04/2024	0		B0004	24041609-11	80	0.08	9	AG	25	21.5985	0.9885	4.58
Planta Ensamble	16/04/2024	31/10/2024	198	C0026	B0004	24041609-11	3	0.08	9	AG	25	24.268	2.0408	8.41
Planta Ensamble	16/04/2024	16/09/2024	153	C0027	B0004	24041609-11	3	0.08	9	AG	25	24.6087	2.2116	8.99
Planta Ensamble	16/04/2024	17/09/2024	154	C0027	B0004	24041609-11	15	0.08	9	AG	25	23.6631	1.4666	6.2
Planta Ensamble	16/04/2024	10/09/2024	147	C0029	B0004	24041609-11	3	0.08	9	AG	25	28.9187	0.3507	1.21
Planta Ensamble	16/04/2024	05/09/2024	142	C0030	B0004	24041609-11	3	0.08	9	AG	25	24.0103	3.3404	13.91
Planta Ensamble	16/04/2024	16/08/2024	122	C0031	B0004	24041609-11	3	0.08	9	AG	25	23.175	0.3317	1.43

4. Comparación de Tiempos Reales y Predichos por Mezcla y Tiempo Nominal

La Tabla 5 (Ver Anexo) compara los tiempos de retardo reales con las predicciones de los tres modelos para el conjunto de prueba, incluyendo el tiempo nominal como referencia. Las predicciones de Random Forest son las más cercanas a los valores reales, con un MAE de 3.6297 ms, lo que indica que las predicciones tienen un error promedio inferior a 4 ms.

La Figura 8 ilustra esta comparación mediante un gráfico de dispersión combinado, donde los puntos de Random Forest se alinean más estrechamente con la línea (y=x). La Tabla 5 (Ver Anexo) agrupa los resultados por mezcla retardante y tiempo nominal, mostrando que las mezclas AG (17–25 ms) y BG (42–125 ms) presentan predicciones más precisas (diferencias < 4 ms) que las mezclas CG (150–600 ms) y DG (700–1000 ms), posiblemente debido a la mayor variabilidad en tiempos nominales largos. La Figura 9 utiliza barras agrupadas para comparar los tiempos reales y predichos, confirmando la alta precisión del sistema en la mayoría de las mezclas.

Figura 8

Comparación de Tiempos Reales vs Predicciones (Todos los Modelos)

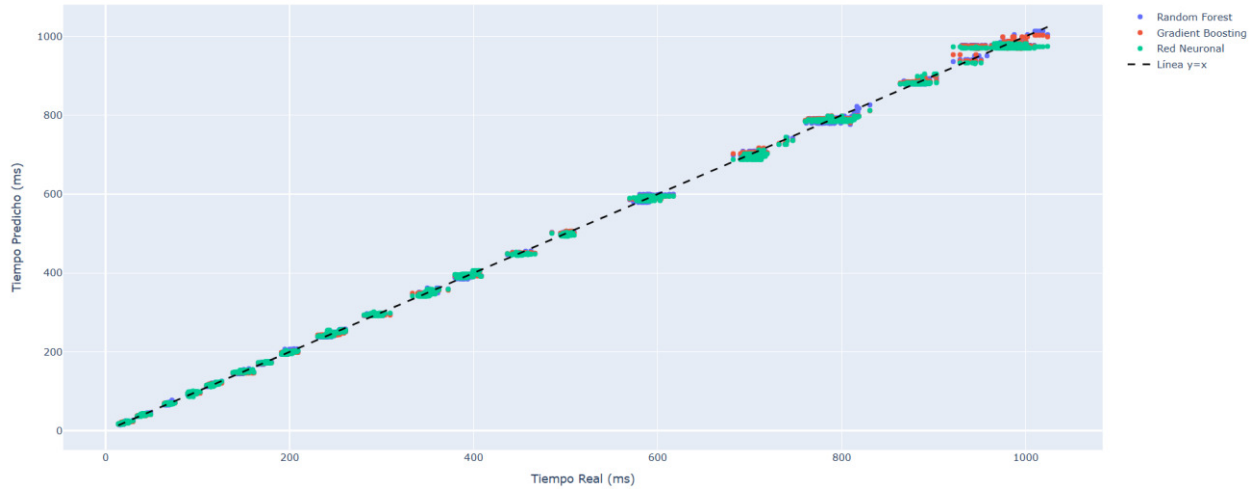
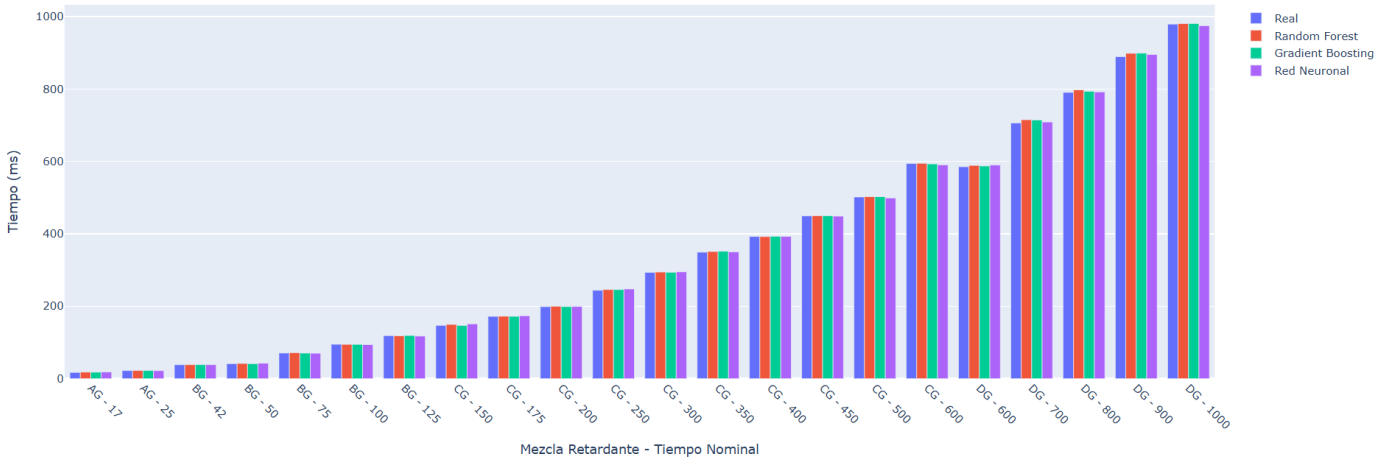


Figura 9

Gráfico de barras agrupadas para comparar los tiempos reales y predichos



5. Análisis de la Variabilidad

El análisis ANOVA multifactorial, con un nivel de significancia del 5%, arrojó los siguientes resultados:

- Tiempo Nominal: ($F = 1.4711$), ($p = 0.2080$). No se encontraron diferencias significativas entre los tiempos nominales ($p > 0.05$).
- Mezcla Retardante: ($F = 0.0000$), ($p = 1.0000$). No hay diferencias significativas entre las mezclas retardantes.
- Máquina: ($F = 4.3622$), ($p = 0.0128$). Existen diferencias significativas entre las máquinas (($p < 0.05$)), sugiriendo que el tipo de máquina influye en los tiempos de retardo.
- Humedad (Categorizada): ($F = 9.2843$), ($p = 0.0001$). La humedad tiene un impacto significativo en los tiempos de retardo ($p < 0.05$).
- Interacciones: Las interacciones entre Tiempo Nominal y Mezcla Retardante (($p = 1.0000$)) y entre Tiempo Nominal y Máquina ($p = 1.0000$) no son significativas.

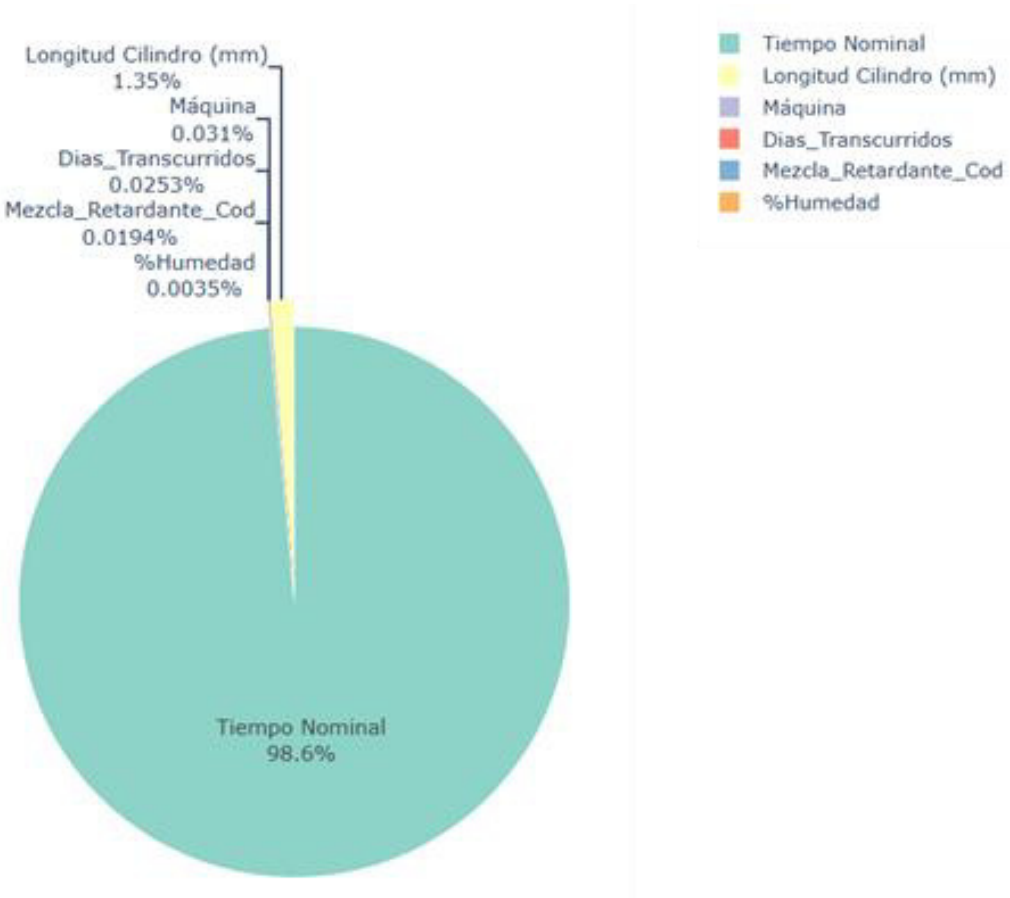
Estos resultados indican que la máquina y la humedad son factores clave que afectan la variabilidad de los tiempos de retardo, mientras que la mezcla retardante y el tiempo nominal no muestran diferencias significativas en este análisis.

6. Importancia de Características

La Tabla 6 describe la Importancia de Características y su gráfico circular (Figura 10) muestran que las variables Tiempo Nominal (importancia: 0.985658) y Longitud Cilindro (mm) (importancia: 0.013550) son las más influyentes en las predicciones de Random Forest, mientras que Máquina (0.000310), Días Transcurridos (0.000253), Mezcla Retardante (0.000194) y %Humedad (0.000035) tienen un impacto menor. Esto indica que los tiempos de retardo están fuertemente determinados por las especificaciones nominales y las dimensiones físicas del detonador. El bajo impacto de %Humedad y Mezcla Retardante en el modelo de Random Forest contrasta con los resultados del ANOVA, donde la humedad mostró significancia estadística, sugiriendo que su efecto podría estar mediado por interacciones con otras variables.

Figura 10

Grafica circular de Importancia de Características.



VI. DISCUSIÓN

La presente investigación desarrolló un sistema integrado de análisis de datos en Python para caracterizar, predecir y reducir la variabilidad en los tiempos de retardo de detonadores no eléctricos en las etapas de fabricación (plantas de mixtos, dosificadora y ensamble) de una empresa de explosivos en Lima, Perú. Los resultados obtenidos, presentados en el apartado VI.1, demuestran que el sistema cumple con los objetivos establecidos, al lograr predicciones precisas de los tiempos de retardo y proporcionar herramientas para monitorear y optimizar el control de calidad en el proceso de fabricación. A continuación, se discuten los hallazgos clave en relación con el objetivo general, su relevancia en el contexto industrial, las implicaciones prácticas, las limitaciones del estudio y su alineación con la literatura existente.

Cumplimiento del Objetivo General

El sistema integrado de análisis de datos en Python, implementado con algoritmos de aprendizaje automático (Random Forest, Gradient Boosting y redes neuronales), demostró una alta capacidad para caracterizar y predecir los tiempos de retardo de los detonadores no eléctricos. Los resultados de las métricas de evaluación, calculadas sobre un conjunto de aproximadamente 10,000 registros, muestran que Random Forest obtuvo el mejor desempeño (Figura 6), con un R^2 de 0.9996, un MSE de 39.2752 y un MAE de 3.6297 ms, seguido de Gradient Boosting (R^2: 0.9995, MSE: 47.9773, MAE: 4.2971 ms) y la red neuronal (R^2: 0.9994, MSE: 55.9307, MAE: 4.4314 ms). Estos valores reflejan un ajuste excepcional de los modelos a los datos, con errores promedio inferiores a 5 ms, lo

que es particularmente significativo considerando el rango de tiempos de retardo (17–1000 ms). Este nivel de precisión permite predecir con confianza los tiempos de retardo en las plantas de mixtos, dosificadora y ensamble.

Relevancia de los Resultados en el Contexto Industrial

La alta precisión de Random Forest (R^2: 0.9996) y su capacidad para identificar las variables más influyentes, como el Tiempo Nominal (importancia: 98.6%) y la Longitud Cilindro (mm) (importancia: 1.4%), tiene implicaciones significativas para la industria de explosivos en Lima, Perú. En el sector minero, donde los detonadores no eléctricos son esenciales para garantizar la sincronización de voladuras, una variabilidad mínima en los tiempos de retardo es crítica para la seguridad y eficiencia (Agrawal y Mishra, 2018). El sistema desarrollado permite a la empresa monitorear en tiempo real la estabilidad de los procesos en las plantas de mixtos, dosificadora y ensamble, identificando lotes anómalos (en DG - 900, según Tabla 3) y optimizando parámetros clave como la longitud del cilindro. La importancia de características revela que el Tiempo Nominal domina las predicciones, lo que sugiere que las especificaciones de diseño de los detonadores son el factor principal en la determinación de los tiempos de retardo, mientras que variables como %Humedad (importancia: 0.0035%) tienen un impacto menor en el modelo de Random Forest. Sin embargo, el análisis ANOVA multifactorial indica que la humedad categorizada tiene un efecto significativo ($p = 0.0$), lo que resalta la necesidad de controlar las condiciones ambientales en las plantas para minimizar la variabilidad.

Comparación con Técnicas Estadísticas

El modelo de regresión lineal múltiple, aplicado a las primeras 50 filas de la base de datos, logró un (R^2) de 0.88 y un MAE de 8.7 ms, reflejando una capacidad limitada para modelar la complejidad de los datos de los detonadores (Figura 7). En contraste, los modelos de machine learning, como Random Forest (R^2: 0.9996, MAE: 3.6297 ms), Gradient Boosting (R^2: 0.9995, MAE: 4.2971 ms) y redes neuronales (R^2: 0.9994, MAE: 4.4314 ms), demostraron una precisión significativamente mayor al capturar no linealidades e interacciones complejas, como (Tiempo Nominal - Días Transcurridos). Esta superioridad se debe a la capacidad de los modelos de machine learning para manejar patrones no lineales superando las restricciones inherentes a la regresión lineal, que asume relaciones lineales y no captura interacciones avanzadas.

El análisis ANOVA multifactorial, aunque útil para identificar la significancia de factores principales se limita a un enfoque exploratorio y no ofrece predicciones directas. La complementariedad entre ANOVA y machine learning radica en que el primero selecciona variables relevantes, mientras que los segundos las modelan con alta precisión, consolidando al sistema como una herramienta robusta para la industria de explosivos.

Puntos por mejorar en el proceso de fabricación de detonadores

Optimización del Proceso de Fabricación: Los gráficos de cajas identifican lotes fuera de los límites de control, especialmente en la mezcla DG (700–1000 ms), lo que sugiere la necesidad de ajustes en la planta de mixtos, como una mayor estandarización en la preparación de mezclas retardantes. La mayor dispersión en DG, observada en Tabla 2, podría mitigarse ajustando la composición química o los parámetros de dosificación.

Control de Calidad: El sistema permite implementar un monitoreo rutinario basado en las predicciones de Random Forest y la detección de anomalías, facilitando la trazabilidad de los lotes en las tres plantas. Esto es particularmente relevante en la planta de ensamble, donde se observaron mayores desviaciones, como en DG - 900 (Tabla 3).

Estandarización de Máquinas: El análisis ANOVA mostró que el factor Máquina es significativo ($p < 0.05$), indicando que las diferencias entre máquinas contribuyen a la variabilidad en los tiempos de retardo. La empresa podría implementar calibraciones más estrictas o estandarizar las configuraciones de las máquinas para reducir este efecto.

Parámetro de Humedad: Aunque la humedad tiene una importancia baja en Random Forest, su significancia en el ANOVA ($p = 0.0001$) sugiere que controlar la humedad de los mixtos en la Planta de Mixtos podría reducir la variabilidad, especialmente en las mezclas CG y DG.

VII. CONCLUSIONES Y RECOMENDACIONES

El sistema integrado de análisis de datos desarrollado en Python, utilizando herramientas estadísticas (ANOVA, coeficiente de variación) y algoritmos de aprendizaje automático (Random Forest, Gradient Boosting y Redes Neuronales), ha demostrado ser una herramienta efectiva para predecir la variabilidad en los tiempos de retardo de detonadores no eléctricos en las plantas de Mixtos, Dosificadora y Ensamble de una empresa de explosivos en Lima, Perú. Los resultados muestran que el modelo de Random Forest logra un R^2 de 0.9996, un MSE de 39.2752 y un MAE de 3.6297 ms, cumpliendo con el objetivo de precisión (MAE < 5 ms) y mejorando la confiabilidad del producto al permitir la detección de lotes no conformes. La generación de archivos CSV con Pandas asegura una trazabilidad completa, registrando tiempos predichos y reales por lote, lo que optimiza el control de calidad y reduce costos asociados a rechazos, alineándose con los principios de la Industria 4.0 (Soori et al., 2024). Además, se encargó de vincular computacionalmente los lotes por planta con su respectiva información, resultados estadísticos y la detección de desviaciones en los procesos de producción.

Se identificaron los factores que generan variabilidad en los tiempos de retardo mediante análisis estadístico y aprendizaje automático. El análisis ANOVA multifactorial (Tabla 4) reveló que Tiempo Nominal ($F = 881265.2556$, $p < 0.0001$), Máquina ($F = 71.3680$, $p < 0.0001$), Humedad ($F = 17.7080$, $p < 0.0001$), Días Transcurridos ($F = 70.7105$, $p < 0.0001$) y Longitud del Cilindro ($F =$

70.6994, $p < 0.0001$) son factores significativos, con interacciones relevantes como Tiempo Nominal × Máquina ($F = 13.2911$, $p < 0.0001$). Además, el análisis de importancia de características en Random Forest (Tabla 6) mostró que Tiempo Nominal (0.985658) y Longitud del Cilindro (0.013550) son los principales contribuyentes a la variabilidad, mientras que Humedad (0.000035) y Mezcla Retardante (0.000194) tienen un impacto menor. La mezcla AG (17–25 ms) presentó menor variabilidad (coeficiente de variación < 5%, Tabla 2) que la mezcla DG (700–1000 ms), lo que indica mayor estabilidad en tiempos cortos y sugiere ajustes en la preparación de mezclas DG en la Planta de Mixtos.

Los algoritmos de aprendizaje automático implementados en Python lograron predecir desviaciones en los tiempos de retardo con alta precisión, cumpliendo el objetivo de MAE < 5 ms. Random Forest obtuvo el mejor desempeño (MAE = 3.6297 ms, $R^2 = 0.9996$), seguido por Gradient Boosting (MAE = 4.2971 ms, $R^2 = 0.9995$) y Redes Neuronales (MAE = 4.4314 ms, $R^2 = 0.9994$), según la Tabla 8. Las predicciones fueron más precisas para las mezclas AG y BG (diferencias < 4 ms, Tabla 3) que, para CG y DG, donde se observaron mayores. Estas métricas se calcularon utilizando un 20 % de los datos totales (aproximadamente 2,000 registros) reservados exclusivamente para la etapa de prueba, lo que garantiza que los resultados reflejan la capacidad real de predicción del modelo sobre datos no vistos durante el entrenamiento. El sistema de trazabilidad, implementado con Pandas, generó archivos CSV que registran tiempos predichos y reales por lote,

permitiendo intervenciones correctivas como ajustes en la humedad o longitud del cilindro en la Planta de Mixtos.

Recomendaciones para Futuras Investigaciones

- Implementar índices de capacidad de proceso Cp y Cpk para monitorear y optimizar la variabilidad de tiempos de retardo.
- Incorporar la intervención del operador. Esto permitirá incluir otra variable en el ANOVA y los modelos de ML evaluado su impacto en la variabilidad en los tiempos de retardo. Además, se puede complementar en el ANOVA con un test post-hoc (tukey) para identificar operadores específicos que afecten la calidad.
- Aumentar el volumen de la muestra para darle robustez a los modelos ML. Optimización de Modelos: Probar otros tipos de algoritmos de machine learning, como Support Vector Machines o Decision Trees, para realizar una comparación con los ML realizados en esta investigación.

VIII. BIBLIOGRAFÍAS Y ANEXOS

- Referencias bibliográficas:

Acosta, J., Quiroz, E., Pereyra, O., y Ventura, A. (2022). *Reducción de vibraciones hacia comunidades mediante control de retardos*. Instituto de Ingenieros de Minas del Perú.

Agrawal, H., & Mishra, A. (2018). Probabilistic analysis on scattering effect of initiation systems and concept of modified charge per delay for prediction of blast induced ground vibrations. *Measurement*, 130, 306–317. https://doi.org/10.1016/j.measurement.2018.08.032.

Breiman, L. Random Forests. *Machine Learning* 45, 5–32 (2001). https://doi.org/10.1023/A:1010933404324

Brunhoferová, H., Trecáková, T., & Kočí, V. (2025). Life cycle assessment of electronic, electric and nonelectric detonators; a site-specific case for Czech Republic. *Heliyon*, e42371. https://doi.org/10.1016/j.heliyon.2025.e42371

Evans, J. R., & Lindsay, W. M. (2020*). Managing for Quality and Performance Excellence (11th ed.)*. Cengage Learning.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning. In *Springer series in statistics*. https://doi.org/10.1007/978-0-387-84858-7

ISO. (2015). *ISO 9001:2015: Sistemas de gestión de la calidad – Requisitos*. International Organization for Standardization.

Ji, X., Liu, Y., Li, Z., Yu, Q., Gao, Y., Zhang, H., & Wang, L. (2017). Thermal behavior of Al/Zr/KClO4 pyrotechnic compositions at high temperature. Thermochimica Acta, 659, 55–58. https://doi.org/10.1016/j.tca.2017.11.006

Khan, A., Malik, A. Q., Lodhi, Z. H., & Abdin, Z. U. (2018). Effect of body material and temperature variation on the performance of the time delay pyrotechnic compositions. *Defence Technology*, *14*(3), 261–265. https://doi.org/10.1016/j.dt.2018.03.006

McKinney, W. (2017). *Python for data analysis: Data Wrangling with Pandas, Numpy, and Ipython*. O'Reilly Media.

Montgomery, D. C. (2020). *Introduction to statistical quality control*. John Wiley & Sons.

Msakni, M. K., Risan, A., & Schütz, P. (2023). Using machine learning prediction models for quality control: a case study from the automotive industry. *Computational Management Science*, *20*(1). https://doi.org/10.1007/s10287-023-00448-0

Müller, A. C., & Guido, S. (2016). *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media.

Pandas Development Team. (2025). *Pandas: Python Data Analysis Library.*

Polis, M., Szydło, K., Zakusylo, R., Hawelek, L., Stolarczyk, A., & Jarosz, T. (2023). Study of the combustion mechanism of ZN/KMNO4 Pyrotechnic Composition. *Molecules*, *28*(15), 5741. https://doi.org/10.3390/molecules28155741

Python Software Foundation. (s.f.). *Welcome to Python.org.* https://www.python.org/

Raven, D. B., Chikkula, Y., Patel, K. M., Ghazal, A. H. A., Salloum, H. S., Bakhurji, A. S., & Patwardhan, R. S. (2024). Machine learning & conventional approaches to process control & optimization: Industrial applications & perspectives. *Computers & Chemical Engineering*, *189*, 108789. https://doi.org/10.1016/j.compchemeng.2024.108789

Revista Tecnología Minera. (2020). *Famesa Explosivos lanza sistema de trazabilidad para prevenir el comercio ilegal de explosivos*. Revista Tecnología Minera.

Romero Paúcar, R. (2016). *Voladura con detonadores electrónicos para optimizar la fragmentación y seguridad en el tajo Toromocho – Minera Chinalco Perú S.A.* [Tesis de licenciatura, Universidad Nacional del Centro del Perú].

Schmitt, J., Bönig, J., Borggräfe, T., Beitinger, G., & Deuse, J. (2020). Predictive model-based quality inspection using Machine Learning and Edge Cloud Computing. Advanced Engineering Informatics, 45, 101101. https://doi.org/10.1016/j.aei.2020.101101

Scikit-learn Developers. (2025). *Scikit-learn: Machine Learning in Python*.

Silva, J., Li, L., & Gernand, J. M. (2018). Reliability analysis for mine blast performance based on delay type and firing time. *International Journal of Mining Science and Technology*, *28*(2), 195–204. https://doi.org/10.1016/j.ijmst.2017.07.004

Soori, M., Jough, F. K. G., Dastres, R., & Arezoo, B. (2024). AI-Based Decision Support Systems in Industry 4.0, a review. *Journal of Economy and Technology*. https://doi.org/10.1016/j.ject.2024.08.005

Taiwo, B. O., Fissha, Y., Hosseini, S., Khishe, M., Kahraman, E., Adebayo, B., Sazid, M., Adesida, P. A., Famobuwa, O. V., Faluyi, J. O., & Akinlabi, A. A. (2024). *Machine learning based prediction of flyrock distance in rock blasting: A safe and sustainable mining approach*. Green and Smart Mining Engineering Journal. https://doi.org/10.1016/j.gsme.2024.08.006

Tao, F., Qi, Q., Wang, L., & Nee, A. (2019). Digital Twins and Cyber–Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison. *Engineering*, *5*(4), 653–661. https://doi.org/10.1016/j.eng.2019.01.014

Tichapondwa, S., Guo, S., & Roux, W. (2021). Performance of MN/BI2O3 pyrotechnic time delay compositions. *Central European Journal of Energetic Materials*, *18*(1), 46–62. https://doi.org/10.22211/cejem/134798

- Anexos:

1. Tabla 5

Comparación de Tiempos Agrupados por Tipo de Mixtos vs Tiempos reales y Predichos.

Mezcla Retardante	Tiempo Nominal (ms)	Tiempo Promedio Real (ms)	Predicción Random Forest (ms)	Predicción Gradient Boosting (ms)	Predicción Red Neuronal (ms)
AG	17	16.97	17.76	17.54	18.05
AG	25	22.03	22.3	22.15	21.79
BG	42	38.17	38.32	38.38	38.23
BG	50	41.05	41.76	41.1	42.25
BG	75	70.42	71.44	70.28	69.98
BG	100	94.76	94.44	94.42	93.9
BG	125	118.55	118.13	118.97	117.33
CG	150	146.95	149.53	146.78	151.06
CG	175	171.86	172.19	171.89	173.42
CG	200	198.7	199.73	198.75	198.99
CG	250	243.9	245.83	245.77	247.43
CG	300	293.4	294.17	293.58	294.79
CG	350	349.16	351	351.61	350
CG	400	392.53	392.19	392.95	392.59
CG	450	449.44	449.94	449.87	448.72
CG	500	501.58	502.46	502.36	498.93
CG	600	594.1	594.71	593	590.39
DG	600	585.24	588.92	587.37	589.9
DG	700	706.12	715.17	714.36	708.91
DG	800	790.69	797.74	793.66	791.96
DG	900	889.51	898.65	899.06	895.45
DG	1000	979.41	980.7	981.1	975.05

2. Tabla 6

Tabla de Importancia de Características

Feature	Importance
Tiempo Nominal	0.985658
Longitud Cilindro (mm)	0.01355
Máquina	0.00031
Dias_Transcurridos	0.000253
Mezcla_Retardante_Cod	0.000194
%Humedad	0.000035

3. Tabla 7

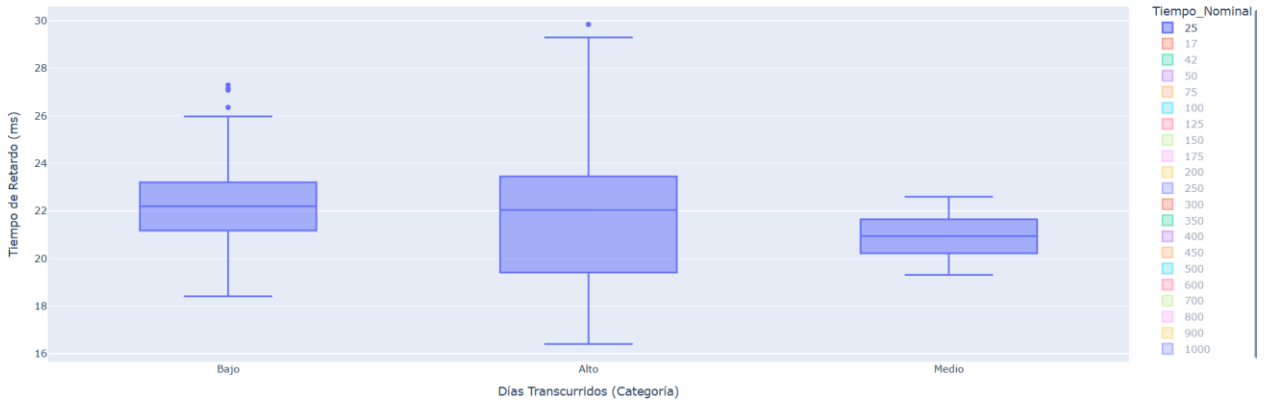
Características de Tiempos aplicados en la planta Mixtos

Tiempo de Retardo (ms)	Longitud Cilindro (mm)	Mezcla Iniciadora	Mezcla Retardante
17	9	-	AG
25	9	-	AG
33	15	-	AG
42	9	-	BG
50	9	-	BG
75	15	-	BG
100	18	-	BG
100	21	-	BG
125	24	-	BG
125	21	-	BG
150	15	BG	CG
175	15	BG	CG
200	15	BG	CG
250	18	BG	CG
300	21	BG	CG
350	24	BG	CG
400	27	BG	CG
450	30	BG	CG
500	30	BG	CG
500	33	BG	CG
600	18	BG	DG
600	24	BG	CG
600	36	BG	CG
700	21	BG	DG
800	18	BG	DG
800	21	BG	DG
900	24	BG	DG
1000	24	BG	DG

Nota: *Tomado de los procedimientos del área de Control de Calidad*

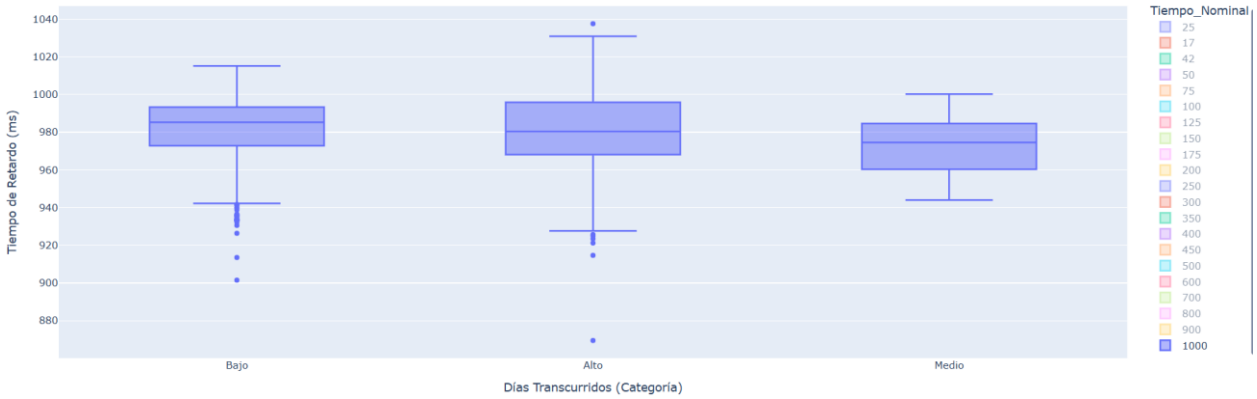
4. **Figura 11**

Gráfico de Cajas para el tiempo 25ms de Mixto AG



5. **Figura 12**

Gráfico de cajas para el tiempo 1000ms de Mixto DG



6. **Tabla 8**

Tabla de Errores de cada Modelo del sistema Integrado en Python.

Modelo	MSE	MAE	R2
Random Forest	39.2752	3.6297	0.9996
Gradient Boosting	47.9773	4.2971	0.9995
Red Neuronal	55.9307	4.4314	0.9994

7. Tabla 9

Comparación de Tiempos Nominales, Reales y Predichos

Tiempo Nominal (ms)	Tiempo Real (ms)	Predicción Random Forest (ms)	Predicción Gradient Boosting (ms)	Predicción Red Neuronal (ms)
75	69.12	68.9	69.73	68.11000061035156
450	447.23	446.95	447.95	445.1199951171875
250	246.08	249.31	243.52	248.88999938964844
1000	997.03	977.24	975.88	970.2000122070312
350	345.14	349.76	348.66	346.8299865722656
75	70.59	71.67	70.43	70.83000183105469
25	24.21	23.84	22.36	23.920000076293945
25	20.66	21.88	22.24	21.020000457763672
42	40.45	39.39	38.46	39.97999954223633
600	586.35	584.55	586.57	588.1099853515625
25	21.18	22.55	22.42	22.309999465942383
350	352.81	346.77	349.19	344.1000061035156
1000	972.91	971.22	970.92	970.3300170898438
1000	975.59	977.24	975.88	970.2000122070312
300	294.21	296.29	293.3	297.6300048828125
17	18.06	16.25	16.99	17.450000762939453
250	258.56	257.13	252.8	255.72999572753906
900	944.86	945.1	939.15	930.5800170898438
800	775.41	779.93	785.98	783.8699951171875
25	23.24	21.41	21.96	21.860000610351562
250	237.43	240.53	242.01	241.24000549316406
250	239.96	240.53	242.01	241.24000549316406
42	37.89	38.1	38.35	38.619998931884766
1000	994.6	977.24	975.88	970.2000122070312
75	70.36	71.67	70.43	70.83000183105469