

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

FACULTAD DE CIENCIAS E INGENIERÍA



**OPTIMIZACIÓN DE LA VOLADURA EN UNA MINA SUPERFICIAL A TRAVÉS DE UN
MODELO PREDICTIVO PARA DETERMINAR EL NIVEL DE FRAGMENTACIÓN
APLICANDO MACHINE LEARNING**

Tesis para obtener el título profesional de Ingeniera de Minas

AUTORA:

Nancy Nicolle Vargas Flores

Tesis para obtener el título profesional de Ingeniero de Minas

AUTOR:

Miguel Angel Marcelo Encarnación

ASESOR:

Luis Alberto Mendieta Britto

Lima, Noviembre, 2024

Informe de Similitud

Yo, Luis Alberto Mendieta Britto, docente de la Facultad de Ciencias e Ingeniería de la Pontificia Universidad Católica del Perú, asesor(a) de la tesis/el trabajo de investigación titulado

OPTIMIZACIÓN DE LA VOLADURA EN UNA MINA SUPERFICIAL A TRAVÉS DE UN MODELO PREDICTIVO PARA DETERMINAR EL NIVEL DE FRAGMENTACIÓN APLICANDO MACHINE LEARNING

del/de la autor(a)/ de los(as) autores(as)


Nancy Nicolle Vargas Flores,

Miguel Ángel Marcelo Encarnación,

dejo constancia de lo siguiente:

- El mencionado documento tiene un índice de puntuación de similitud de 17%. Así lo consigna el reporte de similitud emitido por el software *Turnitin* el 14/03/2025.
- He revisado con detalle dicho reporte y la Tesis o Trabajo de Suficiencia Profesional, y no se advierte indicios de plagio.
- Las citas a otros autores y sus respectivas referencias cumplen con las pautas académicas.

Lugar y fecha: Lima 14 de marzo del 2025

Apellidos y nombres del asesor / de la asesora: <u>Mendieta Britto, Luis Alberto</u>	
DNI:44887003	Firma 
ORCID: 0000-0003-1749-3190	

Resumen

La tesis se centra en la implementación de una metodología de Machine Learning supervisado para desarrollar un programa predictivo que optimice los fenómenos de voladura en una unidad minera a tajo abierto en Perú. Se propone el uso de un algoritmo de regresión lineal generalizado multivariable para analizar diversos parámetros operativos, como la resistencia del macizo rocoso, burden, espaciamiento, factor de potencia y tiempos de retardo, entre otros. A través de la revisión de datos históricos de voladuras desde 2016 hasta 2020, se identificaron variaciones significativas en los resultados, lo que indica que la operación de voladura no es uniforme.

Este estudio destaca la importancia de la preparación de datos, donde se abordaron problemas como datos omisos y desordenados, seleccionando aquellos con menor cantidad de fallas para el desarrollo del modelo. Se realizó un análisis de la fragmentación de las voladuras, enfocándose en los parámetros P80 y porcentaje de finos, que mostraron suficiente variabilidad para desarrollar el modelado predictivo.

Además, se revisan estudios previos que aplican técnicas de regresión y redes neuronales artificiales (ANN) en contextos similares, subrayando la relevancia de estos enfoques en la optimización de procesos mineros. La tesis concluye que la implementación de modelos predictivos puede mejorar la eficiencia en la voladura, contribuyendo a una mejor gestión de los recursos y a la seguridad en las operaciones mineras.

Este trabajo no solo busca mejorar los resultados de voladura, sino también contribuir al desarrollo tecnológico de la minería en Perú, un sector clave para la economía del país.

Dedicatoria

A nuestras queridas familias, quienes, con su amor incondicional, paciencia y apoyo constante han sido la fuerza que nos ha impulsado a llegar hasta aquí. Gracias por enseñarnos a nunca rendirnos y por estar siempre a nuestro lado, brindándonos el aliento necesario en los momentos de duda.

A nuestros seres queridos, por ser el refugio de nuestras inquietudes, la fuente de nuestra motivación y la razón por la cual seguimos adelante con esperanza y determinación. Esta tesis es también suya, porque sin ustedes no habiésemos llegado a este logro.

Con todo nuestro cariño y gratitud,

Nancy Vargas y Miguel Marcelo

Tabla de contenidos

1.	Introducción	1
1.1.	Antecedentes	3
1.2.	Objetivos.....	6
1.1.1.	Objetivo general.....	6
1.1.2.	Objetivos específicos.....	6
1.3.	Hipótesis.....	6
2.	Marco Teórico	7
2.1.	Industria Minera	7
2.1.1.	Actividad minera.....	7
2.1.2.	Cadena de valor en el sector minero.....	8
2.1.3.	Minería superficial.....	10
2.1.4.	Perforación y voladura.....	12
2.1.5.	Fragmentación.....	19
2.2.	Machine Learning	26
2.2.1.	Antecedentes de machine learning.....	26
2.2.2.	Machine learning.....	32
2.2.3.	Análisis de regresión.....	36
3.	Metodología.....	40
3.1.	Recolección e identificación de datos.....	40
3.2.	Preparación de datos	41
3.3.	Selección de algoritmos	43
3.4.	Entrenamiento del modelo	44
3.5.	Evaluación de desempeño.....	45
3.6.	Implementación del modelo.....	45
3.7.	Desarrollo de predicciones.....	46
3.8.	Evaluación de la precisión de predicción.....	46
4.	Desarrollo.....	48
4.1.	Definición del problema	48
4.2.	Desarrollo del programa	49
4.2.1.	Recolección de datos.....	49
4.2.2.	Preparación de datos.....	50

4.2.3. Selección de algoritmos.....	61
4.2.4. Entrenamiento del modelo.....	85
4.2.5. Evaluación del desempeño.....	85
5. Presentación y discusión de resultados.....	86
5.1. P80.....	86
5.2. Porcentaje de finos.....	87
6. Conclusiones y recomendaciones.....	89
7. Bibliografía.....	91
8. Anexos.....	96



Índice de tablas

Tabla 1: Estratificación de la Minería Formal.....	9
Tabla 2: Funciones de enlace.....	39
Tabla 3: Tipo de distribuciones.....	39
Tabla 4: Comparación de los modelos.....	39
Tabla 5: Modelo de estructura inicial de datos extraída.....	51
Tabla 6: Modelo de estructura inicial de datos seleccionada.....	52
Tabla 7: Tabla resumen de las variables.....	60
Tabla 8: Tabla de resultados del modelo univariado T-taladros vs P80.....	68
Tabla 9: Tabla de resultados del modelo univariado T-taladros vs P80.....	69
Tabla 10: Tabla de resultados del modelo con interacciones Resistencia y FP vs P80.....	70
Tabla 11: Tabla de resultados del modelo con interacciones Resistencia y FP vs P80.....	71
Tabla 12: Tabla de resultados del modelo univariado Densidad material vs P80.....	72
Tabla 13: Tabla de resultados del modelo univariado Densidad explosivo final vs P80.....	73
Tabla 14: Tabla de resultados del modelo univariado Densidad explosivo final vs P80.....	73
Tabla 15: Tabla de resultados del modelo multivariado con todas las variables vs P80.....	75
Tabla 16: Tabla de resultados del modelo multivariado vs P80.....	76
Tabla 17: Tabla de resultados del modelo univariado T-taladros vs Porcentaje de finos.....	77
Tabla 18: Tabla de resultados del modelo univariado T-taladros vs Porcentaje de finos.....	78
Tabla 19: Tabla de resultados del modelo univariado Factor de Potencia vs Porcentaje de finos.....	79
Tabla 20: Tabla de resultados del modelo univariado Factor de Potencia vs Porcentaje de finos.....	79
Tabla 21: Tabla de resultados del modelo univariado Resistencia vs Porcentaje de finos.....	80
Tabla 22: Tabla de resultados del modelo univariado Resistencia vs Porcentaje de finos.....	80
Tabla 23: Tabla de resultados del modelo univariado Densidad material vs Porcentaje de finos.....	81
Tabla 24: Tabla de resultados del modelo univariado Densidad de explosivo final vs Porcentaje de finos.....	81
Tabla 25: Tabla de resultados del modelo multivariado con todas las variables vs Porcentaje de finos.....	82
Tabla 26: Tabla de resultados del modelo univariado Burden vs P80.....	83
Tabla 27: Tabla de resultados del modelo univariado Burden vs Porcentaje de finos.....	83
Tabla 28: Tabla de resultados del modelo multivariado con todas las variables vs P80.....	84
Tabla 29: Tabla de resultados del modelo multivariado con todas las variables vs Porcentaje de finos.....	84

Índice de figuras

Figura 1. Flowchart de cadena de valor minera.	8
Figura 2. Descripción gráfica de los ángulos inter-rampa y de cara de banco.	11
Figura 3. Acciones básicas de los métodos rotopercutivos y a rotación con tricono.	13
Figura 5. Variación de velocidad de perforación con el número de varillas en DTH y TH.	14
Figura 6. Influencia del diámetro de carga en la velocidad de detonación de acuerdo al tipo de explosivo.	18
Figura 7. Voladura en minería superficial.	22
Figura 8. Variables de diseño en voladura de banco	24
Figura 9. Ejecución de la voladura.	25
Figura 10. Flowchart de método de Machine Learning.	33
Figura 11. Recolección de datos de 272 erupciones volcánicas.	37
Figura 12. Histograma de distribución de datos de burden(m).	53
Figura 13. Histograma de distribución de datos de espaciamento(m).	54
Figura 14. Histograma de distribución de datos de resistencia.	54
Figura 15. Histograma de distribución de datos de factor de potencia (kg/TN).	55
Figura 16. Histograma de distribución de datos de tiempo entre taladros(ms).	56
Figura 17. Histograma de distribución de datos de densidad de material(gr/cm ³).	56
Figura 18. Histograma de distribución de datos de densidad de explosivo final(gr/cm ³).	57
Figura 19. Histograma de distribución de datos de relación nitrato (%).	57
Figura 20. Histograma de distribución de datos de relación nitrito (%).	58
Figura 21. Histograma de distribución de datos de relación emulsión (%).	58
Figura 22. Histograma de distribución de datos de densidad de P80.	59
Figura 23. Histograma de distribución de datos de porcentaje de finos (%).	60
Figura 24. Explicabilidad vs precisión de algoritmos de Machine Learning	62
Figura 25. Gráfico de diagnóstico de normalidad – P80.	63
Figura 26. Gráfico de diagnóstico de normalidad – Porcentaje de finos.	63
Figura 27. Funciones de densidad beta.	67
Figura 28. Ejemplo de regresión beta con funciones de enlace log-log, logit.	68
Figura 29. Tiempo de retardo entre taladros(ms) vs P80 (pulg).	70
Figura 30. Resistencia- Factor de Potencia(kg/TN) vs P80 (pulg).	72
Figura 31. Densidad final del explosivo (gr/cm ³) vs P80 (pulg).	74
Figura 32. Tabla de resultados del modelo multivariado vs P80.	77
Figura 33. Tiempo de retardo entre taladros(ms) vs Porcentaje de finos.	78
Figura 34. Factor de Potencia vs Porcentaje de finos.	79
Figura 35. Resistencia vs Porcentaje de finos.	81
Figura 36. Tabla de resultados del modelo multivariado vs Porcentaje de finos.	82
Figura 37. Gráfico de regresión lineal – P80 (sin burden).	86
Figura 38. Gráfico de regresión lineal – P80 (con burden).	87
Figura 39. Gráfico de regresión lineal – Porcentaje de finos (sin burden).	88
Figura 40. Gráfico de regresión lineal – Porcentaje de finos (con burden).	88

1. Introducción

La minería ha sido un pilar fundamental en la economía del Perú desde tiempos remotos. Las antiguas civilizaciones peruanas trabajaban metales como el cobre, hierro, estaño y oro, destacándose por su habilidad en la orfebrería. Durante la época colonial, la plata se convirtió en la base económica del territorio, mientras que, en los inicios de la república, la explotación del salitre y la construcción de grandes ferrocarriles consolidaron aún más el sector. En la actualidad, según la *Guía para la Inversión Minera en el Perú 2022/2023*, elaborada por el Ministerio de Relaciones Exteriores y EY Perú, la minería representa el 64 % de las exportaciones nacionales y contribuye con el 10 % del PBI. Además, la exministra de Economía y Finanzas, Claudia Cooper, destaca en su libro *Minería y equilibrio económico* que esta actividad es clave para impulsar otros sectores y lograr una diversificación productiva real. Por su parte, el *Anuario Minero 2023*, publicado por el Ministerio de Energía y Minas, reporta que el sector minero aportó al país S/ 7,665 millones en concepto de canon y regalías durante dicho año.

La minería puede entenderse como un proceso de suministro en el que los minerales pasan de ser recursos geológicos a convertirse en productos comercializables. Este proceso abarca tanto la extracción como el procesamiento de los minerales, mediante técnicas de concentración, fundición y/o refinación, con el objetivo de obtener materias primas metálicas listas para su uso en diversas industrias.

Por lo expuesto, se puede concluir que la inversión en el sector minero asciende a miles de millones de dólares. Sin embargo, el valor de un yacimiento no solo depende de sus características geológicas, sino también de la tecnología empleada en su explotación y de la gestión integral del proyecto, que incluye aspectos ambientales y sociales. Dado que el precio de los metales está sujeto a la ley de oferta y demanda, la optimización de costos se vuelve un factor clave para mejorar los

márgenes de ganancia. En la actualidad, se invierte y se investiga constantemente en la mejora de los procesos dentro de la industria minera, siendo la voladura uno de los más estudiados.

La voladura es una actividad de alto riesgo que puede presentar diversos problemas en su aplicación. En la minería superficial, estos inconvenientes incluyen una fragmentación inadecuada, inestabilidad del macizo rocoso, dilución del mineral, sobrecarga de la chancadora, mayor consumo de energía en la planta y un desgaste acelerado de los equipos de chancado, entre otros. Para minimizar estos efectos y optimizar el proceso de voladura, se han implementado técnicas de voladura controlada, como la reducción de los factores de carga, el uso de taladros de precorte, recorte o amortiguadores, así como la variación en el acoplamiento de cargas explosivas y el contenido de tacos en los taladros.

La presente tesis se centra en la optimización del proceso de voladura en minería superficial mediante la aplicación de algoritmos de *machine learning*. Esta metodología permite establecer una relación entre los mecanismos de voladura controlada y sus posibles resultados, con el objetivo de definir su aplicación óptima.

El *machine learning* es una herramienta relativamente nueva en el sector minero y aún no se encuentra ampliamente implementada a nivel nacional. Se trata de una rama de la inteligencia artificial que permite a un sistema aprender a partir de datos recopilados, con el fin de diseñar modelos predictivos capaces de enfrentar distintos escenarios y mejorar la toma de decisiones.

Se comenzará con una explicación general de los temas relevantes a través del marco teórico, seguido del caso de estudio y la metodología aplicada en esta tesis. Además, se detallará la funcionalidad del modelo predictivo en un capítulo dedicado a su desarrollo y aplicación. Posteriormente, se presentarán los resultados operacionales y las evaluaciones estadísticas obtenidas. Finalmente, se expondrán las conclusiones y recomendaciones del estudio.

1.1. Antecedentes

Uno de los efectos negativos como resultado de una mala voladura en minería es la pobre fragmentación. Este desvío genera varias dificultades como, por ejemplo, el chancado de mineral por encima del tamaño máximo, mayor dilución mineral, problemas en control de estabilidad, caída de rocas, entre otros. Todo esto suele incrementar los costos unitarios, principalmente el costo de planta en el proceso de chancado, ya que el costo de energía es de los más elevados en una operación minera; el costo y tiempo extra en el procesamiento del mineral por encima del tamaño óptimo, así como el costo de sostenimiento y transporte.

Es así como existen distintos modelos predictivos para poder determinar la fragmentación ocasionada bajo diferentes parámetros y poder realizar una comparación de resultados, estos modelos generalmente son del tipo matemáticos y empíricos; por ejemplo:

Fórmula de Larsson:

Propuesta por Larsson (1973), se utiliza para la determinación del K50, que es la abertura de malla cuadrada por la que pasa el 50% del material volado.

$$K50 = s * e^{(0,58 * \ln B - 0,145 * \ln (S/B) - 1,18 * \ln (CE/c) - 0,82)}$$

donde:

B = Burden [m].

S/B = Relación espaciamiento-burden.

CE = Consumo específico de explosivo [kg/m³].

C = Constante de roca. Equivale al consumo específico de explosivo gelatinoso necesario para fragmentar la roca, normalmente varía entre 0,3 y 0,5 [kg/m³].

s = Constante de Volabilidad. Es un factor que tiene en cuenta la heterogeneidad y discontinuidades del macizo rocoso.

Fórmula de SVEDEFO (Swedish Detonic Research Foundation):

La fórmula de Larsson no tiene en cuenta el efecto de la altura de banco y la longitud de retacado, pero la SVEDEFO ha añadido un término de tal manera que la ecuación final es:

$$K50 = s * \left[1 + 4,67 * \left(\frac{T}{L} \right)^{2,5} \right] * e^{(0,29 * \ln B^2 * \sqrt{\frac{S}{1,25}} - 1,18 * \ln \left[\frac{CE}{c} \right]^{-0,82})}$$

Las variables nuevas son:

T = Longitud de retacado [m].

L = Profundidad del taladro [m].

Modelo Kuz-Ram:

Modelo desarrollado por Claude Cunningham en 1983, en base a la curva de distribución granulométrica de Rosín-Rammler y la fórmula empírica del tamaño medio de los fragmentos procedentes de voladuras dada por V.M. Kuznetsov. (López. C, 2003)

a) Fórmula de Kuznetsov

$$T_b = Fr * \left(\frac{VRo}{Q} \right)^{0,8} * Q^{\frac{1}{6}}$$

donde:

T_b = Tamaño medio de los fragmentos de voladuras [cm].

Fr = Factor de roca.

VRo = Volumen de roca fragmentada por taladro [m³].

Q = Cantidad de TNT equivalente a la carga de explosivo por taladro [kg].

Q_b = Carga de explosivo dentro del taladro [kg].

PRP = Potencia Relativa en Peso de este explosivo. (ANFO = 100, TNT = 115)

Q_b * PRP = Q * 115

a) Ecuación de Rosín-Rammler:

$$PC = e^{-\left(\frac{Tb}{Tbc}\right)^*u}$$

donde:

PC = Proporción de material retenido para una abertura de malla «Tb».

Tb = Abertura de malla.

Tbc = Tamaño característico.

u = Índice de uniformidad.

Como la fórmula de Kuznetsov da el valor medio de «Tb» entonces PC = 0,5, por lo que:

$$0,5 = e^{-\left(\frac{Tb}{Tbc}\right)^*u}$$

$$Tbc = \frac{Tb}{(0,693)^{\frac{1}{u}}}$$

La ecuación final del modelo KUZ-RAM es una expresión para el índice de uniformidad:

$$n = \left(2,2 - 0,014 * \frac{B}{\Phi h}\right) * \sqrt{1 - \frac{SD}{B}} * [0,5 * \left(1 + \frac{S}{B}\right)] * \left[\frac{|Lb - Lc|}{Ltot} + 0,1\right]^{0,1} * \left(\frac{Ltot}{H}\right)$$

donde:

B = Burden [m].

S = Espaciamiento [m].

Φh = Diámetro del taladro perforado [m].

Lb = Largo de la carga de fondo [m].

Lc = Largo de la columna cargada [m].

Ltot = Largo de carga total [m].

H = Altura de banco o profundidad del taladro [m].

SD = Desviación estándar de la precisión de perforación [m].

1.2. Objetivos

1.1.1. Objetivo general.

Evaluar la aplicación del Machine Learning Supervisado para desarrollar un modelo predictivo de parámetros de fragmentación en la voladura de una mina superficial; buscando optimizar la voladura y prevenir los resultados no favorables.

1.1.2. Objetivos específicos.

- Analizar las variables relacionadas a la voladura en minería superficial.
- Determinar el algoritmo apropiado para el desarrollo del modelo predictivo con Machine Learning Supervisado.
- Seleccionar los modelos de regresión apropiados para el análisis de las variables dependientes del modelo predictivo.
- Generar el modelo predictivo con uso de las variables independientes y dependientes, mediante un modelo supervisado de Machine Learning.
- Evaluar la explicabilidad del modelo predictivo en base a la comprobación de la fórmula con el 20% del total de datos.

1.3. Hipótesis

La aplicación de algoritmos de Machine Learning permite predecir la calidad de fragmentación de la voladura para su posterior optimización.

2. Marco Teórico

2.1. Industria Minera

2.1.1. Actividad minera.

Es una de las principales actividades que se desarrolla hace miles de años, desde inicios de la humanidad, y consiste en la extracción y explotación de minerales de la corteza terrestre, los cuales pueden estar aflorando en superficie o encontrarse en el subsuelo. La actividad minera representa un proceso que va desde los primeros indicios de un posible yacimiento minero hasta la comercialización; ya sea de concentrados o de commodities.

Su importancia radica en que representa un pilar en la economía de muchos países. Por ejemplo, en Perú, la actividad minera representa el 9.1% del PBI Nacional (2023) y el 65.1% de sus exportaciones totales (2023), principalmente en base a la producción de cobre (Cu), oro (Au), plata (Ag), zinc (Zn), plomo (Pb), hierro (Fe), estaño (Sn) y molibdeno (Mo).

Es importante destacar la amplia variedad de usos que tienen los elementos en la vida cotidiana y en diversas industrias. Metales como el cobre (Cu), el oro (Au), la plata (Ag) y el aluminio (Al) son fundamentales en la tecnología, presentes en dispositivos móviles y equipos electrónicos. En la industria de la salud, la plata se emplea por sus propiedades antimicrobianas, mientras que, en el sector energético, el uranio (U) es utilizado como combustible en plantas nucleares. Estos materiales desempeñan un papel clave en el desarrollo de la sociedad moderna.

La actividad minera siempre debe ejecutarse con base en un desarrollo sostenible que considere aspectos económicos, sociales y ambientales. Solo así se podrá generar una industria capaz de maximizar los beneficios económicos de sus accionistas, fomentar un entorno en dónde pueda surgir el progreso de la sociedad y, al mismo tiempo, minimizar el impacto ambiental que pueda generar.

2.1.2. Cadena de valor en el sector minero.

La cadena de la actividad minera se compone de cuatro etapas: exploración, explotación, beneficio y comercialización, las cuales deben funcionar de forma sincronizada para alcanzar los objetivos de la operación. A continuación, se presenta la siguiente figura.

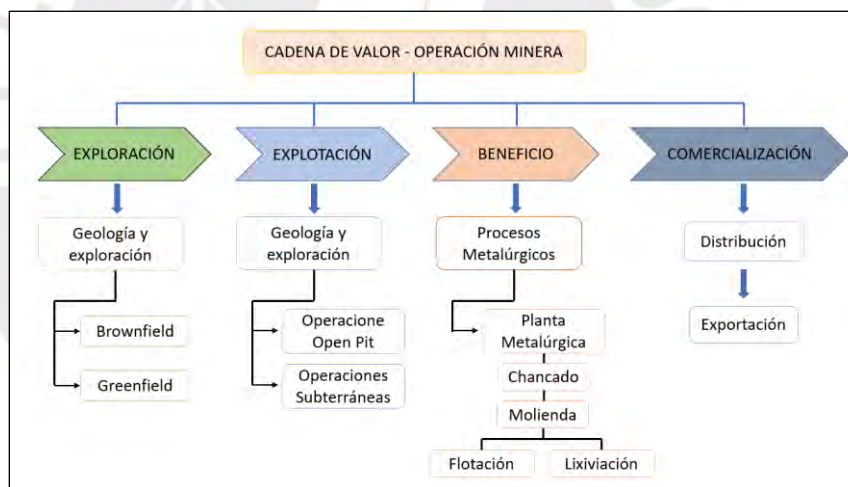


Figura 1. Flowchart de cadena de valor minera.

Fuente: Elaboración propia.

Exploración:

Este es el proceso inicial de la cadena y consiste en determinar o confirmar los minerales que se encuentran dentro del yacimiento minero. Para ello, se realizan desde pruebas fisicoquímicas hasta sondajes diamantinos con la finalidad de aumentar la certeza de los minerales que se encuentra en la zona de interés y, así, determinar la

cantidad de recursos y reservas que posee la mina con base en diversos factores. Las exploraciones de una operación minera en marcha se dividen en dos tipos:

- Exploraciones Greenfield: Se enfocan en la búsqueda de nuevos yacimientos con potencial de extracción en áreas no exploradas.
- Exploraciones Brownfield: Se desarrollan en áreas cercanas de minas actuales o alrededores con la finalidad de ampliar la vida útil de la mina mediante el incremento de los recursos minerales actuales o el aumento de su certeza.

Explotación:

Una vez determinadas las reservas que posee el yacimiento, se procede a extraer los minerales del suelo o subsuelo. Para ello, se diseña y construye la infraestructura de la mina con base en el tipo de minado definido, empleando una variedad de recursos, como equipos mecanizados, material de voladura, material de sostenimiento y personal operativo, entre otros. Todo con la finalidad de obtener el material mineralizado y llevarlo a superficie. Según diversos factores, la operación puede ser subterránea o superficial, cada una con diferentes ratios de tonelaje de producción diaria (TMD). Además, según la normativa peruana, las operaciones se clasifican en tres tipos según su extensión y producción.

ITEM	ESTRATO	HECTAREAS	CAPACIDAD PRODUCTIVA	SECTOR COMPETENTE
1	GRAN Y MEDIANA MINERIA (REGIMEN GENERAL)	MAS DE 2,000 HAS.	MAS DE 350 TMD	MEM - OEFA OSINERGMIN
2	PEQUEÑO PRODUCTOR MINERO (1)	METÁLICO : HASTA 2,000 HAS. NO METÁLICO: SIN LIMITE	METÁLICO: HASTA 350 TMD NO METÁLICO: HASTA 1200 TMD	GOBIERNO REGIONAL
3	PRODUCTOR MINERO ARTESANAL (2)	METÁLICO: HASTA 1,000 HAS. NO METÁLICO: SIN LIMITE	METÁLICO: HASTA 25 TMD NO METÁLICO: HASTA 100 TMD	GOBIERNO REGIONAL
(1) Yacimiento tipo Placer		Hasta 3000 m3/Día		
(2) Yacimiento tipo Placer		Hasta 200 m3/Día		

Tabla 1: Estratificación de la Minería Formal.

Fuente: "MINEM (2013)".

Beneficio:

Esta etapa consiste en el procesamiento de los minerales extraídos de mina para la recuperación de los metales económicos mediante un proceso metalúrgico. Dicho proceso incluye las etapas de chancado, molienda, flotación o lixiviación y filtrado. La efectividad de estos procesos se mide con base en las recuperaciones metalúrgicas, con el objetivo de obtener un producto final, como un concentrado de mineral, metales preciosos fundidos (barra doré, bullion) o refinados de cobre (Cu), zinc (Zn) y estaño (Sn) en forma de cátodo o blíster.

Comercialización:

Es la etapa final de la cadena de valor de una operación minera y consiste en la comercialización y distribución de los productos finales. Una vez obtenidos, estos son transportados por vía marítima hacia los diferentes países consumidores.

Asimismo, se considera la recuperación comercial, en la cual se determinan las pérdidas asociadas a maquilas, penalidades, fletes, entre otros.

2.1.3. Minería superficial.

El desarrollo del presente trabajo de tesis se enfocará en los datos de una mina superficial, por lo cual se explicará brevemente en que consiste este tipo de minería.

La minería superficial, también conocida como open pit, es un método masivo de explotación de un yacimiento minero. Este método posee un conjunto de características propias del método, tales como:

- Usualmente desarrolla cuando el yacimiento mineralizado se encuentra aflorando y posee poca profundidad.

- Extensión horizontal amplia, pues, suelen cubrir grandes áreas; lo que facilita el desarrollo de tajos abiertos y extracción masiva.
- Los depósitos suelen ser diseminados y de gran tonelaje, como pórfidos cupríferos, depósitos lateríticos y menas de hierro (Banded Iron Formations).
- Usualmente posee grandes tonelajes de mineral con bajas leyes; ya que al tener un costo unitario de extracción menor que en minería subterránea, el minado masivo permite compensar la menor ley del mineral.
- Se evalúa para su viabilidad económica, el Strip-Ratio que refiere a la cantidad de TN de material estéril que se necesita extraer para obtener 1 TN de mineral.
- Debido a las grandes dimensiones de los bancos y rampas, permite el uso de equipos mecanizados de grandes capacidades.
- Para su evaluación económica, se requiere el Pit Óptimo Final, cuyo objetivo es maximizar la extracción del material mineralizado (recursos-reservas), así como optimizar el valor actual neto (VAN) y la tasa interna de retorno (TIR) en conjunto con la tasa de descuento (WACC).

A continuación, se muestra una imagen con las principales partes de un Tajo Abierto que usualmente se desarrollan:

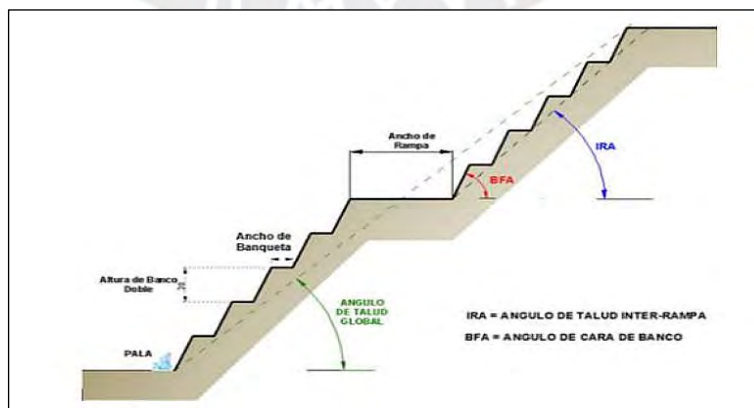


Figura 2. Descripción gráfica de los ángulos inter-rampa y de cara de banco.

Fuente: INGEMMET (2021).

En el Perú las principales compañías mineras que emplean el método Open Pit, dividido por tipo de metal a extraer son:

- Cobre: Cerro Verde, Antamina, Shouthern Copper, Las Bambas, Quellaveco, Chinalco, Antapaccay, Marcobre.
- Oro: Yanacocha, Summa Gold (Operación el Toro), Alto Chicama.
- Polimetálicas: Antamina, Cerro de Pasco, Marcona.

2.1.4. Perforación y voladura.

Representan dos de las actividades críticas de la operación superficial y responsables directas de la obtención del P80 y el Porcentaje de finos, los cuales indicarán los resultados de la perforación y voladura al final del proceso y serán objetivos de la presente investigación.

La perforación siendo el punto de partida de la operación, cumple un importante rol, ya que, si este proceso no se desarrolló, según los parámetros establecidos de longitud de perforación, diámetro de perforación, máxima desviación de taladros, uso de brocas adecuadas y todos los parámetros que implican el método de rotopercusión o rotativo, el resto del proceso no tendrá los resultados esperados, debido a que el paso previo no se desarrolló con éxito.

Los métodos mecánicos para la perforación de rocas que existen actualmente son:

- Método Rotopercutivo: Consiste en utilizar un martillo ya sea en el fondo (DTH) o en la cabeza del barreno (Top Hammer) y es aplicado en la mayoría de los tipos de rocas.

- Método Rotativo: En caso de rocas duras se realiza la penetración por trituración, empleando las brocas tricono y para rocas blandas se utiliza la penetración por corte usando brocas especiales.

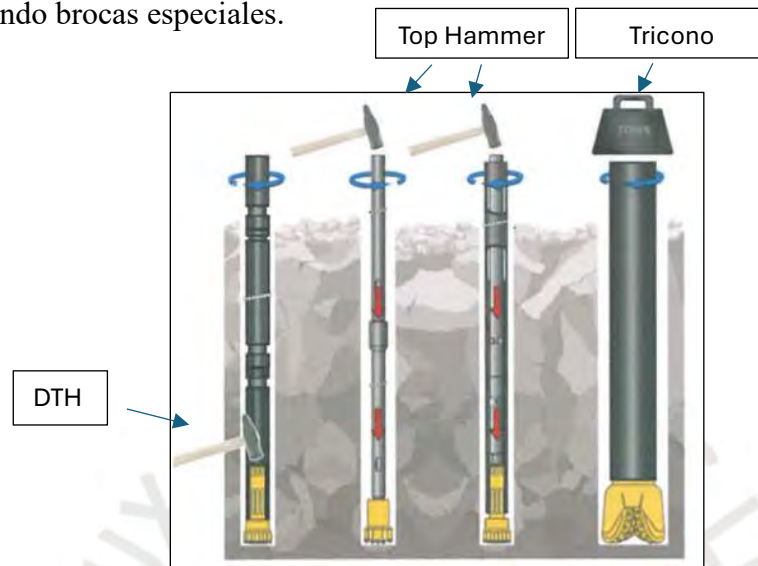


Figura 3. Acciones básicas de los métodos rotopercutivos y a rotación con tricono.

Fuente: "Famesa Explosivos S.A.C" (2019).

En la minería superficial, la perforación se basa principalmente en el método rotopercutivo, el cual utiliza una fuente de energía mecánica, como un martillo, para transmitir la fuerza a través del varillaje o los barrenos. Esta energía se transfiere a la broca, que impacta directamente sobre la roca a perforar. Además, se emplea un fluido de barrido para evacuar y limpiar los detritos dentro del taladro, asegurando un proceso eficiente.

Los equipos de perforación se clasifican según el sistema utilizado: rotopercutivo, ya sea con DTH o Top Hammer, o rotativo, utilizando tricono. Según las necesidades operativas, se pueden emplear equipos sobre orugas con compresor remolcado o equipos autopulsados sobre neumáticos. (Famesa Explosivos S.A.C., 2019).

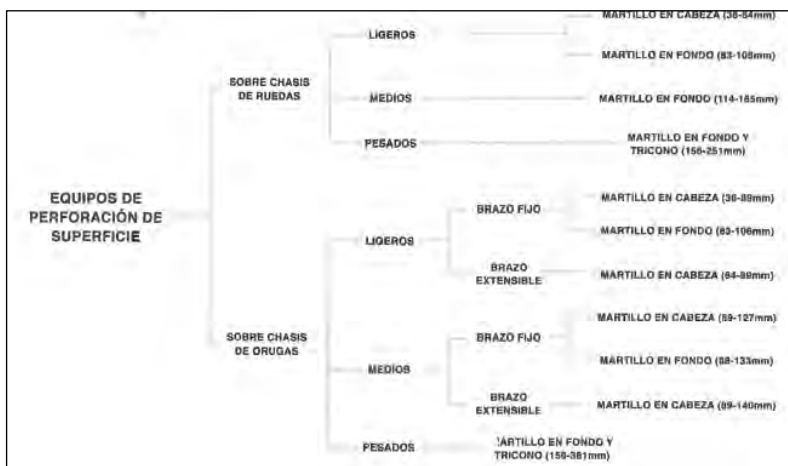


Figura 4. Sistemas de montaje de perforadoras de superficie.

Fuente: "Famesa Explosivos S.A.C" (2019).

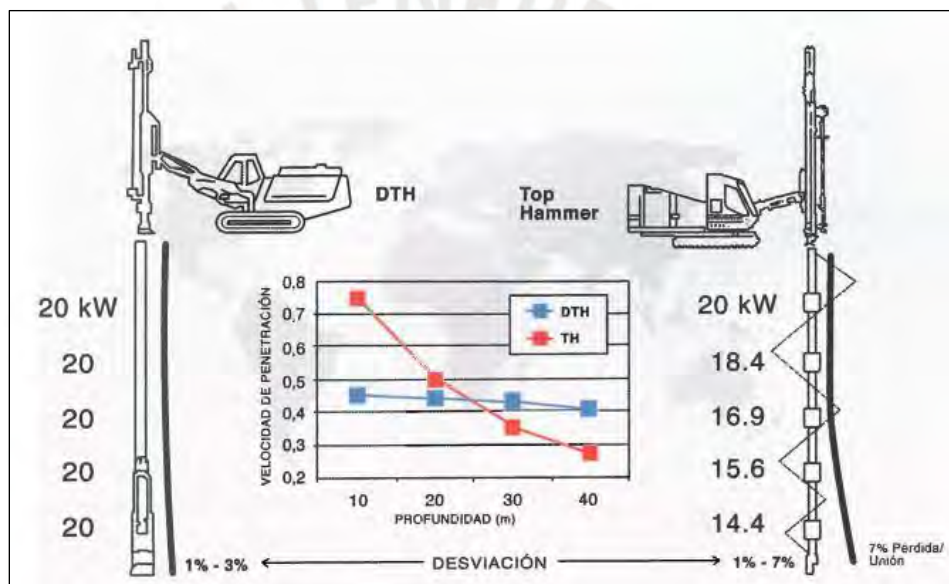


Figura 5. Variación de velocidad de perforación con el número de varillas en DTH y TH.

Fuente: "Famesa Explosivos S.A.C" (2019).

2.1.4.1. Explosivos.

La fragmentación de rocas por voladura comprende a la acción de un explosivo y a la consecuente respuesta de la masa de roca circundante, involucrando múltiples factores como el tiempo, energía termodinámica, ondas de presión, mecánica de rocas y otros; en un rápido y complejo mecanismo de interacción.

En términos generales, un alto explosivo reúne tres características básicas: (1) es un compuesto o mezcla química que se inicia por calor (temperatura), choque, impacto, fricción, o una combinación de estas condiciones; (2) cuando se inicia mediante un detonador o “booster” se descompone rápidamente (del orden de una fracción de microsegundo); y, (3) con la detonación, se produce una liberación rápida de calor y grandes cantidades de gases de alta presión, los que se expanden rápidamente con la potencia suficiente para vencer fuerzas de confinamiento, como por ejemplo, la fuerza de confinamiento ejercida por la masa rocosa que rodea al explosivo (ISEE, 2008).

Tres términos de importante definición son la combustión, deflagración y detonación. Estas constituyen las tres formas de descomposición de una sustancia explosiva y están relacionadas en términos de confinamiento, velocidad de reacción y propagación de la onda de choque. Cuando el mecanismo predominante es la transmisión de calor (como mecanismo de activación), la reacción se da entre la sustancia que se oxida y una sustancia oxidante (oxígeno) y el fenómeno se clasifica como combustión. Cuando la propagación de esta reacción se produce a velocidades superiores a la del sonido, generando una onda de choque, acontece una deflagración. Si esa velocidad toma valores supersónicos, la onda de choque empieza a propagarse en la sustancia explosiva. Esta situación genera el calentamiento debido al choque y, como la transición de la energía tiene velocidades más elevadas que la expansión de los productos de reacción, continúa generando la onda choque, estimulando la columna de explosivo, aconteciendo así, una detonación (Sena Leite, 2013).

Es así como a un explosivo se le atribuye el ser una sustancia capaz de, mediante su detonación o deflagración, liberar energía de forma explosiva, acompañada de gases de expansión.

En las operaciones de voladura comercial, la detonación de explosivos tiene cinco efectos básicos:

- (1) Fragmentación de roca
- (2) Desplazamiento de roca
- (3) Vibración del terreno
- (4) Sobrepresión del aire
- (5) Gases o humos tóxicos y no tóxicos

Cada explosivo tiene ciertas características o propiedades específicas que lo diferencian de los otros; conociendo qué propiedades son críticas para un buen desempeño es que los ingenieros de explosivos pueden escoger el mejor tipo de explosivo para una situación específica de voladura (ISEE, 2008).

Existen diferentes tipos de explosivos; entre los de mayor uso comercial para la minería se encuentran: la dinamita, los hidrogeles o slurries, las emulsiones, el ANFO y la pentrita.

Dinamita:

Resulta de la mezcla de nitroglicerina (NC) con materiales diversos y en variados porcentajes; de esta forma se producen diferentes tipos y grados de dinamitas. La dinamita es un tipo de explosivo sensible al detonador y que contiene nitroglicerina como sensibilizador.

Existen 3 especies de dinamitas: granulares, semi-gelatinosas y gelatinosas. Las dinamitas semi-gelatinosas alcanzan mayores velocidades de detonación y producen menor cantidad de gases que las granulares; asimismo, poseen una mayor resistencia a la presencia de agua; por este motivo son populares en minería subterránea y en voladuras de canteras.

Hidrogeles / Slurries:

- Slurry: consiste en una solución acuosa espesada de un oxidante y/o sales de combustibles (normalmente nitrato de amonio (NA), nitrato de sodio (NS) y/o nitrato de calcio (NC)), en la que se dispersan oxidantes sólidos adicionales y/o combustibles (comúnmente el aluminio, carbón, gilsonita, azúcar, etilenglicol o aceite), así como también sensibilizadores químicos y/o físicos. Es líquido, bombeable y esencialmente fácil de disolver en agua (ISEE, 2008).
- Explosivo en hidrogel: se genera gracias a la adición de un agente de reticulación (S) al slurry; consiguiendo una mayor estabilidad y resistencia al agua. Este es el tipo de explosivo que casi siempre se emplea como la carga de explosivo final al interior del taladro (ISEE, 2008).

Emulsiones:

Una emulsión es un explosivo diseñado para aplicaciones como columna explosiva. Da buenos resultados de fragmentación, reduce los gases post voladura y es altamente resistente al agua.

ANFO:

Es una pequeña esfera porosa llamada “prill” y compuesta principalmente por nitrato de amonio mezclado con aceite combustible.

El producto ANFO de mayor uso es una mezcla balanceada de oxígeno, de libre fluido, conteniendo aproximadamente 94% de nitrato de amonio y 6% de combustible Diesel No 2. No es resistente al agua y suele producir mayor cantidad de gases que los explosivos mencionados anteriormente; sin embargo, es muy utilizado en la industria minera, si las condiciones geológicas e hidrológicas lo permiten, debido a su bajo costo (ISEE, 2017).

Una alternativa al ANFO, que ha ganado mucha popularidad es el Heavy ANFO, familia de explosivos formada por distintas proporciones de mezcla entre emulsión y ANFO; a diferencia del ANFO los Heavy ANFO sí poseen cierta resistencia al agua, producen menos humos y no resultan tan costosos como las emulsiones.

Pentrita/Pentolita:

Los boosters son sensibles al detonador que, por lo general, contienen altos explosivos (pentrita/pentolita); de esta manera funcionan como primer; es decir, tienen como objetivo la iniciación de los agentes de voladura (ENAEX, n.a.).

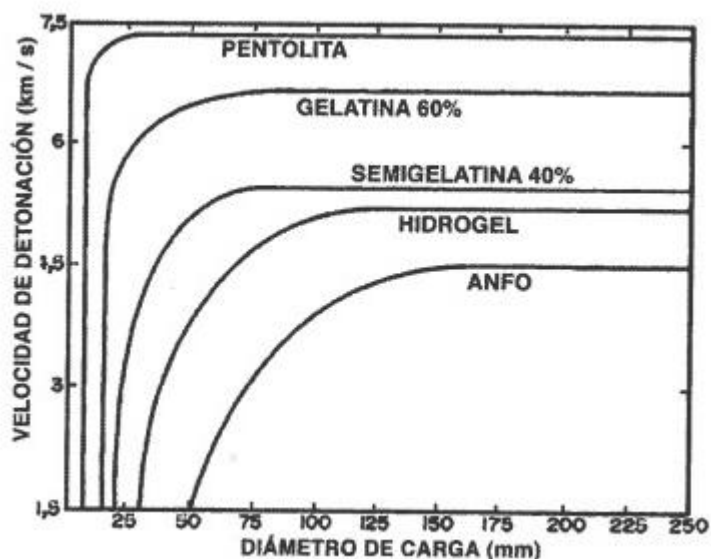


Figura 6. Influencia del diámetro de carga en la velocidad de detonación de acuerdo al tipo de explosivo.

Fuente: "Famesa Explosivos S.A.C" (2019).

2.1.4.2. Dispositivos de iniciación.

Un explosivo puede tener diferentes formas de iniciación: la temperatura, el choque, el impacto y la fricción. Convencionalmente, existían 2 categorías de dispositivos de iniciación: eléctricos y no eléctricos, dependiendo de su fuente de energía principal; asimismo, en las últimas décadas, se ha implementado el uso de detonadores electrónicos que permiten un mayor nivel de control de tiempos, optimizando la voladura.

Los detonadores eléctricos emergieron como el estándar; sin embargo, fueron superados por la introducción de los detonadores no-eléctricos que se activan sin la necesidad de la energía eléctrica.

Avances más recientes en la tecnología de explosivos, ha sido el desarrollo de sistemas de iniciación electrónica. Los detonadores electrónicos alcanzan una precisión de retardos de milisegundos; mejorando la productividad, control medioambiental y la seguridad de la operación.

2.1.5. Fragmentación.

¿Qué es fragmentación?

Como se mencionó previamente, la fragmentación de las rocas corresponde a la acción de un explosivo; ya sabemos que un explosivo es la mezcla de productos: uno combustible y un oxidante dan lugar a una reacción exotérmica que produce altas temperaturas, grandes presiones y gran cantidad de gases; este fenómeno es aprovechado para el rompimiento de las rocas.

En sí, el proceso de fragmentación de un macizo rocoso es una respuesta a la interacción de las propiedades mecánicas de las rocas controladas por la geología del

material, tales como su resistencia, y la aplicación de esfuerzos mediante, por ejemplo, la carga de explosivos.

Una vez ejecutada la voladura, es necesario analizar los resultados obtenidos; ya que, esto permitirá introducir modificaciones sucesivas en los parámetros de diseño de las siguientes voladuras; constituyendo una etapa básica dentro del proceso de optimización. Para evaluar el nivel de fragmentación, en el presente estudio se están utilizando 2 parámetros: P80 y porcentaje de finos.

- P80: Para la determinación de la granulometría observada en una voladura (para evaluar si la voladura fue exitosa en términos de fragmentación), se emplea el término P80, que se define como el 80% de material pasante de una muestra. Este concepto hace referencia al mayor tamaño que alcanzan el 80% de fragmentos de macizo rocoso en la voladura.
- Porcentaje de finos: Es la cantidad porcentual en volumen, del macizo rocoso como resultado de la voladura, que está por debajo del tamaño máximo para considerar a los fragmentos como finos aguas abajo; es decir, en la planta de procesamiento.

Asimismo, por cuestiones de practicidad, en esta tesis, se han propuesto cuatro categorías de clasificación para los factores que podrían intervenir en la fragmentación: (1) Características del macizo rocoso, (2) Características del explosivo, (3) Diseño de la malla y distribución de explosivos y (4) Ejecución de la voladura.

1. Características del macizo rocoso:

La caracterización geomecánica del macizo rocoso es fundamental para entender el comportamiento que tendrá el mismo en interacción con las labores de minado. Hoy en

día, existen diversas clasificaciones de caracterización del macizo rocoso, sin embargo, la presente tesis se enfocará en la clasificación de Bieniawski o índice RMR.

Caracterización por índice RMR:

El índice RMR o Rock Mass Rating, fue desarrollado por Bieniawski para clasificar un macizo rocoso de acuerdo a un puntaje del 0 al 100, en donde 0 representa un macizo de calidad muy mala y 100, de calidad muy buena (Ramírez & Alejano, 2004). Además, la puntuación se desarrolla con respecto a distintas escalas de acuerdo con los parámetros que se evalúen.

Entre los parámetros a evaluar tenemos:

- Resistencia de roca intacta
- RQD (Rock Quality Designation)
- Espaciado entre discontinuidades
- Condición de discontinuidades (Meteorización, apertura, persistencia, etc.)
- Presencia de agua
- Orientación de discontinuidades

Luego de haber analizado y sumado puntos por cada uno de ellos, considerando que casi todos los parámetros suman y solo la orientación tiende a afectar con puntaje negativo, el puntaje final nos permitirá clasificar al macizo según rangos designados en el método, pudiendo reconocerse un macizo rocoso categoría I, II, III, IV o V, siendo la categoría I para los macizos de la mejor calidad.

2. Características de los explosivos:

Las características propias del material explosivo a utilizar se tienden a definir según las condiciones de operación.



Figura 7. Voladura en minería superficial.

Fuente: Revista Tecnología Minera, 2022.e

Entre las características más destacadas encontramos:

- **Potencia explosiva:** Se define como la capacidad del explosivo para fracturar la roca y proyectar los restos, en ese sentido, se referirá a la energía liberada que se aprovecha en la voladura. Se tiende a estimar con aplicación de un péndulo balístico o fórmula de energía relativa por unidad de peso y volumen (Bernaola, 2014)
- **Poder rompedor:** Esta se refiere a la capacidad del explosivo para fracturar las rocas por acción de las ondas de detonación, sin considerar la presión originada por los gases. Se tiende a determinar mediante el método de Hess (Bernaola, 2014).
- **Velocidad de detonación:** Como su nombre lo indica, se define como la velocidad a la que se produce la detonación, esta es la transformación del explosivo en un volumen de gases de gran presión y temperatura. Se tiende a determinar por mediciones en laboratorio o realizadas en campo (Bernaola, 2014).
- **Densidad de explosivo:** La densidad es la unidad de masa por unidad de volumen que ocupa un material, en este caso, el explosivo. Esta dependerá de los componentes utilizados para desarrollar el explosivo. La densidad relativa se comprende entre 0.8 y 1.5 g/cm³.

- **Diámetro crítico:** Se refiere al diámetro mínimo de volumen cilíndrico de la carga de explosivo para el que la onda de detonación se propaga adecuadamente, un diámetro menor originaría que la onda no se propague o con una velocidad muy baja (Bernaola, 2014).
- **Masa crítica:** Es la mínima masa de explosivo necesaria para que se origine una detonación por acción de una llama (Bernaola, 2014).
- **Resistencia al agua:** Se refiere a la habilidad que posee el explosivo por la cual este mantiene sus propiedades durante un tiempo en contacto con el agua, sin necesidad de una cubierta protectora (Bernaola, 2014). Esta característica es esencial para definir el tipo de explosivo a utilizar de acuerdo con la presencia de agua en el frente.
- **Sensibilidad:** Se define como el grado de energía que se necesita transmitir al explosivo para que se inicie la detonación; la sensibilidad se puede considerar por acción del detonador, las ondas explosivas, los choques o el rozamiento (Bernaola, 2014).
- **Estabilidad química:** Se refiere a la característica del explosivo por la cual este puede mantenerse químicamente inalterable en el tiempo (Bernaola, 2014). El correcto almacenamiento del explosivo contribuye con la estabilidad química del mismo, y, con ello, evitar su descomposición.

3. Diseño de la malla y distribución de explosivos:

La malla de perforación se planteará de acuerdo las necesidades de producción y condiciones del frente. Esta manejará distintos parámetros en cuanto a la geometría de este y la distribución de las columnas explosivas, así como la cantidad de explosivo presente. Entre estos parámetros consideramos:

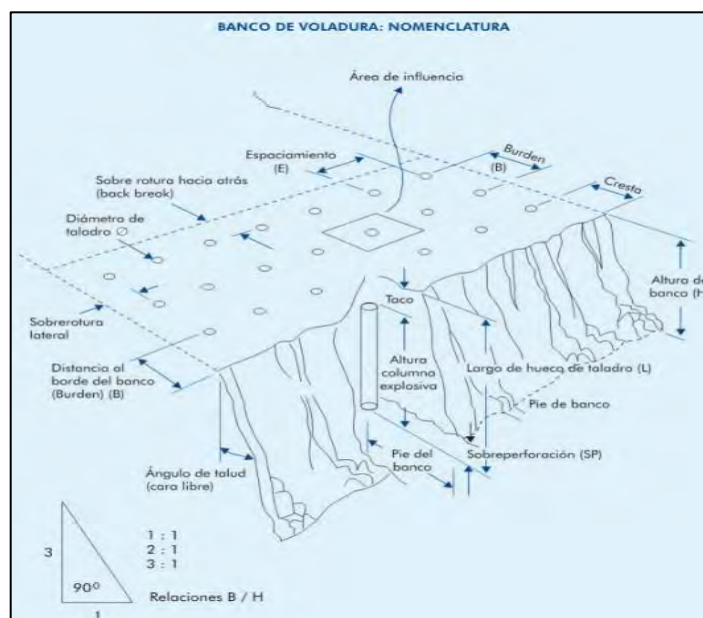


Figura 8. Variables de diseño en voladura de banco.

Fuente: "Manual práctico de voladura EXSA" (2021)

- Factor de carga: Es definida como la relación de la cantidad de explosivo aplicada por cantidad de tonelaje a extraer en acción de esta.
- Diámetro del taladro: Los diámetros de los taladros tenderán a variar de acuerdo con la aplicación del taladro desarrollado. En caso se requiera de taladros de alivio o vacíos, se dispone la utilización de diámetros más extensos en comparación con los diámetros de los taladros de producción. Además, los taladros de recorte o precorte pueden variar sus diámetros en comparación con los de producción de la misma forma.
- Longitud del taladro: Se define como la profundidad de la perforación que se requerirá en cada taladro, usualmente se aplica una longitud de doce pies aproximadamente.
- Burden: Es definida como la distancia perpendicular a la cara libre de las filas de taladros a perforar.
- Espaciamiento: Esta se define como la distancia de separación entre taladros perpendicular a la orientación del burden medido.

- Tiempo de retardo: Para el diseño de la voladura se tiende a hacer uso de retardantes, estos son instrumentos aplicados para reducir la velocidad de la detonación de mechas explosivas y poder sincronizar la voladura de ciertos taladros, además de disponer el orden de salida de estos. Estos tiempos tenderán a variar de acuerdo con el armado planteado para cada operación de voladura.
- Tipo de sistema de iniciación: Este parámetro es determinado por el mecanismo utilizado para iniciar la voladura, pudiendo darse por medios eléctricos, no eléctricos o electrónicos. En el caso de los eléctricos, el tiempo de retardo es más inexacto y presenta mayor peligro en su aplicación, sin embargo, son los sistemas de iniciación más económicos; por otro lado, los sistemas de iniciación electrónicos desarrollan tiempos de iniciación más exactos con la realidad por la aplicación de chips y control programado en computadora; finalmente, los no eléctricos son más seguros y exactos que los eléctricos, pero más inexactos que los electrónicos.

4. Ejecución de la voladura:



Figura 9. Ejecución de la voladura.

Fuente: IIMP, 2017.

Esta categoría de parámetros hace referencia al factor humano presente en todas las labores pertenecientes a la voladura, el no lograr rendimientos perfectos es un accionar

normal del ser humano y como tal, debemos anticipar las posibles deficiencias a ocurrir en cada labor, los parámetros considerados fueron:

- Desviación del taladro:

El taladro puede tender a desviarse en relación directa con la longitud de la perforación e inclusive por la inexperiencia o poca habilidad del perforista, la medición de este permite parametrizar los errores y reanalizar estrategias de preparación de la voladura.

- Eficiencia del perforista: Este parámetro se relaciona más con el factor humano presente en las labores, un perforista puede controlar la perforadora de mejor manera que otro de acuerdo con la experiencia que paisea en la labor, así pudiendo evitar atascos del barreno, taladros desviados o sobre perforaciones innecesarias.

- Eficiencia en el carguío de explosivos: Este parámetro dependerá de la habilidad del cargador de los taladros, se puede cargar con mayor o menor eficiencia de acuerdo con el confinamiento necesario y la fricción que exista entre el explosivo y las paredes del taladro. Un carguío erróneo originará voladuras igual de deficientes.

- Preparación de la armada: Esta se refiere a la labor de atado de todas las columnas explosivas del frente y diseño de la voladura de acuerdo con los tiempos de retardo necesarios para su desarrollo óptimo. Así, controlando el orden de detonación de las columnas explosivas.

2.2. Machine Learning

2.2.1. Antecedentes de machine learning.

En el paper “The Use of a Machine Learning Method to Predict the Real-Time Link Travel Time of Open-Pit Trucks”, Xiaoyu Sun, Hang Zhang, Fengliang Tian and Lei Yang buscan realizar una predicción realista del tiempo de acarreo de los camiones;

uno de los factores críticos en el sistema dispatch de una mina open pit. Los STS (shovel-truck systems o sistemas de camión-pala) representan aproximadamente el 50% de los costos de explotación en minas a cielo abierto. Por lo tanto, casi todas las grandes minas superficiales están tratando de optimizar el transporte de camiones para lograr menores costos y una mayor eficiencia minera.

Se hizo uso de data de Fushun West Open-pit Mine (FWOM), la mina superficial de carbón más grande de Asia (3 millones de toneladas de carbón al año), para diseñar modelos predictivos de tiempo de transporte de los camiones de la flota usando Machine Learning. Como es sabido, el tiempo del ciclo de acarreo se ve influenciado por diferentes factores como las condiciones del camino, la precisión del carguío y las condiciones meteorológicas; es así que en este caso, se decidió clasificar los caminos en fijos y temporales, de tal manera que se puedan evaluar los resultados de acuerdo a esa clasificación y poder entrenar al modelo predictivo de manera independiente para cada tramo, en lugar de para toda la red de caminos; además también se tomaron en cuenta las condiciones meteorológicas y se usaron más de un millón de registros para entrenar al modelo.

Al igual que en la tesis que se está planteando, el LTTP (link travel time prediction o predicción de tiempo de acarreo por tramos) de los open pit trucks corresponde a una tarea que se realizaría por ML supervisado debido a que tanto los inputs como los outputs son conocidos y ya están clasificados.

En este estudio, se decidió utilizar 3 algoritmos: kNN (k nearest neighbours), SVM (support vector machine) y RF (random forest). Llegando a la conclusión que los modelos basados en SVM y RF eran más exactos que los basados en kNN. Asimismo,

la precisión de las predicciones en este estudio resultó aproximadamente 15.8% más exactas que las calculadas con otros métodos tradicionales. El incluir las variables meteorológicas incrementó la exactitud en 5.13%; y, el realizar el estudio por tramos la incrementó en 11.82%. (Sun et al., 2018)

En la investigación titulada “Optimizing overbreak prediction based on geological parameters comparing multiple regression analysis and artificial neural network”, Hyongdoo Jang y Erkan Topal, a partir de la recolección de data en el túnel Bigye, ubicado en la ciudad Gumi, en Corea del Sur, se buscó realizar un modelo predictivo para la generación de overbreak en tunelería. En este artículo, el diseño de voladuras y el parámetro de distribución de explosivos se consideraron constantes en el túnel; por lo que, la investigación se centró en los efectos de los parámetros geológicos en el overbreak. Se aplicó el análisis de regresión múltiple lineal y el no lineal (LMRA y NMRA); además del algoritmo de artificial neural network (ANN) a 49 conjuntos de datos de clasificación de macizo rocoso (RMR) y de overbreak como parámetros de entrada y salida, respectivamente. El desempeño de los modelos optimizados (LMRA, NMRA y ANN) se evaluó comparando coeficientes de correlación; los valores obtenidos fueron 0,694, 0,704 y 0,945, respectivamente, lo que significa que el nivel más alto de precisión se logró con el algoritmo ANN optimizada. (Jang & Topal, 2013)

Rock Engineering Systems o RES fue propuesto por John Hudson en 1992 y es una de las técnicas más poderosas para resolver intrincados problemas de mecánica de rocas en los que todas las variables pueden afectarse entre sí. Es muy imperativo considerar no solo todos los factores primarios, sino también cómo interactúan todos los factores. Este objetivo se logra a través de la matriz de interacción (interaction matrix), que es

el elemento clave en el enfoque de RES. En “Applying Rock Engineering Systems Approach for Prediction of Overbreak Produced in Tunnels Driven in Hard Rock”, Hasan Mohammadi y Adalat Azat propusieron un modelo RES para la predicción de overbreak.

Para este estudio, la data se obtuvo de 4 túneles localizados en el proyecto de la presa de Bakhtiari, en Iran. Para 54 operaciones de voladura realizadas en los túneles investigados, se midieron los valores de 14 parámetros principales y posibles causantes de overbreak; y, los resultados mostraron que el modelo poseía un alto valor de coeficiente de correlación (0.91) lo que sugiere que es un modelo confiable. Asimismo, los resultados del estudio también insinúan que el índice de calidad de roca (Q), el factor de carga de los taladros perimetrales y la ratio burden/espaciamiento juegan un rol importante en los valores de overbreak. (Mohammadi & Azad, 2020)

Manoj Khandelwal y M. Monjezi realizaron un modelo predictivo para el backbreak en “Prediction of backbreak in open pit blasting operations using of machine learning methods”. El backbreak es un fenómeno perjudicial para la voladura en las minas superficiales; ya que, puede generar inestabilidad en los taludes, caída de maquinaria, mala fragmentación, reducción en la eficiencia de la perforación, entre otros. El hecho de que existan muchas variables y que la correlación entre algunas de ellas sea desconocida genera que los modelos empíricos suelen ser inexactos. Por lo que, en este estudio se aplicó el método de SVM (support vector machine) para la predicción de backbreak en las operaciones de Sougun Iron Mine, Iran. Las variables que se toman en cuenta para dicho análisis están relacionadas a las propiedades del macizo rocoso y el diseño de voladura. De acuerdo a los resultados obtenidos en el

estudio, a partir de la aplicación del algoritmo SVM, el COD (coefficient of determination) y el MAE (mean absolute error) fueron de 0.987 y 0.29 respectivamente. (Khandelwal & Monjezi, 2013).

La tesis “Machine Learning aplicado al control de la fragmentación de rocas en la voladura de minas a tajo abierto”, se centra en el diseño de una malla de voladura óptima en minería, utilizando un enfoque cuantitativo y experimental. Se analizan datos de una unidad minera para establecer patrones de comportamiento y probar hipótesis relacionadas con la fragmentación de rocas, empleando técnicas de recolección, limpieza y análisis de datos, así como modelos predictivos basados en redes neuronales artificiales. Es así como se logró establecer una relación significativa entre las variables independientes y la fragmentación, permitiendo optimizar el diseño de la malla de voladura. La implementación del modelo predictivo mejoró la precisión en la predicción del Factor de Carga, contribuyendo a una mayor eficiencia en el proceso de voladura. (Idrogo, 2022).

Un estudio reciente realizado por Sharma et al. (2024) se centró en la predicción de la fragmentación de rocas en una mina de carbón en India, donde las condiciones de alta temperatura, provocadas por la combustión espontánea del carbón, presentan desafíos significativos durante el proceso de voladura.

El estudio llevó a cabo 100 voladuras experimentales, siguiendo estrictas medidas de seguridad, para analizar el impacto de diversas variables de diseño de voladura, como la relación entre el espaciamiento y el burden (S/B), la relación entre la longitud del taladro y el burden (l_d/B), relación entre la longitud del taco y del taladro (l_s/l_d) y la dureza del macizo rocoso medida con el martillo Schmidt. Para ello, se emplearon

técnicas avanzadas de modelamiento, incluyendo el Reponse Surface Method (RSM), Artificial Neural Networks (ANN) y Random Forest Algorithm (RFA), con el objetivo de predecir el tamaño medio de los fragmentos (k_{50}).

Los hallazgos del estudio indicaron que la selección adecuada de parámetros de voladura es crucial para lograr una fragmentación efectiva y segura, especialmente en condiciones adversas. Además, se implementaron medidas de precaución, como el enfriamiento de los agujeros de voladura con agua si la temperatura superaba los 80 °C, para prevenir detonaciones prematuras. Este trabajo proporciona una base valiosa para futuras investigaciones y aplicaciones en el diseño de voladuras en minas de carbón, destacando la importancia de utilizar métodos de modelado avanzados para optimizar la fragmentación de rocas. (Sharma et al., 2024).

Otro estudio realizado en 2024 se centró en el desarrollo de un modelo de predicción para la fragmentación de rocas en una cantera de caliza ubicada en el distrito de Baloda Bazaar, Chhattisgarh, India. Este trabajo aborda los desafíos asociados con la fragmentación, como el tamaño excesivo del material fragmentado y los costos indirectos que esto conlleva, así como problemas relacionados con la vibración del suelo y el transporte de material.

El estudio implementa algoritmos de aprendizaje automático (Random Forest Regression y Support Vector Regression) y deep learning (Neural Network Regression); se consideraron parámetros como número de taladros, espaciamiento, burden y factor de potencia para el entrenamiento de los modelos. El objetivo del estudio fue mejorar la precisión de las predicciones de fragmentación, lo que se traduce en una mayor eficiencia en las operaciones mineras y una reducción de costos asociados

a la voladura. A través de la recopilación de datos y la división de estos en conjuntos de entrenamiento, validación y prueba, los autores logran ajustar los hiperparámetros y evaluar el rendimiento del modelo, contribuyendo así a un enfoque más efectivo y sostenible en la minería a cielo abierto. (Kumar R. et al., 2024).

2.2.2. Machine learning.

En la actualidad, el desarrollo de machine learning, inteligencia artificial (IA) y ciencias de la computación como forma de ayuda para la toma de decisiones es prueba fehaciente de la creciente importancia que está adquiriendo el adecuado análisis de datos en todas las industrias.

No obstante, los algoritmos de inteligencia artificial y machine learning no son del todo recientes, el campo de la IA se remonta a mediados del siglo XX. Arthur Lee Samuels, un investigador de IBM, desarrolló uno de los primeros programas de aprendizaje automático: un programa de autoaprendizaje para jugar a las damas; él fue quien acuñó el término machine learning.

Machine Learning es una forma de inteligencia artificial que asiste en el desarrollo de análisis retrospectivo (enfoque descriptivo) y prospectivo (enfoque predictivo); ya que permite a un sistema generar información a partir de data histórica que se le entrega como input, en lugar de realizar una única programación explícita. Es así que esta herramienta está empezando a adquirir una importante popularidad entre organizaciones que buscan procedimientos innovadores para el manejo y análisis de sus bases de datos, lo que les podría ayudar a ganar cierto nivel de entendimiento sobre su contexto empresarial; y, con el adecuado modelamiento de datos, la capacidad de predecir los futuros cambios en dicho contexto, otorgándoles, de esta forma, una

ventaja competitiva. Asimismo, continuar entregándole información constantemente al modelo, es una forma de asegurar que la solución o el output que resulte sea el más actualizado posible.

Para el desarrollo del modelo, machine learning hace uso de una variedad de algoritmos que, de manera iterativa, generan la respuesta solicitada a partir de un previo “data training” o entrenamiento de datos. El “data training” consiste en entrenar al modelo con datos de entrada o “input” y con el resultado obtenido a partir de esos datos o “output”; una vez terminado el entrenamiento inicial, al proveer el input, el modelo debe devolver, en tiempo real, el output que se generaría; cabe resaltar que, a mayor entrenamiento hay más probabilidad de mayor precisión en el output.

Asimismo, los algoritmos estadísticos más complejos tienen la ventaja de que se pueden ajustar automáticamente basado en cambios rápidos en ciertas variables que se encuentran como input; por ejemplo, datos de sensores, tiempo meteorológico, métricas de sentimiento del cliente, variaciones horarias, entre otras. De esta forma un modelo predictivo del tiempo podría predecir un tornado horas antes de que ocurra y se podría activar una sirena de emergencias; ya que todo el proceso se realiza en tiempo real; no obstante, una limitante es la precisión en el resultado, que depende, en gran parte, del “data training”. Se elaboró la siguiente figura:

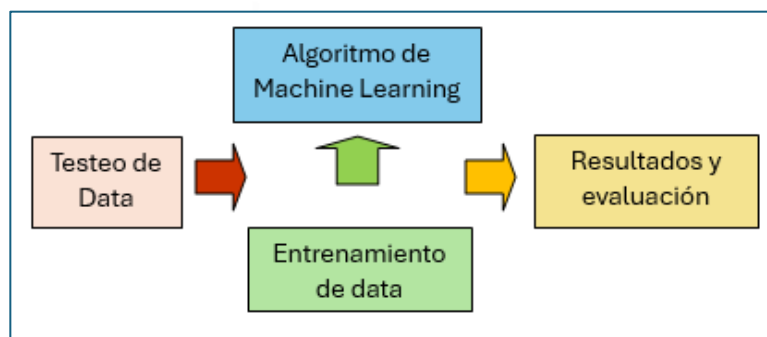


Figura 10. Flowchart de método de Machine Learning.

Fuente: Elaboración propia.

Existen 4 tipos de machine learning:

(1) **Supervised Learning:** Es el más utilizado, se llama de esa forma porque el desarrollador actúa como una guía para enseñar al algoritmo las conclusiones a las que debe llegar; es decir, el output sí es conocido y se incluye en el entrenamiento del modelo. Este es el tipo de machine learning del que se hará uso en la presente tesis.

(2) **Unsupervised Learning:** Está más alineado a la inteligencia artificial y busca que el aprendizaje se centre en la identificación de procesos y patrones complejos en un humano para proporcionar orientación; se basa en algoritmos de clustering y asociación, principalmente; no existe un conjunto de datos de entrenamiento y el output es desconocido, el modelo es guiado por operaciones lógicas.

(3) **Reinforcement Learning:** En este aprendizaje no hay una clave de respuesta, pero el algoritmo aún debe decidir cómo actuar para realizar su tarea, en ausencia de datos de capacitación, el algoritmo aprende de la experiencia; a través de ensayo y error.

(4) **Deep Learning (Neural Networks):** También llamado Aprendizaje profundo, utiliza los principios de los algoritmos básicos de machine learning en conjunto, intentando modelar abstracciones de alto nivel en datos. Este es el algoritmo de mayor precisión, alcance y complejidad; sin embargo, requiere de una mucho mayor cantidad de datos que los anteriores y de tecnología de muy alta calidad para desarrollar los algoritmos.

Los algoritmos son un conjunto de instrucciones que se le entrega a la computadora, en un lenguaje que esta pueda entender, que le indican cómo interactuar y/o transformar la data. Los más utilizados en machine learning supervisado son:

a) **Regression Algorithms:** Los algoritmos de regresión se utilizan comúnmente para análisis estadístico y son algoritmos clave para su uso en el aprendizaje automático. Ayudan a los analistas a modelar relaciones entre los datos y pueden cuantificar la fuerza de la correlación entre diferentes variables en un conjunto de datos. Sin embargo, este tipo de algoritmos asumen que la correlación se vincula con la causalidad, sin comprender el contexto en torno a los datos, es por ello que pueden llevar a predicciones inexactas. Entre ellos están los algoritmos. Linear Regression, Polynomial Regression, Nonlinear Multiple Regression, etc.

b) **Decision Tree Regression:** Utilizan una estructura de ramificación para ilustrar los resultados; se pueden utilizar para mapear los posibles outcomes al tomar una decisión, en el que cada nodo de un árbol de decisión o “decisión tree” representaría un posible resultado, los porcentajes se asignan a los nodos según la probabilidad de que ocurra cada resultado.

c) **Random Forest Regression:** Creado por Ho en 1995, este enfoque incorpora el algoritmo de agregación bootstrap, que es un método para generar múltiples versiones de un predictor y luego usarlos para obtener un predictor agregado.

d) **K-Nearest Neighbors (k-NN):** Es un método no paramétrico usado para clasificación y regresión; este algoritmo calcula la media de los valores en función a sus vecinos más cercanos.

e) **Support Vector Machine (SVM):** El original fue inventado por Cortes y Vapnik, su eficiencia en la clasificación y regresión ha sido verificada en muchos casos de estudio; consiste en la construcción de hiperplano o curvas que modelen la tendencia

de los datos para poder predecirlos en el futuro; en este caso, la curva resultado es la que mejor se adapta a los datos.

El ciclo de machine learning es continuo, elegir un adecuado algoritmo es solo uno de los pasos, el procedimiento para el desarrollo de un modelo de machine learning es el siguiente:

1. Identificar la data
2. Preparar la data
3. Seleccionar el algoritmo
4. Entrenar al modelo
5. Evaluar el desempeño del modelo
6. Implementar el modelo
7. Realizar la predicción
8. Evaluar la precisión de las predicciones

Finalizados estos pasos, es recomendable continuar actualizando al modelo constantemente, como se mencionó líneas arriba.

Como se mencionó anteriormente, es necesario que el lenguaje en el que se escriba el algoritmo sea comprendido por la máquina, es en este sentido que, a continuación, se describirá el lenguaje Python que se decidió utilizar para la presente tesis.

2.2.3. Análisis de regresión.

En la práctica existen situaciones en las cuales se observan un conjunto de variables, por consiguiente, resulta natural pensar en una relación entre ellas. El poder saber si existe una relación entre ellas o una forma funcional que pueda efectuar

predicciones sobre valores previsible de alguna otra variable, se puede realizar mediante un análisis de regresión o criterio estadístico. (Tusell F, 2011)

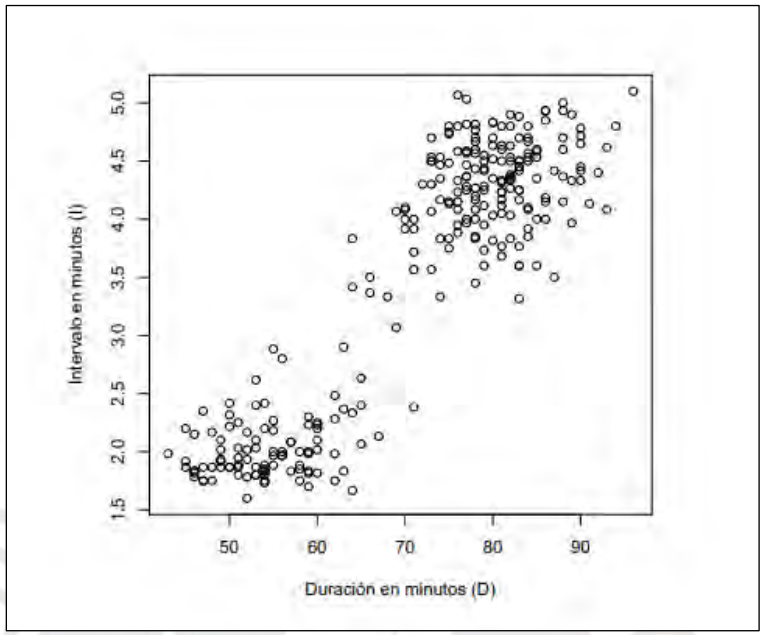


Figura 11. Recolección de datos de 272 erupciones volcánicas.

Fuente: "Tusell F, 2011".

A manera de ejemplo, se usará la Imagen 18, la cual recolectó información de un registro de erupciones volcánicas del geiser Old Faith-full en el parque nacional de Yellowstone (1982) y se plantea la expresión: $I = f(D) \dots(I)$, en dónde:

- I =Intervalo en minutos
- D =Duración en minutos

Debido a que resulta muy complicado desarrollar una función que relacione los 272 datos, se le agregará a la expresión (I), la cantidad aleatoria inobservable ϵ , finalmente se obtendría una función:

$$I = f(D) + \epsilon \dots \dots \dots (II)$$

Por lo cual se reconoce a $f(D)$ como una función de regresión de I sobre D y el objetivo final será determinar la forma de la función.

Modelo lineal generalizado (GLM)

Estos modelos se basan en que las propiedades básicas de los estimadores son insensibles al tipo de distribución asumida y que dependen del grado de independencia de las observaciones. La hipótesis que se asumen para un Modelo Lineal generalizado (GLM) son:

- Existe independencia entre las variables respuesta.
- La media cambia con las condiciones, pero no la forma funcional de la distribución.
- La media o una de sus transformaciones cambia de manera lineal cuando las condiciones varían.

Los GLM vienen expresados de la siguiente manera:

$$g(\mu_i) = y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i + \varepsilon$$

Donde:

$g(\mu_i)$: Función de enlace

y_i : Variable respuesta o dependiente

$x_{(1,2,3,\dots,i)}$: Variables explicativas o independientes

$\beta_{(1,2,3,\dots,i)}$: Coeficientes del modelo

ε : Error de predicción del modelo

Los GLM unifican los modelos en los cuales las variables dependientes mantienen una relación lineal con sus variables independientes y no se fuerzan transformaciones de las variables a escalas no naturales que sean complicadas de interpretar y por último

aplican un mismo algoritmo de estimación, los mínimos cuadrados ponderados iterados. (Martinez, A., & Morales J, 2001).

Se elaboró las siguientes tablas que muestran distribuciones que permiten el reajuste de los algoritmos anteriormente mencionados:

Tabla 2: Funciones de enlace.

Familia	Función de enlace
Gausiana	Identity
Binomial	Logit, Probit, cloglog
Poisson	Log, identity, sqrt
Gamma	Inverse, identity, log
Inverse gaussian	Sqrt inversa
Quasibinomial	Logit, Probit, cloglog
Quasipoisson	Log, identity, sqrt
Binomial Negativa	Log, identity, sqrt

Fuente: Elaboración propia.

Tabla 3: Tipo de distribuciones.

Nombre de la función de enlace	Función $\eta = g(\mu)$
identity	μ
sqrt	$\mu^{0.5}$
log	$\log_e(\mu)$
inverse	μ^{-1}
sqrt Inverse	μ^{-2}
logit	$\mu \log_e(1 - \mu)$
probit	$\Phi^{-1}(\mu)$
log-log	$-\log_e - \log_e(\mu)$
cloglog	$\log_e - \log_e(1 - \mu)$

Fuente: Elaboración propia.

Tabla 4: Comparación de los modelos.

Modelo Lineal General (LM)	Modelo Generalizado Lineal (GLM)
<ol style="list-style-type: none"> 1. Asume homogeneidad de varianza (residuales, errores homogéneos). 2. Los residuales (errores) deben ser estar distribuidos normalmente. Se espera que la variable respuesta tenga distribución normal, ya que esto condicionaría que los residuales- errores sean normales. 	<ol style="list-style-type: none"> 1. Incluye una serie de diferentes opciones para lidiar con la heterogeneidad de varianzas en los datos. 2. El investigador especifica qué distribución teórica sigue la variable respuesta, y por ende, los residuales-errores.

Fuente: Elaboración propia.

3. Metodología

La metodología por aplicar en la presente investigación fue la de Machine Learning Supervisado, que, como ya se mencionó en el marco teórico del presente, consiste en desarrollar una estructura de comandos o programa que pueda predecir los posibles escenarios a ocurrir en función del aprendizaje de resultados preestablecidos y logrados por pruebas en campo real realizadas en el pasado (experiencias pasadas).

Para la aplicación de esta, se ha dispuesto desarrollar el siguiente procedimiento estándar de Machine Learning Supervisado:

3.1. Recolección e identificación de datos

Como fase inicial del proceso de aplicación de Machine Learning, en esta fase se evaluará el desempeño del proceso que se estudiará para la aplicación futura (perforación, voladura, chancado, carguío, etc.) y las posibilidades de optimización de este mediante la predicción a través del programa a desarrollar.

Tras haber seleccionado el proceso a optimizar, se pasa a una etapa de identificación de data presente, en la cual se evaluará la cantidad de datos históricos del proceso con los que se cuenta actualmente y qué tantas variaciones presentan pues, en cuanto a los parámetros objetivo que deseamos predecir, requerimos que haya data con suficiente variación en resultados para poder lograr un programa predictivo que sea suficientemente completo. Por ejemplo, un conjunto de resultados muy parecidos solo logrará que el programa desarrollado cuente con un rango de evaluación muy corto.

Podemos ejemplificar con un programa que prediga qué días puede llover de acuerdo con las condiciones del clima, si la data que ingresamos para el aprendizaje del programa son solo

días en los que no ha llovido o, la gran mayoría son datos de este tipo, entonces nuestro programa no podrá predecir con facilidad resultados de días de lluvia, pues no es algo que se le haya enseñado consistentemente como resultados reales.

El programa debe asemejar el aprendizaje de un ser humano, pero con un entendimiento más rápido, sin embargo, si a una persona no le enseñas algo que deseas que resuelva, no lo logrará. Cabe allí la importancia de esta fase del proceso y resulta entonces necesario recolectar toda la data histórica posible. Así, la evaluación será más efectiva y existirán menores problemas hacia las siguientes fases del proceso.

3.2. Preparación de datos

Dentro de esta segunda fase del desarrollo del programa se tendrá como objetivo la selección de los tipos de parámetros, descripción general de la data histórica de estos y la selección de muestras de datos más apropiadas para el modelado.

La fase inicia con la extracción de conjuntos de datos por parámetro históricos recolectados y examinación de sus estructuras. Se analiza en caso algunos parámetros no tengan mucho que ver en la evaluación del proceso.

Para extraer un agrupamiento de datos, por parámetro, efectivo de la data recolectada, la estructura operativa del proceso debe analizarse e identificar cualquier cambio en las condiciones operativas de acuerdo con la variación de datos. La forma óptima de desarrollar esta evaluación es mediante la aplicación algoritmos de correlación entre parámetros presentes, ello pudiéndose desarrollar en distintos lenguajes de programación como R o Python. Estas correlaciones pueden ser darse por análisis gaussiano, relaciones lineales, relaciones no

lineales en varias variables, correlaciones en el tiempo, entre otras. Logrando la selección de parámetros y descripción general de datos que más impactan en los resultados del proceso.

Una vez seleccionada la data presente esta debe limpiarse de acuerdo al análisis de omisión de datos, puesto que pueden faltar datos en los agrupamientos por parámetros; y el análisis de valores atípicos o errores graves que deben eliminarse pues, en caso se ingresaran en el modelado, deteriorarían el aprendizaje y con ello la efectividad del programa predictivo.

En cuanto a la omisión de datos, ello suele pasar con frecuencia puesto que los administradores y los encargados del llenado de las bases de datos son seres humanos que pueden cometer errores al olvidarse de registrar ingresar algún dato, cometer errores ortográficos, o intencionalmente borrar ciertos datos que indiquen algún problema en los procesos efectuados. Asimismo, los encargados suelen rotar, no logrando una consistencia en la estructura de los datos que puede dificultar la aplicación de la metodología. Sin embargo, la omisión de datos puede ser resulta a través de un algoritmo de llenado de datos, ello pudiéndose aplicar aplicando lenguajes de R, Python, etc.

A través del llenado de datos, el ordenador logra completar los datos omisos con datos acorde a la correlación que se identifique con los demás llenados correctamente. Sin embargo, es recomendable que los datos omisos no representen la mayor proporción de la data de algún parámetro, puesto que puede inducirnos a errores.

Finalmente, los valores atípicos o errores graves pueden ser identificados por diagramas de cajas que nos permita identificar valores muy fuera del rango en cada parámetro. Siendo simplemente eliminados a mano y corriendo de nuevo el llenado de datos para terminar con una data limpia y lista para la aplicación de algoritmos.

3.3. Selección de algoritmos

Una vez que se logra conseguir una estructura de datos limpia, se procede a desarrollar la fase de selección de algoritmos de Machine Learning apropiados para la estructura de datos a evaluar.

Gracias al desarrollo de la fase de preparación de data, la estructura de datos ha sido descrita y caracterizada suficientemente para conocer la complejidad de la misma. Siendo así, la elección de los algoritmos a utilizar se fundamentará en la evaluación de los requerimientos para la aplicación de cada uno de ellos y la complejidad en la descripción de resultados que deseamos, recordando que el programa debe ser aplicable, no sólo por expertos, sino adicionarse como herramienta extra en el desarrollo del proceso.

De esta forma podemos preguntarnos:

- ¿Qué tipo de algoritmo de Machine Learning deberíamos emplear para el entrenamiento o aprendizaje del modelo?
- ¿Cuánta complejidad debería proporcionar la estructura del modelo para describir la información de los datos?
- ¿Es suficiente un modelo de estructura única?
- ¿Se necesita un modelo de estructura múltiple?

De acuerdo a la respuesta de las mismas, podemos elegir distintos algoritmos como, por ejemplo, algoritmos de regresión lineal simple o multivariable, en caso la relación entre parámetros sea lineal; algoritmos de aprendizaje no lineal como las redes neuronales artificiales o un modelo de regresión vectorial de soporte, en caso exista predominancia de

relaciones no lineales; o estructuras de modelo de regresión múltiple, en caso se hayan identificado variedad de condiciones operativas para distinguir el aprendizaje de datos.

Asimismo, aunque no existe un único método de selección de algoritmos, la comunidad científica proporciona varias reglas o criterios útiles que podrían considerarse para la selección, como el criterio de información de Akaike, el criterio de información bayesiano, el criterio de Hannan-Quinn, etc.

3.4. Entrenamiento del modelo

Una vez terminada la fase de selección de algoritmos, la fase de entrenamiento del modelo consistirá en la aplicación de los algoritmos de aprendizaje designados en el porcentaje de datos destinados al aprendizaje del modelo.

El porcentaje de datos de entrenamiento suele significar los dos tercios de la cantidad total de datos de la estructura, y servirán como experiencia de aprendizaje del programa. Este identificará los parámetros operativos y los resultados que se producen de acuerdo a lo mismos, identificando experiencia tras experiencia, estableciendo una correlación para identificar, hacia el futuro, los mismos parámetros y anticiparse a los resultados posibles de los mismos.

Siendo así, los dos tercios de datos serán evaluados por el programa. Es en esta etapa en donde se hace un filtro final de los parámetros a evaluar, puesto que, si existiese algún problema en la incidencia de los parámetros para el resultado, se identificará como problema en el ordenador. Así seleccionando los parámetros finales óptimos y obteniendo un programa preparado para una próxima predicción.

3.5. Evaluación de desempeño

Una vez terminada la fase de entrenamiento del modelo, corresponde seguir con la fase de evaluación de su desempeño en la predicción de resultados.

Hoy en día existen muchos métodos definidos para la validación del programa y la evaluación de su rendimiento como la validación cruzada, análisis de estabilidad del modelo, análisis de robustez, análisis de sensibilidad de parámetros, entre otros. Todos con la finalidad de evaluar la variación de datos predichos con respecto a los datos reales resultado de experiencias pasadas.

En este caso se aplicará el porcentaje de datos de validación, el cual corresponde al uso del 20% de datos para la comprobación de la efectividad del modelo.

Al evaluar los datos de validación, se ingresarán los parámetros incidentes al proceso en el programa, obteniendo los parámetros resultado predichos. Luego se hace un primer análisis de la variación con gráfica de correspondencia de valores predichos y resultados reales a fin de identificar una curva resultado de la relación cercana al tipo “ $x = y$ ”. Finalmente se procede con los métodos mencionados anteriormente.

3.6. Implementación del modelo

Esta fase corresponde a los pasos a seguir tras la demostración de la efectividad del método. Una vez desarrollado y habiendo logrado suficiente precisión para la predicción del programa, se debe implementar el modelo final en un proceso real de la industria a la cual va enfocado, en este caso la industria minera.

La implementación de un modelo de Machine Learning puede significar la necesidad de reformar un sistema de organización de base de datos, para mantener una data ordenada que

permita el desarrollo eficaz del programa, sin trabas de error a solucionar en cada momento; ubicación de instrumentos de medición en tiempo real para identificar parámetros y realizar las predicciones; entre otros que podrían tener la finalidad de automatizar más el proceso y volverlo más efectivo.

Asimismo, con el cambio de las condiciones operativas del proceso, la selección del modelo, la capacitación y las evaluaciones de desempeño deben llevarse a cabo periódicamente, a fin de mantener un modelo satisfactorio para los análisis.

3.7. Desarrollo de predicciones

Una vez implementado el modelo en el proceso productivo, pasa a la fase de desarrollo de las predicciones, en la cual el programa predictivo se pone en práctica constante durante la operación.

Es necesaria la supervisión de la primera etapa de aplicación del sistema a fin de observar que se desarrolle uniformemente el proceso. Al analizar, debemos tener en cuenta cada error que se pueda producir y que impida la predicción, algunos ejemplos pueden ser la inexactitud en el registro de los parámetros incidentes o la variación desbordada de los valores de estos, para los que podría no estar preparado el programa.

La evaluación permitirá volver a fases previas y afinar aún más el modelo, permitiendo optimizar el proceso cada vez más.

3.8. Evaluación de la precisión de predicción

Finalmente, aunque las fases pueden ser repetitivas, se dará la fase de evaluación de la precisión de predicción, en esta se evaluará y comparará los resultados predichos con respecto a los valores reales que vienen resultando del proceso.

Al analizar, podemos concluir si el programa cumple con todas las condiciones necesarias para su aplicación completa en la operación. Una vez validado el programa, se ajustarían los parámetros del proceso de acuerdo con las predicciones de este y así lograr los resultados optimizados que se busca.

Es importante señalar que, en caso aún no se cumpla con la efectividad deseada en esta etapa, se debe volver a las etapas pasadas y seguir afinando el modelo a fin de aumentar esta precisión, el proceso siempre puede ser mejorado.



4. Desarrollo

El presente estudio implementará la metodología de Machine Learning Supervisado para el desarrollo de un programa predictivo de los fenómenos de voladura: P80 y porcentaje de finos, producto de las labores de voladura en una Unidad Minera a tajo abierto que actualmente se encuentra operando en el país.

En este caso se plantea la aplicación de un algoritmo de regresión lineal multivariable para la cantidad de parámetros seleccionados en el estudio: UCS, burden, espaciamiento, factor de potencia, tiempos de retardo entre taladros, tiempos de retardo entre filas, P80 y porcentaje de finos. Dentro del desarrollo, se encontró que la correlación lineal es eficiente para este tipo de predicción.

Por razones de confidencialidad, el nombre de la unidad minera se mantendrá en reserva.

4.1. Definición del problema

Hoy en día, la operación minera en evaluación mantiene un registro de condiciones operativas y resultados de cada voladura que produjeron desde el año 2016 hasta el año 2020. Dentro de los mismos, se pueden apreciar distintas evaluaciones en cuanto a la empresa de explosivos a la que se encuentra encargada la supervisión de esta labor.

Siendo de esta forma, al efectuar la revisión de los datos se pudo constatar que la labor de voladura en la presente unidad minera no es uniforme, existen muchas variaciones en cuanto a los resultados e inclusive las condiciones de operación de voladura. Sin embargo, no se han presentado valores fuera de lo que permiten las normas de seguridad emitidas en los reglamentos del estado peruano ni han sobrepasado los estándares que se ha propuesto la empresa de explosivos a cargo de la operación.

Teniendo en cuenta ello, el modelo predictivo a desarrollar permitirá establecer una uniformidad mayor de los resultados de la voladura, logrando que la planta de procesamiento pueda mantener un ritmo de producción adecuado con entrega de material de alimento uniforme, así disminuyendo los problemas en el procesado del mineral logrando los objetivos planteados por el área de planeamiento de mina.

4.2. Desarrollo del programa

En cuanto al desarrollo del programa predictivo, se ha resuelto seguir el procedimiento estándar desarrollado en el capítulo de Metodología del presente estudio.

4.2.1. Recolección de datos.

Dentro de esta primera fase del desarrollo del programa, se analizó la operación minera en cuestión y se escogieron los procesos con posibilidad de optimización a través de la predicción de resultados anticipados. De esta manera se decidió la optimización de la labor de voladura puesto que es una de las labores que más impacta en el proceso minero. Sabiendo que es el primer tratamiento del mineral, debiendo lograr una granulometría de material adecuada, permitiendo el correcto desempeño de las demás labores de transporte y procesamiento de este. Es importante mencionar, además, que uno de los mayores costos de operación es el gasto de energía por trituración del material volado. Siendo así, resulta imperativo tener una adecuada granulometría post voladura. Una vez seleccionado el proceso que se evalúa en el presente estudio, se procedió a identificar toda la data posible de la labor de voladura de la unidad minera, obteniendo datos históricos de voladuras entre los años 2016 y 2020. Dentro de la misma, se hallaron variedad de problemas como datos omisos, faltas ortográficas, datos repetidos y, principalmente, desorden. Es en ese sentido que se decidió seleccionar la data mejor preparada para un desarrollo de programa de

Statistical Learning, el cual fue el análisis de fragmentación de las voladuras producidas. En esta se encontraron datas con menor número de fallas y una cantidad de datos suficiente para desarrollar un modelo completo. Los valores de granulometría resultante en cuanto a P80 y porcentaje de finos, se encontraban con suficiente variación y nos anticipaba un adecuado comportamiento para el desarrollo del modelo predictivo. Sin embargo, aún era necesario un trabajo de manejo de datos en la siguiente fase.

4.2.2. Preparación de datos.

En cuanto a la fase de preparación de datos, se cumplieron los objetivos de selección de tipos de parámetros, descripción general de la data histórica de estos y selección de muestras de datos más apropiadas para el modelado. En cuanto a la selección de los tipos de parámetros, se extrajeron los conjuntos de datos de todo el análisis de fragmentación presente entre los años que se recolectó. De esta manera, se obtuvo una estructura de datos que tomaba parámetros de fecha de disparo, fecha de toma de fotos, tajo en el cual se produjo la voladura, número de banco que se voló, número del proyecto de voladura, tipo de roca, tipo de material, malla, factor de potencia real, P80, porcentaje de finos, tiempos de retardo entre taladros y filas, los tipos de secuencia de salida, y el tipo de pala que recogió el material volado, etc. Se elaboró la siguiente tabla que muestra la estructura inicial de la data:

Tabla 5: Modelo de estructura inicial de datos extraída.

Fecha de Disparo	Toma de Fotos	Tajo	BANCO	Proy.	Tipo de Roca	Material	Malla	Factor Potencia Real (Kg/ton)	Pasante al 80% P-80 (pulg)	% Finos	TT(ms) / TF(ms)	Secuencia de salida	Pala
2-Ene	3-Ene	TS	3870	89	MZ DIOR	MINERAL	8.2 x 9.4	0.26	3.49	30.39%	7/150	V	61
3-Ene	4-Ene	TS	3720	27	MZ	DESMONTE	8.2 x 9.4	0.18	4.31	23.04%	7/120	V	61
4-Ene	5-Ene	TS	3690	8	MZ	MINERAL	8.2 x 9.4	0.21	3.43	57.19%	7/150	V	51
4-Ene	6-Ene	TS	3720	28	MZ	DESMONTE	8.2 x 9.4	0.21	3.41	47.03%	25/25	ECHELON	51
6-Ene	7-Ene	TS	3690	9	MZ	MINERAL	8.2 x 9.4	0.19	3.78	57.38%	7/120	V	62
7-Ene	8-Ene	TS	3870	94	MZ	MINERAL	8.2 x 9.4	0.24	3.54	26.67%	4/120	V	62
8-Ene	9-Ene	TS	3855	76	MZ	MINERAL	8.2 x 9.4	0.22	1.84	55.42%	7/120	V	Carg. 29
9-Ene	10-Ene	TS	3855	77	MZ	MINERAL	8.2 x 9.4	0.21	3.03	39.02%	7/140	V	Carg. 29
10-Ene	11-Ene	TS	3855	78	MZ	MINERAL	8.2 x 9.4	0.23	2.2	49.52%	7/140	V	60
10-Ene	12-Ene	TS	3855	78	MZ	MINERAL	8.2 x 9.4	0.23	2.65	30.41%	7/140	V	60
12-Ene	13-Ene	TS	3870	100	CZ	DESMONTE	8.2 x 9.4	0.24	4.98	13.16%	12/120	V	61
12-Ene	13-Ene	TS	3855	79	MZ	MINERAL	8.2 x 9.4	0.22	2.98	29.17%	7/120	ECHELON	60
13-Ene	14-Ene	TS	3855	80	MZ	MINERAL	8.2 x 9.4	0.22	4.39	33.32%	8/120	V	60
13-Ene	14-Ene	TS	3870	101	CZ	DESMONTE	8.2 x 9.4	0.20	4.96	19.36%	12/120	V	61
10-Ene	15-Ene		3705	27	MZ	MINERAL	8.2 x 9.4	0.17	3.88	36.84%	7/140	V - ECHELON	51
15-Ene	16-Ene	TS	3705	28	MZ	MINERAL	8.2 x 9.4	0.26	2.31	58.31%	13/120	ECHELON	51
17-Ene	18-Ene	TS	3705	29	MZ	MINERAL	8.2 x 9.4	0.15	2.84	41.24%	13/120	V	50
18-Ene	19-Ene	TS	3705	30	MZ	DESMONTE	8.2 x 9.4	0.18	3.25	34.57%	4/120	ECHELON	51
19-Ene	20-Ene	TS	3855	82	MZ	DESMONTE	8.2 x 9.4	0.20	3.84	28.95%	13/120	V	60
18-Ene	20-Ene	TS	3870	105	MZ	DESMONTE	8.2 x 9.4	0.20	4.76	23.73%	7/120	V	62
21-Ene	22-Ene	TS	3690	12	MZ	MINERAL	8.2 x 9.4	0.21	3.62	50.72%	7/120	V	50
21-Ene	22-Ene	TS	3855	86	MZ	MINERAL	8.2 x 9.4	0.17	3.03	43.26%	4/120	V	62
21-Ene	23-Ene	TS	3855	86	MZ	MINERAL	8.2 x 9.4	0.17	2.96	44.61%	4/120	V	62
4-Ene	24-Ene	TS	3690	14	MZ	MINERAL	8.2 x 9.4	0.17	2.98	34.92%	7/120	V	50
22-Ene	25-Ene	TS	3855	84	CZ	DESMONTE	8.2 x 9.4	0.24	5.33	16.86%	4/120	V	62
22-Ene	26-Ene	TS	3855	84	CZ	DESMONTE	8.2 x 9.4	0.24	5.01	20.52%	4/120	V	62

Fuente: Elaboración propia.

Tras seleccionar la estructura inicial de datos para el modelado del programa, resulta esencial efectuar un análisis de datos omitidos puesto que es muy probable que existan valores en blanco por errores humanos de registro.

El análisis de omisión de datos es posible de efectuar con ayuda de la aplicación de interfaces de programación con lenguaje en R o Python, por nombrar algunas. En este caso, se optó por el uso del lenguaje de programación R para el desarrollo. Siendo así, aplicamos un algoritmo “aggr” o “Aggregations for missing/imputed values”.

El algoritmo “aggr” propio de la librería VIM del lenguaje R es capaz de calcular o graficar la cantidad de valores omitidos en cada variable o parámetro, así como las combinaciones de estos tras ingresar la estructura de datos dentro de los valores de ingreso o argumentos de esta función o algoritmo.

Se realizó un filtrado de los parámetros a considerar de la data total extraída, de los cuales se obtuvo un total de 1401 datos de voladura, incluyendo filas en donde falte algún dato de voladura, luego los parámetros que entrarán a un análisis previo para el

modelo serán: burden, espaciamento, resistencia, factor de potencia, tiempos de retardo entre taladros, densidad de material, densidad de explosivo final, densidad, P80 y porcentaje de finos.

Tabla 6: Modelo de estructura inicial de datos seleccionada.

BURDEN (m)	ESPACIAMIENTO (m)	RESISTENCIA	FP-REAL (kg/TN)	T-TALADROS (ms)	DENSIDAD MATERIAL (g/cm ³)	DENSIDAD EXPLOSIVO FINAL (g/cm ³)	Pasante al 80 % P-80 (µg)	%Finos
8.20	9.40	R3-R4	0.26	7.00	2.60	1.13	3.49	30%
8.20	9.40	R4	0.21	7.00	2.61	1.13	3.43	57%
8.20	9.40	R4	0.19	7.00	2.61	1.14	3.78	57%
8.20	9.40	R3-R4	0.24	4.00	2.64	1.13	3.54	27%
8.20	9.40	R4	0.22	7.00	2.59	1.13	1.84	55%
8.20	9.40	R4	0.21	7.00	2.60	1.12	3.03	39%
8.20	9.40	R4	0.23	7.00	2.64	1.10	2.20	50%
8.20	9.40	R3	0.23	7.00	2.63	1.10	2.65	30%
8.20	9.40	R3	0.22	7.00	2.60	1.13	2.98	29%
8.20	9.40	R3	0.22	8.00	2.58	1.14	4.39	33%
8.20	9.40		0.17	7.00			3.88	37%
8.20	9.40	R3	0.26	13.00	2.60	1.14	2.31	58%
8.20	9.40	R3	0.15	13.00	2.62	1.13	2.84	41%
8.20	9.40	R3	0.21	7.00	2.61	1.14	3.62	54%
8.20	9.40	R3	0.17	4.00	2.60	1.12	3.03	43%
8.20	9.40	R3	0.17	4.00	2.60	1.12	2.96	45%
8.20	9.40	R3-R4	0.17	7.00	2.62	1.12	2.98	35%
8.20	9.40	R3	0.21	8.00	2.60	1.13	2.15	55%
8.20	9.40	R3	0.23	7.00	2.62	1.11	2.32	58%
8.20	9.40	R3	0.21	7.00	2.60	1.13	2.59	54%
8.20	9.40	R3	0.23	7.00	2.60	1.13	2.54	45%
8.20	9.40	R3	0.24	7.00	2.62	1.12	2.96	51%
8.20	9.40	R3	0.24	7.00	2.62	1.12	2.84	50%
8.20	9.40	R3	0.24	7.00	2.62	1.12	4.11	28%
8.20	9.40	R3	0.24	7.00	2.62	1.12	3.14	37%
8.20	9.40	R3	0.23	7.00	2.58	1.13	4.62	32%
8.20	9.40	R3	0.21	4.00	2.66	1.13	2.53	40%
8.20	9.40	R3-R4	0.24	4.00	2.62	1.12	2.85	53%
8.20	9.40	R3	0.19	7.00	2.59	1.13	2.31	51%
8.20	9.40	R3	0.19	7.00	2.59	1.13	2.53	43%
8.20	9.40	R3	0.22	4.00	2.60	1.13	2.12	50%
8.20	9.40	R3	0.20	4.00	2.60	1.12	2.72	31%
8.20	9.40	R3	0.30	4.00	2.60	1.13	2.73	38%

Fuente: Elaboración propia.

Una vez seleccionadas las variables que poseen las mejores características para el modelo predictivo, procederemos a verificar su variabilidad con un histograma con la finalidad de que brinden una estadística necesaria y suficiente para utilizarlas en el desarrollo del modelo.

NOTA: En el caso de la variable resistencia se filtraron los valores con menor frecuencia y que solo harían perder explicabilidad al modelo, por lo cual de los 12 tipos de resistencias encontrados en la data inicial, se decidió trabajar con solo 3 tipos de resistencias: R3; R3-R4 y R4 que juntos conforman el 88% del total de casos de resistencia. El análisis se desarrollará tomando en cuenta solo a esos tres tipos de resistencias; por consiguiente, la cantidad de filas completas finales, con la que se trabajó será de 1233 datos de voladuras.

Una vez que se limpió la gran cantidad de casos de resistencias y valores nulos en las filas se generó los histogramas de cada variable, obteniendo como resultado:

Burden(m): Se registraron en total solo 6 medidas de burden para los 1233 registros del total, siendo la medida de 7.2m el 41% del total de datos, por consiguiente, la poca variabilidad de este dato burden, no favorece al modelo en cuanto a su explicabilidad; por ello, se decidió no usarlo para el desarrollo del modelo predictivo. Se elaboró el siguiente histograma con su cuadro resumen.

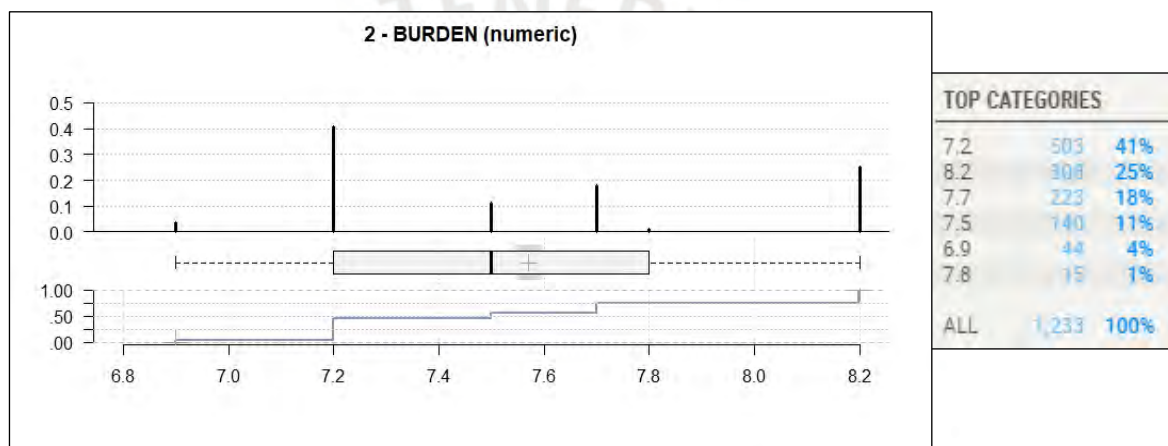


Figura 12. Histograma de distribución de datos de burden(m).

Fuente: Elaboración propia.

Espaciamiento(m): Similar al caso anterior, se registraron en total solo 7 medidas de espaciamento para los 1233 registros del total, siendo la medida de 8.3m el 41% del total de datos, por consiguiente, la poca variabilidad de este dato espaciamento, no favorece al modelo en cuanto a su explicabilidad; por ello, se decidió no usarlo para el desarrollo del modelo predictivo. Se elaboró el siguiente histograma con su cuadro resumen.

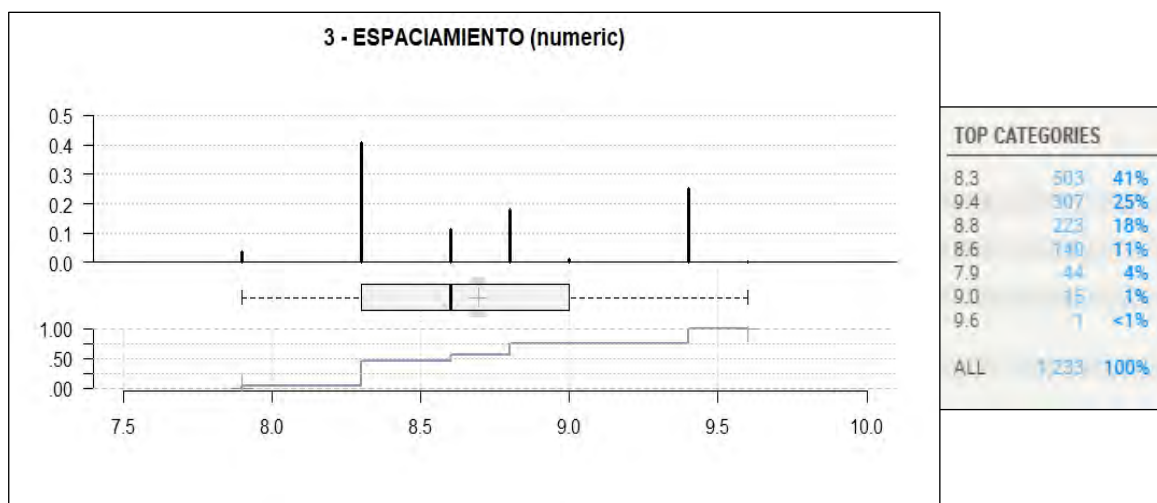


Figura 13. Histograma de distribución de datos de espaciamiento(m). Fuente: Elaboración propia.

Resistencia (m): Para el caso de esta variable categórica, se resumirá los tipos de resistencia a solo 3 en la data final, los cuales son R3, R3-R4 y R4; siendo la roca más suave R3, la más dura R4 y el tipo R3-R4 un intermedio entre ambas resistencias. Todas las resistencias están basadas en la cantidad de MPa que resisten las rocas a la compresión uniaxial (UCS). Siendo la de mayor frecuencia el tipo R4 con 597 casos, representando el 48% del total de casos. Se elaboró el siguiente histograma con su cuadro resumen.

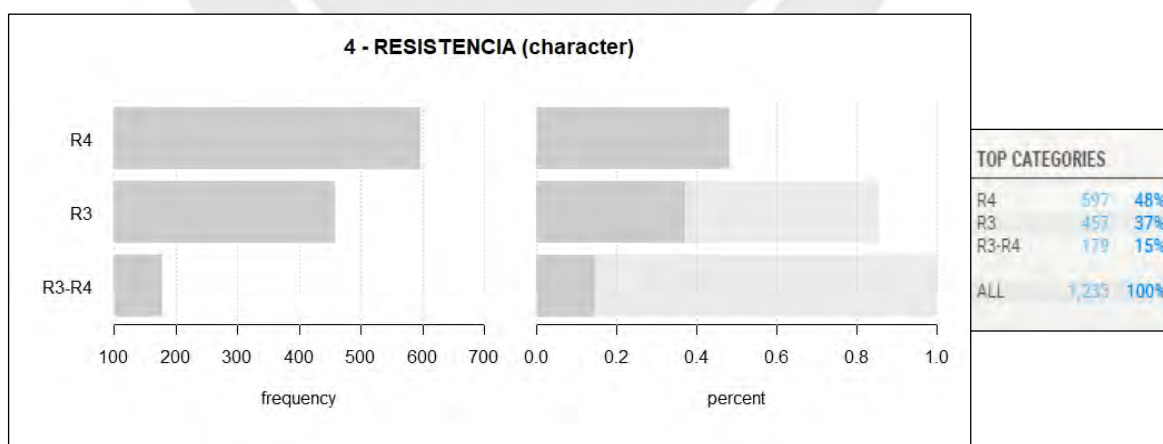


Figura 14. Histograma de distribución de datos de resistencia.

Fuente: Elaboración propia.

Factor de Potencia (kg/TN): Para esta variable Factor de potencia se registraron los 1233 casos, en dónde el registro mínimo es 0.124kg/TN y el máximo 0.513kg/TN, dando como valor promedio 0.303kg/TN; se aprecia una curva dónde los datos se encuentran distribuidos a largo del todo el eje y por ello una variabilidad de datos considerable, por consecuencia la variable factor de potencia formará parte del modelo predictivo. Se elaboró el siguiente histograma con su cuadro resumen.

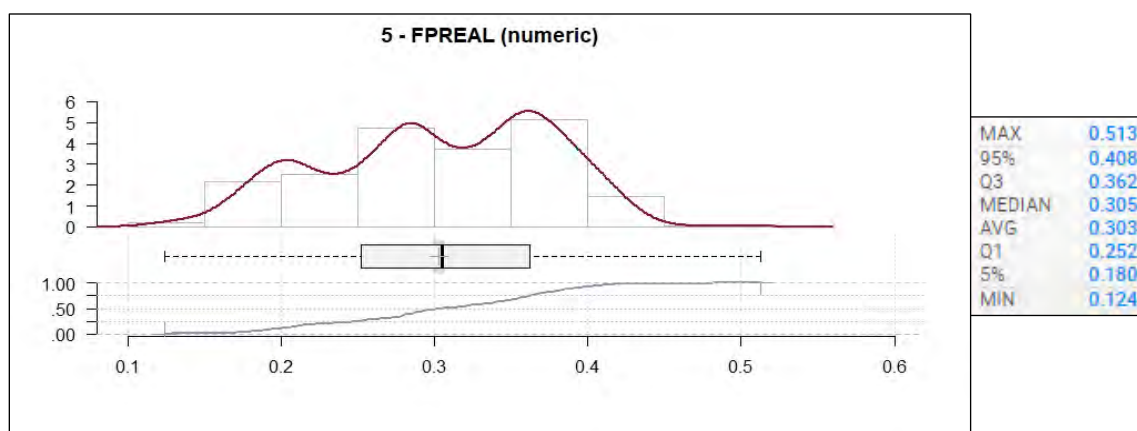


Figura 15. Histograma de distribución de datos de factor de potencia (kg/TN).

Fuente: Elaboración propia.

Tiempo de retardo entre taladros (ms): En el caso de la variable tiempo entre taladros, se registraron 1233 casos, en dónde el tiempo mínimo es 3ms y el mayor 73ms, siendo en total 33 casos diferentes de tiempos; se aprecia una curva bimodal con datos distribuidos a lo largo del eje, por ello una variabilidad de datos considerable, por consecuencia la variable tiempo entre taladros formará parte del modelo predictivo. Se elaboró el siguiente histograma con su cuadro resumen.

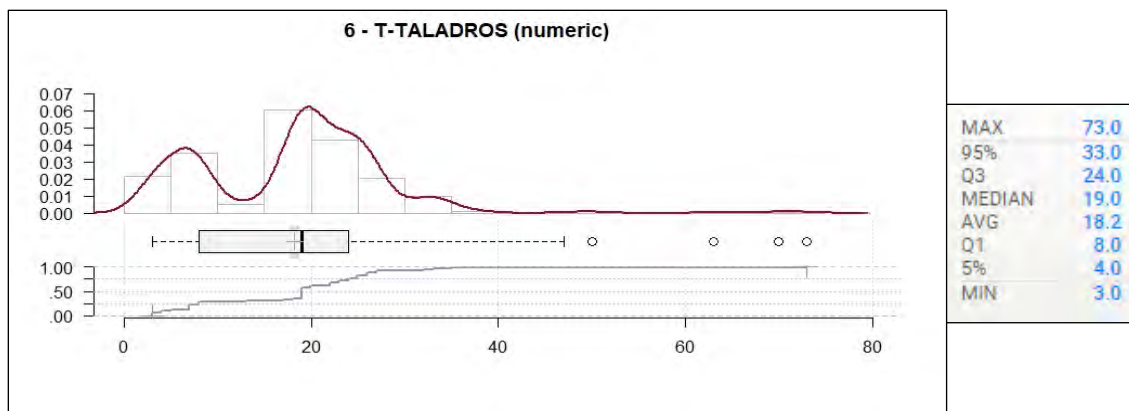


Figura 16. Histograma de distribución de datos de tiempo entre taladros(ms).

Fuente: Elaboración propia.

Densidad de material (gr/cm^3): La variable densidad del material presenta un valor mínimo de $1.72 \text{ gr}/\text{cm}^3$ y un máximo de $3.81 \text{ gr}/\text{cm}^3$, con un promedio de $2.61 \text{ gr}/\text{cm}^3$. Sin embargo, este valor máximo se debe a errores en el llenado de datos, probablemente causados por un error humano, ya que no es posible que la densidad del macizo rocoso alcance valores tan elevados. La mayor parte de los datos se agrupa en torno a los $2.6 \text{ gr}/\text{cm}^3$, reflejando la tendencia real de la variable. Aunque presenta cierta variabilidad debido a diferencias en los decimales registrados, no se consideró en el modelo multivariado predictivo final, ya que se determinó que no era una variable significativa para el análisis. A continuación, se presenta el histograma junto con su cuadro resumen.

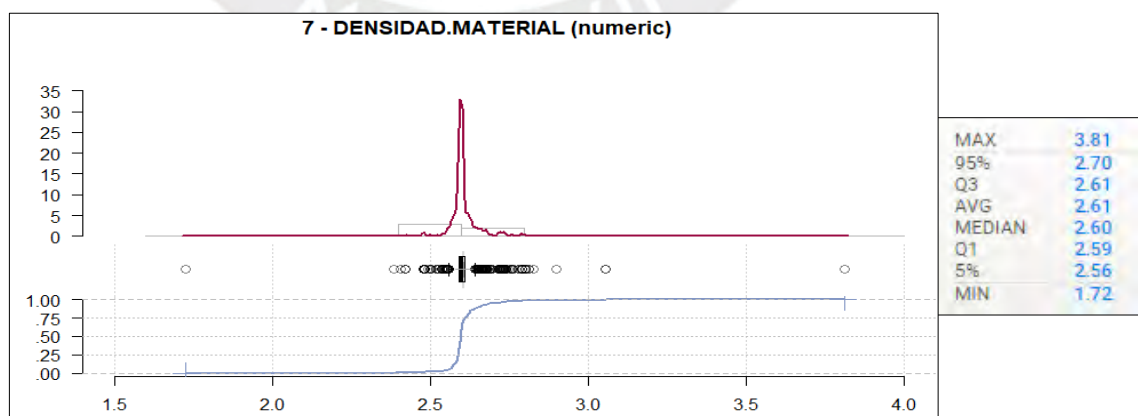


Figura 17. Histograma de distribución de datos de densidad de material(gr/cm^3).

Fuente: Elaboración propia.

Densidad de explosivo final (gr/cm³): La variable densidad de explosivo final posee como valor mínimo 1.087 gr/cm³ y como máximo 1.19 gr/cm³, dando un promedio de 1.125 gr/cm³. Esta variable presenta variabilidad, puesto que poseen cantidades de decimales diferentes en los registros que la hacen diferenciarse, tal como se aprecia en el gráfico que los datos se reparten a lo largo del eje, por lo que califica como una variable óptima para el modelo predictivo. Se elaboró el siguiente histograma con su cuadro resumen.

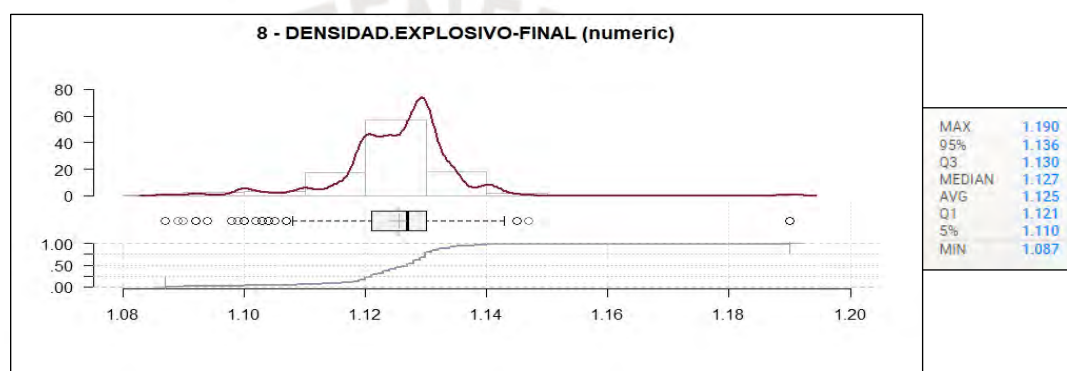


Figura 18. Histograma de distribución de datos de densidad de explosivo final (gr/cm³).

Fuente: Elaboración propia.

Relación nitrato (%): Esta variable resultó ser 29.1% para la totalidad de los datos registrados por lo cual no representa un dato que proporcione explicabilidad al modelo predictivo, por lo tanto, no formará parte del modelo. Se elaboró el siguiente histograma.

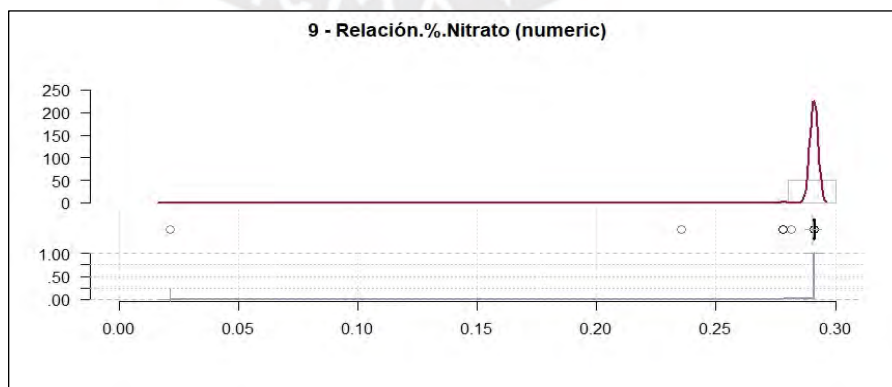


Figura 19. Histograma de distribución de datos de relación nitrato (%). Fuente: Elaboración propia.

Relación nitrito (%): Esta variable resultó ser 0.3% para casi la totalidad de los datos registrados por lo cual no representa un dato que proporcione explicabilidad al modelo predictivo, por lo tanto, no formará parte del modelo. Se elaboró el siguiente histograma.

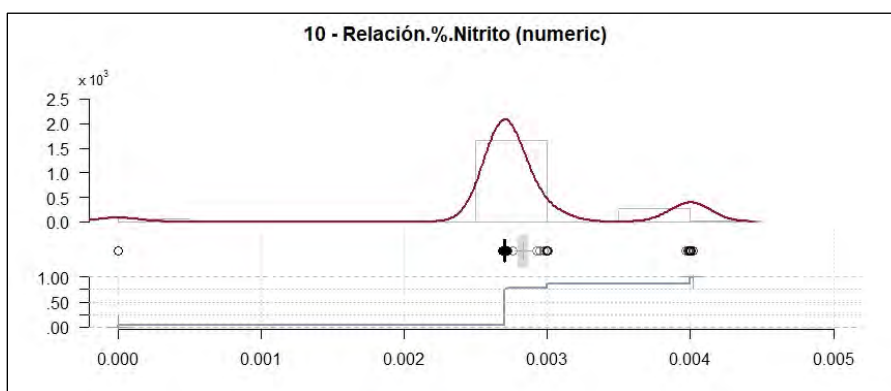


Figura 20. Histograma de distribución de datos de relación nitrito (%). Fuente: Elaboración propia.

Relación emulsión (%): Esta variable resultó ser 70% para la totalidad de los datos registrados por lo cual no representa un dato que proporcione explicabilidad al modelo predictivo, por lo tanto, no formará parte del modelo. Se elaboró el siguiente histograma.

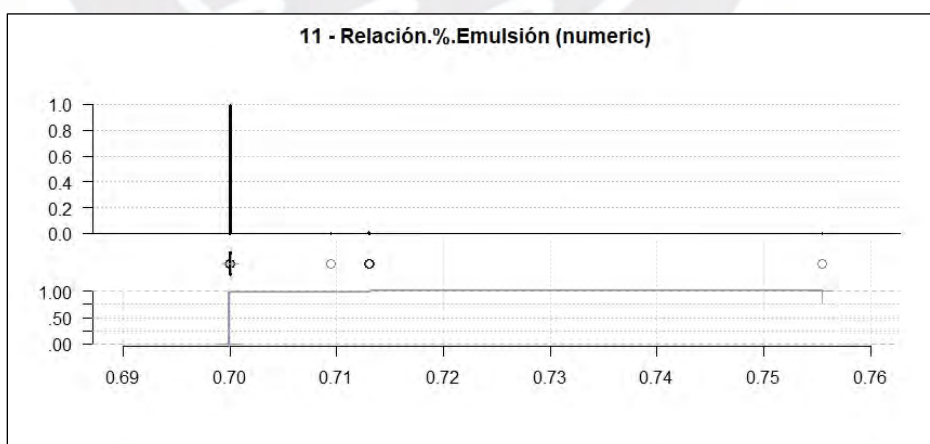


Figura 21. Histograma de distribución de datos de relación emulsión (%).

Fuente: Elaboración propia.

P80 (pulg): Para esta variable P80 se registraron los 1233 casos, en donde el registro mínimo es 1.46 pulg y el máximo 5.37 pulg, dando como valor promedio 2.69 pulg; se aprecia una curva donde los datos se encuentran distribuidos a largo de todo el eje y por ello una variabilidad de datos considerable y robusta, por consecuencia la variable P80 califica para ser parte del modelo predictivo. Se elaboró el siguiente histograma.

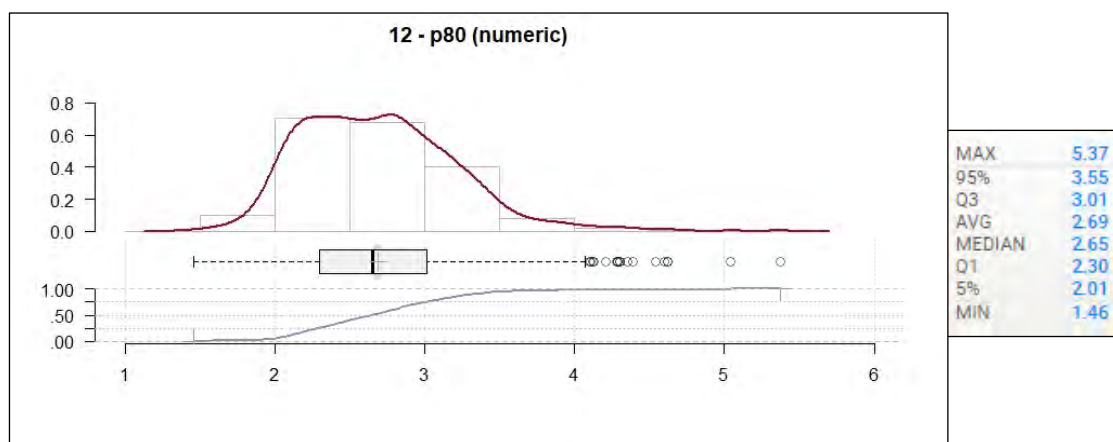


Figura 22. Histograma de distribución de datos de densidad de P80. Fuente: Elaboración propia.

Porcentaje de finos (%): En el caso de la variable porcentaje de finos, se registraron los 1233 casos, en donde el valor mínimo es 17.8% y el máximo 72.5%, dando como valor promedio 49.8%; se aprecia una curva donde los datos se encuentran acumulados y cercanos al valor 50%; sin embargo, a pesar de ello, presenta variabilidad de datos considerable y robusta en cuanto a sus decimales que le permite calificar como variable óptima para el modelo predictivo.

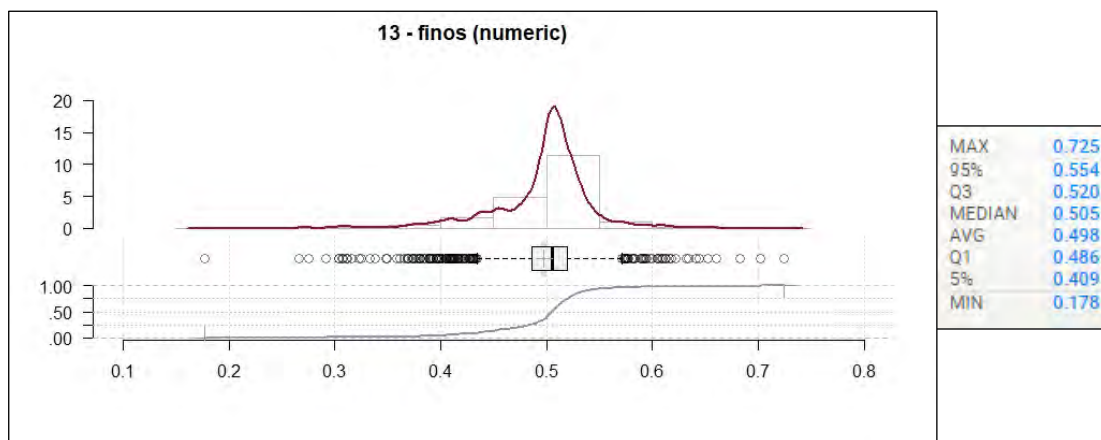


Figura 23. Histograma de distribución de datos de porcentaje de finos (%).

Fuente: Elaboración propia.

Según los histogramas mostrados se define que el burden y espaciamento no formarán parte del modelo predictivo en una evaluación inicial, pues no presenta suficiente variabilidad y no son representativas para el modelo, el resto de las variables pasarán a la siguiente etapa del proceso; no obstante, en un análisis posterior se tomará en cuenta la variable burden. Se muestra a continuación el resumen de las variables evaluadas y la cantidad de vacíos para cada tipo de variable, el cual es cero.

Tabla 7: Tabla resumen de las variables.

```

Describe datos (data.frame):

data frame:      1233 obs. of 13 variables
                  1233 complete cases (100.0%)

  Nr  ColName                Class  NAs  Levels
  --  ---
  1   UCS                    numeric .
  2   BURDEN                  numeric .
  3   ESPACIAMIENTO          numeric .
  4   RESISTENCIA            character .
  5   FPREAL                  numeric .
  6   T-TALADROS             numeric .
  7   DENSIDAD.MATERIAL      numeric .
  8   DENSIDAD.EXPLOSIVO-FINAL numeric .
  9   Relación.%.Nitrato     numeric .
 10  Relación.%.Nitrito     numeric .
 11  Relación.%.Emulsión    numeric .
 12  p80                     numeric .
 13  finos                   numeric .

```

Fuente: Elaboración propia.

Análisis de datos Atípicos: Solo se retiró un valor de dato atípico de 1.2kg/TN, el cual se encontraba en la columna de Factor de Potencia, pues este se encontraba aislado del resto de datos y una vez descartado, quedó como resultado final los 1233 casos mostrados. Nota: No se aplicó los BoxPlot, ya que las dimensiones de las cajas eran muy reducidas, generando falsos valores atípicos que no aportarían al modelo.

4.2.3. Selección de algoritmos.

Como se mencionó anteriormente, existen 2 propiedades esenciales en lo que se refiere a la selección de los algoritmos: la explicabilidad y la precisión. La primera hace referencia a la simplicidad del algoritmo y la facilidad con la que la influencia de cada uno de los parámetros y variables explicativas interactúa con la variable respuesta; estos algoritmos suelen tener menos precisión; ya que le dan más importancia a explicar el por qué, gracias a eso, estos algoritmos nos ayudan más a entender qué cambios se deben realizar en cada parámetro para obtener diferentes resultados. Por otro lado, existen parámetros que le dan mayor importancia a la precisión; esto quiere decir que no le dan mucha importancia al cómo, si no al resultado en sí; es decir, no resulta clara la relación de interacción que establecen entre las variables explicativas y la variable respuesta; pero suelen ser aquellas cuyos modelos predictivos se acercan más a la realidad. A continuación, se presenta un gráfico que nos permite observar el nivel de precisión y explicabilidad de cada algoritmo

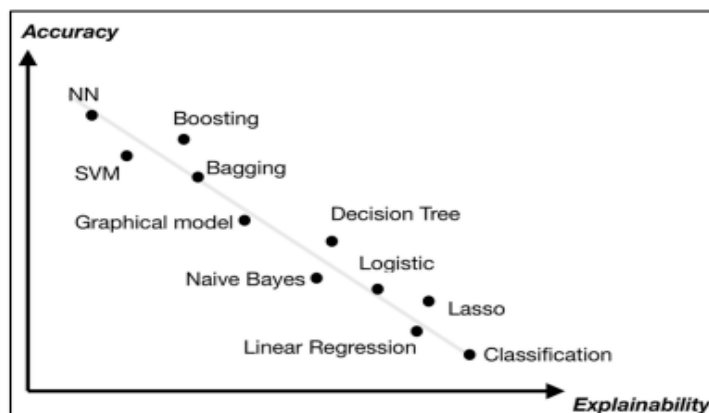


Figura 24. Explicabilidad vs precisión de algoritmos de Machine Learning

Fuente: "Explainable Artificial Intelligence" (2019).

En primera instancia se optó por priorizar la explicabilidad del modelo, por lo cual se decidió realizar modelos de regresión, siendo el de tipo lineal el primero en ser probado para la variable P80; sin embargo, al realizar las pruebas con la data, se encontraron los siguientes resultados:

- No existe normalidad entre los valores residuales del modelo lineal.
- No hay homogeneidad de varianzas en los residuales o errores del modelo lineal.

Para la evaluación de las variables respuestas se realizó un Normality Diagnosis Plot de cada una de ellas, utilizando el método de Anderson Darling debido a que la gran cantidad de data hacía imposible el uso del Shapiro Test.

Cabe resaltar que, un residual es una medida del error que se comete al estimar la variable dependiente (variable respuesta). Se calcula como la diferencia que existe entre el valor observado de la variable dependiente y el valor estimado. Por lo tanto, los residuales indican cual es la parte de Y que no está explicada por el modelo que se estima.

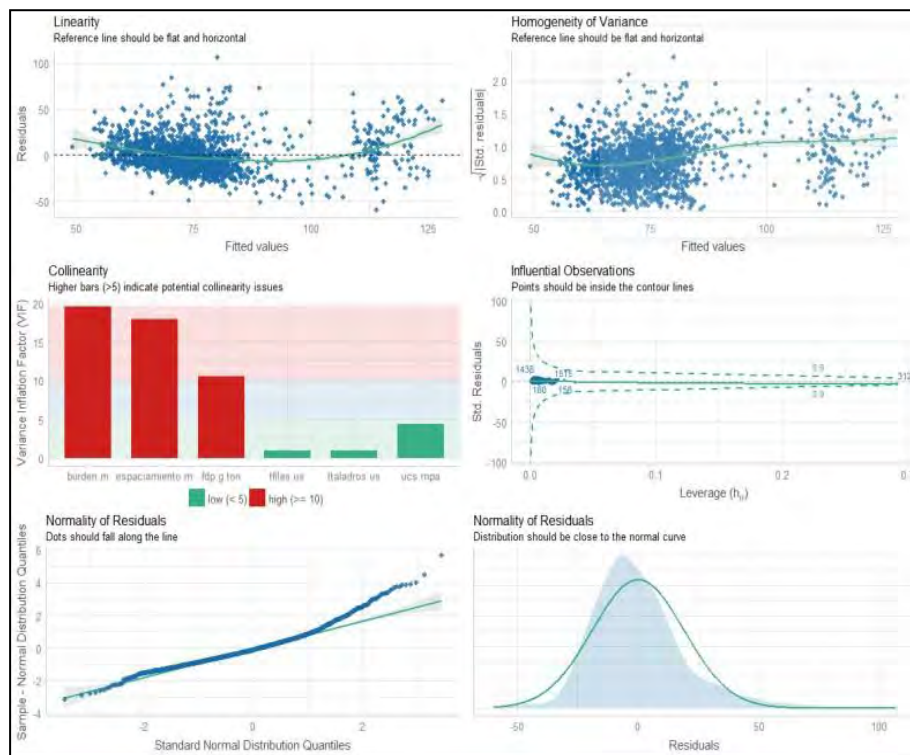


Figura 25. Gráfico de diagnóstico de normalidad – P80.

Fuente: Elaboración propia



Figura 26. Gráfico de diagnóstico de normalidad – Porcentaje de finos. Fuente: Elaboración propia

Los gráficos nos permitieron evaluar la linealidad, colinealidad, homocedasticidad y normalidad de los residuales al evaluar esta variable respuesta. En estadística, se dice que un modelo predictivo presenta homocedasticidad cuando la varianza del error condicional a las variables explicativas es constante a lo largo de las observaciones; es decir, que mientras el gráfico de “Homogeneidad de Varianza” presente menos sesgo, las probabilidades de que el modelo predictivo sea más efectivo son más altas; asimismo la linealidad es útil para la selección del algoritmo a utilizar; pues mientras más “flat” sea la curva de “Linealidad”; las funciones de la familia “Identity” previamente especificadas son más eficaces. Por otro lado, la colinealidad es el problema que se presenta cuando, en un análisis de regresión, alguna variable independiente o explicativa resulta de la combinación de las otras dos; este es un problema complejo, porque al presentar algún grado de correlación, el sistema de ecuaciones normales podría tener infinitas soluciones. En estos casos, se podría optar por disminuir el número de dimensiones del modelo. Finalmente, otra de las propiedades que se pudieron analizar con estos gráficos (los dos últimos) fue la normalidad de los residuales de cada una de las variables respuesta; en general, para asegurar que los residuales presentan una distribución normal, en el Q-Q plot o “Standard Normal Distribution Quantiles”, los puntos azules deben acercarse lo más posible a la recta verde; asimismo, en la curva de densidad la sombra que representa a los residuales de la variable debe asemejarse a la curva verde que representa la distribución normal; si la sombra se sesga hacia la izquierda o derecha, a ese comportamiento se le conoce como “skewness” y si la sombra se extiende por arriba de la curva a eso se le conoce como “kurtosis”. A partir de los gráficos de P80 y

porcentaje de finos; se puede observar que los residuales no presentan una distribución normal para ninguno de los casos.

Debido a que no se cumplen las premisas para optar por un modelo lineal, se procede a realizar el análisis con Modelos Lineales Generalizados (GLM).

MODELO DE REGRESIÓN GAMMA - P80-

Se optó por trabajar con un modelo Gamma, debido a que la variable que se quiere explicar, la variable P80, pertenece al rango mayor que cero hasta el más infinito. Por lo cual se realizó análisis univariados y algunos con interacciones, para cada variable explicativa del modelo.

Teóricamente, la distribución Gamma debería ser la opción correcta cuando la variable dependiente tiene un valor real en un rango de 0 a ∞ . Y la Gamma es adecuada cuando se sospecha que el vínculo entre la media y la varianza es “fijo”. Si espera un valor pequeño de y_i , también debe esperar solo una pequeña cantidad de variabilidad en los valores observados. Por el contrario, si espera un valor enorme de y_i , debería esperar mucha variabilidad.

Una variable aleatoria Y sigue una distribución gamma si su función de densidad de probabilidad está dada por

$$f(y|\alpha, \lambda) = \frac{\lambda^\alpha y^{\alpha-1} e^{-\lambda y}}{\Gamma(\alpha)} I(0, \infty)(y) \quad (1)$$

donde $\alpha, \lambda > 0$, $\Gamma(\cdot)$ denota la función gamma, e $I(\cdot)$ es la función indicadora. Bajo esta parametrización, la media y la varianza de Y vienen dadas por $E(Y) = \alpha/\lambda$ y $\text{Var}(Y) = \alpha/\lambda^2 = \mu^2/\alpha$. Dado que $\lambda = \alpha/\mu$, la función de distribución gamma (1) puede ser reparametrizada en función de los parámetros media (μ) y forma (α), como se propone

en Cepeda-Cuervo (2001) y Cepeda-Cuervo y Gammerman (2005), donde la función de distribución gamma se escribe como en la ecuación (2).

$$f(y|\mu, \alpha) = \frac{1}{y\Gamma(\alpha)} \left(\frac{\alpha y}{\mu}\right)^\alpha e^{-\alpha y/\mu} I(0, \infty)(y) \quad (2)$$

donde $\mu, \alpha > 0$, $\Gamma(\cdot)$ denota la función gamma e $I(\cdot)$ es la función indicadora. La notación $Y \sim G(\mu, \alpha)$ se utiliza para denotar que Y sigue una distribución gamma con $E(Y) = \mu$ y α como parámetro de forma.

Sea $Y_i \sim G(\mu_i, \alpha)$, $i = 1, \dots, n$, una muestra aleatoria de tamaño n . En los modelos de regresión gamma con parámetro de forma constante, la estructura de regresión media se define por

$$g(\mu_i) = \eta_i = x_i' \beta$$

donde g es la función de enlace, $\beta = (\beta_0, \dots, \beta_p)'$ es el vector de los parámetros de regresión media, x_i es el valor del i -ésimo vector de las variables explicativas y η_i es un predictor lineal. Aquí, $g(\cdot): (0, \infty) \rightarrow \mathbb{R}$ es una función de valor real, estrictamente monótona y doblemente diferenciable (Moreno. G, 2021). Algunas funciones de enlace promedio usuales en la regresión gamma son: la función logarítmica, $g(\mu) = \log(\mu)$; la función de identidad, $g(\mu) = \mu$, y la función inversa $g(\mu) = 1/\mu$. En los modelos lineales generalizados (GLM), el enlace canónico de la media es la función inversa.

MODELO DE REGRESION BETA – PORCENTAJE DE FINOS:

Los modelos de regresión comúnmente tienen por objetivo el modelar y analizar las relaciones entre una variable dependiente (y) con otras variables independientes (x_1, x_2, \dots, x_i) en un conjunto de observaciones dado. Por lo cual el modelo lineal es el empleado con frecuencia; sin embargo, no siempre resulta el modelo más ajustado o

apropiado para situaciones en las cuales la variable respuesta o dependiente se encuentra restringida en el intervalo $(0 < y < 1)$ siendo un valor continuo. (Real. A, 2015)

Por lo cual surge la regresión beta que basa su aproximación con apoyo del GLM y una distribución tipo beta de la variable dependiente y que su media esté relacionada con un conjunto de regresores a través de un predictor lineal con coeficientes desconocidos y una función de enlace, lo cual implica proporciones y razones dónde los valores se encuentren en el rango de $(0 < y < 1)$. (Cribari-Neto & Zeileis, 2010)

Se dice que la variable aleatoria y tiene una distribución de probabilidad beta con parámetros $\alpha > 0$ y $\beta > 0$ si y solo si la función de densidad de y cumple:

$$f(x) = \begin{cases} \frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha, \beta)}, & 0 \leq y \leq 1 \\ 0, & \text{en cualquier otro punto} \end{cases}$$

Donde

$$B(\alpha, \beta) = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1} dy = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}$$

En base a los diferentes parámetros $\alpha > 0$ y $\beta > 0$ que se puedan definir, se crean diferentes gráficos.

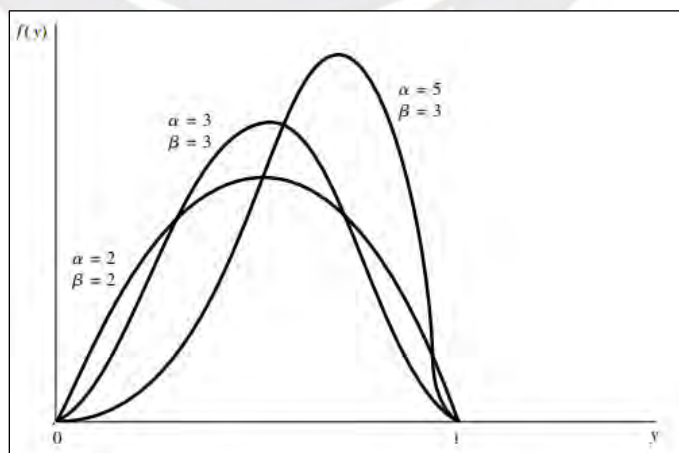


Figura 27. Funciones de densidad beta. Fuente: Wackerly, Mendenhall y Scheaffer (2010, p. 195).

Finalmente, una vez definido los parámetros, se procede aplicar la regresión beta, dando gráficos como el mostrado a continuación.

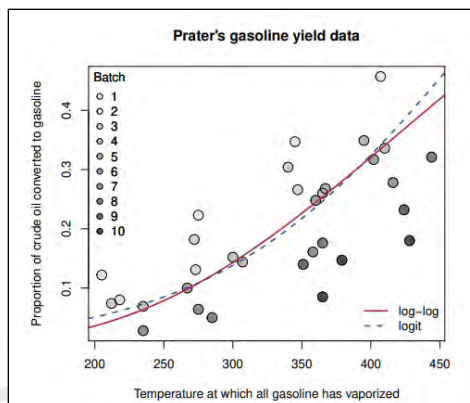


Figura 28. Ejemplo de regresión beta con funciones de enlace log-log, logit.

Fuente: (Cribari-Neto & Zeileis, 2010)

NOTA: Observar que en los resultados modelados para la variable dependiente y solo varía entre 0 y 1.

ANÁLISIS DE VARIABLES EXPLICATIVAS – P80

Tiempo de retardo entre taladro (ms) vs P80 (pulg)

Como primer paso debemos corregir la normalidad aplicando el modelo de regresión GLM gamma con la función de enlace logaritmo, una vez aplicado el modelo de tipo gamma se revisa si la variable Tiempo entre taladro posee explicabilidad respecto a la variable P80.

Tabla 8: Tabla de resultados del modelo univariado T-taladros vs P80.

Predictors	Estimates	p80	
		CI	p
(Intercept)	2.5528	2.4963 – 2.6106	<0.001
T-TALADROS	1.0029	1.0018 – 1.0040	<0.001
Observations	1233		

Fuente: Elaboración propia.

Según los resultados de la tabla, se observa que el p-valor de la prueba de ratio de verosimilitudes (LTR, siglas en inglés: Likelihood Ratio Test) del modelo es inferior a 0.05, el cual es el máximo valor como nivel de significación y por lo que se deduce que el modelo sí es significativo. Adicionalmente, según los valores obtenidos se espera que por cada unidad de aumento en Tiempo de retardo entre taladros(ms), el P80 aumente su promedio esperado en 2.9%. A continuación, se pondrá a prueba el modelo obtenido vs el Modelo nulo, lo cual dirá si existe diferencia alguna en el modelo, en caso se tome en cuenta a la variable T-Taladros y en caso de no hacerlo.

Tabla 9: Tabla de resultados del modelo univariado T-taladros vs P80.

```
> # TEST DE F
> anova(mod4_p80, mod4_p80_NULO, test="LRT")
Analysis of Deviance Table

Model 1: p80 ~ `T-TALADROS`
Model 2: p80 ~ 1
  Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1      1231      41.858
2      1232      42.804 -1  -0.94655 2.068e-07 ***
```

Fuente: Elaboración propia.

Por consiguiente, en los resultados se obtuvo que la diferencia residual del Modelo 1 (Considerando la variable T-Taladros) es 1231 y el Modelo 2 (Con el modelo nulo, sin la variable T-Taladros) posee 1232 como valor, lo que conlleva a la conclusión de que al tener el Modelo 1 menor diferencia residual, genera un impacto positivo a la explicación de la variable P80 en comparación al caso en que no se tome en cuenta a la variable T-Taladros.

Finalmente, se muestra un gráfico que muestra la relación existente entre las variables Tiempo de Retardo entre taladros(ms) vs P80.

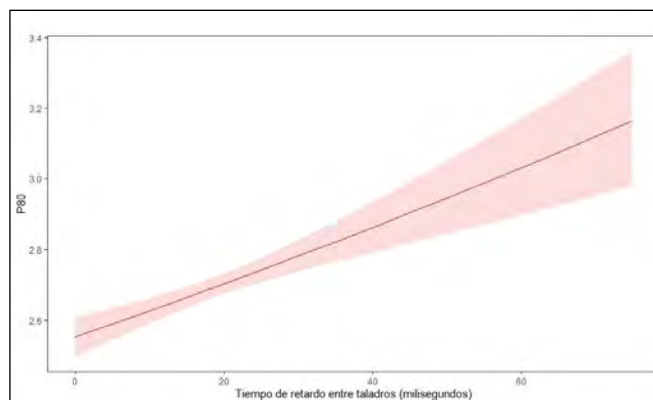


Figura 29. Tiempo de retardo entre taladros(ms) vs P80 (pulg).

Fuente: Elaboración propia.

Resistencia y Factor de potencia vs P80

Para el desarrollo del modelo, fue necesario agrupar la variable de Resistencia a la variable Factor de potencia, ya que de esta manera se podría tomar en cuenta a la Resistencia en el modelo, ya que es un valor cualitativo.

Una vez desarrollado el modelo con interacciones de Resistencia y factor de potencia vs P80, se procede a corregir la normalidad aplicando el modelo de regresión GLM gamma con la función de enlace logaritmo y tomando como nivel de referencia el tipo de resistencia R3, una vez aplicado el modelo de tipo gamma se revisa si el modelo con interacciones de Resistencia y Factor de potencia posee explicabilidad respecto a la variable P80.

Tabla 10: Tabla de resultados del modelo con interacciones Resistencia y FP vs P80.

Predictors	Estimates	p80		p
		CI		
(Intercept)	2.2011	2.0490	- 2.3646	<0.001
FPREAL	2.0828	1.5651	- 2.7745	<0.001
RESISTENCIA [R3-R4]	1.2308	1.0367	- 1.4621	0.017
RESISTENCIA [R4]	1.3862	1.2159	- 1.5809	<0.001
FPREAL × RESISTENCIA [R3-R4]	0.5605	0.3118	- 1.0073	0.050
FPREAL × RESISTENCIA [R4]	0.3347	0.2191	- 0.5109	<0.001
Observations	1233			

Fuente: Elaboración propia.

Según los resultados de la tabla, se observa que el p-valor de la prueba de ratio de verosimilitudes (LTR, siglas en inglés: Likelihood Ratio Test) del modelo es menor igual a 0.05, el cual es el máximo valor como nivel de significación y por lo que se deduce que el modelo sí es significativo. Adicionalmente, según los valores obtenidos se espera que por cada unidad de aumento de Factor de potencia (kg/TN) cuando la resistencia de la roca sea tipo R4, se espera que se genere una disminución del 66.53% del P80 respecto al promedio esperado por cada unidad de aumento de FP-REAL para el sustrato R3.

Tabla 11: Tabla de resultados del modelo con interacciones Resistencia y FP vs P80.

```
> # TEST DE F
> anova(mod5_p80, mod5_p80_NULO, test="LRT")
Analysis of Deviance Table

Model 1: p80 ~ FPREAL * RESISTENCIA
Model 2: p80 ~ 1
  Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1      1227      41.002
2      1232      42.804 -5  -1.8026 5.198e-10 ***
```

Fuente: Elaboración propia.

Por consiguiente en los resultados se obtuvo que la diferencia residual del Modelo 1 (Considerando las variables Resistencia y Factor de potencia) es 1227 y el Modelo 2 (Con el modelo nulo, sin las variables Resistencia y FP) posee 1232 como valor, lo que conlleva a la conclusión de que al tener el Modelo 1 menor diferencia residual, genera un impacto positivo a la explicación de la variable P80 en comparación al caso en que no se tome en cuenta a las variables Resistencia y Factor de Potencia.

Finalmente, se muestra un gráfico que muestra la relación existente entre las variables Resistencia - Factor de potencia(kg/TN) vs P80 (pulg).

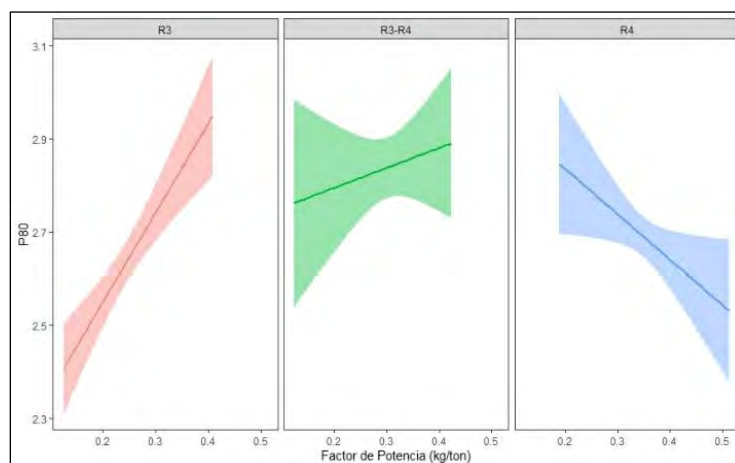


Figura 30. Resistencia- Factor de Potencia(kg/TN) vs P80 (pulg). Fuente: Elaboración propia.

Densidad de material vs P80

Procedemos a corregir la normalidad aplicando el modelo de regresión GLM gamma con la función de enlace logaritmo, una vez aplicado el modelo de tipo gamma se revisa si la variable Densidad de material posee explicabilidad respecto a la variable P80.

Tabla 12: Tabla de resultados del modelo univariado Densidad material vs P80.

<i>Predictors</i>	<i>Estimates</i>	p80		<i>p</i>
		<i>CI</i>		
(Intercept)	2.6506	2.4190 – 2.8883		<0.001
DENSIDAD MATERIAL	1.0057	0.9736 – 1.0412		0.735
Observations	1233			

Fuente: Elaboración propia.

Según los resultados de la tabla, se observa que el p-valor de la prueba de ratio de verosimilitudes (LTR, siglas en inglés: Likelihood Ratio Test) del modelo es mayor a 0.05, el cual es el máximo valor como nivel de significación y por lo que se deduce que

este modelo univariado no es significativo, dando por finalizado el análisis de este modelo y se pasa al siguiente.

Densidad de explosivo final vs P80

Se corrige la normalidad aplicando el modelo de regresión GLM gamma con la función de enlace logaritmo, una vez aplicado el modelo de tipo gamma se revisa si la variable Densidad de material posee explicabilidad respecto a la variable P80.

Tabla 13: Tabla de resultados del modelo univariado Densidad explosivo final vs P80.

<i>Predictors</i>	<i>Estimates</i>	p80	
		<i>CI</i>	<i>p</i>
(Intercept)	0.6301	0.1667 – 2.3787	0.492
DENSIDAD EXPLOSIVO-FINAL	3.6321	1.1157 – 11.8387	0.031
Observations	1233		

Fuente: Elaboración propia.

Según los resultados de la tabla, se observa que el p-valor de la prueba de ratio de verosimilitudes (LTR, siglas en inglés: Likelihood Ratio Test) del modelo es inferior, pero cercano a 0.05, el cual es el máximo valor como nivel de significación y por lo que se deduce que el modelo sí es significativo. Adicionalmente, según los valores obtenidos se espera que por cada 0.01 de aumento de Densidad de explosivo final (gr/cm³), el P80 aumente su promedio esperado en 3.63 %. A continuación, se pondrá a prueba el modelo obtenido vs el Modelo nulo, lo cual dirá si existe diferencia alguna en el modelo, en caso se tome en cuenta a la variable Densidad de explosivo final y en caso de no hacerlo.

Tabla 14: Tabla de resultados del modelo univariado Densidad explosivo final vs P80.

```

> # TEST DE F
> anova(mod7_p80, mod7_p80_NULO, test="LRT")
Analysis of Deviance Table

Model 1: p80 ~ `DENSIDAD.EXPLOSIVO-FINAL`
Model 2: p80 ~ 1
  Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1      1231      42.639
2      1232      42.804 -1 -0.16486  0.0322 *

```

Fuente: Elaboración propia.

Por consiguiente, en los resultados se obtuvo que la diferencia residual del Modelo 1 (Considerando la variable Densidad de explosivo final) es 1231 y el Modelo 2 (Con el modelo nulo, sin las variables Resistencia y FP) posee 1232 como valor, lo que conlleva a la conclusión de que al tener el Modelo 1, menor diferencia residual, genera un impacto positivo a la explicación de la variable P80 en comparación al caso en que no se tome en cuenta a la variable Densidad de explosivo final.

Finalmente, se muestra un gráfico que muestra la relación existente entre la variable Densidad de explosivo final (gr/cm³) vs P80(pulg), dónde se aprecia que existe mayor confianza de predicción entre los valores de densidad de 1.1 a 1.15 gr/cm³, pues la banda en ese intervalo resulta más angosta.

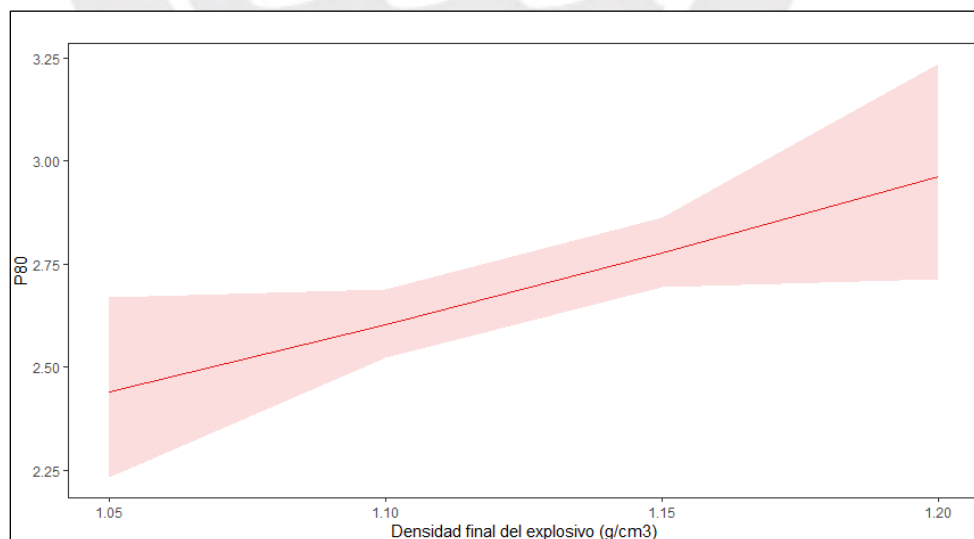


Figura 31. Densidad final del explosivo (gr/cm³) vs P80 (pulg).

Fuente: Elaboración propia.

Modelo multivariado final de P80

Una vez terminado el análisis de las variables Tiempo de retardo entre taladros(ms), densidad de material(gr/cm³), densidad de explosivo final (gr/cm³) mediante un modelo univariado y con interacciones para el caso de Resistencia y Factor de potencia (kg/TN); se realiza el siguiente paso que relaciona al total de las variables mediante un modelo multivariado en el cual se verá la influencia de cada variable en presencia de las otras. Para lo cual se inicia corrigiendo la normalidad aplicando el modelo de regresión GLM gamma con la función de enlace logaritmo, una vez aplicado el modelo de tipo gamma se revisa la explicabilidad de las variables en conjunto con respecto al P80; sin embargo, se excluyó a la variable Densidad de material (gr/cm³) pues esta variable predictora no resultó significativa para el modelo, según su análisis univariado.

Tabla 15: Tabla de resultados del modelo multivariado con todas las variables vs P80.

<i>Predictors</i>	<i>Estimates</i>	p80	
		<i>CI</i>	<i>p</i>
(Intercept)	1.1468	0.2997 – 4.3777	0.841
T-TALADROS	1.0025	1.0013 – 1.0037	<0.001
FPREAL	1.7303	1.2863 – 2.3299	<0.001
RESISTENCIA [R3-R4]	1.1929	1.0048 – 1.4169	0.042
RESISTENCIA [R4]	1.3877	1.2177 – 1.5820	<0.001
DENSIDAD EXPLOSIVO-FINAL	1.8012	0.5435 – 5.9817	0.335
FPREAL × RESISTENCIA [R3-R4]	0.6063	0.3381 – 1.0869	0.089
FPREAL × RESISTENCIA [R4]	0.3356	0.2202 – 0.5110	<0.001
Observations	1233		

Fuente: Elaboración propia.

Según los resultados de la tabla, se observa que el p-valor de la prueba de ratio de verosimilitudes (LTR, siglas en inglés: Likelihood Ratio Test) del modelo multivariado

es inferior a 0.05 en la mayoría de las variables analizadas a excepción de la variable Densidad de explosivo final en dónde se observa un valor de 0.335, lo cual significa que no resulta significativa con presencia del resto de variables; sin embargo, a continuación se evaluará si el modelo en conjunto con todas las variables resulta significativo para la predicción del P80, para lo cual se aplica un análisis de ANOVA y una prueba de F, en dónde el Modelo nulo dirá si existe diferencia alguna en el modelo, en caso se tome en cuenta a las variables y en caso de no hacerlo.

Tabla 16: Tabla de resultados del modelo multivariado vs P80.

```
> # TEST DE F
> anova(mod8_p80_multi, mod8_p80_multi_NULO, test="LRT")
Analysis of Deviance Table

Model 1: p80 ~ `T-TALADROS` + FPREAL * RESISTENCIA + `DENSIDAD.EXPLOSIVO-FINAL`
Model 2: p80 ~ 1
  Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1      1225      40.410
2      1232      42.804 -7    -2.394 1.38e-12 ***
```

Fuente: Elaboración propia.

Por consiguiente, en los resultados se obtuvo que la diferencia residual del Modelo 1 (considerando las variables T-Taladros, FP, Resistencia y densidad de explosivos final) es 1225 y el Modelo 2 (Con el modelo nulo, sin las variables) posee 1232 como valor, lo que conlleva a la conclusión de que al tener el Modelo 1, menor diferencia residual con 7 unidades menos, genera un impacto positivo a la explicación de la variable P80 en comparación al caso en que no se tome en cuenta a las variables.

Finalmente, se muestra un gráfico que muestra las variables del modelo multivariado y su impacto en la predicción del valor P80.

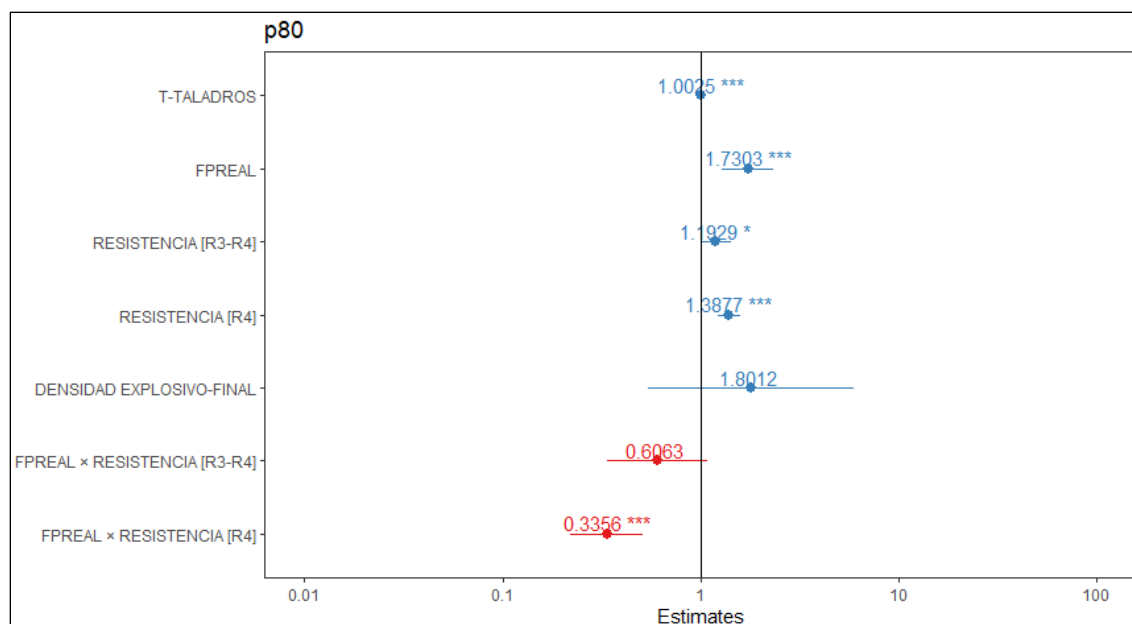


Figura 32. Tabla de resultados del modelo multivariado vs P80.

Fuente: Elaboración propia.

ANÁLISIS DE VARIABLES EXPLICATIVAS – PORCENTAJE DE FINOS

Tiempo de retardo entre taladro (ms) vs Porcentaje de finos

Se puede observar que el modelo sí es significativo; además, se tienen que exponenciar los coeficientes para poder definir las proporciones equivalentes, pues el modelo Beta utiliza una función de enlace (link function) logarítmica. En los resultados se puede observar que, por cada milisegundo de aumento en el tiempo de retardos, el % de finos aumentaría en un 0.15%.

Tabla 17: Tabla de resultados del modelo univariado T-taladros vs Porcentaje de finos.

Predictors	Estimates	finos	
		CI	p
(Intercept)	0.4841	0.4788 – 0.4895	<0.001
T-TALADROS	1.0015	1.0010 – 1.0020	<0.001
Observations	1233		

Fuente: Elaboración propia.

Para modelos betareg, la significancia no puede analizarse con la comparación entre el modelo nulo y el modelo con variables; por lo que, para este caso se utilizó un método de Likelihood ratio test; a través del logaritmo de la verosimilitud, considerando que mientras más alto sea este valor, mejor es el modelo.

Tabla 18: Tabla de resultados del modelo univariado T-taladros vs Porcentaje de finos.

Likelihood ratio test					
Model 1: finos ~ `T-TALADROS`					
Model 2: finos ~ 1					
	#Df	LogLik	Df	Chisq	Pr(>Chisq)
1	3	2009.1			
2	2	1993.3	-1	31.501	1.993e-08 ***

Fuente: Elaboración propia.

Se puede observar que la diferencia entre los LogLik (logaritmo de verosimilitud) es significativa y que el correspondiente al modelo con variable es mayor.

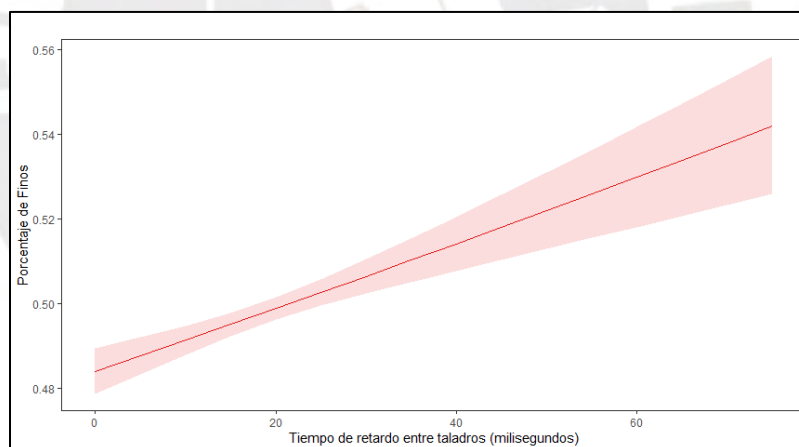


Figura 33. Tiempo de retardo entre taladros(ms) vs Porcentaje de finos. Fuente: Elaboración propia.

Resistencia y Factor de potencia vs Porcentaje de finos

A continuación, se realizó el análisis al modelo bivariado que evaluó la interacción entre las variables resistencia y factor de potencia real; y se pudo observar que no había significancia en las interacciones, por lo que se realizaron los análisis correspondientes a los modelos univariados de cada variable por separado.

El modelo del factor de potencia real resultó significativo; siendo el estimado 0.34189; que al exponenciarlo resulta 1.4076; lo que nos indica que, por cada aumento en la unidad del factor de potencia, el promedio esperado en el porcentaje de finos aumentaría en un 40.76%.

Tabla 19: Tabla de resultados del modelo univariado Factor de Potencia vs Porcentaje de finos.

Predictors	Estimates	finos		p
		CI		
(Intercept)	0.4485	0.4384 – 0.4589		<0.001
FPREAL	1.4076	1.3093 – 1.5133		<0.001
Observations	1233			

Fuente: Elaboración propia.

Tabla 20: Tabla de resultados del modelo univariado Factor de Potencia vs Porcentaje de finos.

```
Likelihood ratio test
Model 1: finos ~ FPREAL
Model 2: finos ~ 1
#Df LogLik Df  Chisq Pr(>Chisq)
1    3 2034.8
2    2 1993.3 -1  82.939 < 2.2e-16 ***
```

Fuente: Elaboración propia.

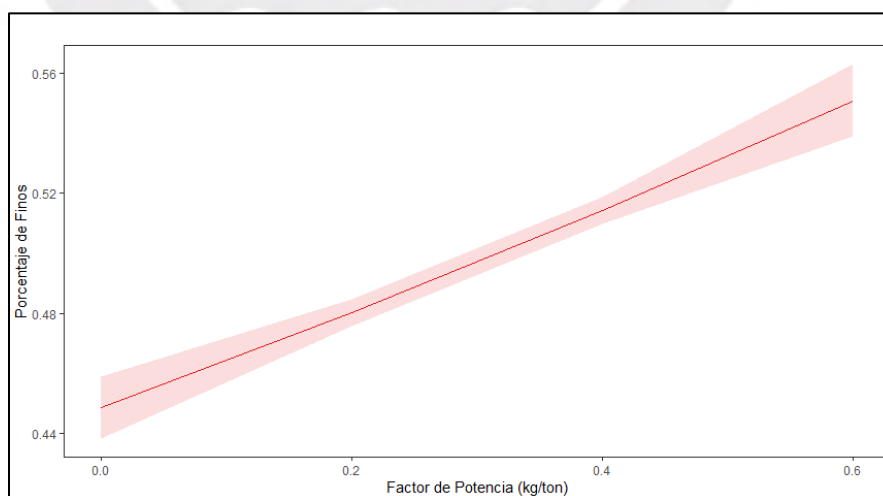


Figura 34. Factor de Potencia vs Porcentaje de finos.

Fuente: Elaboración propia.

Con respecto al modelo univariado utilizando la variable resistencia, considerando que esta es una variable categórica, se identificó que R4 sí se diferencia estadísticamente de R3, mas no de R3-R4; asimismo, pertenecer a R4 genera un estimado positivo; por lo tanto, va a ser un factor de aumento cuando lo exponenciamos. De acuerdo con los resultados obtenidos, se espera que el porcentaje de finos aumente en un 2% cuando la resistencia es R4 en relación con porcentaje de finos para R3; mientras que para la categoría R3-R4, el efecto fue de un aumento en 0.2% (IC95% -1.44% a 1.19%) respecto a R3, pero fue no significativo.

Tabla 21: Tabla de resultados del modelo univariado Resistencia vs Porcentaje de finos.

Predictors	Estimates	finos		p
		CI		
(Intercept)	0.4906	0.4863	-0.4950	<0.001
RESISTENCIA [R3-R4]	1.0023	0.9856	-1.0192	0.792
RESISTENCIA [R4]	1.0289	1.0170	-1.0410	<0.001
Observations	1233			

Fuente: Elaboración propia.

Tabla 22: Tabla de resultados del modelo univariado Resistencia vs Porcentaje de finos.

```

Likelihood ratio test

Model 1: finos ~ RESISTENCIA
Model 2: finos ~ 1
  #Df LogLik Df  Chisq Pr(>Chisq)
1    4 2006.3
2    2 1993.3 -2  25.987  2.275e-06 ***

```

Fuente: Elaboración propia.

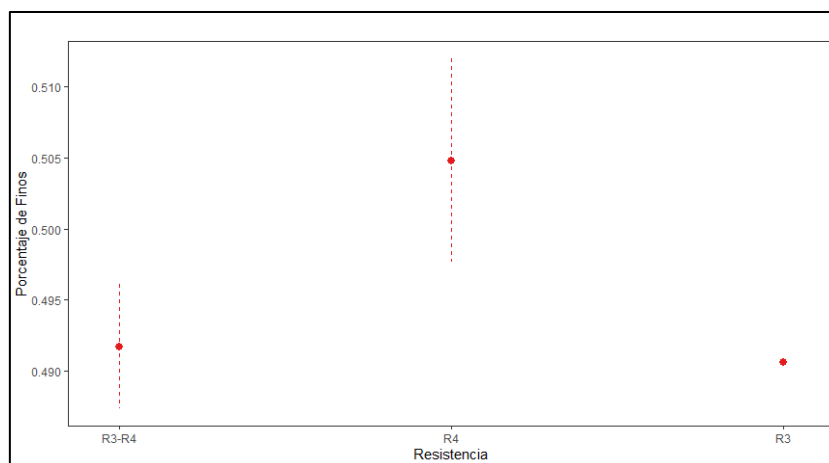


Figura 35. Resistencia vs Porcentaje de finos.

Fuente: Elaboración propia.

Densidad de material vs Porcentaje de finos

Por otro lado, al realizar el análisis con la variable densidad de material, se obtuvo que el modelo no era significativo.

Tabla 23: Tabla de resultados del modelo univariado Densidad material vs Porcentaje de finos.

Predictors	Estimates	finos		p
		CI		
(Intercept)	0.5018	0.4796	- 0.5251	<0.001
DENSIDAD MATERIAL	0.9968	0.9799	- 1.0140	0.714
Observations	1233			

Fuente: Elaboración propia.

Densidad de explosivo final vs Porcentaje de finos

Finalmente, la última variable a analizar fue densidad final del explosivo; esta variable tampoco se identificó como significativa.

Tabla 24: Tabla de resultados del modelo univariado Densidad de explosivo final vs Porcentaje de finos.

Predictors	Estimates	finos		p
		CI		
(Intercept)	0.4000	0.2046	- 0.7823	0.007
DENSIDAD EXPLOSIVO-FINAL	1.2140	0.6690	- 2.2029	0.524
Observations	1233			

Fuente: Elaboración propia.

Modelo multivariado final de P80

Para realizar el modelo multivariado se tomó en cuenta todas las variables previamente mencionadas, se encontró que el tiempo de retardo entre taladros no es significativo a nivel multivariado, el factor de potencia real sí lo es; y, se encontraron diferencias significativas entre los grupos de resistencias R3 y R3-R4.

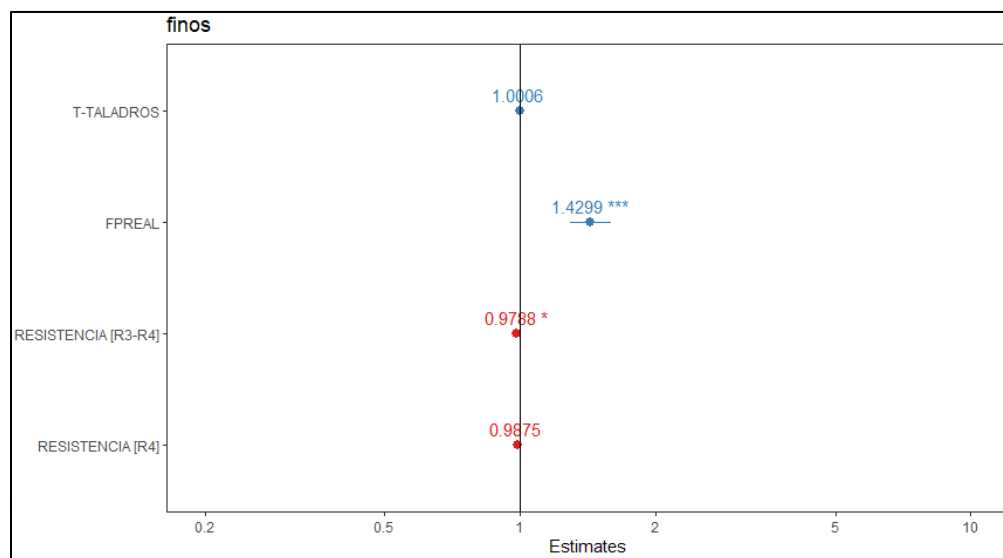


Figura 36. Tabla de resultados del modelo multivariado vs Porcentaje de finos.

Fuente: Elaboración propia

Tabla 25: Tabla de resultados del modelo multivariado con todas las variables vs Porcentaje de finos.

<i>Predictors</i>	finos		
	<i>Estimates</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.4458	0.4346 – 0.4573	<0.001
T-TALADROS	1.0006	1.0000 – 1.0012	0.052
FPREAL	1.4299	1.2880 – 1.5875	<0.001
RESISTENCIA [R3-R4]	0.9788	0.9621 – 0.9958	0.015
RESISTENCIA [R4]	0.9875	0.9725 – 1.0027	0.106
Observations	1233		

Fuente: Elaboración propia.

ANÁLISIS DE MODELOS - Incluyendo al Burden

Con el objetivo de mejorar la explicabilidad de los modelos e incluir la variable burden, que está relacionada a una de las cuatro categorías de clasificación para los factores que se incluyen en la presente tesis (Diseño de malla); se realizó una corrida al modelo incluyendo la variable burden para el P80 y para el porcentaje de finos. En la sección de Presentación y discusión de resultados, se presentarán las diferencias identificadas entre estos modelos.

Modelo univariado de P80 – Incluyendo al burden

Tabla 26: Tabla de resultados del modelo univariado Burden vs P80.

<i>Predictors</i>	<i>Estimates</i>	p80	
		<i>CI</i>	<i>p</i>
(Intercept)	3.2002	2.6355 – 3.8852	<0.001
BURDEN	0.9773	0.9527 – 1.0027	0.075
Observations	1233		

Fuente: Elaboración propia.

Modelo univariado de Porcentaje de finos – Incluyendo al burden

Tabla 27: Tabla de resultados del modelo univariado Burden vs Porcentaje de finos.

<i>Predictors</i>	<i>Estimates</i>	finos	
		<i>CI</i>	<i>p</i>
(Intercept)	0.8667	0.7893 – 0.9516	0.003
BURDEN	0.9293	0.9179 – 0.9409	<0.001
Observations	1233		

Fuente: Elaboración propia.

Modelo multivariado de P80 – Incluyendo al burden

Tabla 28: Tabla de resultados del modelo multivariado con todas las variables vs P80.

<i>Predictors</i>	<i>Estimates</i>	p80	
		<i>CI</i>	<i>p</i>
(Intercept)	0.4928	0.1178 – 2.0562	0.331
T-TALADROS	1.0030	1.0018 – 1.0043	<0.001
FPREAL	2.5802	1.7651 – 3.7730	<0.001
RESISTENCIA [R3-R4]	1.2004	1.0116 – 1.4252	0.035
RESISTENCIA [R4]	1.4742	1.2873 – 1.6889	<0.001
DENSIDAD EXPLOSIVO-FINAL	1.8922	0.5725 – 6.2682	0.295
BURDEN	1.0904	1.0355 – 1.1483	0.001
FPREAL × RESISTENCIA [R3-R4]	0.5935	0.3315 – 1.0624	0.076
FPREAL × RESISTENCIA [R4]	0.2874	0.1868 – 0.4417	<0.001
Observations	1233		

Fuente: Elaboración propia.

Modelo multivariado de Porcentaje de finos – Incluyendo al burden

Tabla 29: Tabla de resultados del modelo multivariado con todas las variables vs Porcentaje de finos.

<i>Predictors</i>	<i>Estimates</i>	finos	
		<i>CI</i>	<i>p</i>
(Intercept)	1.0174	0.8136 – 1.2722	0.880
T-TALADROS	1.0001	0.9995 – 1.0007	0.754
FPREAL	0.9911	0.8598 – 1.1424	0.901
RESISTENCIA [R3-R4]	0.9749	0.9586 – 0.9915	0.003
RESISTENCIA [R4]	0.9740	0.9592 – 0.9890	0.001
BURDEN	0.9119	0.8895 – 0.9348	<0.001
Observations	1233		

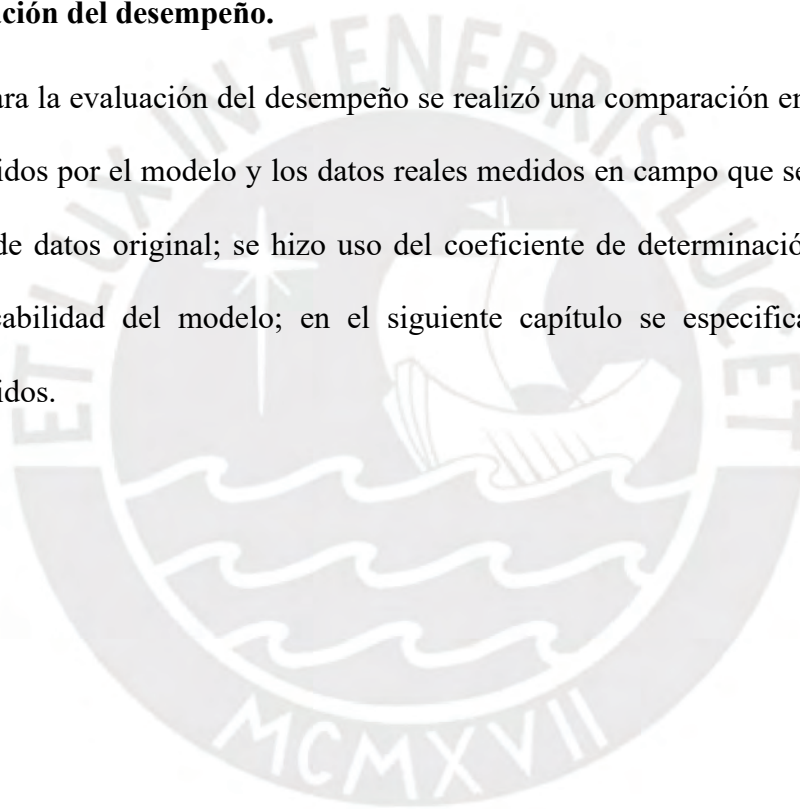
Fuente: Elaboración propia.

4.2.4. Entrenamiento del modelo.

El entrenamiento de data se realiza posterior a la primera corrida del modelo; para ello se necesita añadir la data actualizada de los nuevos resultados de voladura a la base de datos original; en los siguientes pasos se presentará la primera corrida del modelo; según los resultados de estos se decidirá si continuar entrenando el modelo o cambiar de algoritmo para una mejor obtención de resultados.

4.2.5. Evaluación del desempeño.

Para la evaluación del desempeño se realizó una comparación entre los resultados obtenidos por el modelo y los datos reales medidos en campo que se registraron en la base de datos original; se hizo uso del coeficiente de determinación para evaluar la explicabilidad del modelo; en el siguiente capítulo se especifican los resultados obtenidos.



5. Presentación y discusión de resultados

5.1. P80.

Al finalizar la aplicación del modelo de regresión lineal generalizado, para el P80; a partir del intercepto y los coeficientes previamente calculados, se diseñaron los siguientes gráficos comparativos entre los resultados obtenidos por el modelo predictivo y los valores reales; para los modelos multivariados con y sin burden. Para medir el nivel de precisión del modelo se hizo uso del coeficiente de determinación; que es la proporción de la varianza total de la variable explicada por la regresión; en el caso del modelo multivariado en el que no se incluía al burden, el coeficiente resultó 0.7668; lo que indica un buen ajuste, pero que da cabida a planes de mejora, ya sea incluyendo un mayor número de variables explicativas o quitando las variables no significativas y/o las que poseen colinealidad; asimismo, en el modelo en el que sí se incluye al burden, el coeficiente resultó ser 0.9301; lo que indica un ajuste mucho más preciso; esto resulta coherente, pues en la tabla del modelo multivariado para el burden, podemos identificar que el algoritmo identificó esta variable como significativa para el modelo.

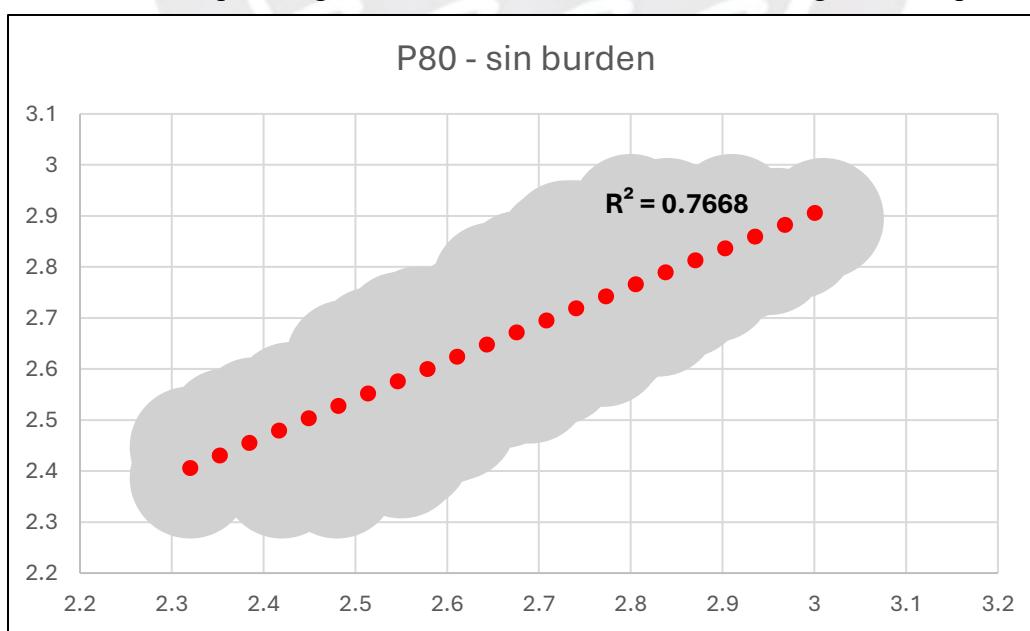


Figura 37. Gráfico de regresión lineal – P80 (sin burden).

Fuente: Elaboración propia

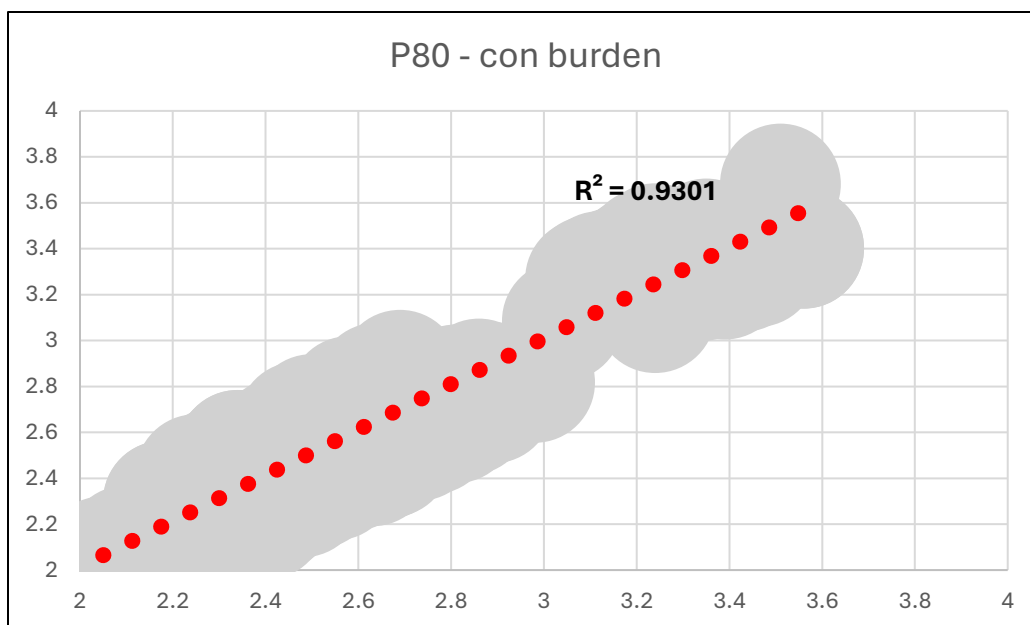


Figura 38. Gráfico de regresión lineal – P80 (con burden).

Fuente: Elaboración propia

5.2. Porcentaje de finos.

Por otro lado, se realizó un análisis similar para el porcentaje de finos, donde con los resultados obtenidos se realizaron las siguientes gráfica y se calculó el coeficiente de determinación para el modelo multivariado sin burden que, en este caso, resultó 0.654, podemos notar que es menor al del caso anterior, y esto puede deberse a que la distribución de los residuales, como se pudo identificar previamente en los Normality Diagnosis Plots, se alejaba más que en los casos anteriores a una distribución normal. Asimismo, se calculó el coeficiente de determinación para el modelo generado incluyendo a la variable burden y este resultó ser de 0.7294; de igual manera, esto va de acuerdo con la significancia detectada por el algoritmo para la variable burden tanto en el modelo univariado como en el multivariado de esta variable respuesta.

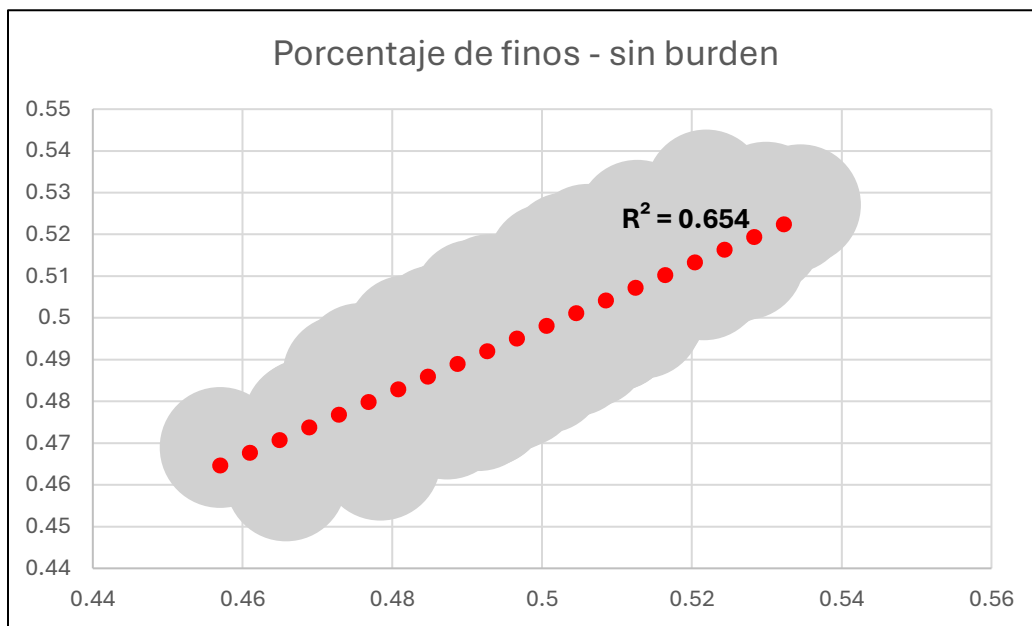


Figura 39. Gráfico de regresión lineal – Porcentaje de finos (sin burden). Fuente: Elaboración propia

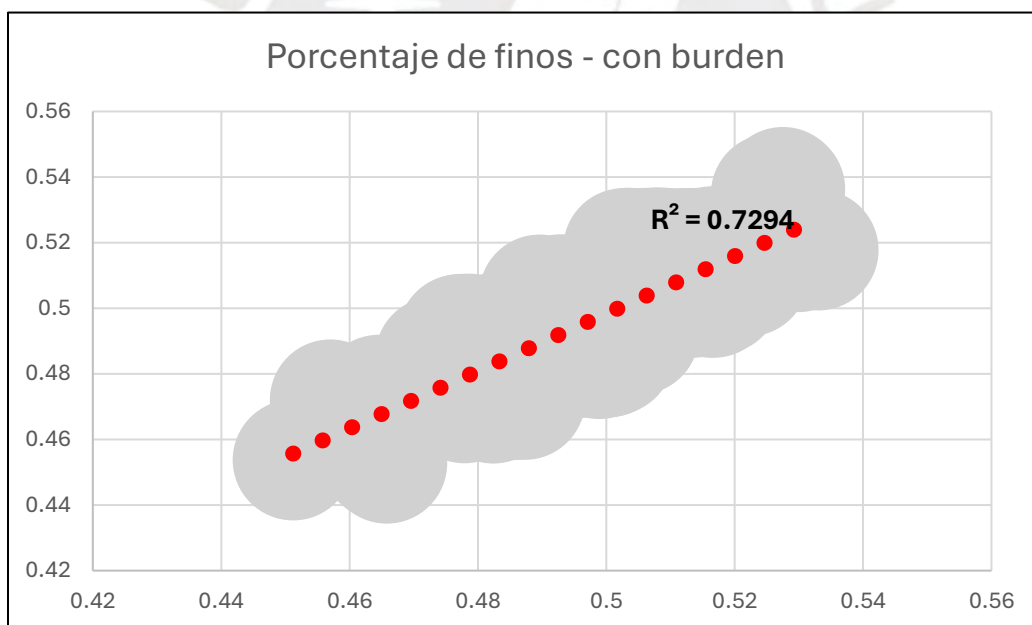


Figura 40. Gráfico de regresión lineal – Porcentaje de finos (con burden). Fuente: Elaboración propia

6. Conclusiones y recomendaciones

- Se ha podido identificar que la organización de los datos recolectados por la mina no es la adecuada y dificulta el proceso de limpieza de datos para la posterior aplicación de algoritmos avanzados.
- El modelo lineal generalizado nos ha permitido identificar las variables significativas que intervienen a mayor escala para mejorar la fragmentación; estos fueron: burden, resistencia del macizo rocoso, factor de potencia y tiempos de retardo.
- Se ha confirmado que el LGM es efectivo para las variables de respuesta cuyos residuales no siguen una distribución normal.
- Se llega a concluir que mediante el método LGM propuesto se obtuvo una explicabilidad del 93.01% para el grupo de datos de validación, en el caso del P80 y una explicabilidad del 72.94% para el grupo de validación de la variable respuesta Porcentaje de finos.
- Se observa que los modelos que incluyen al burden, pese a su poca variabilidad; tienen un mayor nivel de explicabilidad que aquellos que no lo incluyen; demostrando su significancia para la fragmentación.
- Gracias a la evaluación independiente de los productos, se observó que las variables explicativas para cada variable de respuesta cambiaron; a pesar de que todos estaban relacionados con el fenómeno de las voladuras.
- Debido a la cantidad de datos disponibles, se deberían continuar desarrollando corridas y actualizando el modelo para aumentar el nivel de confiabilidad del mismo, con un mayor número de datos de entrenamiento. Asimismo, se podrían realizar corridas para buscar predecir otros parámetros de voladura.

- Tras haber realizado el presente estudio, se propone ampliar la gama de algoritmos de Machine Learning a aplicar y así comparar los resultados e identificar el algoritmo óptimo.
- El LGM ha demostrado ser efectivo para la predicción de la voladura superficial, y sería interesante analizar su aplicación en una mina subterránea en operación.
- La metodología de machine learning supervisado es novedosa en cuanto a su aplicación dentro de la minería. Al ser así, resulta importante estudiar su aplicación en otras áreas además del análisis de fragmentación en voladura, como podría ser la sobrerotura en minería subterránea o la existencia de fly rock en cuanto a la minería superficial.
- Hacia el futuro, resulta importante analizar mayor cantidad de minas para la aplicación del machine learning a fin de conocer si es que se puede establecer algún método estándar que funcione sin importar la mina en la que se encuentre. Es claro que sería mucho más especializado y complejo, pero podríamos estar ante una nueva herramienta que permita el avance exponencial del sector minero, disminuyendo costos y aumentando los beneficios.

7. Bibliografía

- Blasting, R. (1991). *Title and Subtitle Rock Blasting and Overbreak Control*.
- Bernaola, J. (2014). *Perforación y Voladura de Roca en Minería*. Journal of Investment Management.
- Carhuamaca, R. (2018). *Diseño de mallas de perforación y voladura para optimizar avances y sobre rotura N.1225 - Mina Andaychagua - VCM S.A.A.*
- Celi, C. (2013). *Control de un generador de pulsos, una fuente de voltaje- corriente y matriz por medio del protocolo GPIB programado en Python*. (Tesis de licenciatura, Universidad San Francisco de Quito).
- Chakraborty, A. K., Jethwa, J. L., & Paithankar, A. G. (1994). *Overbreak in Tunnel Blasting*. Science, 9(4), 471–482.
- Chancasanampa, W. (2019). *Diseño de la malla de perforación y voladura para incrementar la productividad de tajeos en la Compañía Minera Great Panther Coricancha*.
- Cribari-Neto, F., & Zeileis, A. (2010). *Beta regression in R*. Journal of Statistical Software, 34(2), 1–24. <https://doi.org/10.18637/jss.v034.i02>
- Control, G. (1981). *Underground Excavations*.
- Diéguez, Y., Otaño, J., & Sargentón, G. (2014). *Diseño de voladuras de contorno en túneles*. Minería y Geología, 30(3), 49–66.
- Duval, A. (2019). *Explainable Artificial Intelligence (XAI)*. <https://doi.org/10.13140/RG.2.2.24722.09929>
- Ernst & Young. (2023). *Peru mining & metals investment guide 2023/2024*.
- EXSA. (2021). *Manual de voladuras y explosivos industriales*. EXSA S.A.
- Famesa Explosivos S.A.C (2019). *Manual de perforación y voladura*. Editorial Letras e Imágenes S.A.C
- Ibarra, J. A., Maerz, N. H., & Franklin, J. A. (1996). *Overbreak and underbreak in underground openings Part 2: Causes and implications*. Geotechnical and Geological Engineering, 14(4), 325–340. <https://doi.org/10.1007/BF00421947>

Idrogo, Y. (2022). *Machine learning aplicado al control de la fragmentación de rocas en la voladura de minas a tajo abierto* (Tesis de licenciatura, Universidad Nacional de Ingeniería).

Innaurato, N., Mancini, R., & Cardu, M. (1998). *On the influence of rock mass quality on the quality of blasting work in tunnel driving*. *Tunnelling and Underground Space Technology*, 13(1), 81–89. [https://doi.org/10.1016/S0886-7798\(98\)00027-3](https://doi.org/10.1016/S0886-7798(98)00027-3)

International Society of Explosives Engineers (ISSE). (2008). *Manual del Especialista en Voladura* (R. B. Hopler, Ed.; 17th ed.).

Iverson, S. R., Hustrulid, W. A., & Johnson, J. C. (2013). *A new perimeter control blast design concept for underground metal/nonmetal drifting applications*. DEPARTMENT OF HEALTH AND HUMAN SERVICES Centers for Disease Control and Prevention National Institute for Occupational Safety and Health Office of Mine Safety and Health Research.

Jang, H., & Topal, E. (2013). *Optimizing overbreak prediction based on geological parameters comparing multiple regression analysis and artificial neural network*. *Tunnelling and Underground Space Technology*, 38(September), 161–169. <https://doi.org/10.1016/j.tust.2013.06.003>

Kashnikov, A. V., & Levin, L. (2017). *Applying machine learning techniques to mine ventilation control systems*. *Proceedings of 2017 20th IEEE International Conference on Soft Computing and Measurements, SCM 2017*, 391–393. <https://doi.org/10.1109/SCM.2017.7970595>

Khandelwal, M., & Monjezi, M. (2013). *Prediction of backbreak in open-pit blasting operations using the machine learning method*. *Rock Mechanics and Rock Engineering*, 46(2), 389–396. <https://doi.org/10.1007/s00603-012-0269-3>

Kumar, R. et al. (2024). *Enhancing rock fragmentation assessment in mine blasting through machine learning algorithms: A practical approach*. *Research and Discover Applied Sciences*, 6, 223. <https://doi.org/10.1007/s42452-024-05888-0>

López, C. (2003). *Manual de Perforación y Voladura de Rocas*. In Ingeopress (p. 7).

Mandal, S., & Singh, M. (2009). *Evaluating extent and causes of overbreak in tunnels*. *Tunneling and*

Underground Space Technology, 24(1), 22–36. <https://doi.org/10.1016/j.tust.2008.01.007>

Martinez, A., & Morales J. (2001). *Modelos Lineales Generalizados*.

Mining Services, A. (2014). *Surface blasting Handbook*. 17–25.

Ministerio de Energía y Minas. (2013). *Actividad minera en el Perú: Definiciones*.

Ministerio de Energía y Minas. (2023). *Anuario Minero 2023*.

Mohammadi, H., & Azad, A. (2020). *Applying Rock Engineering Systems Approach for Prediction of Overbreak Produced in Tunnels Driven in Hard Rock*. Geotechnical and Geological Engineering, 38(3), 2447–2463. <https://doi.org/10.1007/s10706-019-01161-z>

Moreno, G. (2021). *Modelos de regresión espacio-temporales bayesianos aplicados al mercado de alojamiento Airbnb y HomeAway en Bogotá para el periodo 2015-2019* (Tesis de maestría, Universidad Santo Tomás).

Muñoz, Castro, J. J., & Correa, Rodriguez, J. M. (2016). *Reduccion De La Sobrerotura en el Crucero 3910 Del Nivel 2360 De La Mina Pec Consorcio Minero Horizonte*. 59.

Namazi, S., Brankovic, L., Moghtaderi, B., & Zanganeh, J. (2020). *Comparative study of data mining techniques for predicting explosions in coal mines*. Proceedings of the Confluence 2020 - 10th International Conference on Cloud Computing, Data Science and Engineering, 738–743. <https://doi.org/10.1109/Confluence47617.2020.9057921>

Organismo Supervisor de la inversión en Energía y Minería. (2017). *La industria de la minería en el Perú - 20 años de contribución al crecimiento y desarrollo económico del país*.

Ouanan, H., & Abdelwahed, E. H. (2019). *Image processing and machine learning applications in mining industry: Mine 4.0*. Proceedings - 2019 International Conference on Intelligent Systems and Advanced Computing Sciences, ISACS 2019, 4–8. <https://doi.org/10.1109/ISACS48493.2019.9068884>

Pan, P., Feng, X. T., Yan, F., Pan, P., He, B. G., Hatzor, Y. H., Cai, M., & Wang, X. (2018) *Understanding of Rockburst Evolution Mechanism Using Numerical Modeling*. Rockburst: Mechanisms, Monitoring,

Warning and Mitigation (Issue1993). <https://doi.org/10.1016/B978-0-12-805054-5.00007-X>

Ramírez, P., & Alejano, L. (2004). *Mecánica de Rocas: Fundamentos e Ingeniería de Taludes*. Journal of Chemical Information and Modeling, 300.

Real, A. (2015). *Modelo de regresión para razones y proporciones: Variantes y aplicaciones socioeconómicas y biosanitarias* (Trabajo de fin de grado, Universidad de Sevilla, Facultad de Matemáticas)

Sena, F. (2013). *Desarrollo De Una Herramienta De Diseño De Voladuras En Túneles*. (Tesis de titulación). E.T.S.I Minas (UPM)

Sharma, M., Choudhary, B. S., Raina, A. K., Khandelwal, M., & Rukhiyar, S. (2024). *Prediction of rock fragmentation in a fiery seam of an open-pit coal mine in India*. Journal of Rock Mechanics and Geotechnical Engineering. <https://doi.org/10.1016/j.jrmge.2023.11.047>

Singh, B., & Goel, R. K. (2011). *Method of Excavation*. Engineering Rock Mass Classification, 1977, 281–285. <https://doi.org/10.1016/b978-0-12-385878-8.00021-5>

Singh, S. P. (2018). *Over break control in underground mines*. MOJ Civil Engineering, 4(5),429–435. <https://doi.org/10.15406/mojce.2018.04.00140>

Sun, X., Zhang, H., Tian, F., & Yang, L. (2018). *The Use of a Machine Learning Method to Predict the Real-Time Link Travel Time of Open-Pit Trucks*. Mathematical Problems in Engineering, 2018. <https://doi.org/10.1155/2018/4368045>

Tapia, F., Gutiérrez, A. (2021). *Estabilización de taludes a doble banco en Sociedad Minera Cerro Verde*. Instituto Geológico, Minero y Metalúrgico-INGEMMET.

Tecnología Minera. (2023). *5 métodos comunes de voladura en minería a cielo abierto*. Tecnología Minera.

Tusell, F. (2011). *Análisis de Regresión. Introducción Teórica y Práctica basada en R*.

Van Eldert, J. (2018). *Analysis of Excavation Damage, Rock Mass Characterisation and Rock Support Design using Drilling Monitoring* (Issue July). <https://doi.org/10.13140/RG.2.2.11423.92329>

Wackerly, D., Mendenhall, W., & Scheaffer, R. (2010). *Estadística matemática con aplicaciones* (7^a ed., pp. 185-196). Cengage Learning.

Wallum P. (1970). *Mining Engineering*. Mining Mag (Vol. 123, Issue 3). <https://doi.org/10.1201/b11072->



8. Anexos

1. Tabla de grupo de datos de validación para el porcentaje de finos.

Finos_mod1	Finos_pred1	Finos_mod2	Finos_pred2	Finos_mod1	Finos_pred1	Finos_mod2	Finos_pred2	Finos_mod1	Finos_pred1	Finos_mod2	Finos_pred2
0.486	0.486	0.503	0.503	0.514	0.512	0.488	0.486	0.491	0.486	0.523	0.518
0.499	0.499	0.504	0.504	0.505	0.503	0.500	0.503	0.499	0.504	0.499	0.503
0.499	0.499	0.492	0.492	0.497	0.499	0.502	0.504	0.480	0.475	0.477	0.472
0.492	0.492	0.504	0.504	0.501	0.499	0.495	0.492	0.529	0.524	0.488	0.492
0.505	0.504	0.503	0.503	0.530	0.528	0.501	0.504	0.466	0.461	0.499	0.503
0.504	0.504	0.485	0.485	0.484	0.486	0.508	0.504	0.491	0.486	0.499	0.503
0.500	0.500	0.473	0.472	0.507	0.504	0.500	0.503	0.513	0.508	0.523	0.518
0.497	0.497	0.502	0.502	0.478	0.475	0.500	0.504	0.491	0.486	0.499	0.503
0.500	0.500	0.504	0.504	0.506	0.504	0.500	0.503	0.499	0.504	0.508	0.503
0.494	0.494	0.504	0.504	0.504	0.507	0.507	0.503	0.481	0.486	0.497	0.492
0.499	0.499	0.504	0.503	0.506	0.504	0.469	0.472	0.505	0.499	0.497	0.492
0.503	0.504	0.486	0.485	0.478	0.481	0.501	0.504	0.488	0.483	0.513	0.518
0.508	0.509	0.502	0.503	0.493	0.496	0.522	0.518	0.513	0.508	0.498	0.503
0.504	0.504	0.519	0.518	0.502	0.499	0.506	0.503	0.495	0.500	0.497	0.492
0.506	0.506	0.502	0.503	0.507	0.510	0.507	0.503	0.507	0.501	0.497	0.492
0.512	0.512	0.492	0.492	0.481	0.478	0.468	0.472	0.481	0.486	0.513	0.518
0.498	0.499	0.491	0.491	0.493	0.490	0.469	0.472	0.502	0.497	0.508	0.503
0.481	0.481	0.517	0.518	0.496	0.499	0.500	0.504	0.512	0.506	0.523	0.518
0.491	0.492	0.491	0.492	0.497	0.500	0.506	0.503	0.529	0.523	0.513	0.518
0.512	0.511	0.505	0.503	0.502	0.499	0.507	0.503	0.512	0.506	0.513	0.518
0.512	0.511	0.517	0.518	0.484	0.481	0.488	0.485	0.508	0.503	0.491	0.485
0.510	0.511	0.516	0.518	0.494	0.498	0.522	0.518	0.505	0.499	0.491	0.485
0.493	0.492	0.504	0.503	0.496	0.493	0.488	0.491	0.505	0.499	0.510	0.504
0.491	0.492	0.493	0.491	0.516	0.520	0.507	0.503	0.513	0.507	0.478	0.472
0.496	0.495	0.505	0.503	0.510	0.514	0.478	0.482	0.508	0.502	0.491	0.485

0.500	0.499	0.503	0.504	0.499	0.495	0.468	0.472	0.493	0.487	0.497	0.503
0.509	0.511	0.484	0.485	0.490	0.493	0.500	0.504	0.506	0.512	0.478	0.472
0.497	0.499	0.501	0.503	0.482	0.478	0.489	0.486	0.505	0.499	0.491	0.485
0.503	0.504	0.501	0.503	0.504	0.500	0.507	0.503	0.491	0.498	0.510	0.504
0.511	0.509	0.502	0.504	0.502	0.506	0.514	0.518	0.486	0.479	0.497	0.491
0.465	0.463	0.502	0.503	0.491	0.487	0.506	0.502	0.497	0.503	0.497	0.503
0.504	0.502	0.493	0.491	0.497	0.501	0.499	0.503	0.509	0.503	0.491	0.485
0.490	0.492	0.520	0.518	0.493	0.489	0.500	0.504	0.493	0.487	0.523	0.517
0.484	0.482	0.506	0.504	0.485	0.489	0.500	0.504	0.502	0.495	0.486	0.492
0.508	0.506	0.517	0.519	0.496	0.492	0.499	0.503	0.499	0.492	0.478	0.472
0.495	0.493	0.476	0.473	0.499	0.503	0.481	0.485	0.504	0.497	0.498	0.492
0.464	0.466	0.483	0.485	0.505	0.509	0.521	0.517	0.497	0.504	0.512	0.518
0.496	0.498	0.488	0.486	0.497	0.493	0.499	0.503	0.503	0.496	0.510	0.503
0.497	0.495	0.507	0.504	0.489	0.485	0.522	0.518	0.508	0.515	0.510	0.503
0.501	0.499	0.451	0.454	0.498	0.503	0.481	0.485	0.508	0.515	0.510	0.503
0.493	0.479	0.533	0.518	0.482	0.496	0.501	0.486	0.493	0.479	0.500	0.485

Finos_mod1	Finos_pred1	Finos_mod2	Finos_pred2	Finos_mod1	Finos_pred1	Finos_mod2	Finos_pred2	Finos_mod1	Finos_pred1	Finos_mod2	Finos_pred2
0.521	0.514	0.512	0.518	0.510	0.519	0.509	0.518	0.475	0.486	0.530	0.518
0.468	0.475	0.523	0.516	0.493	0.502	0.526	0.535	0.501	0.489	0.516	0.503
0.482	0.489	0.478	0.471	0.501	0.492	0.481	0.472	0.484	0.496	0.522	0.510
0.486	0.493	0.510	0.504	0.493	0.502	0.528	0.537	0.492	0.504	0.505	0.517
0.469	0.476	0.492	0.485	0.474	0.483	0.483	0.474	0.488	0.499	0.491	0.503
0.518	0.511	0.496	0.503	0.529	0.520	0.528	0.518	0.491	0.480	0.491	0.503
0.535	0.527	0.510	0.503	0.490	0.481	0.483	0.492	0.494	0.482	0.490	0.503
0.495	0.503	0.492	0.485	0.495	0.485	0.494	0.485	0.457	0.469	0.531	0.518
0.491	0.498	0.525	0.518	0.494	0.504	0.501	0.491	0.508	0.496	0.516	0.503
0.501	0.509	0.511	0.518	0.487	0.497	0.482	0.472	0.523	0.511	0.530	0.517
0.507	0.499	0.493	0.485	0.490	0.499	0.482	0.472	0.525	0.513	0.498	0.485

0.489	0.481	0.511	0.504	0.494	0.503	0.508	0.518	0.498	0.486	0.485	0.472
0.478	0.485	0.511	0.504	0.512	0.502	0.527	0.517	0.501	0.513	0.505	0.492
0.523	0.516	0.496	0.503	0.507	0.497	0.482	0.492	0.514	0.502	0.505	0.492
0.504	0.511	0.496	0.503	0.502	0.492	0.514	0.503	0.494	0.481	0.466	0.453
0.533	0.525	0.526	0.518	0.507	0.497	0.493	0.503	0.503	0.516	0.518	0.504
0.496	0.504	0.493	0.485	0.495	0.505	0.502	0.491	0.486	0.498	0.490	0.503
0.491	0.499	0.510	0.518	0.488	0.478	0.508	0.518	0.492	0.479	0.505	0.518
0.494	0.486	0.494	0.486	0.493	0.483	0.475	0.485	0.522	0.509	0.517	0.503
0.512	0.504	0.480	0.472	0.496	0.486	0.496	0.486	0.478	0.466	0.499	0.486
0.503	0.495	0.481	0.472	0.488	0.478	0.493	0.504	0.493	0.506	0.486	0.472
0.492	0.500	0.512	0.504	0.493	0.503	0.507	0.518	0.490	0.478	0.479	0.492
0.501	0.509	0.512	0.503	0.501	0.491	0.496	0.485	0.481	0.494	0.499	0.485
0.494	0.486	0.512	0.503	0.523	0.512	0.529	0.518	0.496	0.483	0.532	0.518
0.502	0.494	0.512	0.503	0.508	0.497	0.474	0.485	0.513	0.500	0.500	0.486
0.488	0.496	0.528	0.536	0.483	0.493	0.529	0.518	0.483	0.496	0.504	0.518
0.500	0.492	0.495	0.503	0.513	0.523	0.492	0.503	0.510	0.497	0.488	0.474
0.492	0.483	0.500	0.491	0.507	0.496	0.465	0.454	0.510	0.497	0.504	0.518
0.500	0.492	0.494	0.503	0.500	0.489	0.493	0.504	0.473	0.486	0.457	0.471
0.510	0.501	0.512	0.503	0.497	0.486	0.515	0.504	0.502	0.489	0.499	0.485
0.517	0.509	0.512	0.503	0.511	0.500	0.482	0.471	0.500	0.487	0.532	0.518
0.519	0.510	0.481	0.473	0.494	0.483	0.529	0.518	0.505	0.492	0.490	0.504
0.500	0.492	0.494	0.486	0.500	0.489	0.515	0.504	0.527	0.514	0.477	0.492
0.522	0.531	0.509	0.518	0.514	0.502	0.473	0.485	0.497	0.484	0.503	0.518
0.492	0.484	0.508	0.517	0.503	0.492	0.506	0.518	0.476	0.489	0.500	0.485
0.528	0.519	0.501	0.492	0.501	0.512	0.497	0.485	0.504	0.491	0.504	0.518
0.523	0.514	0.465	0.474	0.506	0.518	0.497	0.485	0.506	0.493	0.477	0.492
0.517	0.508	0.494	0.485	0.487	0.476	0.484	0.472	0.496	0.482	0.500	0.485
0.483	0.492	0.494	0.485	0.510	0.499	0.484	0.472	0.513	0.500	0.500	0.485
0.522	0.513	0.493	0.502	0.500	0.488	0.506	0.518	0.491	0.504	0.500	0.485
0.492	0.478	0.487	0.472	0.497	0.483	0.457	0.472	0.500	0.487	0.501	0.486

2. Tabla de grupo de datos de validación para el P80.

P80_mod1	P80_pred1	P80_mod2	P80_pred2	P80_mod1	P80_pred1	P80_mod2	P80_pred2	P80_mod1	P80_pred1	P80_mod2	P80_pred2
2.500	2.500	2.640	2.641	2.450	2.474	2.120	2.097	2.470	2.511	3.380	3.431
2.690	2.689	2.320	2.319	2.730	2.705	2.800	2.775	2.810	2.851	2.680	2.732
2.670	2.671	3.430	3.432	2.750	2.725	3.410	3.385	2.760	2.719	2.460	2.512
2.610	2.608	2.810	2.806	2.460	2.485	2.520	2.495	2.750	2.708	2.390	2.443
2.640	2.643	2.100	2.106	2.580	2.555	2.550	2.575	2.560	2.602	2.470	2.524
2.650	2.646	2.410	2.416	2.640	2.665	2.330	2.355	2.640	2.683	3.280	3.335
2.540	2.545	2.490	2.483	2.910	2.884	3.350	3.376	2.660	2.704	2.740	2.796
2.680	2.685	2.370	2.363	2.730	2.704	3.240	3.267	2.710	2.666	2.330	2.388
2.510	2.505	3.390	3.397	2.520	2.494	2.860	2.832	2.760	2.716	2.840	2.782
2.830	2.824	2.750	2.742	2.580	2.608	2.630	2.658	2.800	2.756	2.400	2.460
2.490	2.483	3.360	3.352	2.540	2.511	2.380	2.408	2.700	2.745	2.440	2.500
2.910	2.903	2.050	2.058	2.780	2.809	3.400	3.371	2.840	2.886	2.460	2.521
2.560	2.553	2.710	2.701	2.550	2.580	2.240	2.211	2.570	2.616	3.230	3.292
2.880	2.873	2.730	2.720	2.700	2.730	2.100	2.070	2.520	2.566	3.300	3.363
2.710	2.718	2.610	2.598	2.840	2.810	3.360	3.391	2.800	2.846	3.250	3.315
2.710	2.718	2.710	2.723	2.570	2.601	3.400	3.431	2.690	2.737	2.800	2.734
2.830	2.821	2.710	2.723	2.410	2.441	2.480	2.512	2.630	2.582	2.800	2.734
2.640	2.651	2.640	2.653	2.480	2.511	2.840	2.806	2.840	2.792	2.590	2.524
2.780	2.769	2.260	2.273	2.530	2.498	2.700	2.734	2.610	2.659	2.580	2.512
2.670	2.681	2.510	2.524	2.640	2.672	2.460	2.495	2.760	2.709	2.770	2.701
2.840	2.851	2.530	2.516	2.530	2.562	2.460	2.495	2.740	2.689	2.290	2.221
2.630	2.642	2.380	2.394	2.460	2.493	2.530	2.495	2.460	2.511	2.150	2.080
2.540	2.553	2.590	2.575	2.460	2.494	2.500	2.535	2.440	2.491	2.450	2.521
2.550	2.537	2.620	2.604	2.460	2.494	3.390	3.428	2.580	2.632	2.850	2.778
2.590	2.576	2.310	2.294	2.420	2.386	3.400	3.362	2.800	2.747	2.860	2.788
2.900	2.886	3.350	3.334	2.890	2.856	2.560	2.598	2.380	2.433	2.320	2.392

2.710	2.695	2.660	2.677	2.750	2.716	2.760	2.721	2.590	2.643	2.890	2.817
2.670	2.654	3.350	3.367	2.620	2.654	3.200	3.240	2.630	2.683	2.400	2.473
2.730	2.747	3.400	3.418	2.620	2.654	2.780	2.740	2.900	2.847	2.430	2.503
2.590	2.572	2.650	2.631	2.480	2.514	3.290	3.330	2.740	2.795	2.400	2.474
2.720	2.702	2.500	2.519	2.470	2.505	2.480	2.521	2.840	2.895	2.600	2.524
2.810	2.792	2.540	2.521	2.730	2.695	2.590	2.548	2.590	2.534	3.340	3.261
2.840	2.821	2.540	2.559	2.530	2.494	2.300	2.257	2.780	2.836	2.290	2.211
2.720	2.739	2.760	2.740	2.730	2.692	2.450	2.493	2.630	2.687	2.640	2.720
2.630	2.650	2.210	2.230	2.540	2.578	2.730	2.685	2.610	2.553	3.200	3.281
2.730	2.709	2.100	2.079	2.660	2.698	2.130	2.085	2.550	2.608	3.330	3.413
2.630	2.609	3.330	3.351	2.690	2.729	3.040	3.085	2.590	2.648	2.410	2.493
2.690	2.712	2.730	2.752	2.770	2.809	2.600	2.553	2.720	2.661	2.540	2.456
2.540	2.563	2.610	2.632	2.590	2.550	3.400	3.352	2.880	2.820	2.410	2.494
2.750	2.773	2.580	2.602	2.750	2.790	2.540	2.589	2.620	2.680	2.520	2.604
2.870	2.846	2.470	2.493	2.860	2.820	2.470	2.521	2.450	2.511	2.190	2.106

P80_mod1	P80_pred1	P80_mod2	P80_pred2	P80_mod1	P80_pred1	P80_mod2	P80_pred2	P80_mod1	P80_pred1	P80_mod2	P80_pred2
2.690	2.751	2.440	2.355	2.830	2.745	2.180	2.300	2.780	2.672	2.550	2.392
2.680	2.619	2.580	2.495	2.720	2.805	2.180	2.301	2.800	2.692	3.560	3.402
2.570	2.632	2.570	2.483	2.720	2.805	3.390	3.267	2.920	2.811	2.500	2.658
2.720	2.658	2.580	2.493	2.590	2.505	2.360	2.485	2.500	2.609	3.200	3.358
2.670	2.608	2.580	2.667	2.560	2.646	2.680	2.806	2.850	2.741	3.560	3.402
2.570	2.633	2.390	2.301	2.760	2.674	2.680	2.806	2.840	2.730	3.170	3.329
2.450	2.513	2.450	2.539	2.750	2.836	2.680	2.806	2.610	2.500	2.480	2.319
2.780	2.716	3.300	3.391	2.580	2.494	2.630	2.503	2.740	2.851	2.890	2.728
2.590	2.525	2.390	2.481	2.790	2.703	3.240	3.367	2.680	2.792	3.190	3.352
2.390	2.455	2.610	2.519	2.890	2.803	2.680	2.808	2.680	2.792	3.240	3.402
2.400	2.465	2.880	2.788	2.570	2.483	3.430	3.301	2.680	2.792	2.621	2.459
2.580	2.514	3.350	3.442	2.730	2.643	2.600	2.732	2.720	2.608	2.420	2.257

2.320	2.386	3.410	3.318	2.790	2.702	2.310	2.442	2.570	2.683	2.840	2.677
2.730	2.664	2.300	2.393	2.540	2.452	2.380	2.512	2.380	2.493	2.980	2.817
2.710	2.644	2.450	2.544	2.740	2.830	2.380	2.512	2.380	2.493	3.150	3.315
2.900	2.833	3.430	3.334	2.710	2.620	2.310	2.443	2.780	2.667	2.730	2.565
2.540	2.608	2.390	2.487	2.730	2.821	2.390	2.524	2.390	2.505	2.250	2.416
2.380	2.449	2.550	2.649	2.390	2.482	2.240	2.106	2.760	2.645	2.760	2.593
2.520	2.589	3.440	3.340	2.390	2.483	2.270	2.405	3.010	2.895	3.200	3.367
2.510	2.580	2.660	2.559	2.600	2.506	2.530	2.394	2.550	2.665	2.810	2.641
2.450	2.520	2.420	2.319	2.730	2.636	2.520	2.657	2.720	2.836	2.480	2.649
2.800	2.730	3.250	3.352	2.780	2.686	2.600	2.463	2.780	2.664	2.330	2.501
2.800	2.730	2.550	2.653	2.480	2.386	3.460	3.322	2.550	2.667	3.510	3.682
2.640	2.711	2.650	2.546	2.850	2.756	2.430	2.569	2.840	2.722	2.640	2.814
2.690	2.763	2.600	2.495	2.580	2.674	2.790	2.650	2.830	2.712	2.470	2.645
2.410	2.484	2.840	2.734	2.920	2.824	2.560	2.420	2.690	2.572	2.660	2.484
2.570	2.645	3.160	3.268	2.490	2.587	3.250	3.390	2.650	2.769	2.580	2.756
2.400	2.477	3.220	3.330	2.660	2.563	2.260	2.401	2.860	2.741	2.340	2.516
2.400	2.478	3.470	3.358	2.530	2.433	2.560	2.416	2.890	2.771	2.630	2.454
2.820	2.739	2.190	2.302	2.510	2.608	2.330	2.474	2.560	2.441	2.670	2.493
2.850	2.769	2.350	2.463	2.480	2.580	2.430	2.575	2.730	2.851	2.500	2.677
2.690	2.771	2.630	2.516	2.890	2.788	2.630	2.484	2.550	2.672	3.090	3.268
2.670	2.588	2.220	2.106	2.600	2.497	3.230	3.376	2.520	2.643	3.110	3.288
2.810	2.728	2.800	2.685	2.800	2.903	2.460	2.314	2.360	2.483	2.250	2.072
2.540	2.623	2.760	2.645	2.950	2.846	2.450	2.300	2.980	2.856	2.400	2.221
2.620	2.704	3.220	3.336	2.430	2.534	2.290	2.443	2.550	2.426	3.220	3.400
2.640	2.724	2.380	2.496	2.700	2.804	2.370	2.524	2.770	2.644	3.120	3.300
2.960	2.876	3.240	3.121	2.430	2.534	2.510	2.355	2.950	2.824	2.690	2.870
2.890	2.805	2.330	2.211	2.350	2.455	2.740	2.585	2.830	2.703	2.340	2.521
2.890	2.805	3.240	3.360	2.710	2.605	2.590	2.435	2.490	2.617	3.100	3.281
2.580	2.665	2.540	2.660	2.600	2.494	2.880	2.723	2.320	2.447	2.340	2.521