REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

**Universidad Católica de Santa María**

**Facultad de Ciencias e Ingenierías Físicas y Formales**

**Escuela Profesional de Ingeniería de Minas**

# Aplicación de machine learning para la estimación de vida útil y reducción de costos en brocas triconicas de 11 en minería superficial

Tesis presentada por el Bachiller:

**Troya Rios, Renzo Paolo**

**ORCID: 0009-0007-1986-8028**

Para optar el Título Profesional de Ingeniero de Minas

Asesor (a)

**Mg. Ojeda Vizarreta, Jorge**

**ORCID: 0009-0008-6717-0879**

Arequipa – Perú

2025

Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

UCSM-ERP

# UNIVERSIDAD CATÓLICA DE SANTA MARÍA

## INGENIERIA DE MINAS

## TITULACIÓN CON TESIS

## DICTAMEN APROBACIÓN DE BORRADOR

Arequipa, 04 de Septiembre del 2025

**Dictamen: 016008-C-EPIM-2025**

Visto el borrador del expediente 016008, presentado por:

**2019101271** - **TROYA RIOS RENZO PAOLO**

Titulado:

**APLICACIÓN DE MACHINE LEARNING PARA LA ESTIMACIÓN DE VIDA ÚTIL Y REDUCCIÓN DE COSTOS EN BROCAS TRICONICAS DE 11 EN MINERÍA SUPERFICIAL**

Nuestro dictamen es:

**APROBADO**

Titulo Profesional/Titulo de Segunda Especialidad/Grado Académico a optar:

**INGENIERO DE MINAS**

**40902304** - **LINARES FLORES CASTRO ANTONIO ERICK**
**DICTAMINADOR**

**40379481** - **PAREDES SALAS OMAR WILLY**
**DICTAMINADOR**

**29635304** - **MORALES VALDIVIA JAVIER ANTONIO**
**DICTAMINADOR**

# APLICACIóN DE MACHINE LEARNING PARA LA ESTIMACIóN DE VIDA úTIL Y REDUCCIóN DE COSTOS EN BROCAS TRICONICAS DE 11 EN MINERíA SUPERFICIAL

INFORME DE ORIGINALIDAD

| 18% | 16% | 8% | 10% |
|---|---|---|---|
| INDICE DE SIMILITUD | FUENTES DE INTERNET | PUBLICACIONES | TRABAJOS DEL ESTUDIANTE |

FUENTES PRIMARIAS

| | | |
|---|---|---|
| 1 | Submitted to Universidad Católica de Santa María<br>Trabajo del estudiante | 2% |
| 2 | tesis.ucsm.edu.pe<br>Fuente de Internet | 2% |
| 3 | Submitted to Universidad Internacional de la Rioja<br>Trabajo del estudiante | 1% |
| 4 | Submitted to Universidad Anahuac México Sur<br>Trabajo del estudiante | 1% |
| 5 | Submitted to Universidad TecMilenio<br>Trabajo del estudiante | 1% |
| 6 | www.coursehero.com<br>Fuente de Internet | 1% |
| 7 | Submitted to University of Hertfordshire<br>Trabajo del estudiante | <1% |
| 8 | Submitted to University of North Texas<br>Trabajo del estudiante | <1% |
| 9 | icobi.nsbm.ac.lk<br>Fuente de Internet | <1% |
| 10 | Submitted to Fundacion Universidad de San Andres<br>Trabajo del estudiante | <1% |

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

## DEDICATORIA

A mi madre Miluzka por su apoyo durante toda mi carrera y el inicio de mi vida profesional.

A mi padre Juan por todo el apoyo que me dio desde pequeño.

A mi tia Lorena por creer en mi y siempre apoyarme.

Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

## AGRADECIMIENTOS

A mi madre por siempre estar a mi lado y apoyarme en todo, así como también mi padre y mi tía, que me brindaron su apoyo y confianza.

Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

## RESUMEN

La presente investigación tiene como finalidad el desarrollo de un modelo predictivo en lenguaje Python, basado en técnicas de machine learning, para estimar la vida útil de brocas tricónicas de 11” utilizadas en una mina superficial. Este enfoque se orienta a resolver una problemática recurrente en minería: el reemplazo prematuro o tardío de brocas de perforación, lo que impacta directamente en los costos operativos, la eficiencia del ciclo de minado y el uso óptimo de insumos. El estudio parte del análisis de una base de datos histórica conformada por parámetros operativos registrados en campo, tales como el pulldown (presión de empuje), revoluciones por minuto (RPM), presión de aire y velocidad de penetración (ROP). A través de una adecuada limpieza, estructuración y análisis de estos datos, se logró determinar la correlación entre dichas variables y la vida útil real de las brocas. Este análisis permitió identificar a la velocidad de penetración como el parámetro con mayor grado de correlación directa con la vida útil, lo cual sirvió de base para el desarrollo posterior del modelo predictivo.

El modelo fue implementado en Python usando Regresión Lineal Múltiple, Random Forest y XGBoost. Se entrenó con 80% de los datos y probó con 20%, evaluándose mediante $R^2$, MAE y RMSE. Los resultados evidencian alta capacidad predictiva, destacando la regresión lineal entre vida útil y velocidad de penetración, con $R^2$ superior a 0.99. Finalmente, se estimó el impacto económico del modelo mediante un análisis de ahorro, considerando el uso eficiente de las brocas en función de su vida útil estimada. Este análisis demuestra que, al anticipar con mayor precisión la duración de cada broca, es posible optimizar los ciclos de cambio, reducir pérdidas por desgaste prematuro y minimizar interrupciones operativas.

**Palabras clave:** Brocas Triconicas, Machine Learning, Vida Útil

Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA


## ABSTRACT

The purpose of this research is to develop a predictive model in Python, based on machine learning techniques, to estimate the lifespan of 11" tricone bits used in a surface mine. This approach aims to solve a recurring problem in mining: the premature or late replacement of drill bits, which directly impacts operating costs, the efficiency of the mining cycle, and the optimal use of inputs. The study is based on the analysis of a historical database consisting of operational parameters recorded in the field, such as pulldown (push pressure), revolutions per minute (RPM), air pressure, and rate of penetration (ROP). Through proper cleaning, structuring, and analysis of this data, it was possible to determine the correlation between these variables and the actual lifespan of the bits. This analysis identified the rate of penetration as the parameter with the highest degree of direct correlation with lifespan, which served as the basis for the subsequent development of the predictive model.

The model was implemented in Python using Multiple Linear Regression, Random Forest, and XGBoost was trained with 80% of the data and tested with 20%, and evaluated using $R^2$, MAE, and RMSE. The results demonstrate a high predictive capacity, highlighting the linear regression between tool life and penetration rate, with $R^2$ greater than 0.99. Finally, the economic impact of the model was estimated through a savings analysis, considering the efficient use of drill bits based on their estimated tool life. This analysis demonstrates that, by more accurately predicting the life of each drill bit, it is possible to optimize change cycles, reduce losses due to premature wear, and minimize operational interruptions.


**Keywords:** Tricone Drill Bits, Machine Learning, Tool Life

Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

# ÍNDICE GENERAL

DEDICATORIA

AGRADECIMIENTOS

RESUMEN

ABSTRACT

INTRODUCCIÓN ..... 1

CAPITULO 1 ..... 3

1. PLANTEAMIENTO DE LA INVESTIGACION ..... 3

1.1. Planteamiento del Problema ..... 3

1.2. Objetivos ..... 4

1.2.1. Objetivo General ..... 4

1.2.2. Objetivos Específicos ..... 4

1.3. Preguntas de Investigación ..... 5

1.3.1. Pregunta Global ..... 5

1.3.2. Preguntas Especificas ..... 5

1.4. Línea de Investigación a la que corresponde el Problema ..... 5

1.5. Palabras Clave ..... 5

1.6. Solución Propuesta ..... 6

1.6.1. Justificación e Importancia ..... 6

CAPITULO 2 ..... 8

2. FUNDAMENTOS TEÓRICOS ..... 8

2.1. Estado del arte ..... 8

2.2. Bases Teóricas de la Investigación ..... 11

2.2.1. Perforación en Minería Superficial ..... 11

2.2.2. Brocas Triconicas 11” ..... 11

2.2.3. Desgaste de Brocas Triconicas ..... 12

2.2.4. Vida Útil de Brocas Triconicas ..... 12

2.2.5. Pulldown ..... 12

2.2.6. Revoluciones por Minuto (RPM) ..... 12

2.2.7. Presión de Aire ..... 12

2.2.8. Velocidad de Penetración (ROP) ..... 13

2.2.9. Costos Operativos Asociados a la Perforación ..... 13

Publicación autorizada con fines académicos e investigativos

En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

2.2.10. Machine Learning ..... 13
2.2.11. Machine Learning en Minería ..... 13
2.2.12. Regresión Lineal Múltiple ..... 14
2.2.13. Random Forest Regressor ..... 14
2.2.14. XGBoost Regressor ..... 14
2.2.15. Coeficiente de Determinación ($R^2$) ..... 14
2.2.16. Error Absoluto Medio (MAE) ..... 15
2.2.17. Raíz del Error Cuadrático Medio (RMSE) ..... 15
2.2.18. Python ..... 15
CAPITULO 3 ..... 16
3. MARCO METODOLOGICO ..... 16
3.1. Alcances y Limitaciones ..... 16
3.1.1. Alcances ..... 16
3.1.1. Limitaciones ..... 16
3.2. Aporte ..... 16
3.3. Tipo de Nivel de la Investigación ..... 16
3.3.1. Tipo de Nivel de la Investigación ..... 16
3.3.2. Nivel de la Investigación ..... 17
3.4. Población y Muestra ..... 18
3.5. Ubicación de la investigación ..... 20
3.6. Descripción de la geología general ..... 21
3.7. Descripción geomecánica y geotécnica del tajo ..... 21
3.7.1. Propiedades Geomecánicas Principales ..... 22
3.8. Parámetros de diseño banco-talud ..... 22
3.9. Tipo de broca seleccionada ..... 23
3.10. Métodos, Técnicas e Instrumentos de Recolección de Dato ..... 25
3.11. Plan de análisis estadístico y recolección de los datos ..... 27
3.12. Plan de análisis de datos ..... 27
3.12.1. Preparación y limpieza de datos ..... 27
3.12.2. Selección de variables ..... 28
3.12.3. Análisis de correlación ..... 28
3.12.4. Visualización de correlaciones ..... 28
3.12.5. Desarrollo de modelos predictivos ..... 28
3.12.6. Comparación de desempeño de modelos ..... 28
3.12.7. Selección del modelo final ..... 29

Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

3.12.8. Validación del modelo ..... 29
3.12.9. Almacenamiento del modelo ..... 29
3.12.10. Estimación del ahorro operativo ..... 29
CAPITULO 4 ..... 30
4. PLAN DE TRABAJO ..... 30
4.1. Preparación y limpieza de datos ..... 30
4.1.1. Carga de los datos ..... 33
4.1.2. Selección de variables relevantes ..... 33
4.1.3. Verificación y calidad de los datos ..... 33
4.2. Análisis de correlación ..... 34
4.2.1. Cálculo del coeficiente de correlación ..... 34
4.2.2. Resultados del análisis ..... 35
4.2.3. Visualización de resultados ..... 37
4.2.4. Generación automática de reporte ..... 37
4.2.5. Interpretación ..... 39
4.3. Desarrollo del modelo predictivo ..... 40
4.3.1. Preparación de los datos para el modelo ..... 40
4.3.2. Regresión simple ..... 41
4.3.3. Random Forest Regressor ..... 43
4.3.4. XGBoost Regressor ..... 44
4.3.5. Comparación de Desempeño de Modelos ..... 44
4.4. Resultados Obtenidos y selección del modelo de predicción ..... 45
4.5. Código de predicción ..... 49
4.5.1. Validación del modelo ..... 49
4.5.2. Almacenamiento del modelo para uso futuro ..... 50
4.6. Estimación del ahorro operativo ..... 50
5. CONCLUSIONES ..... 54
6. REFERENCIAS ..... 57
ANEXOS ..... 62

Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

# ÍNDICE DE TABLAS

Tabla 1 Equipos de perforación ..... 19

Tabla 2 Ubicación de las perforadoras ..... 19

Tabla 3 Características Técnicas de la broca de 11" ..... 24

Tabla 4 Variables de recolección de datos ..... 26

Tabla 5 Data de brocas ..... 30

Tabla 6 Parámetros estadísticos de la vida útil de las brocas de 11" ..... 32

Tabla 7 Resumen de Correlación ..... 36

Tabla 8 Resumen de modelos de predicción ..... 48

Tabla 9 Comparación de metros reales vs metros con el modelo predictivo ..... 51

Tabla 10 Promedios de vida útil de las brocas registradas en una mina superficial ..... 52

Tabla 11 Ahorro anual estimado ..... 53

Tabla 12 Escenario 1 sin el modelo predictivo ..... 53

Tabla 13 Escenario 2 con la aplicación del modelo ..... 54

Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

# ÍNDICE DE FIGURAS

Figura 1 Broca Triconica de 11" ..... 25

Figura 2 Correlación entre vida útil vs Velocidad de penetración ..... 35

Figura 3 Correlación entre vida útil vs RPM ..... 35

Figura 4 Correlación entre vida util vs Pulldown ..... 36

Figura 5 Ejecución del modelo de correlación en Python ..... 38

Figura 6 Ejecución del código para comparar modelos predictivos. ..... 46

Figura 7 Gráfico de coeficiente de determinación de cada algoritmo de regresión ..... 47

Figura 8 Gráfico de Mean Absolute Error ..... 47

Figura 9 Gráfico de Root Mean Squared Error de cada algoritmo de regresión ..... 48

Figura 10 Ejecución del modelo predictivo en Python ..... 50

Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

# INTRODUCCIÓN

En el contexto de la minería moderna, uno de los desafíos más significativos para las operaciones de perforación es garantizar la eficiencia en el uso de herramientas críticas como las brocas tricónicas. Estas herramientas, fundamentales en el proceso de desarrollo de mallas de perforación y voladura, poseen una vida útil limitada que depende de diversos factores operativos y geológicos. La falta de un control preciso sobre su desgaste conduce a frecuentes reemplazos anticipados o fallos prematuros, generando pérdidas económicas considerables, tiempos muertos no planificados y sobreconsumo de insumos. Este escenario se agrava en minas superficiales de gran escala, donde los volúmenes de perforación diarios son elevados y el control del rendimiento de cada broca es clave para el éxito operativo.

La presente tesis aborda esta problemática mediante el desarrollo de un modelo predictivo que permita estimar la vida útil de brocas tricónicas de 11” empleadas en una operación minera superficial. La propuesta se fundamenta en el uso de técnicas de machine learning implementadas en Python, una herramienta versátil y potente para el análisis de datos y modelamiento estadístico. A través del uso de modelos de regresión como Regresión Lineal Múltiple, Random Forest y XGBoost, se busca generar una ecuación capaz de predecir la duración estimada de una broca a partir de variables operativas registradas durante su uso.

El trabajo inicia con la recopilación y depuración de una base de datos operativa real, que incluye parámetros como la velocidad de penetración (ROP), pulldown, RPM y presión de aire. Posteriormente, se realiza un análisis exploratorio para identificar las variables con mayor influencia sobre la vida útil de las brocas, encontrando una fuerte correlación entre esta y la



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

velocidad de penetración. Esta variable, al presentar un comportamiento consistentemente predictivo, es seleccionada como base para la construcción del modelo.

Una vez definido el enfoque, se procede a aplicar técnicas de regresión para evaluar su capacidad predictiva. Cada modelo es evaluado con métricas estándar como $R^2$, MAE y RMSE, permitiendo identificar la opción más precisa y robusta. Adicionalmente, se estima el beneficio económico del uso del modelo en campo, comparando el rendimiento real de las brocas con su rendimiento estimado. De esta forma, se cuantifica el ahorro potencial en insumos y costos operativos.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

# CAPITULO 1

## 1. PLANTEAMIENTO DE LA INVESTIGACION

### 1.1. Planteamiento del Problema

En minería superficial, la fase de perforación constituye un eslabón crítico dentro del ciclo productivo, ya que de su eficiencia dependen el avance de banco, la calidad de la voladura y el cumplimiento de los planes de producción. Dentro de esta etapa, las brocas tricónicas de 11" representan un insumo clave cuya vida útil influye directamente en los costos operativos y en la disponibilidad de los equipos de perforación.

La toma de decisiones sobre el cambio o reemplazo de brocas se realiza, en la mayoría de casos, bajo criterios empíricos o promedios históricos, sin considerar variables operativas específicas como el *pulldown* (presión de empuje) las revoluciones por minuto (RPM), la presión de aire y la velocidad de penetración (ROP). Esta falta de análisis genera escenarios operativos ineficientes, como el cambio prematuro de brocas que implica un mal aprovechamiento de la broca o el uso extendido más allá de su capacidad máxima lo que reduce la eficiencia de perforación, incrementa el riesgo de daños, reduciendo la productividad y elevando el costo de perforación.

A pesar de contar con registros históricos de parámetros de perforación en mina , estos datos no están siendo utilizados para predecir el rendimiento real de las brocas. Esto representa una oportunidad clara para aplicar herramientas tecnológicas basadas en machine learning, capaces de procesar múltiples variables y generar modelos de predicción más precisos y ajustados a las condiciones de campo.

La falta de un modelo predictivo implementado limita la capacidad de gestión técnica y económica del área de perforación. Una herramienta que permita anticipar con confiabilidad la



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

vida útil de cada broca no solo facilitaría una mejor planificación del reemplazo, sino que además permitiría reducir costos operativos y aumentar la productividad de los equipos.

Por lo tanto, el problema central radica en la ausencia de un sistema analítico que transforme los datos operativos en una herramienta de apoyo para la toma de decisiones, permitiendo estimar de forma precisa cuándo una broca triconica de 11” debe ser reemplazada, maximizando su rendimiento y optimizando el uso de los recursos.

## 1.2. Objetivos

### 1.2.1. Objetivo General

Desarrollar un modelo predictivo en Python, basado en técnicas de machine learning para estimar la vida útil de brocas triconicas de 11” en una mina superficial y reducir los costos operativos asociados a su reemplazo en operaciones.

### 1.2.2. Objetivos Específicos

- Estructurar y depurar la base de datos de registros operativos de las brocas tricónicas de 11”
- Analizar la relación entre la vida útil de las brocas y los parámetros operativos (pulldown, RPM, presión de aire y ROP) mediante análisis de correlación.
- Evaluar y comparar tres modelos de regresión en Python, Regresión Lineal Múltiple, Random Forest Regressor y XGBoost Regressor para generar una ecuación predictiva de la vida útil.
- Medir el desempeño de los modelos utilizando métricas estadísticas ($R^2$, MAE y RMSE) y seleccionar el más preciso.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

- Estimar el ahorro operativo generado por la implementación del modelo predictivo optimizando el reemplazo de brocas.

## 1.3. Preguntas de Investigación

### 1.3.1. Pregunta Global

¿Qué modelo de machine learning, implementado en Python, permite estimar con mayor precisión la vida útil de las brocas tricónicas de 11" en minería superficial, optimizando su reemplazo y reduciendo costos operativos?

### 1.3.2. Preguntas Especificas

¿Cuál de los parámetros operativos (Pulldown, RPM, presión de aire y ROP) tiene mayor correlación con la vida útil de las brocas triconicas de 11"?

¿ Qué modelo predictivo ofrece el mayor valor de $R^2$ y el menor error (MAE y RMSE) al estimar la vida útil de las brocas a partir de datos históricos?

¿Qué impacto económico tiene la aplicación del modelo predictivo en la planificación del reemplazo de brocas y en los costos operativos de perforación?

## 1.4. Línea de Investigación a la que corresponde el Problema

Optimización de Procesos Mineros

## 1.5. Palabras Clave

Machine learning, brocas tricónicas, ROP, RPM, pulldown.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

## 1.6. Solución Propuesta

### 1.6.1. Justificación e Importancia

Esta aplicación de la tecnología y análisis se realiza para resolver un problema operativo frecuente en minería superficial, el cual es la falta de precisión al estimar la vida útil de las brocas triconicas de 11" ya que actualmente, su reemplazo se basa en criterios empíricos y generales, lo que genera sobrecostos por cambios prematuros o pérdidas de eficiencia por desgaste excesivo.

El objetivo es aprovechar la data operativa disponible y aplicar modelos de machine learning en Python para predecir el desgaste de brocas, optimizar su uso y reducir costos de perforación.

El principal beneficiario será el área de perforación al contar con una herramienta de apoyo técnico y económico para la toma de decisiones ya que la investigación es factible porque se desarrollará con datos reales de la mina y herramientas de software libre.

Descripción de la Solución

La solución propuesta consiste en el desarrollo de un modelo predictivo en Python, basado en técnicas de machine learning, que permita estimar con precisión la vida útil de las brocas triconicas de 11" utilizadas en operaciones de perforación en minería superficial. El modelo será alimentado con datos operativos reales como pulldown, RPM, presión de aire y velocidad de penetración (ROP), seleccionando solo la variable que más influya mediante análisis de correlación.

El producto final será una herramienta analítica funcional y replicable, capaz de predecir el metraje estimado que una broca puede alcanzar bajo condiciones específicas. Este modelo permitirá al área de perforación planificar de forma técnica el reemplazo de brocas, evitando



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

cambios prematuros o desgaste excesivo, lo que generará reducción de costos operativos y mejora en la eficiencia del proceso de perforación.

Se espera que la herramienta creada en Python para la predicción permita tomar decisiones basadas en datos y no en criterios empíricos, de esta manera la solución puede extenderse fácilmente a brocas de distinto diámetro, siempre que se cuente con registros operativos adecuados y mejorar el proceso.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

# CAPITULO 2

## 2. FUNDAMENTOS TEÓRICOS

### 2.1. Estado del arte

Durante la última década, el estudio del desgaste de brocas triconicas en minería superficial ha evolucionado de enfoques empíricos hacia metodologías avanzadas de machine learning, destacando la creciente importancia de integrar datos operativos rutinarios como pulldown, RPM, presión de aire y velocidad de penetración (ROP) en la estimación predictiva de vida útil de consumibles críticos. En 2020, Alajmi y Almeshal llevaron a cabo un enfoque híbrido denominado XGBoost-SDA, que combinó XGBoost con un algoritmo de optimización dinámica espiral para predecir el desgaste lateral (“flank wear”) en herramientas de corte empleadas en metalmecánica (cobre, hierro fundido), logrando un coeficiente de determinación $R^2$ de 0,997 y un MAE próximo al 5 % . Este estudio, aunque realizado en un ambiente industrial controlado, demostró la precisión ejemplar de XGBoost cuando se garantiza homogeneidad de datos y condiciones estables. Un año después, Sotubadi et al. (2021) publicaron una investigación en arXiv en la que aplicaron Random Forest junto con SHAP como herramienta de inteligencia artificial explicable (XAI) para revelar las variables más influyentes en trayectorias de torneado, mostrando que la interpretabilidad es clave para adoptar modelos en entornos operativos reales. Aunque su campo no se centró en brocas rotativas, la metodología XAI resulta prometedora para validar decisiones basadas en machine learning en contextos con variabilidad en condiciones de terreno. En el ámbito concreto de perforación industrial, en 2021 investigadores del International Journal of Mining Science and Technology usaron redes neuronales convolucionales (CNN) y un análisis por wavelet de señales de vibración para clasificar diferentes grados de desgaste de brocas triconicas, logrando alta



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

confiabilidad en la detección temprana de fallas. Pese a su éxito, este modelo requiere sensores de vibración costosos, lo que limita su aplicabilidad dentro de operaciones de gran escala donde la simplicidad operativa es crucial.

Hacia 2022–2023, se destacaron estudios en entornos subterráneos y petroleros: investigadores en túneles aplicaron XGBoost combinado con metaheurísticos como Bayesian Optimization, HHO (Harris Hawks Optimization) y DA (Dragonfly Algorithm) para predecir la ROP y condiciones de sobreexcavación, alcanzando $R^2$ superiores a 0,93 incluso con variabilidad geológica significativa. Zhang et al. (2022), en un análisis del comportamiento de desgaste de rodajas en vetas de carbón, integraron atributos físicos-mecánicos y técnicas de aprendizaje automático, estableciendo correlaciones sólidas que mejoran la precisión de modelos predictivos en terrenos heterogéneos. Estos avances refuerzan la idea de que es viable construir modelos predictivos robustos aun en ausencia de condiciones controladas, siempre que el dataset incluya variables representativas del entorno.

En 2023, Yumani et al. presentaron un modelo que combinó visión artificial (YOLO) con Random Forest para identificar visualmente causas de falla en brocas de perforación, logrando un F1-score macro promedio de 0,94. Este enfoque demuestra que la integración entre métodos de visión por computadora y machine learning puede ser útil en mantenimiento predictivo, aunque su dependencia de sistemas de captura de imagen lo hace menos viable para aplicaciones puramente operativas en campo, especialmente en entornos polvorientos o con visibilidad limitada. Paralelamente, en 2023–2024 se observan investigaciones de previsión en tiempo real de desgaste en maquinaria pesada como TBM (túneles de perforación), donde se implementaron XGBoost y Random Forest para estimar el desgaste de discos de corte, obteniendo $R^2$ de hasta 0,942 , lo que



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

indica que estos algoritmos son aplicables a herramientas de gran escala y condiciones variables a cielo abierto.

Dentro del ámbito más cercano a nuestro estudio, Khalifa et al. (2024) aplicaron Random Forest, con balanceo de base de datos mediante K-means y SMOTE, sobre un dataset de 743 registros para clasificar cinco niveles de desgaste en brocas triconicas, alcanzando una precisión del 87 % en condiciones reales de perforación superficial. Aunque su trabajo sebasó en clasificación en lugar de regresión continua, demuestra que los datos recopilados en campo son suficientes para generar herramientas predictivas útiles en entornos operativos reales como la perforación superficial. Desde otra perspectiva, Senjoba et al. (2024) introdujeron técnicas basadas en Explainable AI utilizando Grad-CAM aplicado a CNN para rehabilitar la interpretación de patrones de desgaste en brocas rotativas a través de mapas de atención, mejorando la transparencia y confianza en las predicciones. Este avance ofrece un camino para combinar precisión y explicabilidad, clave en contextos donde los operadores deben confiar en los modelos de IA.

En resumen, la literatura de los últimos diez años muestra una clara tendencia hacia el uso de machine learning, especialmente algoritmos como XGBoost y Random Forest, reforzados con técnicas de optimización metaheurística, redes neuronales o componentes de visión artificial. Si bien los estudios destacan altos niveles de precisión, su mayoría depende de condiciones controladas, sensores especializados o se desarrollan en escenarios industriales o subterráneos; pocos concilian el equilibrio entre precisión, simplicidad operativa y uso de variables estándar de perforación. Es aquí donde la presente investigación se posiciona como un aporte significativo: desarrolla un modelo predictivo en Python para estimar la vida útil de brocas triconicas de 11” en una mina superficial, utilizando únicamente variables registradas rutinariamente en operación como el Pulldown, RPM, presión de aire y ROP e implementando técnicas como regresión lineal



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

múltiple, Random Forest y XGBoost, optimizadas para condiciones reales de minería superficial. Esta combinación permitirá una herramienta de alto valor técnico y económico ya que ofrece precisión validada por $R^2$, MAE o RMSE en condiciones reales, sin necesidad de sensores adicionales costosos, y facilitará decisiones oportunas en campo sobre reemplazo de brocas y generando reducción de costos. Como conclusión, se puede confirmar que esta investigación cierra una brecha técnica entre los modelos predictivos existentes y su aplicación operativa en minería superficial, aportando una solución adaptable y económica que esta alineada con las necesidades del sector.

## 2.2. Bases Teóricas de la Investigación

### 2.2.1. Perforación en Minería Superficial

La perforación constituye una de las etapas más relevantes en el ciclo productivo de una mina a cielo abierto, ya que tiene un impacto directo en la eficiencia de la voladura, el avance de banco y la continuidad de las operaciones. Esta actividad implica la realización de taladros mediante equipos rotativos o de percusión, y su desempeño depende de factores como la litología, parámetros operativos, condición de las herramientas y estrategia de perforación (Smith, 2015).

### 2.2.2. Brocas Triconicas 11”

Las brocas tricónicas son herramientas de corte rotatorio utilizadas ampliamente en la perforación rotativa para fracturar rocas duras. Estas brocas están compuestas por tres conos giratorios con insertos de carburo de tungsteno o dientes cortantes que giran de forma independiente para triturar la roca. La broca de 11 pulgadas es común en la perforación de producción en minería superficial debido a su balance entre penetración y resistencia al desgaste (Scholars Mine, 2016).



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

### 2.2.3. Desgaste de Brocas Triconicas

El desgaste de las brocas se refiere al deterioro progresivo de sus componentes, como los dientes o insertos, debido al uso continuo en la perforación, este desgaste afecta la eficiencia de la broca, disminuyendo su velocidad de penetración y aumentando el costo por metro perforado y está influenciado por variables operativas y geomecánicas, como la dureza del terreno, la presión aplicada, la velocidad de penetración, entre otras.

### 2.2.4. Vida Útil de Brocas Triconicas

La vida útil de la broca se mide en metros perforados hasta que la broca deja de ser eficiente o se vuelve inutilizable por desgaste extremo.

### 2.2.5. Pulldown

Representa la velocidad angular de rotación de la broca. Valores elevados incrementan el número de impactos de los dientes sobre la roca, pero también pueden generar calor excesivo o desgaste prematuro si no se combinan adecuadamente con el pulldown y el tipo de terreno (Martínez et al., 2021).

### 2.2.6. Revoluciones por Minuto (RPM)

epresenta la velocidad angular de rotación de la broca. Valores elevados incrementan el número de impactos de los dientes sobre la roca, pero también pueden generar calor excesivo o desgaste prematuro si no se combinan adecuadamente con el pulldown y el tipo de terreno (Martínez et al., 2021).

### 2.2.7. Presión de Aire

El aire comprimido tiene una doble función: evacuar los detritos del fondo del pozo y enfriar la broca durante la perforación. Una presión insuficiente puede causar acumulación de



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

material, provocando recirculación, ralentización y desgaste; por otro lado, una presión excesiva puede dañar el sistema neumático.

### 2.2.8. Velocidad de Penetración (ROP)

La ROP (Rate of Penetration) se refiere al metraje perforado por unidad de tiempo, normalmente en metros por hora. Es un indicador directo de eficiencia operativa y está altamente correlacionado con las condiciones del terreno y los parámetros aplicados. Fluctuaciones bruscas en ROP pueden indicar cambios litológicos o problemas con la broca.

### 2.2.9. Costos Operativos Asociados a la Perforación

Las brocas representan un insumo significativo dentro del costo por metro perforado. Una estrategia ineficiente de reemplazo puede derivar en el uso prematuro de herramientas aún funcionales o en prolongar su uso más allá de lo recomendable, afectando el rendimiento del equipo y aumentando el costo total de operación. Por ello, mejorar la estimación de vida útil mediante modelos predictivos tiene un impacto económico directo.

### 2.2.10. Machine Learning

El aprendizaje automático (machine learning) es una disciplina de la inteligencia artificial que permite a los algoritmos generar conocimiento a partir de datos, sin requerir programación explícita para cada tarea. Se basa en el entrenamiento de modelos con datos históricos para que reconozcan patrones y realicen predicciones sobre nuevas observaciones (Mohri, Rostamizadeh & Talwalkar, 2018).

### 2.2.11. Machine Learning en Minería

En minería, el machine learning ha cobrado importancia en la última década por su capacidad para mejorar la toma de decisiones en condiciones de incertidumbre. Se ha utilizado en



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

mantenimiento predictivo, estimación de leyes, clasificación de minerales, y recientemente en la predicción de rendimiento de equipos como perforadoras, cargadores o palas eléctricas (Martínez et al., 2021; Khalifa et al., 2024).

### 2.2.12. Regresión Lineal Múltiple

Es un modelo estadístico clásico que permite explicar la relación lineal entre una variable dependiente (como la vida útil de una broca) y múltiples variables independientes (pulldown, RPM, etc.). Se valora por su simplicidad y capacidad interpretativa (James et al., 2013).

### 2.2.13. Random Forest Regressor

Es un modelo basado en ensambles de árboles de decisión que mejora la precisión y evita el sobreajuste. Es útil cuando las relaciones entre variables no son estrictamente lineales y cuando existe ruido o heterogeneidad en los datos (Breiman, 2001).

### 2.2.14. XGBoost Regressor

Es un modelo de boosting que combina árboles secuenciales optimizados mediante gradiente. Tiene gran capacidad para detectar relaciones complejas, funciona bien con datos heterogéneos y ha sido ampliamente adoptado en competencias de ciencia de datos por su alto rendimiento (Chen & Guestrin, 2016).

### 2.2.15. Coeficiente de Determinación ($R^2$)

Evalúa la proporción de la variabilidad explicada por el modelo. Un valor de $R^2$ cercano a 1 indica que las predicciones se ajustan bien a los datos observados.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

2.2.16. Error Absoluto Medio (MAE)

Calcula la media de las diferencias absolutas entre los valores reales y los predichos, lo que permite interpretar el error promedio en las mismas unidades de la variable dependiente.

2.2.17. Raíz del Error Cuadrático Medio (RMSE)

Mide la raíz cuadrada del promedio de los errores al cuadrado, penalizando más los errores grandes. Es útil para comparar modelos en términos de precisión.

2.2.18. Python

Es un lenguaje de programación de alto nivel, interpretado y de propósito general, caracterizado por una sintaxis clara y legible, lo que facilita el desarrollo rápido de aplicaciones en diversas áreas como análisis de datos, inteligencia artificial, desarrollo web y automatización de procesos (Van Rossum & Drake, 2009).



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

# CAPITULO 3

## 3. MARCO METODOLOGICO

### 3.1. Alcances y Limitaciones

#### 3.1.1. Alcances

Este estudio se centra en desarrollar un modelo predictivo en Python para estimar la vida útil de brocas triconicas de 11” usadas en minería superficial, tomando como base datos históricos de perforación de una mina superficial. Se analizarán los parámetros operativos (Pulldown, RPM, presión de aire y ROP) para seleccionar el de mayor influencia mediante correlación, y se aplicarán modelos de regresión (MLR, Random Forest y XGBoost). El objetivo es predecir con mayor precisión el metraje alcanzable por broca y reducir los costos asociados a su reemplazo prematuro

#### 3.1.1. Limitaciones

La investigación se limita a data histórica ya registrada, no se contempla medición en campo ni validación en tiempo real. El estudio se enfoca solo en brocas triconicas de 11” dentro de una operación minera superficial.

### 3.2. Aporte

Aportará una herramienta predictiva practica que permitirá predecir la vida útil de las brocas de 11” y reducir los costos de perforación.

### 3.3. Tipo de Nivel de la Investigación

#### 3.3.1. Tipo de Nivel de la Investigación

La presente investigación se clasifica como aplicativa y tecnológica porque su finalidad no es únicamente generar conocimiento teórico, sino resolver un problema operativo específico en



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

minería superficial mediante el diseño de una herramienta predictiva funcional. En este caso, el problema identificado es la falta de precisión en la estimación de la vida útil de brocas triconicas de 11”, ya que actualmente se reemplazan según criterios empíricos, generando sobrecostos o pérdidas operativas.

Desde el enfoque aplicativo, el estudio utiliza conocimientos existentes de perforación y técnicas de machine learning para proponer una solución concreta, adaptada a las condiciones reales de operación en mina. Esto se traduce en la implementación de un modelo predictivo en Python, entrenado con registros históricos operativos, cuyo propósito es optimizar la toma de decisiones en el reemplazo de brocas y reducir costos.

Desde el enfoque tecnológico, la investigación implica el desarrollo de un producto funcional, es decir, un modelo que pueda integrarse a la operación como herramienta técnica. Se hace uso de herramientas de programación, procesamiento de datos, validación estadística y algoritmos avanzados de predicción como XGBoost y Random Forest, lo que representa una innovación tecnológica aplicable al contexto de minería superficial.

### 3.3.2. Nivel de la Investigación

La presente investigación se enmarca en el nivel correlacional y experimental, dado que busca analizar la relación existente entre variables operativas de perforación y la vida útil de las brocas tricónicas de 11” en operaciones de minería superficial, y probar el efecto que tienen dichas variables al aplicar un modelo predictivo en condiciones controladas mediante simulación computacional.

En primer lugar, el nivel correlacional se justifica porque uno de los objetivos centrales de la tesis es determinar el grado de relación entre la variable dependiente (vida útil de la broca en



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

metros perforados) y las variables independientes: pulldown, RPM, presión de aire y velocidad de penetración (ROP). A través de técnicas estadísticas como el coeficiente de correlación de Pearson, se establecerá qué tan fuerte es la asociación entre cada parámetro y el desgaste de la broca, permitiendo seleccionar las variables más significativas para construir el modelo predictivo. Este tipo de análisis es típico de investigaciones correlacionales, donde se busca entender el comportamiento conjunto de las variables sin alterar el entorno original en campo.

Por otro lado, la investigación también se considera experimental, en tanto se aplica un proceso sistemático de prueba y validación de diferentes algoritmos de regresión (como Regresión Lineal Múltiple, Random Forest y XGBoost) sobre un conjunto de datos históricos, lo cual implica intervención activa del investigador para modificar las condiciones analíticas mediante técnicas computacionales, evaluar los efectos y comparar los resultados utilizando métricas como $R^2$, MAE y RMSE. Aunque no se ejecutan pruebas físicas en campo, la experimentación se realiza dentro de un entorno simulado controlado (en Python), donde se prueban diferentes configuraciones del modelo hasta obtener una ecuación óptima que prediga con precisión la vida útil de las brocas.

### 3.4. Población y Muestra

Para el estudio se tomarán datos históricos de las brocas de 11” de los meses de enero a junio del 2025, en la fase 05 de la mina superficial, los cuales se instalaron en 3 modelos de perforación rotativa (Bucyrus 49HR, Bucyrus 49RIII y CAT6640) y marca de triconos (Cenerg CN71).



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM
UNIVERSIDAD CATÓLICA DE SANTA MARÍA

Tabla 1 Equipos de perforación

| Código | Marca o Modelo | Diámetro (pulg) | Empuje máximo Kg | Longitud de Barra (m) |
|---|---|---|---|---|
| **DR04** | BUCYRUS 49 RIII | 9.6-16 | 49895 | 16.5 – 18.9 |
| **DR05** | BUCYRUS 49 RIII | 9.6-16 | 49895 | 16.5 – 18.9 |
| **DR07** | BUCYRUS 49 HR | 9.6-16 | 49895 | 16.5 – 18.9 |
| **DR08** | BUCYRUS 49 HR | 9.6-16 | 49895 | 16.5 – 18.9 |
| **DR09** | BUCYRUS 49 HR | 9.6-16 | 49895 | 16.5 – 18.9 |
| **DR12** | CAT MD6640 | 10.63 – 15.98 | 64000 | 16.5 – 18.9 |
| **DR18** | CAT MD6640 | 9.6-16 | 64000 | 16.5 – 18.9 |

Nota: Elaboración propia en base a la información obtenida en campo

Tabla 2 Ubicación de las perforadoras

| Perforadora | Nivel | Fase | Litología | Tipo de material |
|---|---|---|---|---|
| **DR04** | 2740-2695 | Fase 05 | Tonalita-Brecha Hidrotermal, Brecha de dacita Porfiritica | Mineral |
| **DR05** | 2740-2695 | Fase 05 | | |
| **DR07** | 2740-2695 | Fase 05 | | |
| **DR08** | 2740-2695 | Fase 05 | | |
| **DR09** | 2740-2695 | Fase 05 | | |
| **DR12** | 2740-2695 | Fase 05 | | |
| **DR18** | 2740-2695 | Fase 05 | | |

Nota: Elaboración propia en base a la información obtenida en campo



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

Las brocas que se usaron en dichas perforadoras y corrieron en las fases mencionadas representan la población de la cual se hará el estudio.

Para la presente investigación, se cuenta con una base de datos histórica compuesta por 130 registros operativos de brocas triconicas de 11" utilizadas en minería superficial. Aplicando la fórmula de muestreo para poblaciones infinitas, con un nivel de confianza del 95% y un margen de error del 5%, una proporción esperada de 5% se determinó que el tamaño mínimo de muestra requerido para garantizar resultados estadísticamente válidos es de:

$$Z = 1.96 \; para \; una \; confianza \; del \; 95\%$$

$$p = 0.05$$

$$q = 1 - p = 0.95$$

$$e = 0.05$$

$$n = \frac{z^2 * p * q}{e^2}$$

$$n = \frac{1.96^2 * 0.05 * 0.95}{0.05^2}$$

$$n = 72.99 = 73$$

Se necesita entonces una muestra de 73 registros de brocas triconicas de 11".

## 3.5. Ubicación de la investigación

La presente investigación se desarrolla en una operación minera a tajo abierto ubicada en la región sur del Perú, caracterizada por una explotación a gran escala de minerales metálicos. La



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

zona se encuentra dentro de un contexto geográfico de altitud considerable, con condiciones climáticas propias de la sierra sur andina, donde las temperaturas pueden variar significativamente entre el día y la noche. El acceso a la mina se realiza a través de vías pavimentadas y caminos afirmados que conectan la operación con centros poblados cercanos y con la red vial nacional. El área de estudio presenta infraestructura moderna para las operaciones de perforación y voladura, incluyendo talleres de mantenimiento, áreas de acopio, sistemas de abastecimiento de aire y agua, y plataformas de perforación adecuadamente acondicionadas para equipos de gran capacidad.

### 3.6. Descripción de la geología general

El yacimiento se encuentra emplazado en una zona de intrusiones ígneas de composición predominantemente tonalítica y granodiorítica, con mineralización principalmente de calcopirita, bornita y, en menor proporción, molibdenita. La alteración hidrotermal ha generado zonas de potásica, fílica y propilítica, que definen las variaciones en dureza y abrasividad de la roca. Estas características geomecánicas influyen de manera significativa en la selección de parámetros de perforación y en la vida útil de las brocas. El macizo rocoso presenta un rango de resistencia a la compresión uniaxial (UCS) entre 120 y 220 MPa, con sectores de mayor abrasividad asociados a la presencia de cuarzo y feldespato.

### 3.7. Descripción geomecánica y geotécnica del tajo

El tajo abierto en estudio está emplazado en una zona geológicamente compleja del sur del Perú, donde predominan formaciones ígneas intrusivas de origen pórfido ricos en minerales como calcopirita, bornita y molibdenita. Estas intrusiones, sobre una base de rocas sedimentarias y volcanosedimentarias alteradas, han sido sujeto de procesos hidrotermales que generan alteraciones potásicas, propilíticas y fíclicas—características que condicionan significativamente las propiedades mecánicas del macizo rocoso



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

### 3.7.1. Propiedades Geomecánicas Principales

Desde el punto de vista geomecánico, la roca masiva muestra, en general, una alta competencia con valores de resistencia compresiva uniaxial que superan los 120 MPa en zonas frescas, mientras que en sectores alterados estos valores disminuyen de manera significativa. El macizo rocoso se encuentra atravesado por sistemas de diaclasas orientados principalmente en dirección noreste-suroeste y noroeste-sureste, con buzamientos que varían entre 50° y 70°. Esta fracturación controla la estabilidad de los taludes y la fragmentación del material volado.

En términos geotécnicos, el diseño del tajo considera bermas de seguridad y ángulos de talud ajustados a las características de cada sector, aplicando factores de seguridad que garantizan la estabilidad global. La presencia de zonas de debilidad asociadas a fallas locales y planos de alteración arcillosa obliga a reforzar algunos taludes con sistemas de drenaje y monitoreo geotécnico continuo.

La caracterización geomecánica, apoyada en clasificaciones como RMR y GSI, ubica al macizo rocoso en rangos de calidad media a buena en la mayor parte del tajo, con sectores puntuales de calidad regular a baja debido a la presencia de zonas alteradas. Esta variabilidad incide en la selección de parámetros de perforación, donde se requiere ajustar el pulldown y la velocidad de penetración para optimizar la vida útil de las brocas tricónicas y reducir costos operativos.

## 3.8. Parámetros de diseño banco-talud

En el diseño de muros de mina a cielo abierto, los parámetros geométricos tales como altura de banco, ángulo de cara, bermas, rampas y pendiente general son esenciales para garantizar



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

estabilidad y eficiencia operativa. Para el caso estudiado, se utilizan valores respaldados por documentación técnica disponible:

- Altura del banco (BH): 15 m
- Ángulo de cara de banco (Bench Face Angle): 65°
- Anchura de berma (Catch Bench Width): Variable según el dominio estructural, estimada entre 8 y 8,5 m en condiciones generales operativas.
- Pendiente de rampas internas (Ramps): entre 8 % y 10 % (equivalente a ángulo de pendiente aproximado de 5° a 6°)

## 3.9. Tipo de broca seleccionada

Para el presente estudio se seleccionó una broca tricónica de 11 pulgadas de diámetro, de la marca de Cenerg, diseñada para operaciones de perforación rotativa en formaciones de alta resistencia y abrasividad. Este tipo de broca se caracteriza por contar con conos equipados con insertos de carburo de tungsteno, dispuestos para maximizar la fragmentación de la roca y optimizar la velocidad de penetración. Su diseño incorpora sistemas de circulación de aire para la evacuación de detritos y la reducción del calor generado durante la perforación, lo que contribuye a prolongar su vida útil. La elección de este modelo responde a su adaptabilidad a condiciones geológicas variables y a su desempeño comprobado en operaciones de gran escala en minería a tajo abierto.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

Tabla 3 Características Técnicas de la broca de 11"

| Parámetro | Cenerg |
|---|---|
| **Diámetro en pulgadas** | 11” |
| **Código IADC** | 7-1-5 |
| **Peso de la broca (kg)** | 70 kg |
| **Nro. fila del cono (calibre/internos)** | 3 (Gage) / 11 (Inner) |
| **Insertos en el cono (calibre/internos)** | 59 (Gage) / 127 (Inner) |
| **Forma de insertos** | Gage: Cónicos, Inner: Cónicos y Shirttail: Esféricos |
| **Rango de empuje (Kg)** | 19958 – 39916 kg |
| **Rango de velocidad de rotación (RPM)** | 50 – 90 RPM |
| **Resistencia a la compresión uniaxial (Mpa)** | 241 – 397 MPa |
| **Aleación de la estructura de corte** | Carburo de tungsteno |

Nota: Elaborado a partir de la información del catálogo de brocas Cenerg.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

Figura 1 Broca Triconica de 11"

Nota: Herramienta de perforación empleada en operaciones de minería superficial para la ejecución de taladros, foto tomada en campo.

### 3.10. Métodos, Técnicas e Instrumentos de Recolección de Dato

Para el desarrollo de esta investigación, los datos fueron proporcionados por el área de Perforación y Voladura de una operación minera superficial, donde se realiza un monitoreo constante de las condiciones y parámetros de perforación con brocas triconicas de 11”.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

Se empleó una base de datos histórica, compuesta por 130 registros, cada uno correspondiente a los parámetros de perforación registrados por cada broca, en los que se consignan parámetros técnicos obtenidos de los equipos de perforación.

Las variables recolectadas incluyen:

Tabla 4 Variables de recolección de datos

| Variable | Tipo | Unidad de medida | Descripción | Escala |
|---|---|---|---|---|
| **Vida útil de broca** | Dependiente | m | Metros perforados hasta alcanzar desgaste crítico | Razón |
| **Pulldown** | Independiente | kN | Fuerza de empuje sobre la broca | Razón |
| **RPM** | Independiente | rev/min | Velocidad de rotación de la broca | Razón |
| **Presión de aire** | Independiente | psi | Presión de aire para enfriamiento y evacuación de detritos | Razón |
| **Velocidad de penetración (ROP)** | Independiente | m/h | Tasa de avance de perforación | Razón |
| **Tipo de roca** | Control | - | Clasificación geomecánica (duro, medio, blando) | Nominal |

Nota: El registro de datos se realizó mediante hojas de cálculo en formato Excel (.xlsx), las cuales fueron exportadas del sistema de control y supervisión operativa del área.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

### 3.11. Plan de análisis estadístico y recolección de los datos

Para el procesamiento de los datos se utilizó el lenguaje de programación Python, apoyado con bibliotecas como pandas y openpyxl para la lectura y tratamiento de los archivos.

- Depuración de registros: eliminación de duplicados y registros con errores evidentes.
- Validación cruzada: contraste de datos del sistema con bitácoras de campo.
- Normalización de unidades: conversión a sistema internacional (SI) cuando fue necesario.
- Codificación de variables categóricas: el tipo de roca se codificó numéricamente para su uso en modelos.
- Análisis exploratorio: detección y tratamiento de valores atípicos mediante criterios estadísticos (±3 desviaciones estándar).

### 3.12. Plan de análisis de datos

- El procedimiento metodológico se estructuró en seis etapas:

#### 3.12.1. Preparación y limpieza de datos

- Organización y depuración de la base de datos de brocas tricónicas, eliminando valores nulos, corrigiendo inconsistencias y verificando rangos operativos válidos para asegurar la calidad del análisis.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

3.12.2. Selección de variables

- Evaluación de la correlación de Pearson entre cada parámetro operativo y la vida útil.
- Retención de variables con correlación significativa o relevancia técnica comprobada.

3.12.3. Análisis de correlación

- Evaluación de la relación lineal entre la vida útil y las variables operativas mediante el coeficiente de Pearson, para determinar su grado de influencia.

3.12.4. Visualización de correlaciones

- Generación de gráficos de barras que muestran la magnitud y dirección de la correlación de cada variable respecto a la vida útil, facilitando la interpretación visual.

3.12.5. Desarrollo de modelos predictivos

- Entrenamiento y evaluación de tres algoritmos (Regresión Lineal, Random Forest y XGBoost) para estimar la vida útil de las brocas, usando métricas de desempeño como $R^2$, MAE y RMSE.

3.12.6. Comparación de desempeño de modelos

- Construcción de tablas y gráficos comparativos de resultados para seleccionar el modelo con mejor capacidad predictiva y menor error.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

3.12.7. Selección del modelo final

- Elección del modelo Random Forest Regressor como el más adecuado por su mayor $R^2$ y menores MAE y RMSE, optimizando la predicción de la vida útil.

3.12.8. Validación del modelo

- Prueba del modelo con datos simulados dentro del rango operacional para verificar su coherencia y capacidad de generalización.

3.12.9. Almacenamiento del modelo

- Guardado del modelo entrenado en formato. pkl para su futura implementación en entornos operativos sin necesidad de reentrenarlo.

3.12.10. Estimación del ahorro operativo

- Cálculo del impacto económico de aplicar el modelo, estimando la reducción del consumo anual de brocas y el ahorro monetario asociado.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM
UNIVERSIDAD CATÓLICA DE SANTA MARÍA

# CAPITULO 4

## 4. PLAN DE TRABAJO

### 4.1. Preparación y limpieza de datos

Para el desarrollo del modelo predictivo orientado a estimar la vida útil de brocas triconicas de 11” en una operación minera superficial, el primer paso consistió en la recopilación, estructuración y limpieza de la base de datos proporcionada por el área de Perforación y Voladura. Este conjunto de datos incluía registros históricos de diferentes parámetros operativos de las brocas, tales como:

- Vida útil (expresada en metros perforados)
- Velocidad de penetración (ROP)
- Pulldown (presión ejercida hacia abajo)
- RPM (revoluciones por minuto)

Tabla 5 Data de brocas

| VIDA UTIL | VELOCIDAD DE PENETRACIÓN | PULLDO KN | RPM |
|---|---|---|---|
| **940.46** | 18.22 | 72.00 | 61.00 |
| **984.08** | 19.04 | 74.00 | 67.00 |
| **1037.16** | 20.12 | 71.00 | 65.00 |
| **1041.38** | 19.03 | 74.00 | 60.00 |
| **1041.77** | 20.89 | 76.00 | 64.00 |
| **1077.80** | 21.40 | 74.00 | 69.00 |
| **1088.66** | 20.88 | 73.00 | 63.00 |
| **1120.16** | 20.98 | 76.00 | 63.00 |
| **1124.27** | 22.24 | 74.00 | 58.00 |
| **1134.19** | 21.58 | 73.00 | 59.00 |



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

| | | | |
|---|---|---|---|
| **1150.96** | 20.50 | 75.00 | 62.00 |
| **1157.95** | 21.34 | 79.00 | 66.00 |
| **1167.39** | 21.60 | 72.00 | 63.00 |
| **1170.70** | 22.90 | 70.00 | 62.00 |
| **1187.83** | 20.63 | 73.00 | 63.00 |
| **1188.99** | 20.99 | 74.00 | 58.00 |
| **1231.34** | 24.34 | 73.00 | 58.00 |
| **1234.72** | 24.03 | 75.00 | 64.00 |
| **1240.74** | 22.23 | 75.00 | 68.00 |
| **1242.66** | 24.00 | 70.00 | 62.00 |
| **1246.18** | 23.87 | 71.00 | 58.00 |
| **1257.92** | 24.29 | 73.00 | 58.00 |
| **1268.08** | 22.41 | 77.00 | 64.00 |
| **1271.33** | 24.56 | 77.00 | 66.00 |
| **1279.50** | 22.49 | 77.00 | 63.00 |
| **1291.67** | 23.18 | 75.00 | 64.00 |
| **1405.62** | 27.14 | 73.00 | 60.00 |
| **1414.32** | 26.02 | 70.00 | 59.00 |
| **1417.43** | 25.19 | 83.00 | 66.00 |
| **1417.94** | 26.65 | 73.00 | 61.00 |
| **1421.29** | 26.65 | 79.00 | 70.00 |
| **1429.39** | 27.81 | 71.00 | 66.00 |
| **1448.20** | 28.28 | 79.00 | 67.00 |
| **1449.73** | 27.85 | 67.00 | 63.00 |
| **1457.60** | 26.44 | 79.00 | 70.00 |
| **1482.07** | 28.25 | 75.00 | 71.00 |
| **1503.71** | 26.79 | 74.00 | 59.00 |
| **1504.75** | 26.84 | 76.00 | 70.00 |
| **1515.74** | 26.90 | 81.00 | 66.00 |
| **1573.35** | 28.82 | 78.00 | 63.00 |
| **1585.28** | 29.56 | 72.00 | 65.00 |
| **1599.19** | 29.40 | 80.00 | 67.00 |
| **1599.62** | 31.11 | 75.00 | 74.00 |
| **1690.31** | 33.46 | 71.00 | 68.00 |
| **1724.51** | 33.28 | 71.00 | 68.00 |
| **1736.81** | 33.88 | 78.00 | 66.00 |
| **1766.14** | 34.39 | 70.00 | 64.00 |
| **1845.12** | 34.26 | 77.00 | 65.00 |
| **1907.64** | 36.51 | 78.00 | 67.00 |
| **1908.66** | 35.29 | 69.00 | 63.00 |
| **1909.30** | 35.30 | 71.00 | 65.00 |
| **1911.62** | 35.31 | 70.00 | 59.00 |
| **1955.60** | 35.63 | 73.00 | 61.00 |
| **2083.98** | 40.27 | 71.00 | 61.00 |



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

| | | | |
|---|---|---|---|
| **2094.04** | 40.58 | 71.00 | 69.00 |
| **2172.00** | 39.86 | 78.00 | 65.00 |
| **2237.89** | 41.28 | 75.00 | 69.00 |
| **2287.05** | 42.95 | 72.00 | 61.00 |
| **2358.35** | 43.78 | 67.00 | 62.00 |
| **2406.30** | 44.70 | 65.00 | 66.00 |
| **2420.03** | 46.99 | 82.00 | 68.00 |
| **2487.28** | 48.24 | 76.00 | 65.00 |
| **2489.33** | 46.65 | 63.00 | 69.00 |
| **1176.17** | 22.03 | 79.80 | 66.15 |
| **1180.48** | 23.35 | 77.70 | 60.90 |
| **1190.90** | 22.65 | 76.65 | 61.95 |
| **1208.51** | 21.52 | 78.75 | 65.10 |
| **1215.85** | 22.41 | 82.95 | 69.30 |
| **1225.76** | 22.68 | 75.60 | 66.15 |
| **1229.23** | 24.05 | 73.50 | 65.10 |
| **1247.22** | 21.66 | 76.65 | 66.15 |
| **2514.22** | 47.12 | 63.63 | 69.69 |

Nota: La tabla muestra los registros de vida Util, velocidad de penetración, Pulldown y RPM de las brocas de 11”

Tabla 6 Parámetros estadísticos de la vida útil de las brocas de 11"

| **ítem** | **vida útil (m)** |
|---|---|
| **Promedio** | 1536.4 |
| **Mínimo** | 940.46 |
| **Máximo** | 2514.22 |
| **Mediana** | 1415.88 |

Nota: Elaboración propia en base a los datos recolectados de campo, del cual se trabajará únicamente con el promedio de metraje.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

#### 4.1.1. Carga de los datos

Se utilizó la biblioteca panda para la carga del archivo Excel denominado DATA DE BROCAS.xlsx, que contenía la información estructurada por columnas. El primer paso fue estandarizar los nombres de columnas para asegurar consistencia en el análisis, convirtiéndolos a mayúsculas y eliminando espacios innecesarios:

```
import pandas as pd
# Cargar el archivo Excel
archivo = 'DATA DE BROCAS.xlsx'
df = pd.read_excel(archivo)
# Estandarización de nombres de columnas
df.columns = [col.strip().upper() for col in df.columns]
```

#### 4.1.2. Selección de variables relevantes

Posteriormente, se filtraron las variables directamente relacionadas con el objetivo del estudio. Se seleccionaron aquellas que influyen en la vida útil de las brocas, y se descartaron registros incompletos para garantizar la calidad del análisis:

```
# Seleccionar las columnas relevantes y eliminar valores nulos
variables = ['VIDA UTIL', 'VELOCIDAD DE PENETRACIÓN', 'PULLDOWN',
'RPM']
df = df[variables].dropna()
```

#### 4.1.3. Verificación y calidad de los datos

Se realizó un análisis preliminar de los datos para detectar valores atípicos, inconsistencias, errores de digitación o registros fuera del rango operativo normal. Aunque esta parte puede variar según el contexto de la operación minera, algunos criterios aplicados fueron:



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

- Verificación de que la vida útil no sea negativa ni superior a valores técnicamente inviables (por ejemplo, >1000 metros sin recambio).

- Comprobación de que la velocidad de penetración, el pulldown y las RPM estén dentro de los rangos operativos definidos por los fabricantes o la operación minera.

Este paso es crucial para evitar sesgos en el modelo de machine learning y garantizar que las predicciones futuras se basen en datos válidos y representativos de las condiciones reales.

## 4.2. Análisis de correlación

Una vez depurada la base de datos, se procedió al análisis de correlación entre la vida útil de las brocas tricónicas de 11” y los parámetros operativos seleccionados: velocidad de penetración (ROP), pulldown y RPM. Este análisis tiene como objetivo identificar qué variables presentan mayor influencia sobre la vida útil, y por tanto, cuáles deben ser priorizadas en la construcción del modelo predictivo.

### 4.2.1. Cálculo del coeficiente de correlación

Para este análisis se utilizó la función corr() de la biblioteca pandas, la cual calcula el coeficiente de correlación de Pearson. Este coeficiente cuantifica la relación lineal entre dos variables, oscilando entre -1 (correlación negativa perfecta) y +1 (correlación positiva perfecta).

```
# Calcular correlaciones con VIDA UTIL
correlaciones = df.corr()['VIDA UTIL'].drop('VIDA
UTIL').sort_values(ascending=False)

# Mostrar resultados en consola
print("Correlaciones con VIDA UTIL:")
print(correlaciones)
```



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

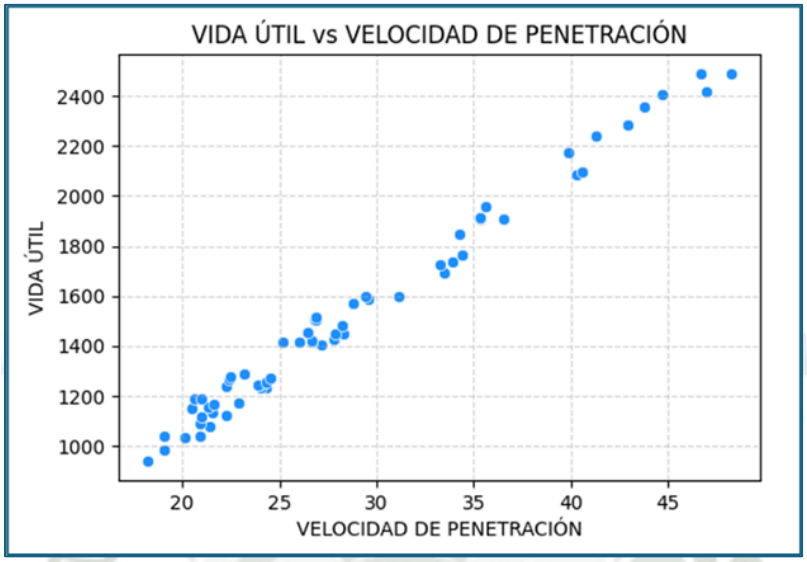
REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

### 4.2.2. Resultados del análisis

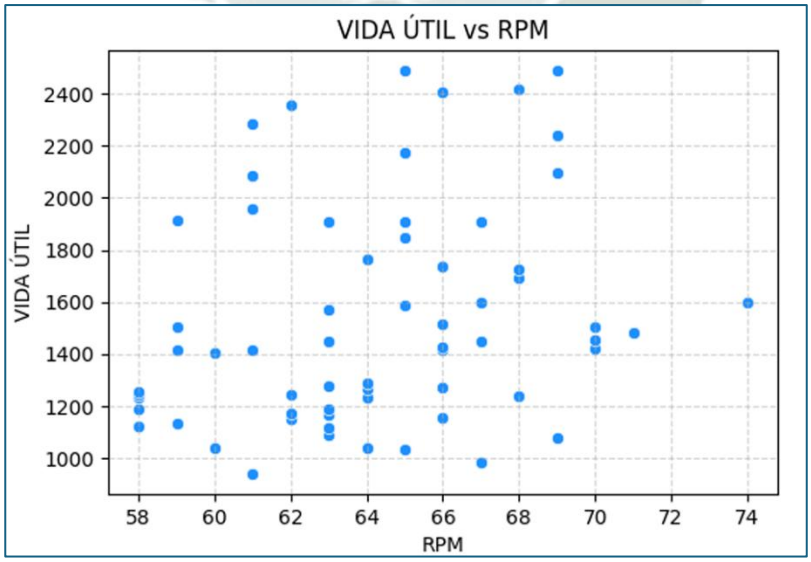
El resultado mostro los siguientes coeficientes de correlación respecto a la vida útil de la broca triconica de 11”:

Figura 2 Correlación entre vida útil vs Velocidad de penetración




Nota: La agrupación de los datos se concentra en una tendencia lineal positiva, con puntos distribuidos cercanos a la recta de ajuste, lo que indica una relación fuerte y consistente.

Figura 3 Correlación entre vida útil vs RPM



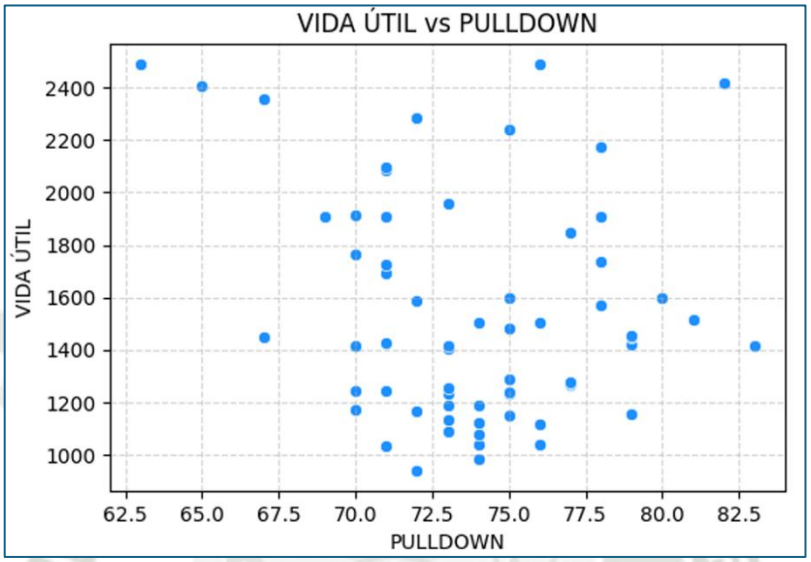



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM
UNIVERSIDAD CATÓLICA DE SANTA MARÍA

Nota: Los datos muestran una agrupación más dispersa, con concentración en valores intermedios de RPM, lo que refleja una correlación débil y poco determinante.

Figura 4 Correlación entre vida util vs Pulldown




Nota: Se observa una agrupación dispersa con tendencia negativa ligera, donde un mayor pulldown tiende a concentrar datos de menor vida útil, evidenciando un desgaste acelerado de la broca.

Tabla 7 Resumen de Correlación

| CORRELACIÓN | |
|---|---|
| **Velocidad de Penetración** | 0.992833 |
| **RPM** | 0.255610 |
| **Pulldown** | -0.189334 |

Nota: La tabla muestra la correlación entre las variables independientes y la variable dependiente (Vida útil de las brocas).



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

#### 4.2.3. Visualización de resultados

Para facilitar la interpretación, se generó un gráfico de barras utilizando seaborn, donde se representa la magnitud y dirección de la correlación de cada variable con respecto a la vida útil:

```
# Graficar las correlaciones
import seaborn as sns
import matplotlib.pyplot as plt
plt.figure(figsize=(8, 4))
sns.barplot(x=correlaciones.values, y=correlaciones.index, palette='viridis')
plt.title("Correlación con VIDA UTIL")
plt.xlabel("Coeficiente de correlación")
plt.tight_layout()
plt.savefig("correlaciones.png")
plt.close()
```

#### 4.2.4. Generación automática de reporte

Como parte de la documentación del proceso, se generó un reporte en formato PDF que contiene tanto los resultados numéricos del análisis como el gráfico correspondiente. Esto se realizó mediante la biblioteca fpdf:

```
from fpdf import FPDF

pdf = FPDF()
pdf.add_page()
pdf.set_font("Arial", size=12)
pdf.cell(200, 10, "Reporte de Correlación: Vida Útil de Brocas Tricónicas", ln=True,
align='C')
pdf.ln(10)

pdf.set_font("Arial", size=11)
for var, corr in correlaciones.items():
    pdf.cell(0, 10, f"- {var}: {corr:.3f}", ln=True)

pdf.ln(10)
pdf.cell(0, 10, "Gráfico de correlaciones:", ln=True)
pdf.image("correlaciones.png", x=30, w=150)
```



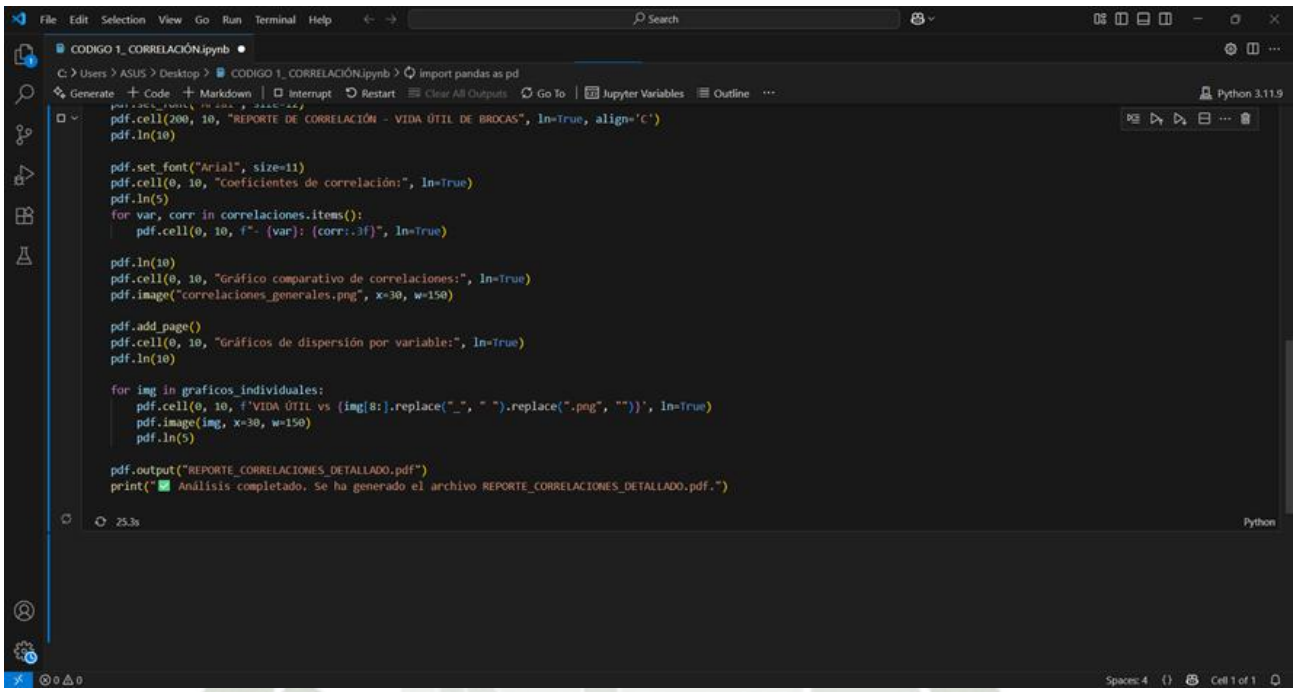
Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

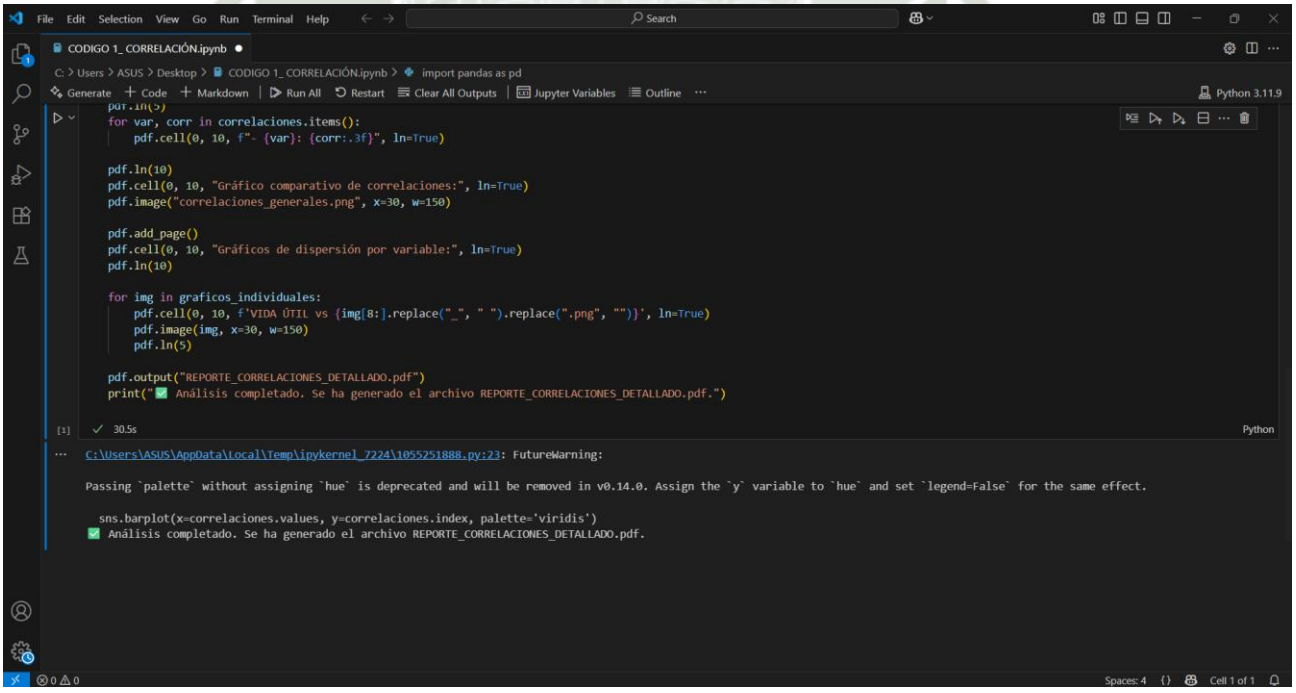
REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

pdf.output("REPORTE_CORRELACIONES_DETALLADO.pdf")

Figura 5 Ejecución del modelo de correlación en Python






Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

Nota: Evidencia del desarrollo computacional realizado para estimar la variable que presenta mayor correlación con la vida útil y generar el reporte.

### 4.2.5. Interpretación

Los resultados del análisis de correlación entre la variable dependiente VIDA ÚTIL y las variables operativas VELOCIDAD DE PENETRACIÓN, RPM y PULLDOWN, permitieron identificar el grado de asociación lineal entre estas variables. A continuación, se interpretan los coeficientes de correlación obtenidos:

- Velocidad de Penetración: presentó una correlación positiva muy fuerte con la vida útil de la broca, con un valor de 0.993. Este resultado indica que, a mayor velocidad de penetración, mayor tiende a ser la vida útil de la broca. Este hallazgo puede atribuirse a que una buena eficiencia de penetración refleja condiciones óptimas de operación (como roca menos abrasiva o buena configuración de los parámetros), lo cual favorece el rendimiento de la herramienta.
- RPM (Revoluciones por minuto): mostró una correlación positiva baja con la vida útil, con un valor de 0.256. Esto sugiere una relación directa pero débil. Es posible que en determinadas condiciones el aumento de RPM contribuya al desempeño, pero no es un factor determinante por sí solo en la durabilidad de la broca.
- Pulldown: obtuvo una correlación negativa débil de -0.189. Esto sugiere que un incremento excesivo en la fuerza de empuje podría tener un efecto adverso sobre la vida útil, posiblemente por generar un desgaste acelerado debido al exceso de presión ejercida sobre la broca.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

Estos resultados permiten concluir que la variable más influyente en la vida útil de las brocas tricónicas es la Velocidad de Penetración, la cual será considerada como uno de los principales insumos en el diseño del modelo predictivo. Las demás variables también serán incluidas en el entrenamiento de los modelos de regresión, pero con la comprensión de su menor nivel de influencia directa.

## 4.3. Desarrollo del modelo predictivo

Con el objetivo de predecir la vida útil de las brocas tricónicas de 11”, se procedió al desarrollo de modelos de regresión utilizando la variable que mostró mayor correlación durante el análisis, la cual fue la velocidad de penetración.

A continuación, se describirá detalladamente la construcción de los modelos, el procedimiento seguido en Python y la evaluación del desempeño de cada enfoque predictivo.

### 4.3.1. Preparación de los datos para el modelo

Se utilizo la única variable que tuvo mayor correlación con la vida útil, la cual fue la velocidad de penetración, teniendo un coeficiente de correlación de 0.993 que representa una correlación bien fuerte entre ambas variables.

Para ello utilizamos dos variables cables:

- Variable independiente (X): Velocidad de penetración
- Variable dependiente (y): Vida útil (en metros perforados)

Se utilizó la biblioteca train_test_split de scikit-learn para dividir los datos en dos subconjuntos:

- 70% para entrenamiento del modelo



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

- 30% para prueba y validación

Esta división permite evaluar el rendimiento del modelo sobre datos no vistos, evitando el sobreajuste y asegurando la generalización.

```
from sklearn.model_selection import train_test_split

# Variables predictora y objetivo
X = df[['VELOCIDAD DE PENETRACIÓN']]
y = df['VIDA UTIL']

# División del dataset en entrenamiento y prueba (70% - 30%)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3,
random_state=42)
```

4.3.2. Regresión simple

Se aplicó un modelo de Regresión Lineal Simple para modelar la relación directa entre la velocidad de penetración y la vida útil, la cual obedece al siguiente formula de la regresión lineal:

$$Y = a * X + b$$

Donde:

- Y es la vida útil estimada
- X la velocidad de penetración
- a la pendiente de la recta
- b el intercepto

El modelo fue entrenado usando el conjunto de datos de entrenamiento, y evaluado con el conjunto de prueba. Se utilizaron métricas estadísticas para cuantificar el rendimiento:



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

$R^2$ (Coeficiente de determinación): mide el porcentaje de la variabilidad de la vida útil que puede ser explicada por la velocidad de penetración.

MAE (Mean Absolute Error): error absoluto promedio entre los valores reales y los predichos.

RMSE (Root Mean Squared Error): error cuadrático medio, sensible a errores grandes.

Se generó también una gráfica que compara la dispersión de los datos reales con la línea de regresión obtenida, permitiendo visualizar la calidad del ajuste del modelo.

**CODIGO**:

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score, mean_absolute_error, mean_squared_error

# Modelo
modelo_lr = LinearRegression()
modelo_lr.fit(X_train, y_train)

# Predicciones
y_pred_lr = modelo_lr.predict(X_test)

# Evaluación
r2_lr = r2_score(y_test, y_pred_lr)
mae_lr = mean_absolute_error(y_test, y_pred_lr)
rmse_lr = mean_squared_error(y_test, y_pred_lr, squared=False)

print(f"R² Regresión Lineal Simple: {r2_lr:.4f}")
print(f"MAE: {mae_lr:.2f}")
print(f"RMSE: {rmse_lr:.2f}")
```

**visualización del modelo**

```
import matplotlib.pyplot as plt

plt.scatter(X_test, y_test, color='blue', label='Datos reales')
plt.plot(X_test, y_pred_lr, color='red', linewidth=2, label='Predicción Lineal')
plt.title('Regresión Lineal Simple: Vida Útil vs Velocidad de Penetración')
```



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

```
plt.xlabel('Velocidad de Penetración')
plt.ylabel('Vida Útil (m)')
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.savefig('regresion_lineal_simple.png')
plt.show()
```

#### 4.3.3. Random Forest Regressor

Como segundo modelo, se utilizó un enfoque no lineal basado en árboles de decisión: el Random Forest Regressor. Este modelo está conformado por múltiples árboles que aprenden diferentes patrones del conjunto de entrenamiento. La predicción final se obtiene promediando los resultados de todos los árboles, logrando así un modelo más robusto frente a ruidos o variabilidad.

Este enfoque es especialmente útil cuando la relación entre las variables no es estrictamente lineal, y permite capturar interacciones más complejas.

Se aplicaron las mismas métricas de evaluación ($R^2$, MAE, RMSE), con el fin de comparar su desempeño frente al modelo lineal.

**CODIGO:**

```
from sklearn.ensemble import RandomForestRegressor
modelo_rf = RandomForestRegressor(n_estimators=100, random_state=42)
modelo_rf.fit(X_train, y_train)
y_pred_rf = modelo_rf.predict(X_test)
r2_rf = r2_score(y_test, y_pred_rf)
mae_rf = mean_absolute_error(y_test, y_pred_rf)
rmse_rf = mean_squared_error(y_test, y_pred_rf, squared=False)
print(f"R² Random Forest: {r2_rf:.4f}")
print(f"MAE: {mae_rf:.2f}")
print(f"RMSE: {rmse_rf:.2f}")
```



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

### 4.3.4. XGBoost Regressor

El tercer enfoque fue el modelo XGBoost (Extreme Gradient Boosting), una técnica avanzada de aprendizaje automático que combina múltiples árboles de forma secuencial. Cada árbol corrige los errores cometidos por el anterior, permitiendo alcanzar una mayor precisión.

Este modelo ha sido ampliamente validado en la literatura técnica y se utiliza comúnmente en competencias de modelado debido a su alta efectividad y rendimiento computacional.

XGBoost fue entrenado con los mismos conjuntos de datos y evaluado con las métricas mencionadas. Se buscó verificar si este modelo ofrecía una mejora significativa en la capacidad predictiva respecto a los anteriores.

```
from xgboost import XGBRegressor
modelo_xgb = XGBRegressor(n_estimators=100, learning_rate=0.1,
random_state=42)
modelo_xgb.fit(X_train, y_train)
y_pred_xgb = modelo_xgb.predict(X_test)


r2_xgb = r2_score(y_test, y_pred_xgb)
mae_xgb = mean_absolute_error(y_test, y_pred_xgb)
rmse_xgb = mean_squared_error(y_test, y_pred_xgb, squared=False)
print(f"R² XGBoost: {r2_xgb:.4f}")
print(f"MAE: {mae_xgb:.2f}")
print(f"RMSE: {rmse_xgb:.2f}")
```

### 4.3.5. Comparación de Desempeño de Modelos

Una vez obtenidas las métricas de desempeño para los tres modelos (Regresión Lineal, Random Forest y XGBoost), se construyó una tabla comparativa que resume los valores obtenidos de $R^2$, MAE y RMSE.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

Asimismo, se elaboró un gráfico de barras que permite visualizar el valor de $R^2$ para cada modelo, facilitando así la identificación del enfoque con mejor capacidad explicativa sobre la variable de interés.

Esta comparación resulta fundamental para la elección del modelo final que será utilizado en la estimación de la vida útil de brocas en base a la velocidad de penetración, así como para la posterior estimación de los ahorros operativos que ello implicaría.

```
import pandas as pd
resultados = pd.DataFrame({
   'Modelo': ['Regresión Lineal', 'Random Forest', 'XGBoost'],
   'R²': [r2_lr, r2_rf, r2_xgb],
   'MAE': [mae_lr, mae_rf, mae_xgb],
   'RMSE': [rmse_lr, rmse_rf, rmse_xgb]})
Visualización de la comparación de R^2
print(resultados)
plt.figure(figsize=(8,4))
plt.bar(resultados['Modelo'], resultados['R²'], color='lightgreen')
plt.title('Comparación del R² entre modelos')
plt.ylabel('R²')
plt.ylim(0, 1)
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.tight_layout()
plt.savefig('comparacion_r2_modelos_velocidad.png')
plt.show()
```

## 4.4. Resultados Obtenidos y selección del modelo de predicción

La visualización de resultados representa una etapa crítica en el análisis predictivo, ya que permite verificar gráficamente la calidad del ajuste de los modelos utilizados y su capacidad para



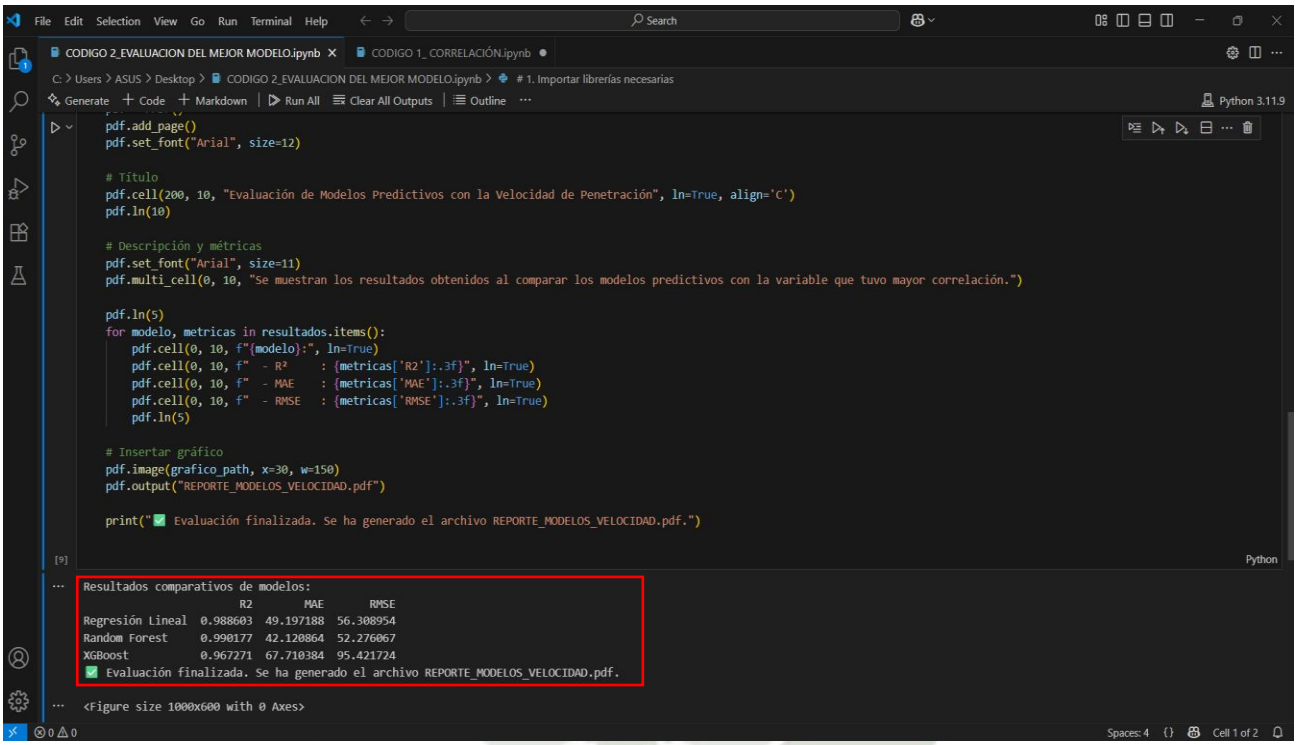
Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

generalizar patrones sobre la vida útil de las brocas triconicas de 11". En este caso, se utilizaron tres algoritmos de regresión: Regresión Lineal Múltiple, Random Forest Regressor y XGBoost Regressor, cuya comparación cuantitativa fue complementada con visualizaciones gráficas.

Figura 6 Ejecución del código para comparar modelos predictivos.



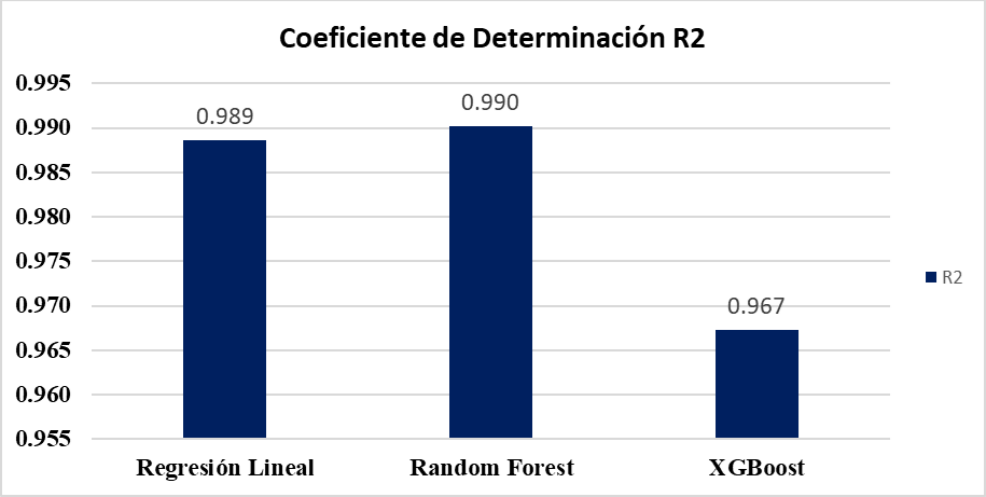

Nota: Evidencia del desarrollo computacional realizado para estimar y comparar el mejor modelo de predicción.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

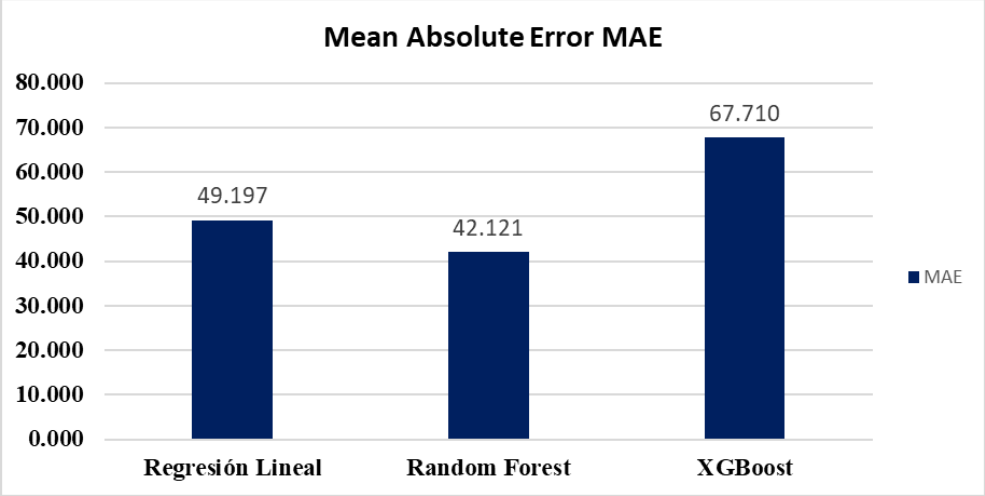
REPOSITORIO DE TESIS UCSM
UNIVERSIDAD CATÓLICA DE SANTA MARÍA

Figura 7 Gráfico de coeficiente de determinación de cada algoritmo de regresión




Nota: El grafico muestra los modelos de regresión y cuál de ellos presenta mayor coeficiente de determinación.

Figura 8 Gráfico de Mean Absolute Error




Nota: El grafico muestra los modelos de regresión y cuál de ellos presenta mayor error absoluto.

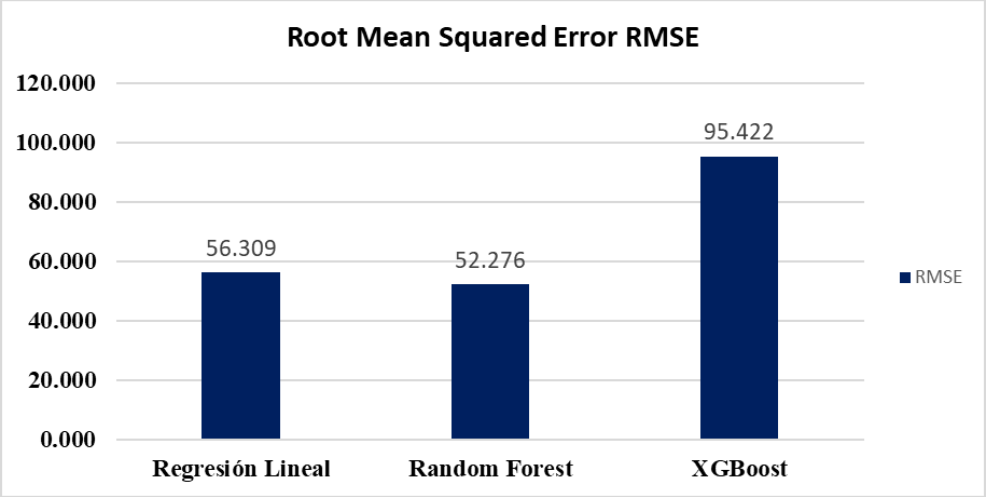


Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

Figura 9 Gráfico de Root Mean Squared Error de cada algoritmo de regresión




Nota: El grafico muestra los modelos de regresión y cuál de ellos presenta mayor Raíz del Error Cuadrático Medio.

Tabla 8 Resumen de modelos de predicción

| Modelo | R2 | MAE | RMSE |
|---|---|---|---|
| **Regresión Lineal** | 0.989 | 49.197 | 56.309 |
| **Random Forest** | 0.990 | 42.121 | 52.276 |
| **XGBoost** | 0.967 | 67.710 | 95.422 |

Nota: La tabla muestra los resultados obtenidos al comparar los 3 modelos bajo diferentes metodologías que nos permitan saber cuál de ellos es más adecuado para la predicción de la vida útil de las brocas de 11".

Los resultados muestran que el modelo de Regresión Lineal presenta una R2 fuerte cuando se utiliza exclusivamente la variable velocidad de penetración, lo cual es coherente con el alto coeficiente de correlación (0.993) obtenido previamente. Sin embargo, los modelos Random Forest y XGBoost demostraron una capacidad superior para capturar variaciones más complejas,



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

por ello es que so escogió el modelo Random Forest junto a la velocidad de penetración para desarrollar el modelo predictivo ya que presenta un R2 mas alto que la regresión lineal un MAE y RMSE bajos a comparación de los otros modelos.

## 4.5. Código de predicción

Para la realización del código de predicción se tomó el modelo de Random Forest Regressor.

Para maximizar el aprendizaje del modelo y aprovechar todos los datos disponibles, el modelo final fue entrenado utilizando la totalidad de la base de datos, es decir, sin particionar en conjunto de entrenamiento y prueba. Este entrenamiento completo permite al modelo incorporar la mayor cantidad de información histórica posible para mejorar su capacidad predictiva futura.

```
modelo_final = RandomForestRegressor(n_estimators=100, random_state=42)
modelo_final.fit(X, y)
```

### 4.5.1. Validación del modelo

Debido a que el modelo se entrenó con toda la base de datos, se decidió por validar su comportamiento con un conjunto de nuevas observaciones simuladas, compuestas por valores de velocidad de penetración dentro del rango operacional de los datos originales. Las predicciones obtenidas se compararon con los valores esperados, y se generó un gráfico de predicción para evaluar visualmente su coherencia.

```
Nuevas_velocidades = pd.DataFrame({'VELOCIDAD DE PENETRACIÓN': [18.22, 19.04,
20.12, 19.03, 20.89, 21.40, 20.88, 20.98, 22.24, 21.58, 20.50, 21.34, 21.60, 22.90, 20.63,    20.99,
24.34, 24.03, 22.23, 24.00, 23.87, 24.29, 22.41, 24.56, 22.49, 23.18, 27.14, 26.02, 25.19, 26.65,
26.65, 27.81, 28.28, 27.85, 26.44, 28.25, 26.79, 26.84, 26.90, 28.82, 29.56, 29.40, 31.11, 33.46,
33.28, 33.88, 34.39, 34.26, 36.51, 35.29, 35.30, 35.31, 35.63, 40.27, 40.58, 39.86, 41.28, 42.95,
43.78, 44.70, 46.99, 48.24, 46.65]})
```



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

predicciones = modelo_final. predict(nuevas_velocidades)

4.5.2. Almacenamiento del modelo para uso futuro

Para facilitar su uso operativo y su implementación en entornos externos como aplicaciones, interfaces o sistemas de monitoreo, el modelo entrenado fue guardado mediante la librería joblib, lo que permitira reutilizarlo sin tener que reentrenarlo cada vez.

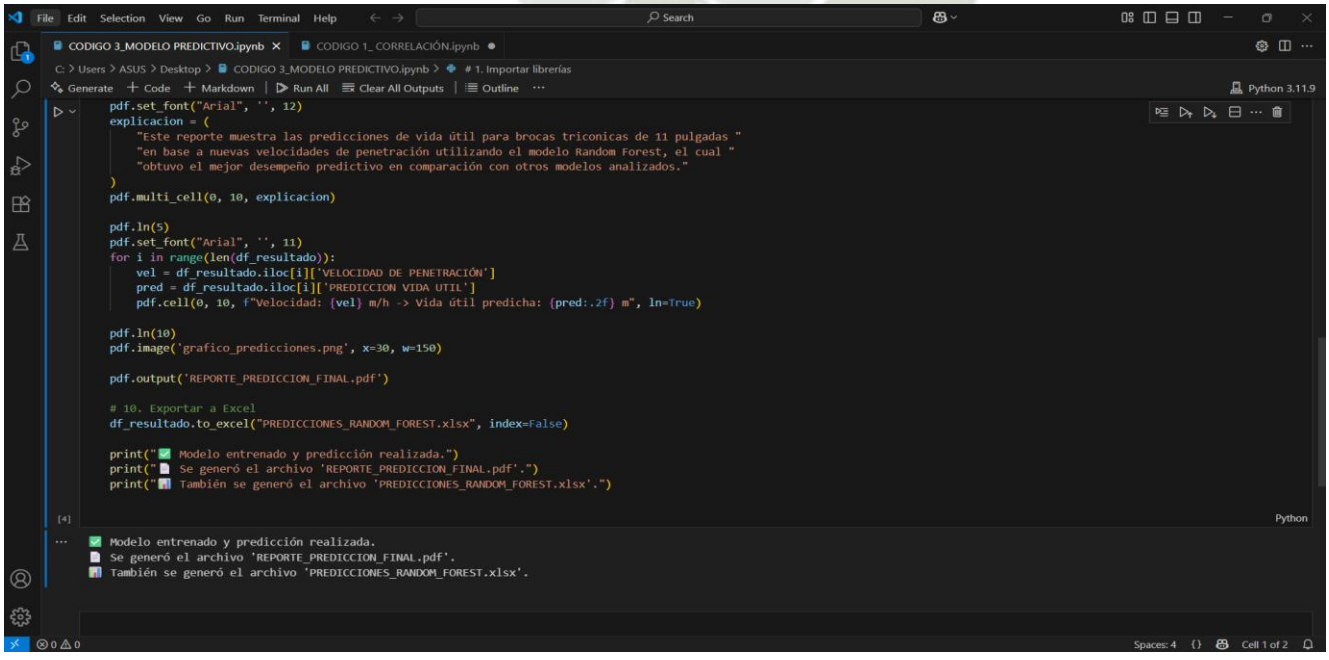
```
import joblib
joblib.dump(modelo_final, 'modelo_final_random_forest.pkl')
modelo_cargado = joblib.load('modelo_final_random_forest.pkl')
```

4.6. Estimación del ahorro operativo

Se realizo la prueba con 64 brocas, de las cuales se realizó el cálculo a través del modelo de predicción en base a sus velocidades de penetración registrada, se registró su vida útil en metros alcanzada al momento del cambio y se comparó con la calculada por el modelo.

De ella se obtuvieron los siguientes datos:

Figura 10 Ejecución del modelo predictivo en Python






Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

Nota: Evidencia del desarrollo computacional realizado para estimar la vida útil de la broca con el modelo predictivo.

Tabla 9 Comparación de metros reales vs metros con el modelo predictivo

| Velocidad de penetración | Metros Reales | Metros Predecidos |
|---|---|---|
| **18.2** | 940.5 | 1041.1 |
| **19.0** | 984.1 | 1074.0 |
| **20.1** | 1037.2 | 1136.9 |
| **19.0** | 1041.4 | 1091.3 |
| **20.9** | 1041.8 | 1129.6 |
| **21.4** | 1077.8 | 1176.2 |
| **20.9** | 1088.7 | 1153.2 |
| **21.0** | 1120.2 | 1187.6 |
| **22.2** | 1124.3 | 1239.5 |
| **21.6** | 1134.2 | 1199.9 |
| **20.5** | 1151.0 | 1210.7 |
| **21.3** | 1157.9 | 1204.9 |
| **21.6** | 1167.4 | 1211.2 |
| **22.9** | 1170.7 | 1277.7 |
| **20.6** | 1187.8 | 1221.7 |
| **21.0** | 1189.0 | 1215.2 |
| **24.3** | 1231.3 | 1302.3 |
| **24.0** | 1234.7 | 1303.8 |
| **22.2** | 1240.7 | 1273.1 |
| **24.0** | 1242.7 | 1310.8 |
| **23.9** | 1246.2 | 1303.3 |
| **24.3** | 1257.9 | 1309.0 |
| **22.4** | 1268.1 | 1330.2 |
| **24.6** | 1271.3 | 1335.7 |
| **22.5** | 1279.5 | 1323.9 |
| **23.2** | 1291.7 | 1316.1 |
| **27.1** | 1405.6 | 1492.7 |
| **26.0** | 1414.3 | 1495.6 |
| **25.2** | 1417.4 | 1446.1 |
| **26.7** | 1417.9 | 1494.7 |
| **26.6** | 1421.3 | 1504.7 |
| **27.8** | 1429.4 | 1503.9 |
| **28.3** | 1448.2 | 1518.5 |
| **27.9** | 1449.7 | 1517.8 |
| **26.4** | 1457.6 | 1516.9 |
| **28.2** | 1482.1 | 1529.8 |
| **26.8** | 1503.7 | 1559.2 |
| **26.8** | 1504.8 | 1565.7 |
| **26.9** | 1515.7 | 1580.4 |



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

| | | |
|---|---|---|
| **28.8** | 1573.3 | 1604.5 |
| **29.6** | 1585.3 | 1656.2 |
| **29.4** | 1599.2 | 1658.1 |
| **31.1** | 1599.6 | 1655.7 |
| **33.5** | 1690.3 | 1777.1 |
| **33.3** | 1724.5 | 1787.8 |
| **33.9** | 1736.8 | 1828.3 |
| **34.4** | 1766.1 | 1857.1 |
| **34.3** | 1845.1 | 1873.7 |
| **36.5** | 1907.6 | 1979.9 |
| **35.3** | 1908.7 | 1981.0 |
| **35.3** | 1909.3 | 1966.3 |
| **35.3** | 1911.6 | 1984.7 |
| **35.6** | 1955.6 | 2016.3 |
| **40.3** | 2084.0 | 2165.6 |
| **40.6** | 2094.0 | 2174.6 |
| **39.9** | 2172.0 | 2201.1 |
| **41.3** | 2237.9 | 2248.4 |
| **42.9** | 2287.1 | 2378.9 |
| **43.8** | 2358.4 | 2400.7 |
| **44.7** | 2406.3 | 2449.5 |
| **47.0** | 2420.0 | 2505.5 |
| **48.2** | 2487.3 | 2533.8 |
| **46.7** | 2489.3 | 2417.9 |

Nota: La tabla muestra los resultados obtenidos al usar el modelo de predicción, dándonos a conocer el metraje real con el que debieron morir las brocas de 11".

Tabla 10 Promedios de vida útil de las brocas registradas en una mina superficial

| **Resultados** | | |
|---|---|---|
| **Promedio de vida útil real** | 1536.4 | Metros |
| **Promedio de vida útil estimada** | 1598.5 | Metros |

Nota: Se muestran los promedios de las 64 brocas, el calculado por el modelo de predicción y el real obtenido.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

Posteriormente a ello se calculó el porcentaje de aumento de vida útil y el impacto económico que su implementación tendría sobre la operación minera, específicamente en términos de reducción de costos operativos asociados al reemplazo de brocas.

Tabla 11 Ahorro anual estimado

| Parámetros | | |
|---|---|---|
| **Promedio de vida útil real** | 1536.4 | Metros |
| **Promedio de vida útil estimada** | 1598.5 | Metros |
| **Aumento Vida** | 4% | Porcentaje |
| **Precio de broca** | $3,400.00 | Dólares |
| **Ahorro** | $137.43 | $/Broca |
| **Metros Anuales** | 1140000 | Metros |
| **Brocas Anuales** | 742 | Unidad |

Nota: La tabla muestra el consumo de brocas anuales y los promedios de vida útil de las brocas utilizadas durante el año y el estimado con el modelo.

Realizando el cálculo a partir del promedio de vida útil dado por el modelo tenemos que:

Tabla 12 Escenario 1 sin el modelo predictivo

| ESCENARIO 1 SIN MODELO | | |
|---|---|---|
| **Promedio de vida útil real** | 1536.4 | Metros |
| **Precio de broca** | $3,400.00 | Dólares |
| **Metros Anuales** | 1140000 | Metros |
| **Brocas Anuales** | 742 | Unidad |
| **Costo total Anual** | $2,522,780.53 | Dólares |



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

Nota: La tabla muestra que el precio total del consumo de brocas anualmente es 2,522,780.53 dólares.

Tabla 13 Escenario 2 con la aplicación del modelo

| ESCENARIO 2 CON EL MODELO | | |
|---|---|---|
| **Promedio de vida útil real** | 1598.5 | Metros |
| **Precio de broca** | $3,400.00 | Dólares |
| **Metros Anuales** | 1140000 | Metros |
| **Brocas Anuales** | 713 | Unidad |
| **Costo total Anual** | $2,424,773.22 | Dólares |

Nota: La tabla muestra que el precio total del consumo de brocas anualmente es de 2,424,773.22 dólares con la aplicación del modelo de predicción.

Ahorro operativo

$$Ahorro\ Operativo = \$2{,}522{,}780.53 - \$2{,}424{,}773.22$$

$$Ahorro\ Operativo = \$\ 98{,}007.30\ dolares$$

Se logro una reducción del 3.9% en el consumo anual de brocas, gracias a una mejor planificación de cambio de brocas debido al modelo predictivo.

## 5. CONCLUSIONES

- El modelo predictivo implementado logró estimar con alta precisión la vida útil de las brocas tricónicas, alcanzando un coeficiente de determinación máximo de $R^2$ = 0.990 con el algoritmo Random Forest. La aplicación del modelo permitió proyectar un incremento promedio de 62.1 metros en la vida útil de cada broca, lo que se tradujo en una reducción anual de 29 unidades adquiridas y un ahorro



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

operativo estimado de USD 98,007.30, cumpliendo plenamente con el objetivo general de optimizar el uso de insumos y disminuir los costos operativos.

- Se consolidó y depuró una base de datos confiable con los registros históricos de perforación, eliminando inconsistencias y valores atípicos. Esto garantizó la calidad de la información utilizada en el modelado y permitió un análisis estadístico robusto de los parámetros operativos, sentando las bases para el desarrollo de modelos predictivos precisos.

- El análisis de correlación evidenció que la velocidad de penetración presentó la mayor relación con la vida útil (r = 0.9928), mientras que las variables RPM y pulldown mostraron correlaciones débiles. Esto fue clave para priorizar las variables más influyentes en la predicción y para ajustar la estrategia de operación de las perforadoras.

- Se evaluaron tres modelos: Regresión Lineal, Random Forest y XGBoost. El modelo Random Forest obtuvo el mejor desempeño ($R^2 = 0.990$, MAE = 42.12 m, RMSE = 52.27 m), superando ligeramente a la regresión lineal y de forma considerable a XGBoost. Esto confirmó la capacidad de Random Forest para manejar no linealidades y optimizar la predicción.

- El uso de métricas como $R^2$, MAE y RMSE permitió cuantificar objetivamente la precisión de cada modelo. Los resultados demostraron un ajuste sobresaliente entre las predicciones y los valores reales, lo que valida la aplicabilidad del modelo en condiciones operativas reales.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

- La simulación del escenario con el modelo predictivo mostró una disminución de 4% en el consumo anual de brocas, representando un ahorro de USD 98,007.30. Este resultado muestra que la implementación de herramientas de machine learning puede generar impactos económicos directos y significativos en la operación minera.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

## 6. REFERENCIAS


- Alajmi, M. S., & Almeshal, A. M. (2020). Predicting the tool wear of a drilling process using novel machine learning XGBoost-SDA. Materials, 13(21), 4952. https://doi.org/10.3390/ma13214952

- Sotubadi, S. V., Ariyajunya, S., & Zhang, Z. (2021). Explainable AI for tool wear prediction in turning. arXiv preprint. https://arxiv.org/abs/2308.08765

- Senjoba, L., Wamalwa, G., & Wafula, C. (2024). Enhancing interpretability in drill bit wear analysis through explainable AI. Applied Sciences, 14(9), 3621. https://doi.org/10.3390/app14093621

- Yamani, A., Al-Otaiby, N., Al-Shemmeri, H., & Boudellioua, I. (2023). You Only Look Twice! For failure causes identification of drill bits. arXiv preprint. https://doi.org/10.48550/arXiv.2410.14282

- Zhang, X., Han, C., Luo, M., & Zhang, D. (2022). New insights into drill cutting characteristics in coal seams. Geofluids, 2022, Article 3201581. https://doi.org/10.1155/2022/3201581

- Khalifa, H., Osman, M. M., & Gawad, M. A. (2024). Machine learning-based drill bit wear prediction using field data. Proceedings of the American Rock Mechanics Association Conference.




Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA


- Wen, H., Li, S., & Xu, Q. (2023). Hybrid XGBoost model for predicting rate of penetration in rock drilling. Arabian Journal for Science and Engineering, 48, 559–571. https://doi.org/10.1007/s13369-022-07379-w

- Anon. (2021). Drill bit wear monitoring and failure prediction for mining automation. International Journal of Mining Science and Technology, 31(5), 731–742. https://doi.org/10.1016/j.ijmst.2021.06.003

- Yumani, A., Ahmed, H., & Benaicha, M. (2023). A hybrid CNN-XGBoost approach for classifying drill bit wear states using vibration data. Journal of Petroleum Science and Engineering, 230, 110045. https://doi.org/10.1016/j.petrol.2023.110045

- Zhang, Y., & Wang, L. (2022). Predictive maintenance for rotary drilling tools using machine learning and signal analysis. Journal of Manufacturing Processes, 77, 600–612. https://doi.org/10.1016/j.jmapro.2022.03.015

- Camargo, J., & Gómez, R. (2019). Brocas de perforación en minería superficial: diseño, selección y aplicación. Revista Técnica de Minería, 25(1), 33–42.

- Van Rossum, G., & Drake, F. L. (2009). Python 3 Reference Manual. Python Software Foundation. https://docs.python.org/3/reference/

- Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324




Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA


- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD, 785–794. https://doi.org/10.1145/2939672.2939785

- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning. Springer.

- Martínez, L., Torres, R., & Yucra, L. (2021). Aplicación de algoritmos de machine learning en minería. Revista Tecnológica del Sur, 29(2), 45–56.

- Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). Foundations of Machine Learning (2nd ed.). MIT Press.

- Scholars Mine. (2016). Drill Bit Selection: Engineering Principles. Missouri University of Science and Technology. https://scholarsmine.mst.edu/

- Smith, R. (2015). Drilling Optimization in Open Pit Mining. International Journal of Mining Engineering, 22(3), 105–114.

- Zhang, H., Li, X., & Wang, Y. (2020). Predicting Bit Wear with Machine Learning Techniques. Journal of Petroleum Science, 35(1), 22–34.

- Yumani, A., Cho, Y., & Lee, S. (2023). Visual Fault Detection in Drill Bits Using YOLO and Random Forest. Mining Technology Journal, 41(1), 60–75.

- Khalifa, M., Aboelela, M., & El-Deeb, H. (2024). Predictive Maintenance of Drill Bits Using AI. Journal of Mining Science and Technology, 39(2), 110–124.

- Smith, R. (2015). Drilling Optimization in Open Pit Mining. International Journal of Mining Engineering, 22(3), 105–114.




Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA


- Zhang, H., Li, X., & Wang, Y. (2020). Predicting bit wear with machine learning techniques. Journal of Petroleum Science, 35(1), 22–34. https://doi.org/10.1016/j.petrol.2020.105

- Yumani, A., Cho, Y., & Lee, S. (2023). Visual fault detection in drill bits using YOLO and random forest. Mining Technology Journal, 41(1), 60–75.

- Khalifa, M., Aboelela, M., & El-Deeb, H. (2024). Predictive maintenance of drill bits using AI. Journal of Mining Science and Technology, 39(2), 110–124. https://doi.org/10.1007/s10913-024-0123-9

- Scholars Mine. (2016). Drill bit selection: Engineering principles. Missouri University of Science and Technology. https://scholarsmine.mst.edu

- Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). Foundations of Machine Learning (2nd ed.). MIT Press.

- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning: With applications in R. Springer. https://doi.org/10.1007/978-1-4614-7138-7

- Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785




Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

- Martínez, L., Torres, R., & Yucra, L. (2021). Aplicación de algoritmos de machine learning en minería. Revista Tecnológica del Sur, 29(2), 45–56.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

# ANEXOS



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

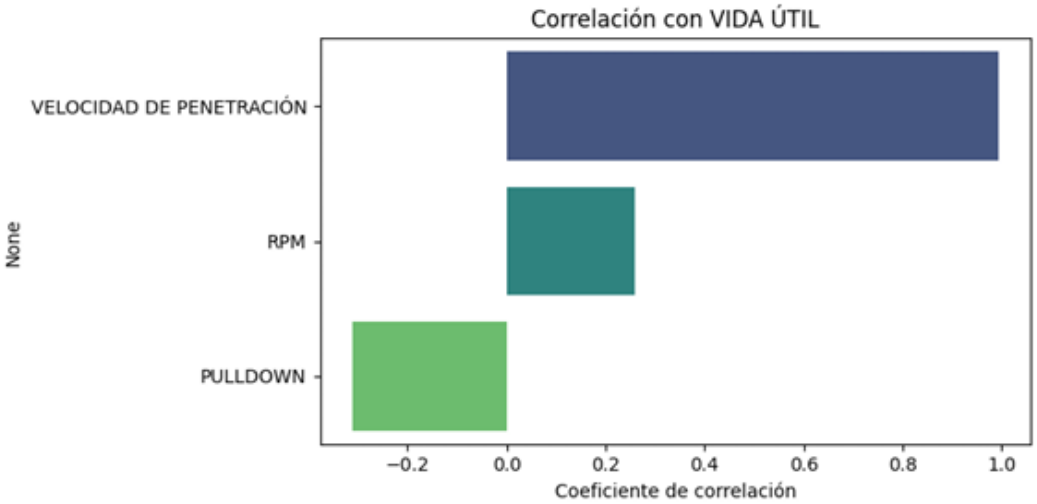
ANEXO 1

## REPORTE DE CORRELACIÓN - VIDA ÚTIL DE BROCAS

Coeficientes de correlación:

- VELOCIDAD DE PENETRACIÓN: 0.993
- RPM: 0.259
- PULLDOWN: -0.312

Gráfico comparativo de correlaciones:





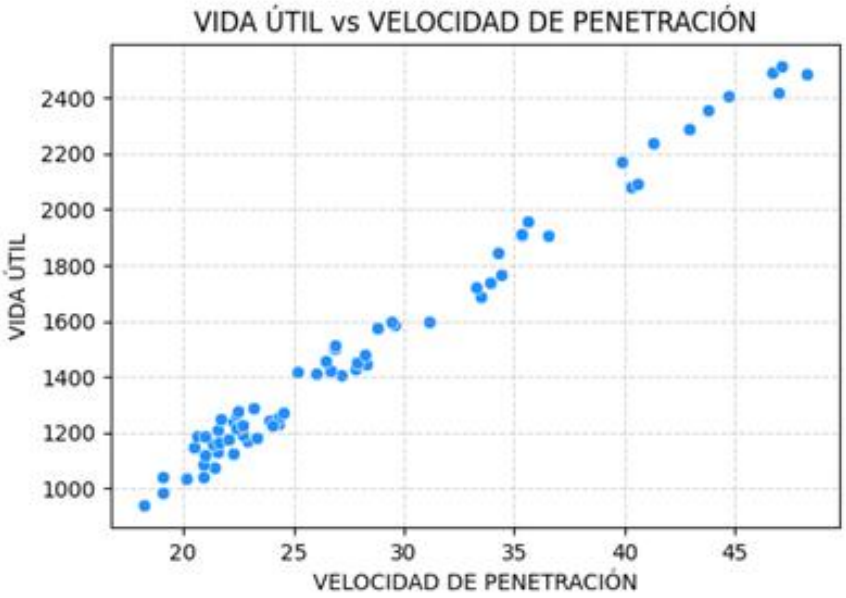

Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

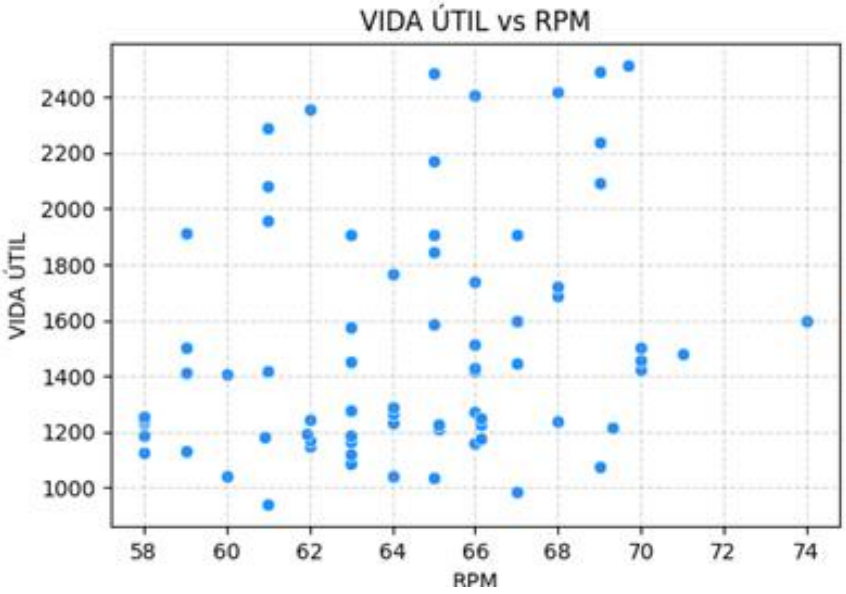
UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

Gráficos de dispersión por variable:

VIDA ÚTIL vs VELOCIDAD DE PENETRACIÓN




VIDA ÚTIL vs RPM



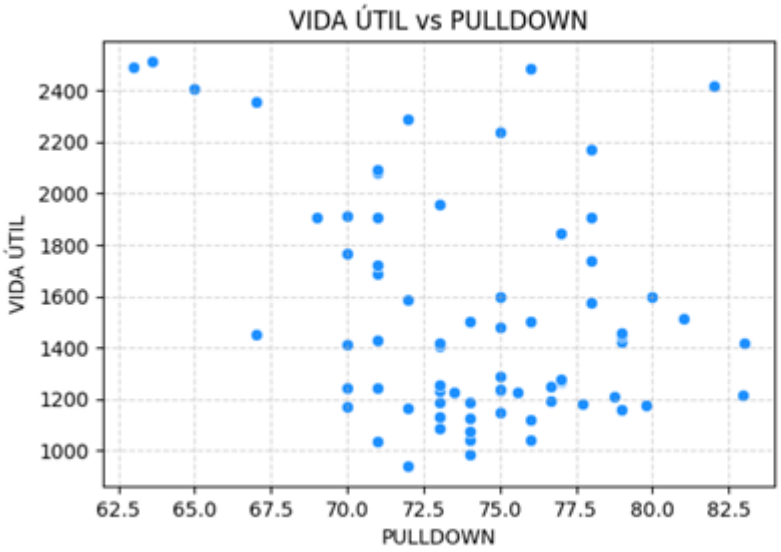



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

VIDA ÚTIL vs PULLDOWN






Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

ANEXO 2

Evaluación de Modelos Predictivos con la Velocidad de Penetración

Se muestran los resultados obtenidos al comparar los modelos predictivos con la variable que tuvo mayor correlación.

Regresión Lineal:

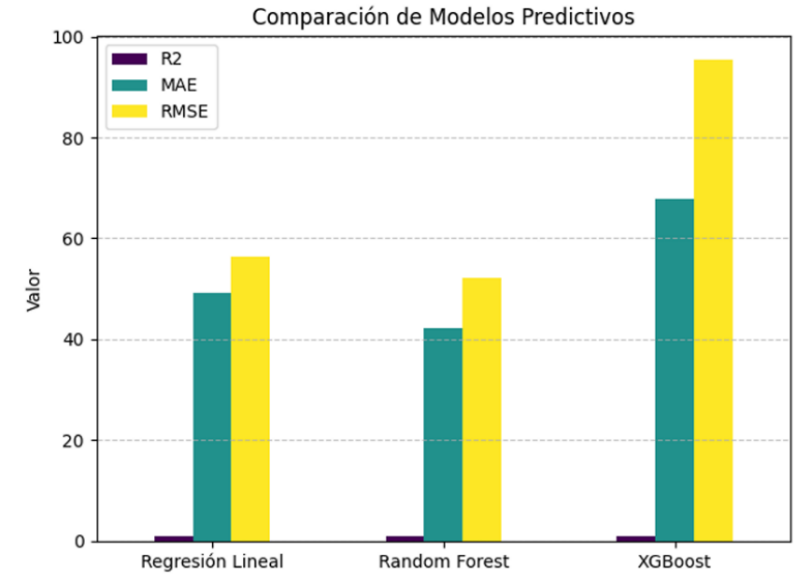
- $R^2$ : 0.989
- MAE : 49.197
- RMSE : 56.309

Random Forest:

- $R^2$ : 0.990
- MAE : 42.121
- RMSE : 52.276

XGBoost:

- $R^2$ : 0.967
- MAE : 67.710
- RMSE : 95.422




Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE TESIS UCSM

UNIVERSIDAD CATÓLICA DE SANTA MARÍA

ANEXO 3

## REPORTE FINAL DE PREDICCIÓN - MODELO RANDOM FOREST

En este anexo se presenta el reporte en formato PDF obtenido tras la ejecución del código de predicción basado en el modelo Random Forest Regressor, desarrollado en Python como parte de la investigación. Dicho reporte contiene las velocidades de penetración ingresadas como variables de prueba y los valores de vida útil estimada (en metros) que resultan de aplicar el modelo entrenado con la base de datos histórica de la operación.

La inclusión de este anexo cumple una doble función:

1. Validación técnica del modelo: demuestra de manera práctica cómo la herramienta desarrollada genera predicciones coherentes con la data real de la operación.
2. Aplicación operativa: evidencia la utilidad del modelo como apoyo a la toma de decisiones en perforación, ya que permite anticipar la duración de las brocas y, en consecuencia, optimizar la programación de reemplazos y la gestión de costos.

De esta forma, el reporte no solo es un producto informático, sino también una prueba del alcance aplicado de la investigación, confirmando que el modelo predictivo puede integrarse como una herramienta de soporte en operaciones de minería superficial.



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

| Velocidad (m/hr) | Vida útil Predicha |
|---|---|
| **18.2** | 1041.1 |
| **19.0** | 1074.0 |
| **20.1** | 1136.9 |
| **19.0** | 1091.3 |
| **20.9** | 1129.6 |
| **21.4** | 1176.2 |
| **20.9** | 1153.2 |
| **21.0** | 1187.6 |
| **22.2** | 1239.5 |
| **21.6** | 1199.9 |
| **20.5** | 1210.7 |
| **21.3** | 1204.9 |
| **21.6** | 1211.2 |
| **22.9** | 1277.7 |
| **20.6** | 1221.7 |
| **21.0** | 1215.2 |
| **24.3** | 1302.3 |
| **24.0** | 1303.8 |
| **22.2** | 1273.1 |
| **24.0** | 1310.8 |
| **23.9** | 1303.3 |
| **24.3** | 1309.0 |
| **22.4** | 1330.2 |
| **24.6** | 1335.7 |
| **22.5** | 1323.9 |
| **23.2** | 1316.1 |
| **27.1** | 1492.7 |



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

| | |
|---|---|
| **26.0** | 1495.6 |
| **25.2** | 1446.1 |
| **26.7** | 1494.7 |
| **26.6** | 1504.7 |
| **27.8** | 1503.9 |
| **28.3** | 1518.5 |
| **27.9** | 1517.8 |
| **26.4** | 1516.9 |
| **28.2** | 1529.8 |
| **26.8** | 1559.2 |
| **26.8** | 1565.7 |
| **26.9** | 1580.4 |
| **28.8** | 1604.5 |
| **29.6** | 1656.2 |
| **29.4** | 1658.1 |
| **31.1** | 1655.7 |
| **33.5** | 1777.1 |
| **33.3** | 1787.8 |
| **33.9** | 1828.3 |
| **34.4** | 1857.1 |
| **34.3** | 1873.7 |
| **36.5** | 1979.9 |
| **35.3** | 1981.0 |
| **35.3** | 1966.3 |
| **35.3** | 1984.7 |
| **35.6** | 2016.3 |
| **40.3** | 2165.6 |
| **40.6** | 2174.6 |
| **39.9** | 2201.1 |



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis

REPOSITORIO DE
TESIS UCSM

UNIVERSIDAD
CATÓLICA
DE SANTA MARÍA

| | |
|---|---|
| **41.3** | 2248.4 |
| **42.9** | 2378.9 |
| **43.8** | 2400.7 |
| **44.7** | 2449.5 |
| **47.0** | 2505.5 |
| **48.2** | 2533.8 |
| **6.7** | 2417.9 |



Publicación autorizada con fines académicos e investigativos
En su investigación no olvide referenciar esta tesis