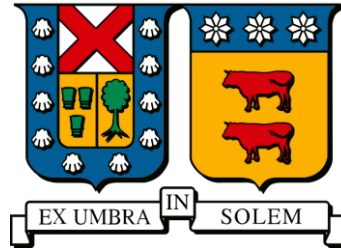


UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INGENIERÍA QUÍMICA Y AMBIENTAL

VALPARAÍSO - CHILE



**“DESARROLLO DE HERRAMIENTA PARA LA
DEFINICIÓN DEL CARGUÍO DE BOLAS EN LA
MOLIENDA SEMI-AUTÓGENA DE CODELCO
CHUQUICAMATA”**

CAMILA FRANCISCA MARÍN GARCÍA

Memoria de Titulación como requisito para optar al título de

INGENIERO CIVIL QUÍMICO

Profesor Guía

Luis Vinnett Peralta

Ingeniero Correferente

Jaime Seguel Herrera

Agosto, 2024

AGRADECIMIENTOS

En primer lugar, quiero agradecer a Dios por darme fortaleza durante toda mi vida, guiarme durante toda mi etapa universitaria, y darme valor ante cada nuevo desafío.

Quiero agradecer a mis padres Lilian y Edmundo, por su apoyo y amor incondicional, y por haberme dado las bases de la persona que soy y la profesional que quiero ser. A mis hermanos, Josefa y Diego, por las risas y la compañía a lo largo de la vida. También agradecer a toda mi familia por su constante apoyo y siempre creer en mí.

Infinitas gracias a Cielo por su cariño y amistad a lo largo de los años, gracias por escucharme y apoyarme, espero seguir creciendo a tu lado. Gracias a Matilda e Ignacio por las risas, consejos y compañía. Gracias a Consuelo, si bien el tiempo juntas fue breve, sé que siempre seguiremos juntas de algún modo. Todo el camino recorrido no sería el mismo sin ustedes, estoy muy agradecida que sean parte de mi vida.

Estoy muy agradecida de Luis, por su comprensión y paciencia, por brindarme apoyo constante y ayudarme en los momentos difíciles. Este proceso hubiera sido mucho más complejo sin su compañía, cariño y ayuda.

Quiero agradecer al *team* materia por las risas, los lindos momentos y las anécdotas que atesoraré por siempre, en especial a Magdalena, Nicolás y Sebastián por su cariño y amistad. También a Valeska, por su apoyo y compañía en los momentos más decisivos.

Muchas gracias al profesor Luis Vinnett, quién me guio y apoyó en todo este proceso con dedicación, compromiso y conocimiento.

También agradecer a los chicos del laboratorio, Alex, Gonzalo y Jorge, con quienes compartí el último tiempo en la universidad.

Finalmente, quiero agradecer al equipo de Codelco División Chuquicamata, en especial a mi tutor Jaime, quién aclaró todas mis dudas y me guio en todo el proceso de memoria. También, quiero agradecer al equipo del área de Analítica Avanzada, en especial Valentina M., Valentina E. y Gonzalo por acogerme en el equipo, aclarar todas mis dudas y siempre estar dispuestos a ayudarme.

RESUMEN

Codelco División Chuquicamata se dedica al procesamiento de mineral de cobre para obtener concentrados de cobre, utilizando técnicas de conminución como la molienda SAG para la liberación de minerales. Para asegurar un adecuado compromiso entre liberación de minerales, eficiencia operacional y optimización de recursos, se utiliza una carga específica de bolas de acero dentro de los molinos SAG.

En este trabajo de memoria se desarrolla una herramienta para definir el carguío de bolas en la molienda SAG, con el fin de contribuir en la eficiencia del proceso y controlar la operación.

Se realizó un análisis de datos históricos y operacionales para caracterizar el comportamiento de un circuito de molienda SAG. Se analizaron variables que afectan el desgaste de las bolas junto con modelos de “*Machine Learning*”, específicamente “*Support Vector Regression*” (SVR) y “*Extreme Gradient Boosting*” (XGBoost). Estos modelos permitieron definir una tasa de desgaste dinámica que se ajustara a las condiciones operacionales del molino, respondiendo a las limitaciones respecto el nivel de llenado de bolas.

Los resultados demostraron que un carguío adecuado de bolas y el ajuste dinámico de la tasa de desgaste contribuyen a tener un mayor control del desgaste de los medios de molienda y su reposición, lo que afectan directamente en el proceso de molienda. El modelo XGBoost mostró un mejor rendimiento que el modelo SVR, por lo que fue seleccionado para su implementación.

Se recomienda utilizar la herramienta desarrollada para ajustar el carguío de bolas en tiempo real, además de continuar alimentando el modelo con nuevos datos para mejorar su precisión y adaptabilidad. Además, es importante capacitar al personal en el uso de la herramienta y la interpretación de los resultados. Este trabajo contribuye a la mejora continua de los procesos de molienda en la División Chuquicamata, alineándose con los objetivos de innovación y productividad de Codelco.

GLOSARIO

CODELCO	Corporación Nacional del Cobre
CSV	Valores Separados por Comas
DCH	División Chuquicamata
<i>Decision trees</i>	Árboles de decisión
<i>Feeder</i>	Alimentador
<i>Grind Out</i>	Condición operacional en la molienda SAG donde el molino se vacía de material, dejando sólo las bolas de molienda en su interior. Se realiza para determinar el J_b
J_b	Nivel de llenado de bolas dentro del molino
Joblib	Librería de Python
<i>Kernel</i>	Núcleo
<i>Lifter bars</i>	Barras levantadoras
Machine Learning	Aprendizaje de máquina
MSE	<i>Mean Squared Error</i> (error cuadrático medio)
<i>Patterns</i>	Patrones
R	Coefficiente de correlación de Pearson
SAG	<i>Semi-Autogenous Grinding</i>
Streamlit	Librería de Python
SVR	Support Vector Regression
TMF	Toneladas Métricas de Fino
XGBoost	<i>Extreme Gradient Boosting</i>

ÍNDICE

AGRADECIMIENTOS	2
RESUMEN	3
GLOSARIO	4
ÍNDICE DE FIGURAS	7
ÍNDICE DE TABLAS	9
1. INTRODUCCIÓN	10
1.1. OBJETIVOS GENERAL	11
1.2. OBJETIVOS ESPECÍFICOS	11
2. ANTECEDENTES.....	12
2.1. CODELCO DIVISIÓN CHUQUICAMATA.....	12
2.2. PROCESO PRODUCTIVO	12
2.2.1. MOLIENDA A2	13
2.3. MOLIENDA SEMIAUTÓGENA	15
2.4. TAMAÑO Y CARGA DE BOLAS	17
2.5. VARIABLES DETERMINANTES EN LA MOLIENDA SEMIAUTÓGENA	18
2.6. MACHINE LEARNING	20
2.6.1. PRINCIPIOS BÁSICOS.....	21
2.6.2. SUPORT VECTOR REGRESSION	22
2.6.3. EXTREME GRADIENT BOOSTING.....	24
3. METODOLOGÍA	25
3.1. ANÁLISIS DE LA PROBLEMÁTICA	25
3.2. PRE-TRATAMIENTO DE DATOS	26
3.2.1. SELECCIÓN DE VARIABLES	26
3.2.2. BASE DE DATOS	27

3.2.3.	LIMPIEZA Y ORGANIZACIÓN DE DATOS	28
3.2.4.	TASA DE DESGASTE	30
3.2.5.	FUENTES DE DATOS	31
3.3.	MODELOS DE MACHINE LEARNING	31
3.3.1.	DISTRIBUCIÓN DE DATOS	31
3.3.2.	SELECCIÓN DEL MODELO	32
3.3.3.	ENTRENAMIENTO Y PRUEBA DEL MODELO	32
3.3.4.	VALIDACIÓN DE LOS MODELOS	33
3.3.5.	ANÁLISIS E IMPLEMENTACIÓN DEL MODELO.....	34
3.4.	APLICACIÓN	34
4.	RESULTADOS	36
4.1.	RESULTADOS DE ENTRENAMIENTO.....	36
4.2.	RESULTADOS DE PRUEBA DE MODELOS	36
4.3.	RESULTADOS DE VALIDACIÓN DE MODELOS	37
4.4.	IMPLEMENTACIÓN	40
5.	CONCLUSIONES Y RECOMENDACIONES.....	42
5.1.	CONCLUSIONES.....	42
5.2.	RECOMENDACIONES	42
	REFERENCIAS	44
	ANEXOS	46
A.	APLICACIÓN VIRTUAL	46

ÍNDICE DE FIGURAS

Figura 2.1 - Diagrama del circuito de molienda A2	14
Figura 2.2 - Secciones de un molino SAG (Wills & Napier-Munn, 2006)	15
Figura 2.3 - (a) Fuerza de impacto aplicada, (b) Fuerza de abrasión aplicada, (d) Impacto, (e) Abrasión, (f) Compresión, (Wills & Napier-Munn, 2006).....	16
Figura 2.4 - Regiones dentro de un Molino SAG, (Wills & Napier-Munn, 2006).....	17
Figura 2.5 - Flujo de operación de un modelo en Machine Learning (Chicoma, 2023)	20
Figura 2.6 - Modelo SVR (Sethi, 2024)	23
Figura 2.7 - Ejemplo de ajuste del Modelo SVR.....	23
Figura 2.8 - Algoritmo XGBoost (Espinosa-Zúñiga, 2020).....	24
Figura 3.1 - Datos anómalos presentes en las bases de datos.....	29
Figura 3.2 - Distribución temporal de los conjuntos de datos.	32
Figura 4.1 - Tasa de desgaste de bolas gr/ton en función del tiempo en días, prueba modelo SVR.	37
Figura 4.2 - Tasa de desgaste de bolas gr/ton en función del tiempo en días, prueba modelo XGBoost.	37
Figura 4.3 - Tasa de desgaste de bolas en función del tiempo, período de validación, modelo SVR	38
Figura 4.4 - Tasa de desgaste de bolas en función del tiempo, período de validación, modelo XGBoost.....	38
Figura 4.5 - Nivel de llenado de bolas Jb en función del tiempo, validación modelo SVR.....	39
Figura 4.6 - Nivel de llenado de bolas Jb en función del tiempo, validación modelo XGBoost.	39
Figura 4.7 - Comportamiento de la tasa de desgaste, implementación modelo XGBoost.	40
Figura 4.8 - Comportamiento de la reposición de bolas, implementación modelo XGBoost ...	41
Figura 4.9 - Comportamiento del Jb, implementación del modelo XGBoost	41
Figura A.1 - Pestaña "Página Inicial" de la aplicación virtual	46
Figura A.2 - Pestaña "Tablero SAG" de la aplicación virtual sección 1	47
Figura A.3 - Pestaña "Tablero SAG" de la aplicación virtual sección 2	47
Figura A.4 - Pestaña "Tablero SAG" de la aplicación virtual sección 3	48
Figura A.5 - Pestaña "Tablero SAG" de la aplicación virtual sección 4	49

Figura A.6 - Pestaña "Cambio data SAG" de la aplicación virtual sección 1	49
Figura A.7 - Pestaña "Cambio data SAG" de la aplicación virtual sección 2	50
Figura A.8 - Pestaña "Cambio data SAG" de la aplicación virtual sección 3	50

ÍNDICE DE TABLAS

Tabla 3.1 - Variables o condiciones operacionales del molino SAG	26
Tabla 3.2 - Mineralogía de alimentación.....	27
Tabla 3.3 - Variables no consideradas en la modelación	27
Tabla 3.4 - Variables extraídas del PI System.....	28
Tabla 3.5 - Variables extraídas de MDT	28
Tabla 3.6 - Organización y tratamiento de datos.....	30
Tabla 3.7 - Conjunto de datos entrenamiento, prueba y validación, y número de datos	32
Tabla 3.8 - Hiperparámetros del modelo SVR	33
Tabla 3.9 - Hiperparámetros del modelo XGBoost	33
Tabla 4.1 - Métricas de bondad de ajuste en el entrenamiento.....	36

CAPITULO 1

1. INTRODUCCIÓN

Codelco División Chuquicamata (DCH) es parte de la Corporación Nacional del Cobre de Chile, CODELCO, empresa estatal pionera en la explotación minera cuprífera a nivel mundial. El yacimiento de Chuquicamata comprende una mina a rajo abierto y una mina subterránea, cuyas reservas alcanzan 1,028 millones de toneladas de mineral mena con una ley de cobre del 0.82%, siendo una de las mayores reservas del mundo de este metal (CODELCO, 2024). El sólido es procesado en distintas etapas de conminución, incluyendo la molienda semi-autógena (SAG), la que emplea bolas de acero y el mismo sólido en la reducción del tamaño de partícula. El molino SAG es un equipo de gran tamaño y capacidad, lo que permite simplificar los circuitos de chancado-molienda.

Codelco DCH dispone de 2 molinos SAG, los cuales procesan menas provenientes del yacimiento de Chuquicamata y del yacimiento de la División Radomiro Tomic con el fin reducir su tamaño, favoreciendo el avance de la liberación de los minerales valiosos. La carga de bolas dentro del molino (J_b) tiene un impacto significativo en la operación y en la eficiencia de la molienda SAG, lo que determina la generación de *pebbles* y la granulometría del producto. El carguío de bolas se define en función del comportamiento del J_b producto del desgaste de las bolas, el que varía según el tonelaje de sólido procesado, su origen, mineralogía y dureza.

En la última década, Codelco DCH ha impulsado tecnologías avanzadas para potenciar los procesos productivos y mejorar su eficiencia. Dentro de éstas, se destaca la implementación de modelos predictivos para propósitos de mantención y proyección de las condiciones operacionales de los equipos. Este trabajo de memoria presenta una estrategia de modelación para definir el carguío de bolas en la molienda SAG. El modelo incorpora las condiciones y variables operacionales del molino, así como las características del sólido procesado. El carguío de bolas se ajusta de manera de obtener condiciones esperadas de J_b (13% - 15%), alineándose con los resultados de los procesos de *Grind Out*. El objetivo es asegurar una operación normal y estable en la molienda SAG.

1.1. OBJETIVOS GENERAL

Desarrollar una herramienta que permita definir y proyectar el carguío de bolas en la molienda SAG.

1.2. OBJETIVOS ESPECÍFICOS

- Realizar un levantamiento de información histórica que permita relacionar variables operacionales con la eficiencia de la molienda SAG.
- Integrar información obtenida de balances de masa con modelos empíricos asociados a la molienda SAG.
- Analizar la sensibilidad de la operación de la molienda SAG ante cambios en los parámetros operacionales y en el carguío de bolas.
- Aplicar herramientas estadísticas y/o de aprendizaje de máquina que permitan definir la carga de bolas en los molinos SAG y que contribuyan a la mejora continua de la reducción de tamaño.

CAPITULO 2

2. ANTECEDENTES

2.1. CODELCO DIVISIÓN CHUQUICAMATA

Codelco DCH es una mina de cobre y oro a rajo abierto perteneciente a Codelco, la cual se ubica a 15 km al norte de la ciudad de Calama, Provincia de El Loa en la Región de Antofagasta, a una altura de a 2,870 m.s.n.m. Fue fundada en mayo de 1915 y considerada una de las más grandes minas del mundo en su tipo. Después de 104 años de operación continua del rajo, en 2019 se inauguró Chuquicamata Subterránea, un proyecto estructural que explota la reserva ubicada bajo el yacimiento de Chuquicamata.

La línea productiva de Codelco DCH se enfoca en la producción de cátodos electro-refinados y electro-obtenidos, con una pureza de 99.9% de cobre, junto con concentrado de cobre y contenido fino de molibdeno. Adicionalmente, como subproducto, produce barros anódicos y ácido sulfúrico. Sólo en 2023, Codelco DCH produjo 248,495 toneladas métricas de cobre fino (CODELCO, 2023).

2.2. PROCESO PRODUCTIVO

El objetivo principal del proceso productivo es obtener cobre de alta concentración en formato de cátodos de cobre electro-refinados y electro-obtenidos, los cuales alcanzan una pureza de 99.9%. Este proceso está conformado por distintas etapas, las que dependen del tipo de mineral. Los minerales oxidados atraviesan un proceso de chancado, de lixiviación, extracción por solventes y finalmente, un proceso de electro-obtención para obtener el producto final. Los minerales sulfurados atraviesan procesos de chancado (primario, secundario y terciario), molienda, flotación, fundición y finalmente, un proceso de electro-refinación para obtener el producto final. Adicionalmente, se suman otros procesos como la planta de molibdeno, planta de filtros, planta de tratamiento de escorias, y espesadores de concentrado y relave.

En Codelco DCH, la etapa de molienda se divide en tres áreas: molienda A0, molienda A1 y molienda A2. Tanto la molienda A0 como la molienda A1 se caracterizan por ser moliendas del tipo convencional, mientras que la molienda A2 corresponde a una molienda del tipo semi-

autógena. Los molinos son alimentados mediante *feeders* desde pilas de acopio o *stockpiles*, donde llega el mineral desde cada origen.

2.2.1. MOLIENDA A2

En la Figura 2.1 se ilustra el proceso de Molienda A2 simplificado. Este proceso está constituido por 2 molinos SAG, una planta de *pebbles* y un molino de bolas o quinto molino. Los molinos SAG, a su vez, constituyen dos sistemas independientes, cada uno con 2 harneros, 1 cuba, 2 molinos de bolas y 1 batería de hidrociclones.

Los molinos SAG de Codelco DCH son alimentados continuamente con agua y mineral mena. Este último ingresa desde 3 alimentadores o *feeders*, los que descargan el sólido desde la Pila A2 por correas transportadoras hacia los molinos. Dentro del SAG, el sólido es molido por acción del impacto producido por las bolas de acero y por las rocas que se encuentran dentro del molino. El producto es enviado a los harneros (Norte y Sur), donde se clasifica el material en *pebbles* y material fino. Los *pebbles* son redirigidos a su respectiva planta, mientras que el material fino cae por los harneros a una cuba donde se adiciona agua, y la pulpa resultante es transportada mediante un sistema de 8 bombas a la batería de hidrociclones de la molienda secundaria.

El proceso de carguío de bolas se realiza de forma periódica según las condiciones de operación del molino SAG, el cual se alimenta de manera discreta, mediante baldadas de aproximadamente 12.5 toneladas. Estas baldadas son descargadas por una tolva que dirige el carguío de bolas a las correas transportadoras de alimentación fresca al molino SAG. Una vez que las bolas han alcanzado un tamaño crítico producto del desgaste, estas son descargadas junto con el producto del SAG, donde son separadas producto de la acción de un campo magnético a la salida de los harneros.

La planta de *pebbles* está constituida por 3 tolvas que alimentan los *pebbles* provenientes de los molinos SAG a 3 chancadores de cono. A la salida del proceso de chancado, se tienen dos correas transportadoras: una que recircula el producto hacia los molinos SAG, y otra que lleva el producto al quinto molino. El producto recirculado a la molienda SAG es transportado por correas hacia un cajón de acopio, desde donde se distribuye el producto a dos correas transportadoras. Cada correa, a su vez, alimenta a la corriente de alimentación del molino SAG

respectivo, mezclándose el producto recirculado de la planta de *pebbles* junto con el mineral fresco proveniente de la Pila A2.

La batería de hidrociclones de la molienda secundaria está compuesta por 14 hidrociclones, de los cuales 11 están dispuestos para clasificación y los 3 restantes para transferir material directamente al Cajón 501. En la batería de hidrociclones, el producto fino u *overflow* es trasladado al Cajón 501, mientras que el material grueso o *underflow* es enviado a los molinos de bolas, cuyo producto es devuelto a la cuba inicial.

El quinto molino es parte de un circuito que está compuesto por el molino, el Cajón 501 y una batería de hidrociclones. El Cajón 501 es alimentado con agua y por las corrientes provenientes de la etapa anterior. De este cajón se impulsa pulpa mediante 2 bombas hacia la batería de hidrociclones. El *overflow* de la clasificación se dirige a la etapa de flotación, mientras que el *underflow* es enviado al quinto molino. El quinto molino se alimenta también con el producto del chancado de *pebbles* y agua. El producto de este molino ingresa nuevamente al Cajón 501.

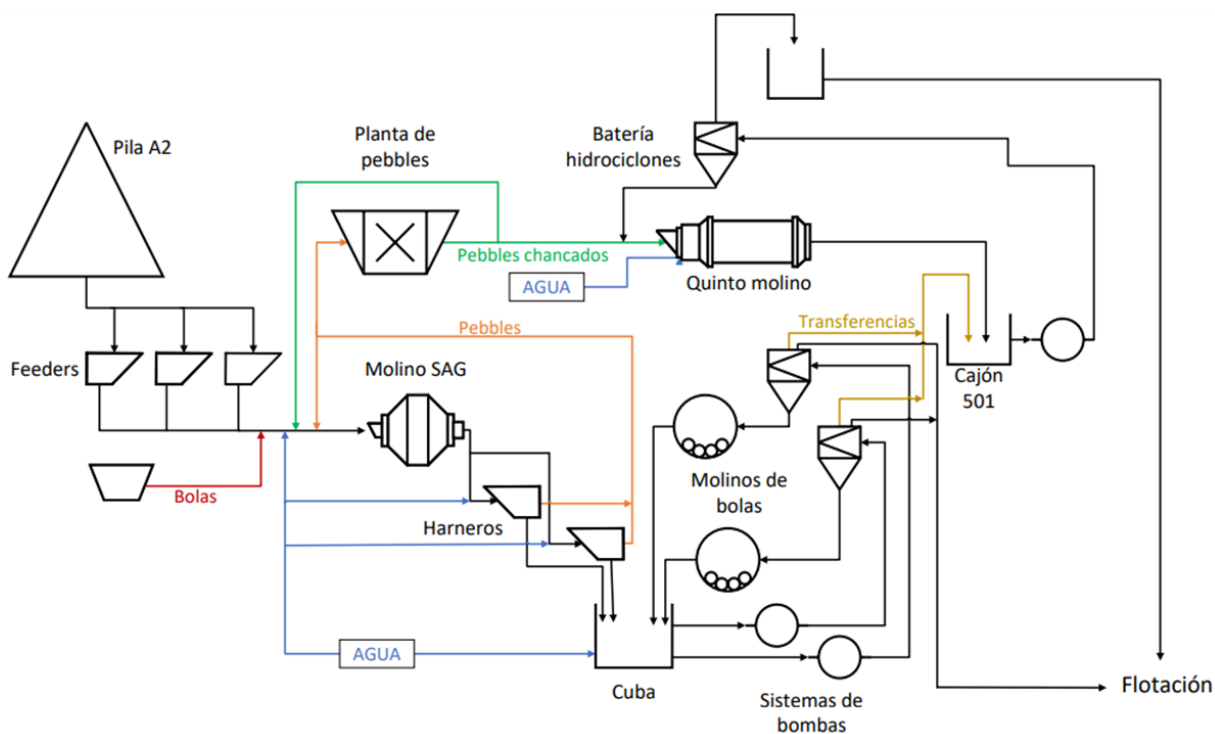


Figura 2.1 - Diagrama del circuito de molienda A2.

2.3. MOLIENDA SEMIAUTÓGENA

El molino SAG utiliza bolas de acero y las rocas presentes en el mineral mena como medios de molienda. Este equipo posee una geometría cilíndrica y rota alrededor de su eje. Su interior está cubierto por un revestimiento compuesto por placas que protegen la estructura interna y facilitan el movimiento de la carga. Esto último se logra mediante barras levantadoras o *lifter bars*, las que elevan la carga con la rotación. Los molinos SAG, típicamente poseen una rejilla en la descarga, la que permite la salida de la pulpa, manteniendo las bolas de acero al interior (Wills & Napier-Munn, 2006). En la Figura 2.2 se observan las secciones de un molino semi-autógeno.

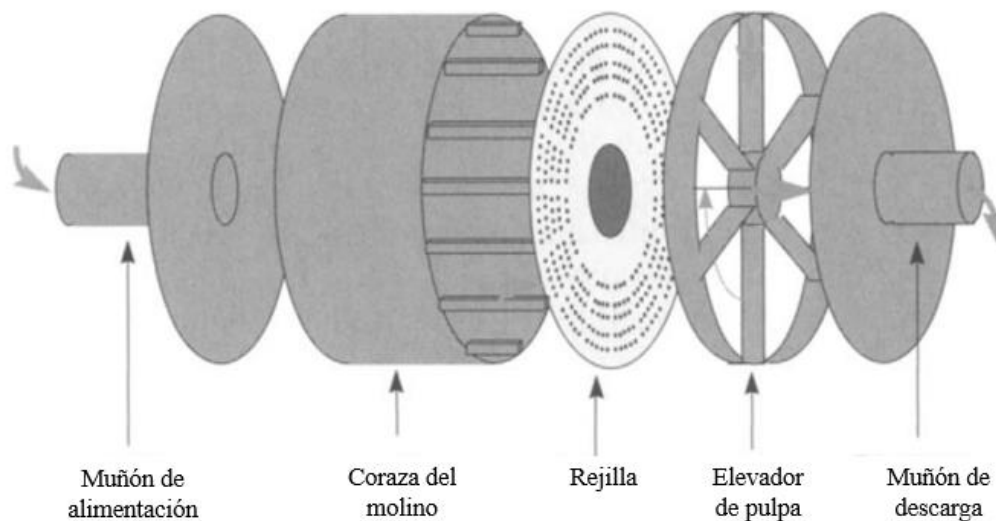


Figura 2.2 - Secciones de un molino SAG (Wills & Napier-Munn, 2006).

La molienda se realiza en húmedo, y la reducción de tamaño se debe principalmente a los mecanismos de impacto, abrasión y compresión. La fractura por impacto se produce cuando la energía aplicada es mayor a la necesaria para fracturar la partícula, lo que resulta en fragmentos de diferentes tamaños. La reducción de tamaño por abrasión ocurre cuando las fuerzas son aplicadas paralelamente a la superficie de las partículas, siendo la energía insuficiente para fracturar la partícula, lo que resulta en partículas finas debido a los esfuerzos localizados sobre la superficie. La fractura por compresión se produce cuando la energía aplicada sólo alcanza pocas regiones de la partícula alcanzando un punto localizado de fractura, lo que resulta en que la partícula se fracture en unos pocos fragmentos gruesos (Kawatra, 2006). En las Figuras 2.3(a), 2.3(b) y 2.3(c) se ilustra la aplicación de las fuerzas de impacto, abrasión y compresión,

respectivamente, mientras que en las Figuras 2.3(d), 2.3(e) y 2.3(f) se presentan los resultados en la reducción de tamaño.

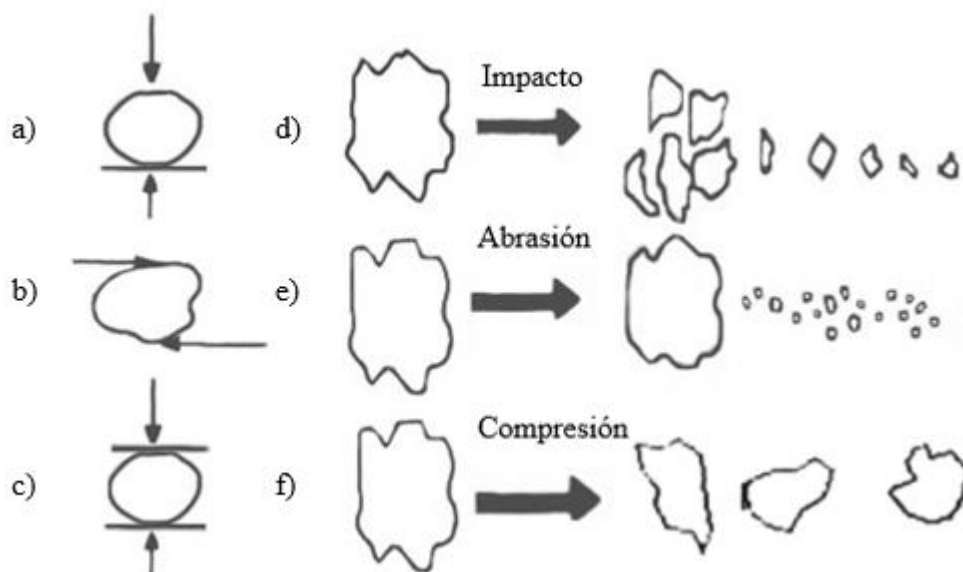


Figura 2.3 - (a) Fuerza de impacto aplicada, (b) Fuerza de abrasión aplicada, (d) Impacto, (e) Abrasión, (f) Compresión, (Wills & Napier-Munn, 2006).

Dentro del molino, y debido a la rotación y a la presencia de los *lifters*, las bolas se elevan hasta alcanzar una posición de equilibrio dinámico, cayendo en cascada o en cataratas hasta la base de la carga. El efecto de cascada se produce a velocidades bajas o con revestimientos desgastados (lisos), donde las bolas no tienden a elevarse, lo que reduce tamaño principalmente por abrasión, produciendo material más fino. El efecto catarata se produce a velocidades altas, donde las bolas siguen una trayectoria parabólica antes de caer por gravedad a la base del molino, llevando a cabo fracturas por impacto y produciendo un material más grueso. En general, ambos efectos coexisten dentro de un molino SAG. En la Figura 2.4, se identifican las regiones dentro del molino. La zona de abrasión es donde las partículas entran en contacto con el revestimiento del molino, produciéndose una conminución abrasiva. La zona con efecto cascada es donde tanto las bolas como el material caen en cascada por la superficie libre, produciéndose también una molienda por abrasión. La zona de la base del molino es donde las bolas y el sólido se acumulan, produciéndose una molienda por impacto. La zona de impacto es donde se produce la conminución por impacto, luego que las bolas tras ser levantadas por la rotación del molino caen de nuevo a la base. La zona de efecto catarata es donde el material y

las bolas siguen una trayectoria parabólica hacia la zona vacía, lo que lleva a la molienda por impacto en la base del molino.

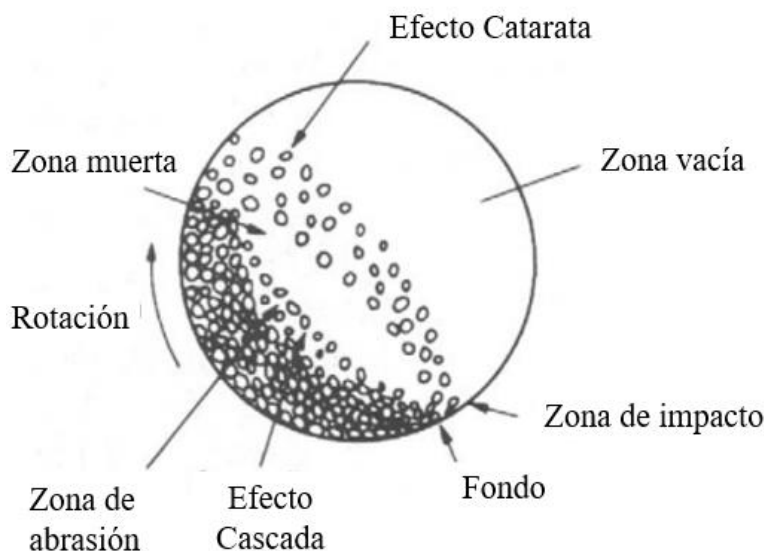


Figura 2.4 - Regiones dentro de un Molino SAG, (Wills & Napier-Munn, 2006).

2.4. TAMAÑO Y CARGA DE BOLAS

Las bolas son uno de los medios de molienda del molino SAG, las que participan críticamente en las acciones de abrasión e impacto en la reducción de tamaño. Esto conlleva, al mismo tiempo, a su desgaste, siendo necesaria su reposición o carguío para mantener una carga de bolas que garantice una adecuada operación del molino. El tamaño y la carga de bolas son factores determinantes en la reducción de tamaño del mineral mena, e influyen directamente en la eficiencia del proceso.

El tamaño y aleación de las bolas determinan la fractura del sólido. A mayor tamaño, las bolas presentan una mayor capacidad de fractura por impacto dada su mayor masa, mientras que bolas pequeñas presentan una mayor capacidad de abrasión debido a su mayor superficie de contacto. Las bolas de alta aleación, compuestas normalmente por cromo, níquel y molibdeno, producen impactos de mayor energía, generando fracturas que permiten la fragmentación de partículas de mayor tamaño. Por otro lado, las bolas de aleación menos densa generan abrasión, desgastando la superficie del sólido sin generar fracturas muy profundas. Un equilibrio en el tamaño y aleación de las bolas evita un desgaste acelerado del revestimiento del molino y un mayor

consumo de energía, lo que impacta no sólo la eficiencia del proceso si no también los costos de operación (Kawatra, 2006).

La carga de bolas se refiere a la cantidad de bolas presentes dentro del molino. Esta se mide como el porcentaje del volumen interno del molino utilizado por las bolas, J_b . Una correcta carga de bolas, así como también un correcto carguío de bolas es crucial para asegurar un adecuado proceso de molienda. Una baja carga de bolas implica menos impactos entre el medio de molienda y el sólido, reduciendo la eficiencia del proceso e incrementando el desgaste del revestimiento por acción del material. Por otro lado, una alta carga de bolas impide que el choque entre el medio de molienda y el sólido se lleve a cabo, provocando que las bolas choquen entre sí, y se fracturen, reduciendo también la eficiencia de la molienda y aumentando el consumo de bolas en el proceso (Wills & Napier-Munn, 2006).

2.5. VARIABLES DETERMINANTES EN LA MOLIENDA SEMIAUTÓGENA

La molienda SAG depende de diversas variables de diseño y operación, las cuales impactan en la eficiencia del proceso y en la vida útil del equipo, así como de sus componentes.

✓ Velocidad de giro.

La velocidad incide directamente en el efecto cascada que ocurre dentro del molino. Existe una velocidad crítica que representa la menor velocidad a la cual las fuerzas centrífugas que actúan sobre las bolas son iguales a las fuerzas gravitatorias; es decir, las bolas se adhieren a las paredes internas del molino y no caen, impidiendo que se genere el efecto cascada. Una velocidad mayor a la velocidad crítica impide la molienda (Cárdenas-Valderrama, 2024).

✓ Mineralogía.

La mineralogía abarca la dureza y la composición del mineral mena. Esta última indica el porcentaje de las componentes presentes en la matriz, tales como calcopirita, calcosina, pirita, cuarzo, entre otros. La dureza del sólido depende generalmente de su origen. Esta dureza determina la energía necesaria para llevar a cabo la molienda, y condiciona un mayor o menor desgaste de los revestimientos del molino y de los medios de molienda (Codelco, 2021).

✓ Distribución granulométrica de la alimentación.

La distribución granulométrica de la alimentación determina el comportamiento de la carga del molino e incide en la eficiencia de la molienda. Un material grueso permite una carga menos densa, lo que se traduce en un mayor flujo de pulpa y capacidad de tratamiento. Además, actúa como medio de molienda, aumentando la acción de fractura por impacto y compresión. Por otro lado, el material fino reduce la acción de impacto y compresión necesaria para llevar a cabo una molienda efectiva, reduciendo la capacidad de tratamiento. Un equilibrio en la distribución granulométrica en la alimentación permite una liberación óptima del mineral y una capacidad de tratamiento adecuada.

✓ Carga de bolas.

La carga de bolas define la eficiencia de la molienda. Existe una carga óptima para maximizar la eficiencia, la que depende del molino y del mineral mena. La distribución de tamaño de bolas dentro del molino es también un factor influyente. Distintos tamaños de bolas en el equipo favorecen la combinación de impactos con el sólido (Codelco, 2021).

✓ Potencia.

La potencia está directamente relacionada con el nivel de llenado del molino; esto es, la proporción del volumen interno ocupado por los medios de molienda y el sólido procesado. Indica la energía necesaria para mantener la carga en movimiento dentro del molino y vencer los desplazamientos internos de las bolas y el mineral.

✓ Porcentaje de sólidos.

El porcentaje de sólidos determina el flujo y la viscosidad de la pulpa dentro del molino. Una adecuada concentración de sólidos dentro del molino favorece la interacción entre el sólido y los medios de molienda, mejorando la eficiencia del proceso (Codelco, 2021).

✓ Presión de los descansos.

La presión de los descansos se relaciona con la carga total del molino. Depende de la carga de sólido y de los medios de molienda dentro del equipo. Una sobrecarga del molino conlleva a un mayor desgaste de los revestimientos del equipo, así como también a dificultades en el proceso de molienda, lo que compromete las acciones de impacto y abrasión (Codelco, 2021).

2.6. MACHINE LEARNING

Machine Learning o aprendizaje de máquina corresponde a una subcategoría de la inteligencia artificial, la que es ampliamente utilizada en el análisis de conjuntos de datos masivos. Consiste en el desarrollo de algoritmos para llevar a cabo una tarea a partir de un conjunto de datos, con el fin de obtener resultados y tomar decisiones de forma autónoma. Este conjunto de datos atraviesa previamente un tratamiento de limpieza y estructuración para evitar sesgos. Su ejecución va dirigida al entrenamiento sucesivo de un algoritmo, detectando patrones recurrentes dentro de un conjunto de datos (*patterns*), para luego realizar análisis predictivos. De esta forma, los algoritmos aprenden y mejoran su rendimiento en la ejecución de la tarea asignada.

Su relevancia en la industria minera y en los procesos de molienda radica en que permite ajustar parámetros de operación en tiempo real, logrando un mejor control en la operación.

El desarrollo de un modelo de *machine learning* consiste en 5 etapas (Bishop, 2006):

1. Recolección, limpieza y estructuración del conjunto de datos.
2. Selección de un algoritmo de *machine learning*.
3. Entrenamiento del modelo.
4. Evaluación del modelo y análisis de los resultados.
5. Implementación y monitoreo (seguimiento) del modelo.

En la Figura 2.5 se ilustra la línea de operación y desarrollo de un modelo de *machine learning*, desde la entrada de información hasta su implementación y monitoreo.

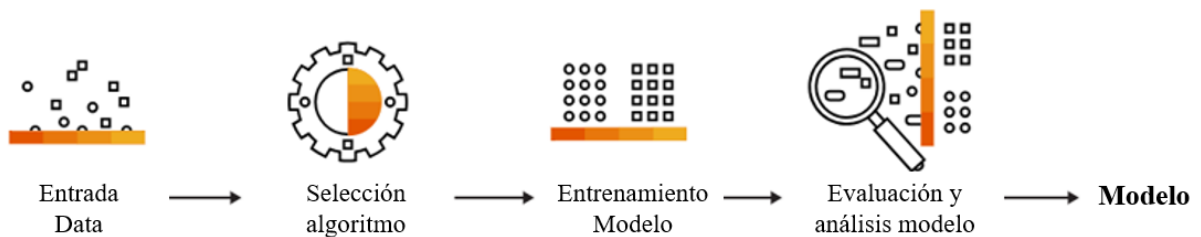


Figura 2.5 - Flujo de operación de un modelo en Machine Learning (Chicama, 2023).

2.6.1. PRINCIPIOS BÁSICOS

Las herramientas de *machine learning* están compuestas por distintos modelos de aprendizaje automático. Dependiendo de las características del conjunto de datos y la tarea a realizar, es posible utilizar uno de los siguientes esquemas de aprendizaje:

- Aprendizaje supervisado.

Los algoritmos que se rigen por el aprendizaje supervisado son entrenados usando un conjunto de datos etiquetados, por lo que cada entrada está asociada a una salida. En este caso, el algoritmo, luego de ser entrenado, predecirá una salida correspondiente a cada nueva entrada que se le adjudique (Bishop, 2006).

- Aprendizaje no supervisado.

Los algoritmos que se rigen por el aprendizaje no supervisado son entrenados usando un conjunto de datos sin etiquetar. En este caso, el algoritmo debe detectar automáticamente *patterns* o patrones presentes en el conjunto de datos y organizarlos (Bishop, 2006).

- Aprendizaje por refuerzo.

Los algoritmos que se rigen por aprendizaje por refuerzo se entrenan al interactuar con el entorno, realizando acciones y aprendiendo mediante un proceso de prueba y error, donde se recompensan las decisiones correctas (Bishop, 2006).

Existe una gran variedad de modelos de *machine learning*, dentro de los que se destacan:

- Árboles de decisión o *decision trees*.
- Redes neuronales.
- Máquina de vectores de soporte o *Support Vector Machine* (SVM).
- *Extreme gradient boosting* (XGBoost).

El modelo *decision trees* clasifica los datos en secciones o nodos a través de reglas de decisión en forma de árbol, por lo que es útil para relaciones simples. Sin embargo, tiende a sobreajustarse ante datos complejos o con ruido, lo que impide un buen manejo de relaciones no lineales complejas (Kelleher, Mac-Namee, & D'Arcy, 2015).

Los modelos de redes neuronales están formados por capas de neuronas artificiales, conformadas a su vez por arreglos de neuronas que comparten información e interactúan entre sí, por lo que es útil para aplicaciones que necesitan precisión ante datos complejos. Sin

embargo, necesita una gran cantidad de datos calidad para poder identificar patrones de manera efectiva, lo que implica un alto costo computacional y una mayor complejidad en la aplicación industrial (Steele, Chandler, & Reddy, 2016).

En este estudio, se seleccionaron los modelos SVR (extensión del modelo SVM) y XGBoost dado a que presentan un mejor rendimiento ante conjuntos de datos complejos y relaciones no lineales. Ambos métodos utilizan técnicas de regularización para evitar el sobreajuste, lo que suele ser un problema común al utilizar modelos de redes neuronales y *decision trees*. Si bien las redes neuronales pueden implementar regularización, los modelos seleccionados suelen ser más efectivos debido a su control específico de complejidad, es decir, los modelos SVR y XGBoost manejan de manera estructurada su capacidad para ajustarse a los datos sin volverse demasiado complejos.

2.6.2. SUPORT VECTOR REGRESSION

Support Vector Regression o regresión de vectores de soporte (SVR) corresponde a una extensión del modelo SVM utilizada en problemas de regresión. Está compuesto por un conjunto de algoritmos de aprendizaje supervisado, pudiendo ser empleado tanto en regresiones lineales como no lineales. Para trabajar con regresiones no lineales, el modelo utiliza *kernels* para situar los datos en un espacio de dimensionalidad más alta (Han, Kamber, & Pei, 2012).

El modelo SVR tiene como objetivo encontrar una función que aproxime la relación entre las variables de entrada y la variable objetivo continua. Su operación radica en situar las variables de entrada en un espacio de dimensionalidad muy alta mediante una función *kernel* y trazar un hiperplano que minimice la distancia entre éste y los datos más cercanos, manteniendo el error dentro de un margen tolerable. Además, se establecen límites de decisión, que se ubican a una distancia ϵ del hiperplano, los que determinan un margen de tolerancia dentro del cual deben encontrarse los datos (Sethi, 2024). En la Figura 2.6 se ilustra como operaría el modelo SVR, con sus respectivos componentes: el hiperplano y los límites de decisión.

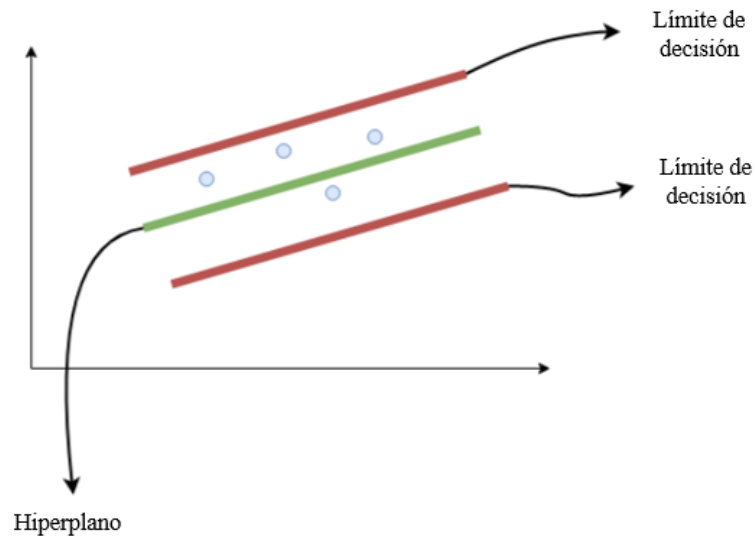


Figura 2.6 - Modelo SVR (Sethi, 2024).

El mejor ajuste corresponderá al hiperplano que incluya la mayor cantidad de puntos dentro del margen de tolerancia, es decir, aquellos con menor error. En la Figura 2.7, se muestra un ejemplo de hiperplano con mejor ajuste (de color azul). Los datos ubicados dentro del margen de tolerancia se representan como puntos rojos, mientras que el dato fuera del margen se presenta como punto rojo enmarcado en un recuadro naranja.

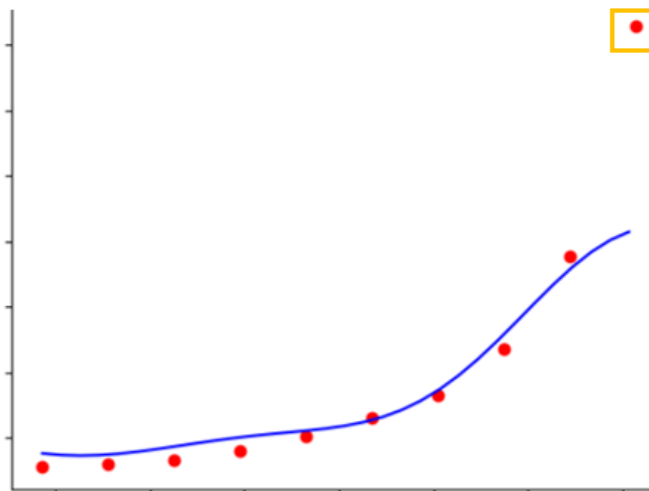


Figura 2.7 - Ejemplo de ajuste del Modelo SVR.

2.6.3. EXTREME GRADIENT BOOSTING

Extreme Gradient Boosting es una implementación de algoritmos de *Boosting* de gradientes, donde el método *Boosting* combina secuencialmente varios modelos débiles, como lo son los árboles de decisión, con el fin de crear un modelo fuerte. Este modelo construye un conjunto de árboles de decisión, donde cada árbol clasifica los datos en secciones o nodos mediante reglas. El modelo *XGBoost* sigue un aprendizaje supervisado para regresión.

El modelo *XGBoost* crea árboles de decisión de forma aditiva, donde cada nuevo árbol responde a corregir los errores residuales de los árboles anteriores. Su operación radica en ajustar un árbol inicial a los datos, para luego construir un segundo árbol que se enfoque en realizar las predicciones que el primer árbol predijo erróneamente, y así sucesivamente. El proceso se lleva a cabo hasta llegar a dos modelos consecutivos con diferencias insignificantes, o hasta que el número máximo de iteraciones previamente establecidas se haya cumplido (Distributed Machine Learning Community, 2024). De esta forma, la combinación de modelos será mejor que cada modelo individual. La Figura 2.8 se muestra cómo opera el algoritmo *XGBoost*, combinando secuencialmente n árboles de decisión para alcanzar un modelo más fuerte para Y .

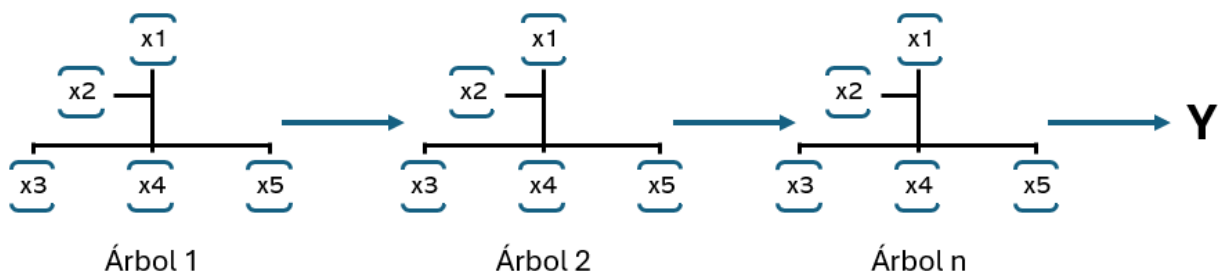


Figura 2.8 - Algoritmo *XGBoost* (Espinosa-Zúñiga, 2020).

CAPITULO 3

3. METODOLOGÍA

3.1. ANÁLISIS DE LA PROBLEMÁTICA

La molienda SAG es una etapa crucial en procesamiento de minerales, la que incide en la reducción del tamaño del mineral mena, la capacidad de procesamiento de una planta y la eficiencia del proceso. Por esta razón, el carguío o reposición de bolas es de suma importancia, siendo un factor determinante para mantener un nivel adecuado de bolas dentro del molino. Este nivel debe encontrarse entre el 12% y un 14%, dependiendo del desgaste del revestimiento.

Actualmente, la reposición de bolas en Codelco División Chuquibambilla se realiza diariamente en lotes de aproximadamente 12.5 ton cada uno, dado que la planta no cuenta con un contador de que permita controlar y cuantificar con exactitud las bolas que ingresan al molino SAG. El carguío de bolas se realiza en base al desgaste de las bolas, cuya estimación se obtiene a partir del producto entre el tonelaje de alimentación y la tasa de desgaste. La tasa de desgaste de bolas es medida en gramos de acero consumido por toneladas de sólido procesado. En este caso, se asume un valor fijo de 328 g/ton.

En el desgaste de las bolas inciden una gran cantidad de variables, como la velocidad y la potencia del molino. Por lo tanto, limitar su cálculo sólo al tonelaje de alimentación y a una tasa de desgaste constante, reduce la confiabilidad del carguío de bolas, afectando la eficiencia en la molienda. En este contexto, resulta crítico poder estimar una tasa dinámica de desgaste de bolas, que se ajuste a la operación del molino y a las características del sólido procesado, respetando los límites para el nivel de llenado. La implementación de un modelo de *Machine Learning* representa una solución atractiva para abordar la problemática del desgaste de las bolas y, en consecuencia, de su carguío.

3.2. PRE-TRATAMIENTO DE DATOS

Se utilizó un conjunto de datos que incorporara todas las variables que inciden en el desgaste de bolas, y de tamaño suficiente grande para entrenar el modelo.

3.2.1. SELECCIÓN DE VARIABLES

Se seleccionaron 36 variables, de las cuales 14 se asocian a la mineralogía de alimentación y 22 a la operación del molino. De estas, se consideraron 32 variables para modelación, 14 correspondientes a la mineralogía y 18 correspondientes a la operación. Las 4 variables restantes no se consideraron en la modelación, pero se utilizaron en la generación de gráficos y análisis posteriores. En la Tabla 3.1 se enlistan las variables relacionadas a la operación, donde la variable “Operación *feeders*” hace referencia a la operación de los 3 *feeders* que posee el molino. En la Tabla 3.2 se enlistan las componentes mineralógicas principales del mineral mena.

Tabla 3.1 - Variables o condiciones operacionales del molino SAG.

Tonelaje acumulado.	Presión de descansos.
<i>Pebbles</i> no chancados recirculados.	SPI.
<i>Pebbles</i> chancados recirculados.	Operación <i>feeders</i> .
Tonelaje de tratamiento.	Mineral mena de Radomiro Tomic.
%Sólido.	Mineral mena de mina Rajo.
Alimentación fresca.	Mineral mena de mina Subterránea.
%Material grueso.	Velocidad.
%Material intermedio.	Potencia.
%Material fino.	Reposición de bolas.
Nivel de llenado de bolas.	Tasa de desgaste.

Tabla 3.2 - Mineralogía de alimentación.

%Muscovita.	%Blenda.
%Bornita.	%Pirita.
%Calcosina.	%Esmectita.
%Fierro.	%Diogenita.
%Covelina.	%Enargita.
%Caolinita.	%Cobre.
%Calcopirita.	%Molibdeno.

En la Tabla 3.3 se enlistan las 4 variables restantes. Las variables “%Material fino” y “Mineral mena de mina Rajo” no se consideraron como datos de entrada del modelo dado que tienen una relación lineal con otras variables. La primera con las variables “%Material grueso” y “%Material intermedio”, y la segunda con las variables “Mineral mena de Radomiro Tomic” y “Mineral mena de mina Subterránea”.

Tabla 3.3 - Variables no consideradas en la modelación.

%Material fino.	Alimentación fresca.
Mineral mena de mina Rajo.	Nivel de llenado de bolas.

3.2.2. BASE DE DATOS

Los datos se recolectaron de 3 fuentes distintas. Las variables extraídas desde la base de datos *PI System* se encuentran tabuladas en la Tabla 3.4, mientras que las variables extraídas de la base de datos *Master Data Table* (MDT) se presentan en la Tabla 3.5. Los datos referentes a la “Reposición de bolas” y al “Nivel de llenado de bolas” se extrajeron de una planilla en formato Excel utilizada por los operadores. La “tasa de desgaste” corresponde a un valor fijo.

Tabla 3.4 - Variables extraídas del *PI System*.

Variables PI System.	
<i>Pebbles</i> no chancados recirculados.	Presión de descansos.
<i>Pebbles</i> chancados recirculados.	Operación feeders.
Alimentación fresca.	Velocidad.
%Material fino.	Potencia.
%Material grueso.	Tonelaje acumulado.
%Material intermedio.	%Sólido.
Tonelaje de tratamiento.	

Tabla 3.5 - Variables extraídas de MDT.

Variables MDT.	
%Muscovita.	%Blenda.
%Bornita.	%Pirita.
%Calcosina.	%Esmectita.
%Fierro.	%Diogenita.
%Covelina.	%Enargita.
%Caolinita.	%Cobre.
%Calcopirita.	%Molibdeno.
SPI.	Mineral mena de mina Subterránea.
Mineral mena de Radomiro Tomic.	Mineral mena de mina Rajo.

3.2.3. LIMPIEZA Y ORGANIZACIÓN DE DATOS

La limpieza de datos consistió en la eliminación de datos duplicados, valores negativos y datos anómalos. La Figura 3.1 ilustra un caso ejemplo. Este proceso de limpieza garantizó que todos los datos correspondieran a períodos de tiempo con el molino en operación. Además de los procesos de limpieza mencionados, se restringió el conjunto de datos a los percentiles 5% al 95% para cada variable.

En el caso específico de las variables “Presión de descansos” y “Tonelaje acumulado”, se realizó una limpieza adicional. En el caso de la primera variable, la limpieza consistió en limitar los datos al rango 40 a 80 kg/cm², eliminando los que no cumplieran esta condición. Esto se realizó debido a datos anómalos que resultan de la sensibilidad de los instrumentos ante variaciones de las condiciones del equipo. En el caso del “Tonelaje acumulado”, la limpieza consistió en limitar los datos a aquellos en que los valores de la variable “Alimentación fresca” fuesen mayores a cero. Los datos que no cumplieron esta condición indicaron que el molino no se encontraba en operación.

10-01-2022 13:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O
10-01-2022 14:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O
10-01-2022 15:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O
10-01-2022 16:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O
10-01-2022 17:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O
10-01-2022 18:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O
10-01-2022 19:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O
10-01-2022 20:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O
10-01-2022 21:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O
10-01-2022 22:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O
10-01-2022 23:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O
11-01-2022 0:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O
11-01-2022 1:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O
11-01-2022 2:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O
11-01-2022 3:00	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O Timeout	I/O

Figura 3.1 - Datos anómalos presentes en las bases de datos.

Los datos se reorganizaron de manera de emplear una tasa de muestreo única. De las bases de datos utilizadas, dos de ellas proporcionaron los datos cada una hora (*PI System* y *MDT*), y una cada un día (planilla de operadores). Por esta razón, los datos se estructuraron en formato de una muestra por día, correspondiente al turno de 5am a 5am.

El proceso de tratamiento de datos consistió en calcular la media, la sumatoria y el máximo diario de las variables de interés. En la Tabla 3.6 se tabulan todas las variables recolectadas, y su respectivo tratamiento.

Tabla 3.6 - Organización y tratamiento de datos.

Variable.	Tratamiento.	Variable.	Tratamiento.
Pebbles no chancados recirculados.	Sumatoria.	%Muscovita.	Media.
Pebbles chancados recirculados.	Sumatoria.	%Bornita.	Media.
Tonelaje de tratamiento.	Sumatoria.	%Calcosina.	Media.
%Material fino.	Media.	%Fierro.	Media.
%Material grueso.	Media.	%Covelina.	Media.
%Material intermedio.	Media.	%Caolinita.	Media.
Presión de descansos.	Media.	%Calcopirita.	Media.
Operación feeders.	Media.	SPI.	Media.
Velocidad.	Media.	%Molibdeno.	Media.
Potencia.	Media.	%Blenda.	Media.
Tonelaje acumulado.	Máximo.	%Pirita.	Media.
%Sólido.	Media.	%Esmectita.	Media.
Mineral mena de mina Subterránea.	Sumatoria.	%Diogenita.	Media.
Mineral mena de mina Rajo.	Sumatoria.	%Enargita.	Media.
Mineral mena de Radomiro Tomic.	Sumatoria.	%Cobre.	Media.
Alimentación fresca.	Sumatoria.		

3.2.4. TASA DE DESGASTE

La tasa de desgaste utilizada por la empresa corresponde a un valor fijo histórico que representa la media de la tasa de desgaste. Esta aproximación se debe a la dificultad técnica de poder obtener un valor real. Por esta razón, para este estudio se optó por obtener el valor diario de la tasa de desgaste según el tonelaje de alimentación, la reposición de bolas y el nivel de llenado de bolas diario, utilizando la función *Solver* de Excel. Para llevar a cabo el cálculo de la tasa de desgaste, se utilizó la Ecuación (3.1).

$$J_{bi-1} + ((R_i - ((T_i * TD_i)/1000))/1000/F_a) - J_{bi} = 0 \quad (3.1)$$

donde J_b corresponde al nivel de llenado de bolas, R la reposición de bolas, T el tonelaje de tratamiento diario, F_a el factor de ajuste e i hace referencia al i -ésimo día. El factor de ajuste indica la relación entre la variación de nivel de llenado y la masa de bolas, donde 19 kg de bolas representa la variación de 1% en el nivel de llenado de bolas. A partir de los resultados derivados de la Ecuación (3.1) se obtuvieron valores de tasa de desgaste para cada día estudiado, los que difieren del valor fijo histórico establecido. Esto puede ser explicado debido a la inexactitud del tonelaje de reposición de bolas, efectos de *Ramp Up* en la alimentación, condiciones de *Grind Out* y posibles errores en la toma de datos de alimentación.

3.2.5. FUENTES DE DATOS

Para alimentar al modelo, se creó una nueva fuente de datos que alberga todas las variables mencionadas anteriormente. Las bases de datos utilizadas consideraron información de las variables desde el año 2020. Durante agosto de 2022, la planta sufrió el colapso de la correa transportadora que alimentaba a la Pila Mina, generando alteraciones en la medición y adquisición de datos. En el contexto de la contingencia, se realizaron modificaciones en los sistemas de control y monitoreo de la planta, lo que permitió tener datos de mejor calidad en comparación a los ya existentes previo al incidente. Los datos seleccionados para alimentar la fuente de datos comprenden el período desde septiembre de 2022 a junio de 2024.

3.3. MODELOS DE MACHINE LEARNING

A continuación de la selección y organización de los datos, se aplicó la selección, entrenamiento y ejecución de los modelos de *Machine Learning*. Luego, una vez obtenidos los primeros resultados, se llevó a cabo un análisis y la prueba del modelo. Finalmente, se realizó el reentrenamiento y mejora del modelo.

3.3.1. DISTRIBUCIÓN DE DATOS

Previo al desarrollo del modelo, se seleccionaron y asignaron los datos a utilizarse en el entrenamiento, prueba y validación. En la Tabla 3.7 se detallan los períodos usados para entrenamiento, prueba y validación, junto con el número de datos asociados a cada conjunto. En la Figura 3.2 se esquematiza la distribución temporal de cada conjunto de datos.

Tabla 3.7 - Conjunto de datos entrenamiento, prueba y validación, y número de datos.

Conjunto.	Período.	Número de datos.
Entrenamiento.	2022/11/18 – 2024/04/03	14,892
Prueba.	2022/09/01 – 2022/11/17	2,414
Validación.	2024/04/04 – 2024/06/01	1,972



Figura 3.2 - Distribución temporal de los conjuntos de datos.

3.3.2. SELECCIÓN DEL MODELO

Dada la naturaleza de las variables y el número de datos, se seleccionaron los modelos *Support Vector Regression* (SVR) y *Extreme Gradient Boosting* (XGBoost) para ser utilizados en el análisis y resolución del problema. Estos modelos poseen robustez ante datos ruidosos y tienen la capacidad de ser aplicados en sistemas multivariados. Una vez entrenados y evaluados los modelos, se analizaron los resultados de manera de determinar el mejor ajuste y la mejor capacidad predictiva.

3.3.3. ENTRENAMIENTO Y PRUEBA DEL MODELO

Para el entrenamiento del modelo, se filtraron los datos según las fechas indicadas en la Tabla 3.8. Los hiperparámetros utilizados se obtuvieron mediante el método *Bayes Search*, los que se presentan en la Tablas 3.8 y 3.9 para cada modelo.

Tabla 3.8 - Hiperparámetros del modelo SVR (Romero-Triana & Muñoz-Cañón, 2021).

Hiperparámetro.	Valor.	Definición técnica.
C	78.30	Coefficiente de regulación. Controla el equilibrio entre el ajuste del modelo y la penalización de errores.
γ	0.000116	Gamma. Define la influencia de cada punto de datos en la curva de decisión.
ϵ	0.000366	Épsilon. Define el margen de tolerancia dentro del cual los errores no son penalizados.

Tabla 3.9 - Hiperparámetros del modelo XGBoost (Yang & Shami, 2020).

Hiperparámetro.	Valor.	Definición técnica.
Learning rate.	0.036	Reduce la contribución de cada árbol. Evita el sobreajuste.
Max Depth.	3	Controla la profundidad máxima de los árboles de decisión en el modelo. Limita la complejidad del modelo.
Subsample.	0.83	Controla la proporción de muestras usadas para entrenar cada árbol. Reduce el sobreajuste.
Colsample bytree.	0.56	Define la fracción de características seleccionadas al construir cada árbol
γ	0	Gamma. Define la reducción mínima de la pérdida necesaria para una partición adicional en un nodo del árbol. Evita divisiones innecesarias.

Una vez entrenados los modelos, se continuó con la etapa de prueba. En esta etapa, los modelos se alimentaron con los datos del período de prueba, ejecutándose y entregando resultados de tasa de desgaste.

3.3.4. VALIDACIÓN DE LOS MODELOS

La validación de los modelos consistió en su ejecución, alimentándolos con nuevos datos, y evaluando su rendimiento. Esto es, se evaluó la consistencia de la tasa de desgaste con el nivel de llenado de bolas estimado.

Para llevar a cabo la validación, los modelos se almacenaron como archivos *pkl*, utilizando la librería *Joblib* de *Python*. Esto se realizó con el objetivo de utilizar los modelos sin la necesidad de entrenarlos nuevamente.

Luego, los modelos fueron evaluados con un nuevo conjunto de datos, correspondiente a aquellos asociados al periodo validación. Para evaluar el rendimiento de los modelos, se utilizaron los resultados de tasa de desgaste, junto con los datos de carguío de bolas para estimar un nivel de llenado de bolas mediante la Ecuación (3.2). Los resultados obtenidos de nivel de llenado de bolas se compararon con los datos disponibles de esta variable. A partir de la capacidad de predicción, se seleccionó el mejor modelo.

$$J_b = J_{bi-1} + ((R_i - ((T_i * TD_i)/1000))/1000/F_a) \quad (3.2)$$

3.3.5. ANÁLISIS E IMPLEMENTACIÓN DEL MODELO

El análisis consistió en comparar los resultados de validación, en términos de la predicción del nivel de llenado de bolas, lo cual se priorizó por sobre el comportamiento de la tasa de desgaste.

La implementación se llevó a cabo una vez seleccionado el modelo. El procedimiento consistió en ejecutar el modelo, establecer límites para el nivel de llenado de bolas, y estimar la reposición de bolas que permita obtener un J_b dentro de este rango. Esto se logró evaluando el carguío de bolas (los 3 posibles carguíos) según las condiciones de la planta, eligiendo aquel que permitiera cumplir con las restricciones de J_b .

3.4. APLICACIÓN

Se desarrolló una aplicación con el objetivo de mejorar la accesibilidad y visualización de los datos de operación del molino y los resultados de la modelación. Se buscó que la aplicación facilitara el acceso a la información y que fuese intuitiva y de fácil ejecución. Al mismo tiempo, debía permitir la validación y modificación de los resultados que surgen de la ejecución del modelo, de manera que se ajusten a la realidad de planta, contribuyendo también al reentrenamiento del modelo.

Para el desarrollo de la aplicación se utilizó la librería *streamlit* de *Python*. Esta librería permitió generar pestañas para separar tareas de modificación y validación de datos, mostrar gráficos, incorporar botones de validación, entre otras herramientas. La aplicación se alimentó de la base

de datos detallada en la Sección 3.2.2 y su objetivo fue cumplir con 3 funciones principales. La primera consistió en la visualización del comportamiento del tratamiento del molino, la tasa de desgaste y el J_b mediante gráficos. La segunda consistió en la opción de validación y modificación de los límites del J_b , la tasa de desgaste entregada por el modelo, y el carguío de bolas únicamente del día de ejecución. La tercera función consistió en dar acceso a la base de datos y proporcionar la opción de ingresar manualmente modificaciones a las variables de reposición, J_b y tasa de desgaste para cualquier fecha seleccionada, independiente del día en que se ejecute la acción. Una vez validada la acción, la base de datos se actualiza automáticamente con los cambios realizados. Esta aplicación se presenta en el Anexo A.

CAPITULO 4

4. RESULTADOS

4.1. RESULTADOS DE ENTRENAMIENTO

En el entrenamiento de los modelos se utilizaron las métricas *Mean Squared Error* (MSE) y el Coeficiente de Correlación de Pearson (R) entre la tasa de desgaste obtenida a partir del procedimiento descrito en la Sección 3.2.4 y los resultados de modelación. Los resultados se tabulan en la Tabla 4.1 para las modelaciones mediante SVR y XGBoost.

Tabla 4.1 - Métricas de bondad de ajuste en el entrenamiento.

Métrica.	Modelo SVR.	Modelo XGBoost.
MSE.	90.27	11.15
R .	0.65	0.56

El MSE indica el nivel de ajuste del modelo a los datos; por lo tanto, valores bajos de MSE indican un mejor ajuste (Wackerly *et al.*, 2008). El Coeficiente R indica la linealidad entre las variables analizadas. Su valor oscila entre -1 y 1, valores positivos como los obtenidos en la Tabla 4.1 indican una correlación lineal positiva, y su cercanía a 1 indica cómo el modelo se ajusta a la tendencia de los datos (Hernández-Lalinde *et al.*, 2018).

En términos de entrenamiento, ambos modelos presentaron una respuesta relativamente errática, con el modelo XGBoost presentando un menor MSE, pero un peor seguimiento a los datos (menor R). Esto se asocia a una baja correlación, pero menor *offset* en los residuos.

4.2. RESULTADOS DE PRUEBA DE MODELOS

Una vez entrenados los modelos, se graficó la tasa de desgaste de bolas (TDB) en función del tiempo. En la Figura 4.1 se muestran los resultados del modelo SVR y en la Figura 4.2 se muestran los resultados del modelo XGBoost. Las predicciones del modelo SVR presentan menos ruido en comparación al modelo XGBoost. No obstante, los resultados del modelo XGBoost presentan un mejor seguimiento a la tendencia de los datos. En general, los resultados

obtenidos en la prueba no mostraron diferencias suficientemente significativas que posicionaran a un modelo por sobre el otro.

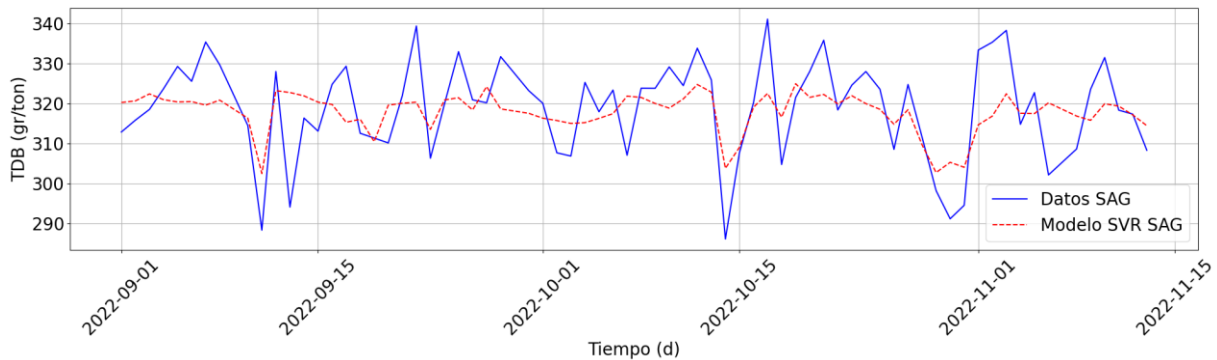


Figura 4.1 - Tasa de desgaste de bolas gr/ton en función del tiempo en días, prueba modelo SVR.

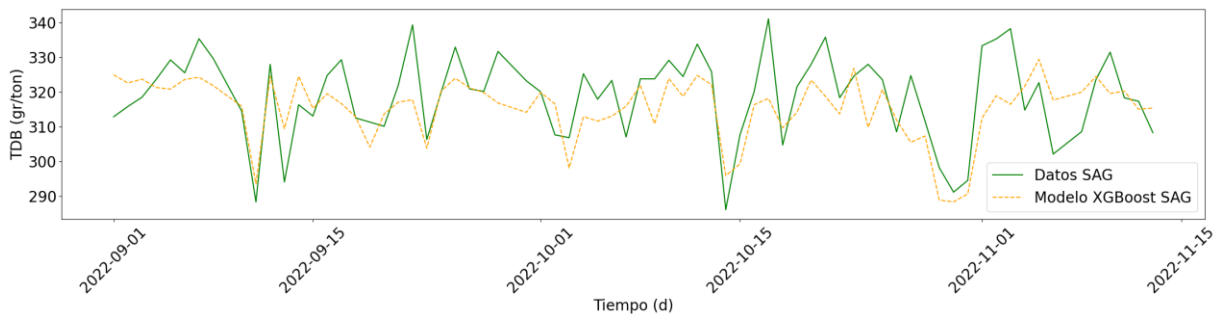


Figura 4.2 - Tasa de desgaste de bolas gr/ton en función del tiempo en días, prueba modelo XGBoost.

4.3. RESULTADOS DE VALIDACIÓN DE MODELOS

Dada la respuesta errática de los modelos en la etapa de prueba, se prosiguió con un proceso de validación. Esta consistió en ejecutar el modelo en un período de tiempo independiente, según lo indicado en la Sección 3.3.1, y así evaluar los modelos a partir de sus predicciones. En las Figuras 4.3 y 4.4 se muestran los resultados de la tasa de desgaste de bolas entregada por los modelos SVR y XGBoost, respectivamente, en función del tiempo durante el período de validación. Se aprecia claramente que las predicciones entregadas por el modelo SVR tiene menor ruido que las entregadas por el modelo XGBoost. Adicionalmente, se analizó la capacidad de predicción de cada modelo en la estimación del nivel de llenado de bolas, y no sólo el ruido que puedan presentar sus predicciones de desgaste.

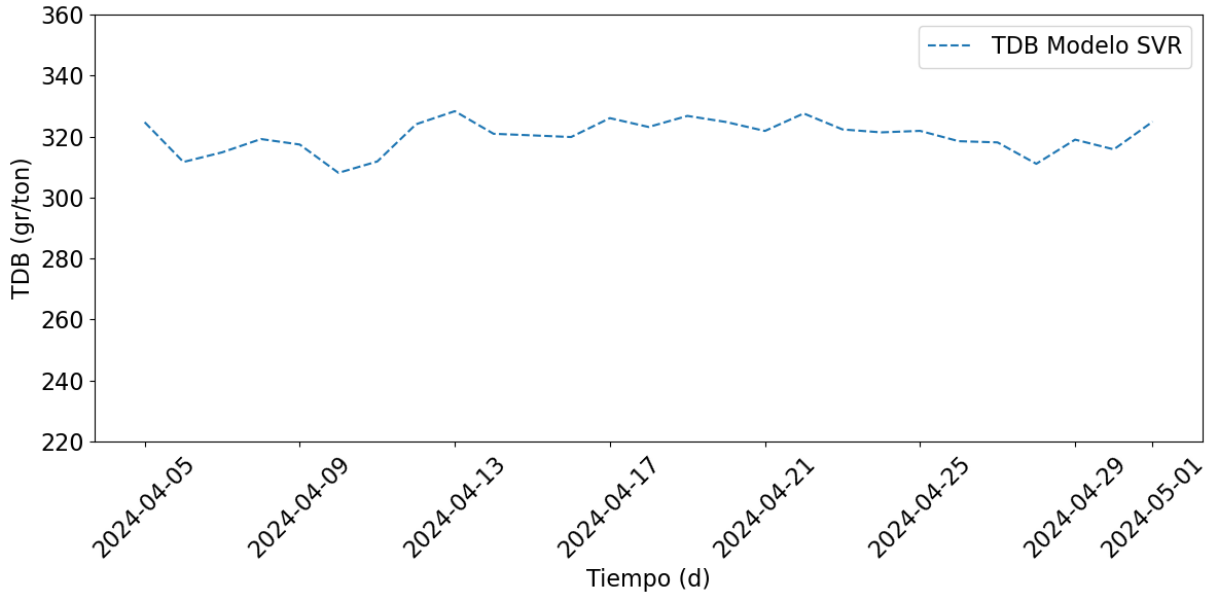


Figura 4.3 - Tasa de desgaste de bolas en función del tiempo, período de validación, modelo SVR.

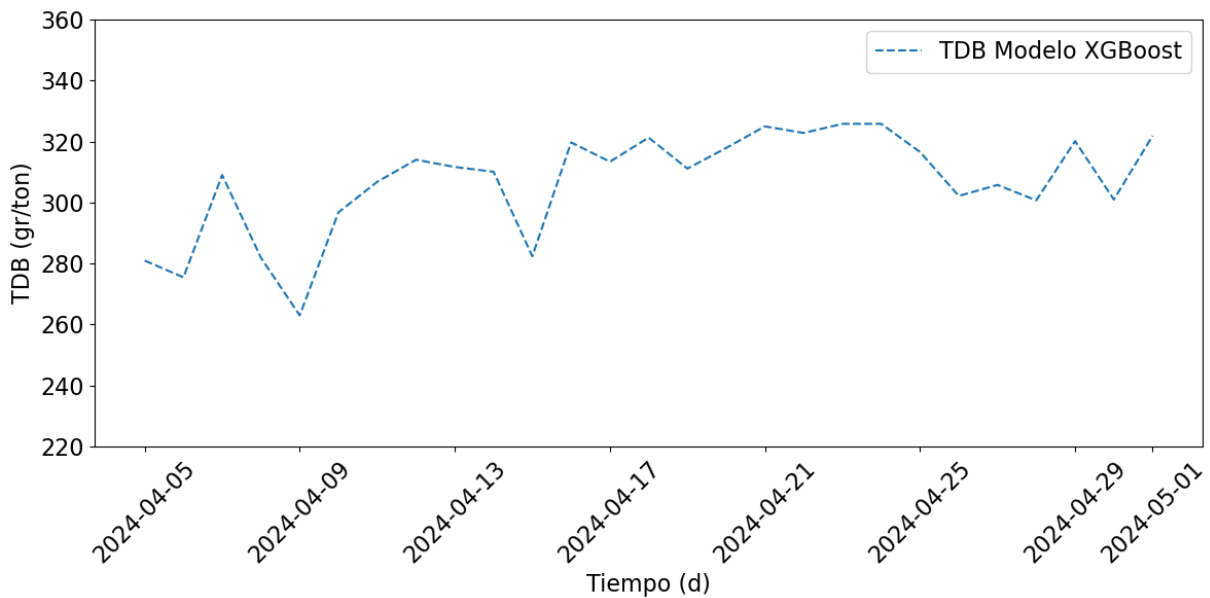


Figura 4.4 - Tasa de desgaste de bolas en función del tiempo, período de validación, modelo XGBoost.

Para evaluar el rendimiento de los modelos, se compararon los resultados de nivel de llenado de bolas en función del tiempo, junto a las estimaciones *Grind Out*. Estos resultados se obtuvieron utilizando la Ecuación (3.1), considerando los descriptores presentados en la Sección 3.2, y las tasas de desgaste modeladas y las estimadas de acuerdo con la Sección 3.2.4.

Las Figuras 4.5 y 4.6 muestran el J_b estimado a partir de las tasas de desgaste entregadas por los modelos SVR y XGBoost, respectivamente, versus el comportamiento del J_b estimado en planta. Además, se presentan los valores de J_b en condición de *Grind Out*. De la Figura 4.5, el modelo SVR no se acerca a los valores de J_b de *Grind Out* ni muestra una tendencia a acercarse a la información disponible en tiempo real para el J_b . Por otro lado, la Figura 4.6 muestra que las predicciones del modelo XGBoost permiten estimar un J_b más cercano a las condiciones de *Grind Out*, mostrando una tendencia más cercana a las estimaciones en tiempo real de J_b .

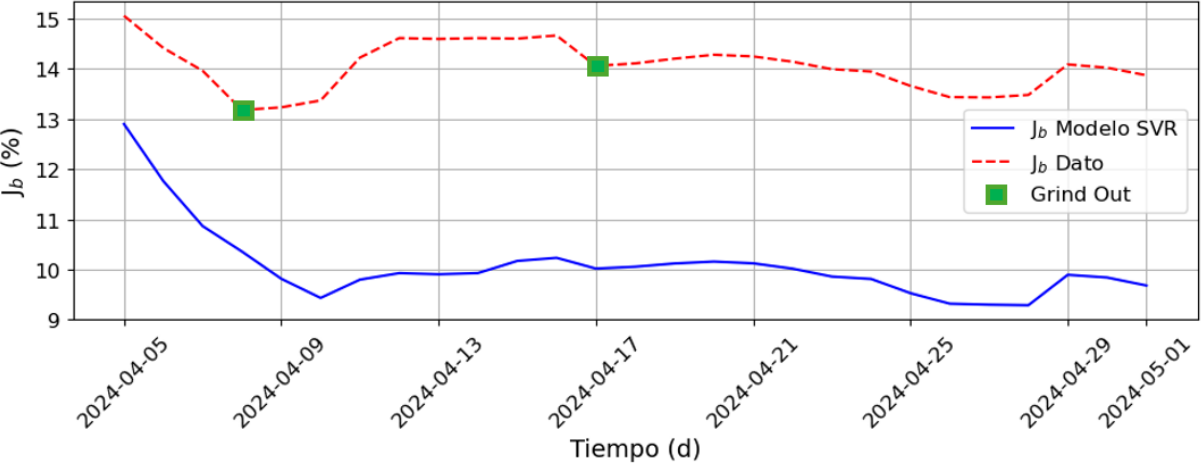


Figura 4.5 - Nivel de llenado de bolas J_b en función del tiempo, validación modelo SVR.

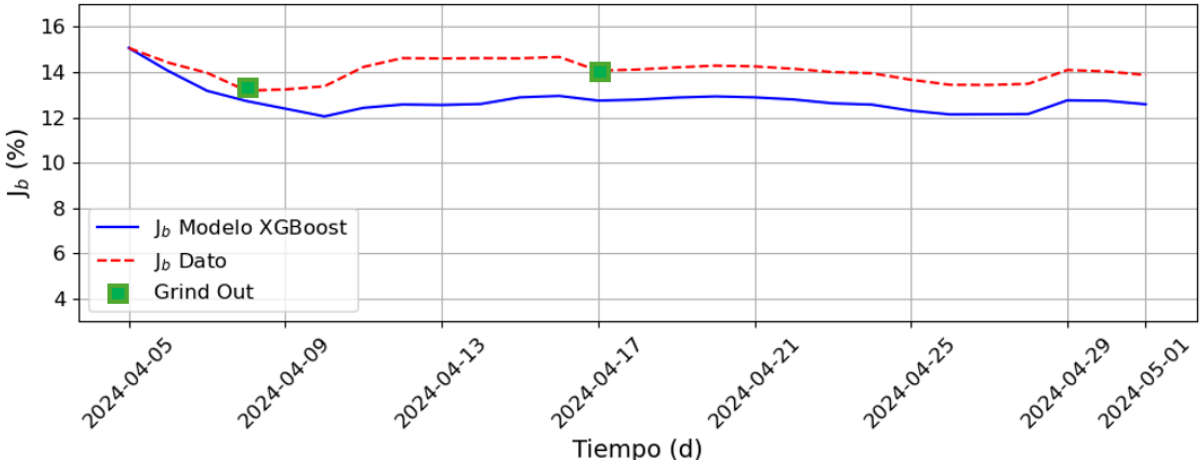


Figura 4.6 - Nivel de llenado de bolas J_b en función del tiempo, validación modelo XGBoost.. A partir del análisis de resultados, se seleccionó el modelo *XGBoost* para estimar la tasa de desgaste de bolas, debido a su mayor consistencia con el nivel de llenado J_b y menor MSE.

Específicamente, las predicciones entregadas por el modelo XGBoost permiten tener un mejor acercamiento a los valores de J_b en las condiciones de *Grind Out*. De este análisis, se infiere que el modelo XGBoost permite tener un mejor rendimiento en la modelación de la molienda SAG.

4.4. IMPLEMENTACIÓN

Una vez seleccionado el modelo, se continuó con su implementación. Esta consistió en ejecutar el modelo diariamente con la base de datos descrita en la Sección 3.2, entregando la tasa de desgaste de bolas, y luego aplicarla para obtener la reposición de bolas diaria junto con la estimación del nivel de llenado de bolas.

Con el fin de mostrar el rendimiento final del modelo seleccionado, se determinó un período de tiempo de 1 mes, dada la disponibilidad de datos al momento de realizar la implementación. Se obtuvieron los resultados del modelo representados a través del J_b y la reposición de bolas.

En la Figura 4.7 se muestran los resultados de tasa de desgaste de bolas entregada por el modelo XGBoost en el período de implementación. Se compara con el valor de tasa de desgaste que se asume en planta, la que corresponde a 328 gr/ton. La tasa utilizada en planta sobre-estima levemente el desgaste de bolas respecto del modelo, lo que puede repercutir en un sobre-consumo del recurso, aunque también entrega una tolerancia ante cambios en la operación.

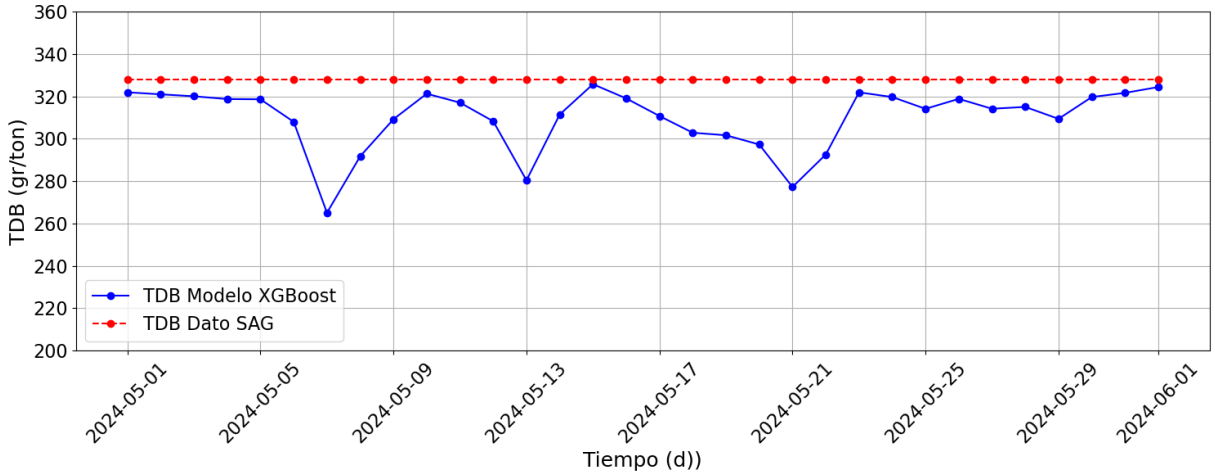


Figura 4.7 - Comportamiento de la tasa de desgaste, implementación modelo XGBoost.

En la Figura 4.8 se muestran los resultados de la reposición de bolas en base al desgaste de bolas entregado por el modelo en el período de implementación. En general, se observa que los resultados del modelo tienden a ser consistentes con la reposición aplicada en planta.

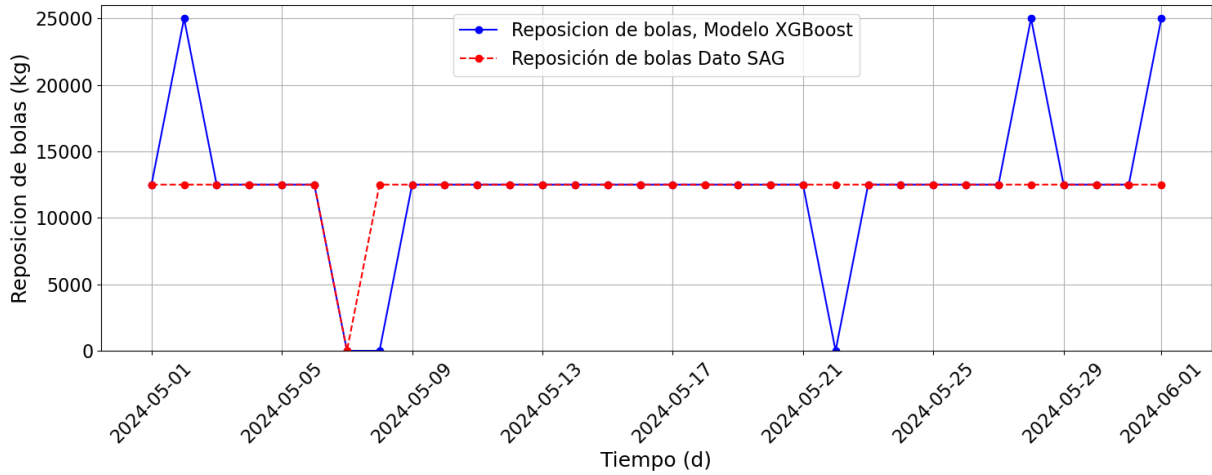


Figura 4.8 - Comportamiento de la reposición de bolas, implementación modelo XGBoost.

En la Figura 4.9 se muestra el comportamiento del J_b estimado a partir de los resultados entregados por el modelo en el período de implementación. Se aprecia que la estrategia de modelación corrige el comportamiento del J_b ante nuevas condiciones de *Grind Out*. Así, la respuesta del modelo tiende a ajustarse a los cambios y volver a entregar una tendencia aproximada de J_b .

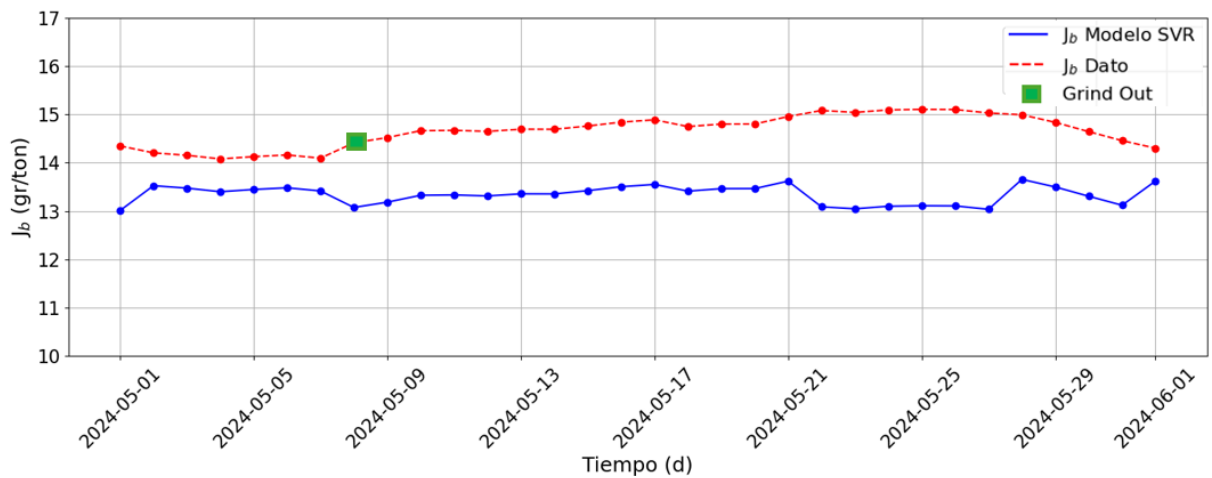


Figura 4.9 - Comportamiento del J_b , implementación del modelo XGBoost.

El modelo XGBoost destaca por su capacidad para proporcionar estimaciones precisas y adaptativas de la tasa de desgaste de bolas. Este enfoque permite ajustar la reposición al desgaste de las bolas, adaptarse a los cambios operacionales de la molienda SAG y tener una estimación del nivel de llenado de bolas.

CAPITULO 5

5. CONCLUSIONES Y RECOMENDACIONES

5.1. CONCLUSIONES

Se desarrolló una herramienta para la definición del carguío de bolas en la molienda semi-autógena, con el fin de mejorar la eficiencia en la etapa de molienda de División Chuquicamata de Codelco. Se analizaron las variables que afectan el desgaste de las bolas y la eficiencia de la molienda primaria. Se realizó un levantamiento de información histórica, y se analizaron datos operacionales para caracterizar el comportamiento de la molienda SAG.

Utilizando herramientas de Machine Learning, específicamente, los modelos Support Vector Regression (SVR) y Extreme Gradient Boosting (XGBoost), se definió una tasa de desgaste de bolas dinámica que permite definir el carguío de bolas y estimar el nivel de llenado de bolas. De la implementación se concluye:

- El modelo XGBoost mostró mejor rendimiento en comparación con el modelo SVR en términos de la estimación de la tasa de desgaste.
- El modelo XGBoost llevó a una mejor aproximación del J_b estimado en tiempo real en planta. Además, se adecuó más rápido ante cambios en el J_b a partir de las mediciones de *Grind Out*, corrigiendo su comportamiento y siguiendo la tendencia de los datos.
- El ajuste dinámico de la tasa de desgaste de bolas permitió definir su carguío con el objetivo de mantener el nivel de llenado de bolas del molino dentro de los rangos esperados, mejorando la estabilidad y eficiencia del proceso de molienda.

5.2. RECOMENDACIONES

Teniendo en cuenta las oportunidades de mejora, se recomienda lo siguiente:

- Seguir implementando la herramienta desarrollada para ajustar el carguío de bolas en tiempo real, optimizando la operación del molino SAG.

- Continuar alimentando el modelo XGBoost con nuevos datos operacionales para mejorar su exactitud y adaptabilidad a nuevas condiciones operacionales.
- Capacitar al personal en el uso de la herramienta y en la interpretación de los resultados del modelo, asegurando una correcta implementación y operación.
- Monitorear continuamente el rendimiento del modelo y aplicar los ajustes necesarios que garanticen una operación eficiente.

REFERENCIAS

- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning* (1st ed.). New York: Springer.
- Cárdenas-Valderrama, E. (2024). *Operación y Mantenimiento de Molinos SAG*. Segunda Especialización en Plantas Industriales de Cominuación y Concentración. Chile. Retrieved Mayo 2024
- Chicama, D. H. (2023). *Comunidad e Innovación en Minería*. Retrieved Julio 2024, from *Introducción a Machine Learning*: <https://www.codeauni.com/comunidad/blog/196/>
- Codelco. (2021). *Curso Avanzado de Molienda SAG y Bolas*. Capacitación para operadores - Curso Avanzado de Molienda SAG y Bolas (distribución restringida). Chile. Retrieved Marzo 2024, from Codelco.
- CODELCO. (2023). *Memoria Anual 2023*. Retrieved Marzo 2024, from Codelco: https://www.codelco.com/prontus_codelco/site/docs/20240404/20240404191040/memoria_codelco_2023.pdf
- CODELCO. (2024, Marzo). *División Chuquicamata*. Retrieved 2024, from CODELCO: <https://www.codelco.com/operaciones/chuquicamata/division-chuquicamata>
- Distributed Machine Learning Community. (2024). *Distributed Machine Learning Community*. Retrieved Julio 2024, from *XGBoost Documentation*: <https://xgboost.readthedocs.io/en/stable/tutorials/model.html>
- Espinosa-Zúñiga, J. (2020). *Aplicación de algoritmos Random Forest y XGBoost en una base de solicitudes de tarjetas de crédito*. *Ingeniería Investigación y Tecnología*, XXI, 1-16. doi:<https://doi.org/10.22201/fi.25940732e.2020.21.3.022>
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques* (Third ed.). Waltham, USA: ELSEVIER.
- Hernández-Lalinde, J., Peñaloza-Tarazona, M., Espinoza-Castro, J., & Rodríguez, J. (2018). *Sobre el uso adecuado del coeficiente de correlación de Pearson: definición, propiedades y suposiciones*. *AVFT*, 587-595.
- Kawatra, S. (2006). *Advances in Comminution*. Littleton: SME.

Kelleher, J. D., Mac-Namee, B., & D'Arcy, A. (2015). *Fundamentals of Machine Learning for Predictive Data Analytics*. Cambridge, Massachusetts, USA: Massachusetts Institute of Technology.

Romero-Triana, J., & Muñoz-Cañón, N. (2021). Optimización de los hiperparámetros de una máquina de regresión de soporte vectorial utilizando enjambre de partículas para el pronóstico de caos de COVID-19. *UIS Ingenierías*, 181-196.

Sethi, A. (2024). *Analytics Vidhya*. Retrieved Julio 2024, from Support Vector Regression Tutorial for Machine Learning: [https://www.analyticsvidhya.com/blog/2020/03/support-vector-regression-tutorial-for-machine-learning/#:~:text=Support%20Vector%20Regression%20\(SVR\)%20is,while%20minimizing%20the%20prediction%20error.](https://www.analyticsvidhya.com/blog/2020/03/support-vector-regression-tutorial-for-machine-learning/#:~:text=Support%20Vector%20Regression%20(SVR)%20is,while%20minimizing%20the%20prediction%20error.)

Steele, B., Chandler, J., & Reddy, S. (2016). *Algorithms for Data Science*. Missoula, USA: Springer.

Wackerly, D., Mendenhall, W., & Scheaffer, R. (2008). *Mathematical Statistics with Applications* (7th ed.). Belmont, USA: THOMSON.

Wills, B., & Napier-Munn, T. (2006). *Mineral Processing Technology* (7th ed.). Oxford: BH.

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms. *Neurocomputing*, 295-316.

ANEXOS

A. APLICACIÓN VIRTUAL

La aplicación se tituló “Tablero Recomendaciones Carguío de bolas SAG” y se elaboró con 3 pestañas para responder a las 3 funciones: “Página Inicial”, “Tablero SAG” y “Cambio Data SAG”. Los cambios realizados en esta aplicación actualizan la base de datos generada anteriormente.

En la Figura A.1 se muestra la primera pestaña de la aplicación “Página Inicial”, señalando las partes que la componen. Esta pestaña da entrada a la aplicación y se muestra el comportamiento del tratamiento del molino de los últimos 10 días, además, se muestra su desglose según origen: Mina Rajo, Mina Subterránea y Mina Radomiro Tomic.

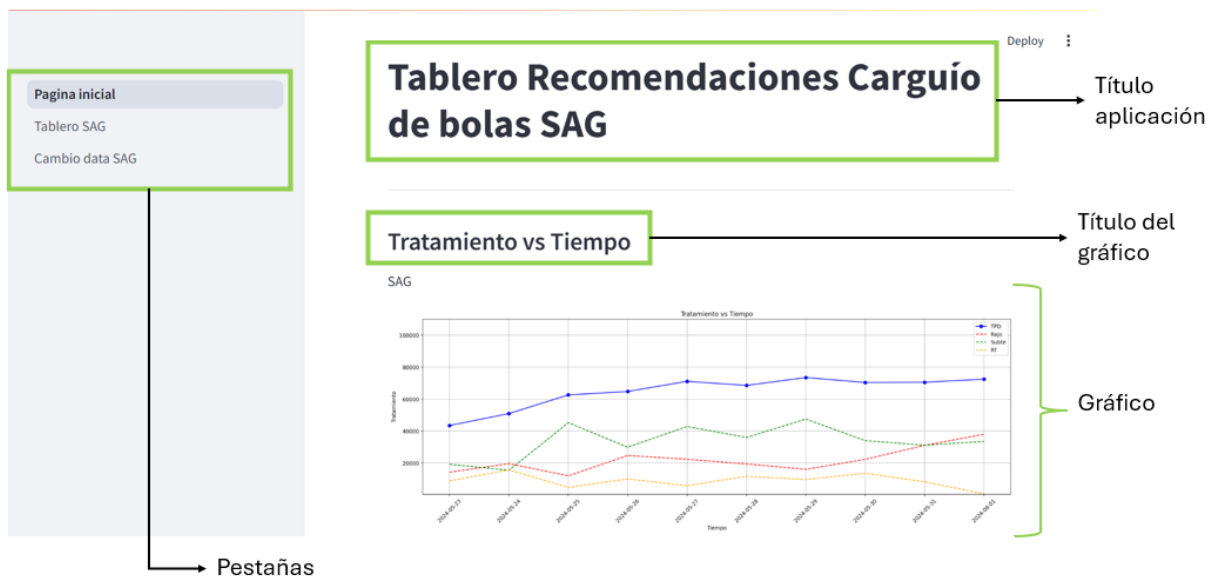


Figura A.1 - Pestaña "Página Inicial" de la aplicación virtual.

En las Figura A.2, Figura A.3, Figura A.4 y Figura A.5 se muestran las partes de la pestaña “Tablero SAG”. Esta pestaña permite cumplir la función de permitir validar y modificar los límites en que se comporta J_b , la tasa de desgaste entregada por el modelo y el carguío de bolas únicamente del último día.

La pestaña se puede dividir en 4 grandes secciones. La primera sección se muestra en la Figura A.2 y se compone por una tabla de datos, la cual contiene los datos de las variables del molino de los último 10 días, además, se señala si el molino está en operación o no.

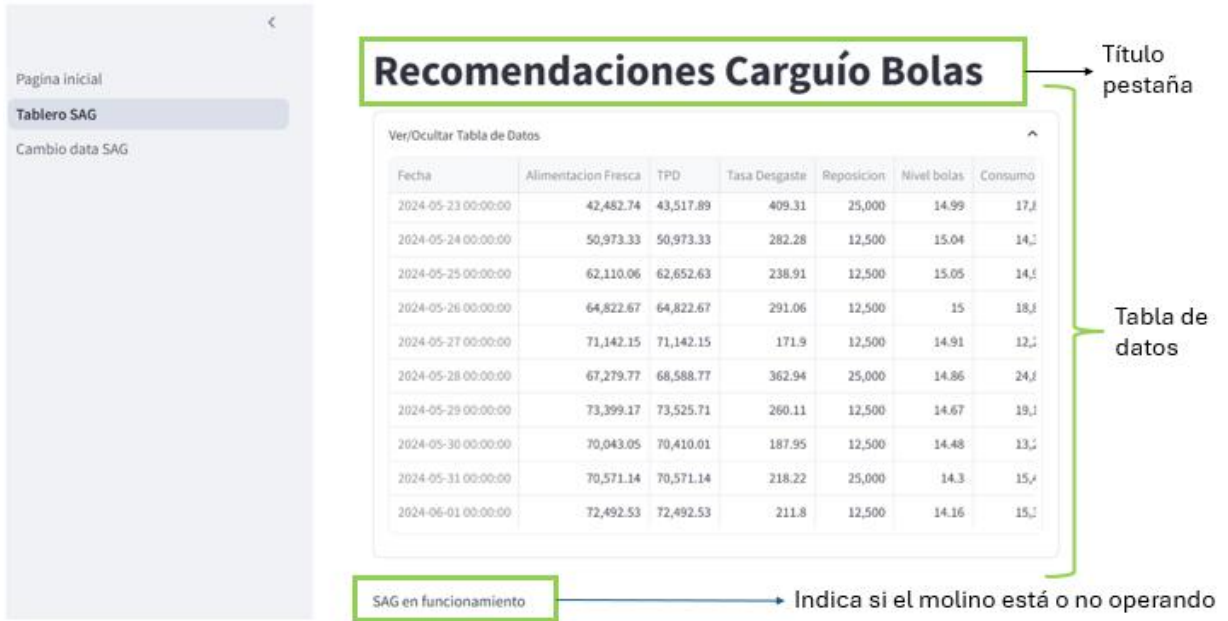


Figura A.2 - Pestaña "Tablero SAG" de la aplicación virtual sección 1.

La segunda sección se muestra en la Figura A.3. En esta sección se muestra los límites superior e inferior para el J_b por defecto y entrega la tasa de desgaste obtenida por el modelo, estos valores pueden ser modificados editando los valores entregados. Luego, mediante un botón de validación, se busca aceptar los valores entregados y/o editados anteriormente, además de mostrar si la validación se puede aceptar (verde) o no (rojo).



Figura A.3 - Pestaña "Tablero SAG" de la aplicación virtual sección 2.

La tercera sección se muestra en la Figura A.4. En esta sección se muestra la fecha, el tratamiento del molino del día anterior, esto permite decidir el carguío de bolas a partir de las condiciones de alimentación del molino.

Luego, se despliegan 3 recomendaciones, 4 opciones de selección de reposición y un botón de validación. Las recomendaciones consisten en entregar el valor de J_b estimado en caso de no hacer una reposición, en reponer 12.5 y 25 ton, donde se destacarán en verde en caso de que el J_b estimado se encuentre dentro de los límites seleccionados anteriormente, y se destacará en rojo en caso de que no cumpla. Las 4 opciones de selección consisten en no reponer bolas, reponer 12.5 o 25 ton, o ingresar una recomendación, de elegir la última opción, se desplegará una casilla donde se debe ingresar manualmente las toneladas de reposición. El botón de validación busca validar los valores seleccionados, además de mostrar si la validación se puede aceptar (verde) o no (rojo).



Figura A.4 - Pestaña "Tablero SAG" de la aplicación virtual sección 3.

La cuarta sección se muestra en la Figura A.5. En esta sección se muestra el comportamiento del J_b y el comportamiento de la tasa de desgaste de los últimos 10 días. Estos gráficos responden a los cambios hechos en la segunda y en la tercera sección una vez validados mediante los botones de validación.



Figura A.5 - Pestaña "Tablero SAG" de la aplicación virtual sección 4.

En las Figura A.6, Figura A.7 y Figura A.8 se muestran las partes de la pestaña “Cambio Data SAG”. Esta pestaña permite cumplir la función de dar acceso a la base de datos y entregar la opción de modificar las variables de reposición, J_b y tasa de desgaste.

La pestaña se puede dividir en 3 grandes secciones. La primera sección se muestra en la Figura A.6 y se compone por una casilla para ingresar una fecha y una tabla de datos. La casilla permite ingresar la fecha en la cual se quiere modificar las variables y la tabla despliega los datos desde un día previo a la fecha ingresada hasta el día actual.



Figura A.6 - Pestaña "Cambio data SAG" de la aplicación virtual sección 1.

La segunda sección se muestra en la Figura A.7 y se compone de 3 opciones de selección, casilla y un botón de validación. Las 3 opciones de selección permiten elegir cuál de las variables se busca modificar y se compone de las variables reposición, tasa de desgaste y Jb, La casilla es donde se debe ingresar el valor designado de la variable que se quiere modificar y el botón de validación busca validar los valores seleccionados.



Figura A.7 - Pestaña "Cambio data SAG" de la aplicación virtual sección 2.

La tercera sección se muestra en la Figura A.8 y se compone por una tabla de datos, la cual contiene los datos de las variables del molino desde un día previo a la fecha ingresada hasta el día actual, con los datos actualizados a partir de las modificaciones realizada previamente.



Figura A.8 - Pestaña "Cambio data SAG" de la aplicación virtual sección 3.