Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América
Facultad de Química e Ingeniería Química
Escuela Profesional de Ingeniería Química

Desarrollo de un modelo predictivo basado en aprendizaje automático para la optimización del consumo de xantato en la flotación de minerales con presencia de pirrotita, en una planta de beneficio ubicada en Marcona – Ica, Perú

TESIS

Para optar el Título Profesional de Ingeniero Químico

AUTOR

Giancarlo GLORIA MARCATINCO

Galan Daniel LAGO VIGILIO

ASESOR

Dra. Juana SANDIVAR ROSAS

Lima, Perú

2025



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

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

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

Referencia bibliográfica

Gloria, G. & Lago, G. (2025). *Desarrollo de un modelo predictivo basado en aprendizaje automático para la optimización del consumo de xantato en la flotación de minerales con presencia de pirrotita, en una planta de beneficio ubicada en Marcona – Ica, Perú*. [Tesis de maestría, Universidad Nacional Mayor de San Marcos, Facultad de Química e Ingeniería Química, Escuela Profesional de Ingeniería Química]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor 1	
Nombres y apellidos	Giancarlo Gloria Marcatinco
Tipo de documento de identidad	DNI
Número de documento de identidad	73754032
URL de ORCID	https://orcid.org/0009-0007-9854-3917
Datos de autor 2	
Nombres y apellidos	Galan Daniel Lago Vigilio
Tipo de documento de identidad	DNI
Número de documento de identidad	70382679
URL de ORCID	https://orcid.org/0009-0005-5359-9189
Datos de asesor	
Nombres y apellidos	Juana Sandivar Rosas
Tipo de documento de identidad	DNI
Número de documento de identidad	8667298
URL de ORCID	https://orcid.org/0000-0001-8494-2940
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Víctor Manuel Fernández Guzmán
Tipo de documento	DNI
Número de documento de identidad	06690674
Miembro del jurado 1	
Nombres y apellidos	José Manuel García Pantigozo

Tipo de documento	DNI
Número de documento de identidad	6078851
Datos de investigación	
Línea de investigación	C.0.4.17. Flotación de Minerales
Grupo de investigación	Optimización de Procesos de Transformación y Manufactura - OPTM
Agencia de financiamiento	Sin financiamiento.
Ubicación geográfica de la investigación	Edificio: Shougang Hierro Peru S.A.A. País: Perú Departamento: Ica Provincia: Nazca Distrito: San Juan de Marcona Latitud: -15.36528 Longitud: -75.16222
Año o rango de años en que se realizó la investigación	Enero 2024 – Diciembre 2024
URL de disciplinas OCDE	Ingeniería de procesos https://purl.org/pe-repo/ocde/ford#2.04.02

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

ESCUELA PROFESIONAL DE INGENIERÍA QUÍMICA

ACTA DE SUSTENTACIÓN DE TESIS

Los suscritos miembros del Jurado Calificador nombrados por el señor Presidente del Comité Ejecutivo del "IV Programa Ciclo Taller de Titulación Profesional por la Modalidad de Tesis y Trabajo de Suficiencia Profesional para la Facultad de Química e Ingeniería Química 2025-I", bajo la Presidencia del **Dr. Víctor Manuel Fernández Guzmán ; Ing. José Manuel García Pantigozo (Miembro)** y la **Dra. Juana Sandivar Rosas (Asesora);** habiendo presentado para tal efecto la **TESIS,** titulada **Desarrollo de un modelo predictivo basado en aprendizaje automático para la optimización del consumo de xantato en la flotación de minerales con presencia de pirrotita, en una planta de beneficio ubicada en Marcona – Ica, Perú,** después de **SUSTENTADA Y APROBADA** la tesis elaborada por el bachiller en Ingeniería Química: **GLORIA MARCATINCO, GIANCARLO;** para optar el **TÍTULO PROFESIONAL** de **INGENIERO QUÍMICO, acordaron** calificarlo con la **NOTA** de:

DIECINUEVE 19

(LETRAS) (NÚMEROS)

Ciudad Universitaria, 16 de agosto del 2025

Dr. VÍCTOR MANUEL FERNÁNDEZ GUZMÁN
Presidente

Ing. JOSÉ MANUEL GARCÍA PANTIGOZO
Miembro

Dra. JUANA SANDIVAR ROSAS
Asesora

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

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

ESCUELA PROFESIONAL DE INGENIERÍA QUÍMICA

ACTA DE SUSTENTACIÓN DE TESIS

Los suscritos miembros del Jurado Calificador nombrados por el señor Presidente del Comité Ejecutivo del "IV Programa Ciclo Taller de Titulación Profesional por la Modalidad de Tesis y Trabajo de Suficiencia Profesional para la Facultad de Química e Ingeniería Química 2025-I", bajo la Presidencia del **Dr. Víctor Manuel Fernández Guzmán; Ing. José Manuel García Pantigozo (Miembro)** y la **Dra. Juana Sandivar Rosas (Asesora);** habiendo presentado para tal efecto la **TESIS,** titulada **Desarrollo de un modelo predictivo basado en aprendizaje automático para la optimización del consumo de xantato en la flotación de minerales con presencia de pirrotita, en una planta de beneficio ubicada en Marcona – Ica, Perú,** después de **SUSTENTADA Y APROBADA** la tesis elaborada por el bachiller en Ingeniería Química: **LAGO VIGILIO, GALAN DANIEL;** para optar el **TÍTULO PROFESIONAL** de **INGENIERO QUÍMICO,** acordaron calificarlo con la **NOTA** de:

DIECINUEVE (19)

(LETRAS) (NÚMEROS)

Ciudad Universitaria, 16 de agosto del 2025

Dr. VÍCTOR MANUEL FERNÁNDEZ GUZMÁN
Presidente

Ing. JOSÉ MANUEL GARCÍA PANTIGOZO
Miembro

Dra. JUANA SANDIVAR ROSAS
Asesora

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

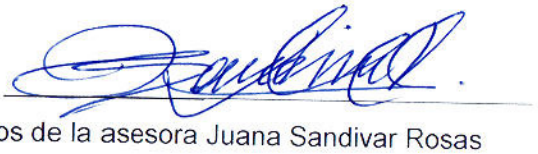
UNIVERSIDAD NACIONAL MAYOR DE
SAN MARCOS
Universidad del Perú, Decana de América

CERTIFICADO DE SIMILITUD

Yo, **Juana Sandivar Rosas** en mi condición de asesora acreditado con el Dictamen N°000001-2025-EPIQ-FQIQ/UNMSM de la tesis, cuyo título es: **Desarrollo de un modelo predictivo basado en aprendizaje automático para la optimización del consumo de xantato en la flotación de minerales con presencia de pirrotita, en una planta de beneficio ubicada en Marcona – Ica, Perú**, presentado por los bachilleres: **Gloria Marcatinco, Giancarlo y Lago Vigilio, Galan Daniel,** para optar el título Profesional de Ingeniería Química, CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante la plataforma PIDS, el documento evaluado cuenta con el 2% de similitud, nivel **PERMITIDO** para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional.**

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del título correspondiente.

Lima, 26 de agosto del 2025

Firma del Asesora 

Nombres y Apellidos de la asesora Juana Sandivar Rosas

DNI: 8667298

Agradecimiento

Expresamos nuestra profunda gratitud a nuestra asesora, Dra. Juana Sandivar Rosas, quien con su guía metodológica, rigor científico, capacidad crítica y experiencia han sido fundamentales para dar forma a este trabajo.

A nuestras familias, por su comprensión durante las largas jornadas de estudio y su respaldo constante en los momentos de incertidumbre. Su presencia ha sido el cimiento que nos permitió perseverar.

A nuestros colegas y compañeros de estudio, por los debates enriquecedores y el intercambio de ideas que nutrieron nuestra investigación. El pensamiento colectivo ha sido una fuente invaluable de inspiración.

Finalmente, reconocemos el valor del conocimiento científico y el pensamiento racional que han guiado nuestro enfoque investigativo.

Resumen

La presente investigación desarrolla un modelo predictivo basado en regresión lineal para estimar el consumo específico de xantato (kg/día) en función del contenido de pirrotita (%) en la alimentación de la planta Shougang Hierro Perú. La pirrotita, sulfuro de hierro altamente reactivo, presenta una afinidad elevada por el xantato, generando flotabilidad no deseada y obligando a sobredosificaciones empíricas que elevan costos y cargas químicas en efluentes. Utilizando datos históricos de operación y análisis mineralógicos, se caracterizó la variabilidad del contenido de pirrotita y del consumo de reactivo, identificando una correlación positiva significativa. La regresión lineal fue seleccionada por su simplicidad, interpretabilidad y facilidad de implementación en sistemas operativos de planta. El modelo final explicó aproximadamente el 56 % de la variabilidad del consumo de xantato, con errores absolutos moderados, demostrando su aplicabilidad para soportar decisiones operativas basadas en datos. La adopción de esta herramienta permitirá optimizar la dosificación del colector, reducir costos operativos y contribuir a la sostenibilidad ambiental del proceso de flotación, alineándose con los objetivos estratégicos de minería responsable y eficiente en la operación minera de Marcona.

Palabras clave: flotación de minerales, pirrotita, xantato, regresión lineal, minería inteligente.

Abstract

This research develops predictive model based on linear regression to estimate the specific consumption of xanthate (kg/día) as a function of pyrrhotite content (%) in the feed to the Shougang Hierro Perú plant. Pyrrhotite, an iron sulfide with high surface reactivity, exhibits strong affinity for xanthate, causing undesired flotation and necessitating empirical overdosing that increases operational costs and chemical loads in effluents. Using historical operational data and mineralogical analyses, the variability of pyrrhotite content and reagent consumption was characterized, revealing a significant positive correlation. The linear regression model was selected for its simplicity, interpretability, and ease of implementation in existing plant control systems. The final model explained approximately 56% of the variability in xanthate consumption, with moderate absolute errors, demonstrating its applicability as a decision-support tool for data-driven operational planning. Adoption of this predictive approach will enable optimization of collector dosage, reduction of operating costs, and enhancement of the environmental sustainability of the flotation process, aligning with strategic goals for responsible and efficient mining at the Marcona operation.

Keywords: mineral flotation, pyrrhotite, xanthate, linear regression, smart mining.

Tabla de contenido

Resumen 2

Abstract 3

Introducción 9

Descripción de la realidad problemática 11

Importancia Industrial del Control de Reactivos en Flotación 13

Impacto de la Variabilidad Mineralógica en la Eficiencia del Proceso 14

Limitaciones de los Enfoques Tradicionales de Control Operativo 15

Propuesta de Solución Basada en Aprendizaje Automático 16

Justificación del Enfoque Adoptado 17

Aportes Esperados de la Investigación 18

Definición del Problema 20

Problema General 21

Problemas Específicos 21

Objetivos 21

Objetivo General 21

Objetivos Específicos 22

Justificación y Relevancia de la Investigación 22

Justificación Teórica 22

Justificación Metodológica 23

Justificación Práctica 24

Pertinencia Industrial y Tecnológica 26

Marco Teórico 27

Fundamento de la Flotación en Minerales de Hierro 27

Comportamiento Reactivo de la Pirrotita 28

Relación entre Contenido de Pirrotita y Consumo de Xantato 30

Aplicaciones de Machine Learning en Procesos Metalúrgicos 32

Ventajas de la Inteligencia Artificial para el Control de Reactivos 34

Antecedentes del Estudio 36

Estudios Nacionales Sobre la Influencia de la Pirrotita 36

Investigaciones Latinoamericanas sobre Activación Superficial y Consumo de Xantato 37

Modelos Estadísticos Para Estimar el Consumo de Reactivos 38

Aplicaciones Internacionales de Machine Learning en Flotación 39

Aportes y Vacíos que Motivan el Presente Estudio 41

Bases Teóricas 43

Pirrotita: Naturaleza Mineralógica y Comportamiento en Flotación 43

Consumo de Xantato como Variable Operativa 45

Modelos de Predicción en Flotación y Comparativa de Algoritmos 46

Definición de Categorías de Análisis 48

Variable Independiente: Contenido de Pirrotita 48

Definición Conceptual 48

Definición Operativa 49

Categoría de Análisis 49

Variable Dependiente: Consumo de Xantato 50

Definición Conceptual 50

Definición Operativa 50

Categoría de Análisis 51

Relación entre las Variables 52

Hipótesis 54

Hipótesis General 54

Hipótesis Específicas 54

Sistemas y Categorías de Análisis 55

Dimensión Mineralógica 55

Dimensión Reactiva 56

Dimensión Predictiva 57

Materiales y Métodos 60

Enfoque de la Investigación 60

Tipo de Investigación 60

Diseño de Investigación 61

Credibilidad de la Investigación 62

Sujetos de Estudio 62

Características del Sujeto de Estudio 63

Procedimiento de Selección de los Sujetos de Estudio 63

Recolección de Datos de los Sujetos de Estudio 64

Justificación de la Selección de los Sujetos de Estudio 64

Procedimientos, Técnicas e Instrumentos de Recolección de Información 64

Procedimiento 65

Técnicas 66

Instrumentos de Recolección de Información 66

Evaluación Comparativa de Algoritmos 66

Métricas de Evaluación del Modelo 67

Selección del Modelo de Evaluación 68

Validación Operativa y Aplicación del Modelo 70

Resultados y Discusión de Resultados 72

Resultados 72

Descripción de la Base de Datos Empleada y Preprocesamiento 72

Análisis Exploratorio de las Variables 73

Análisis de la Correlación entre Variables 82

Comparativa de Modelos de Aprendizaje Automático 85

Modelamiento del Modelo Seleccionado y Métricas 88

Evaluación del Modelo de Regresión Lineal Múltiple 91

Análisis estadístico y validación de hipótesis 94

Discusión 101

Conclusiones y Recomendaciones 104

Conclusiones 104

Recomendaciones 106

Bibliografía 108

Introducción

La minería moderna enfrenta el reto constante de optimizar sus procesos productivos para mejorar la rentabilidad y reducir los impactos ambientales. Entre los procesos críticos se encuentra la flotación de minerales sulfurados, técnica ampliamente empleada para concentrar especies metálicas de valor separándolas de la ganga. En el caso del hierro, la presencia de sulfuros como la pirrotita ($Fe_{1-x}S$) implica un desafío particular debido a su elevada reactividad y afinidad por colectores como el xantato, generando flotabilidad no deseada y elevando el consumo de reactivos (Baldeón Hurtado & Oropeza Huamán, 2015; Ramos Oxa, 2014).

El control eficiente de reactivos en flotación no solo incide en los costos operativos, sino también en la estabilidad metalúrgica, la calidad del concentrado final y la reducción de la carga química en los efluentes. La pirrotita, al ser un sulfuro parcialmente magnético con elevada reactividad superficial, tiende a consumir colector en exceso de forma empírica, obligando a sobredosificaciones que incrementan el costo por tonelada tratada (Cayunao Collio, 2014; García, Ríos & Alva, 2020). En un contexto global de creciente conciencia ambiental y exigencias regulatorias más estrictas, la industria minera se ve presionada a minimizar su huella ambiental mediante estrategias como la optimización del uso de reactivos químicos, el reciclaje de agua de proceso y la disminución de contaminantes en los efluentes (Azadi et al., 2020; Lottermoser, 2010). Estas prácticas son esenciales para cumplir con estándares internacionales de sostenibilidad y lograr licencias sociales para operar.

A pesar de la importancia del control de reactivos, en muchas plantas concentradoras se carece de modelos predictivos operativos robustos que permitan estimar con precisión el consumo específico de xantato en función de la mineralogía de la alimentación. Las decisiones de dosificación suelen basarse en la experiencia del operador y en ajustes manuales, lo cual

genera resultados variables y frecuentemente ineficientes. La flotación de sulfuros es reconocida como una de las tecnologías más efectivas, pero más sensibles a la variabilidad mineralógica, por lo que la industria busca herramientas basadas en datos reales para mejorar la predictibilidad y la eficiencia de sus circuitos (Laskowski, 2010; Wills & Finch, 2016).

En este contexto, la aplicación de técnicas de aprendizaje automático (machine learning) ofrece la posibilidad de desarrollar modelos predictivos informados por datos históricos de operación, capaces de capturar relaciones estadísticas relevantes y apoyar la toma de decisiones en planta (Breiman, 2001; Kuhn & Johnson, 2013). Sin embargo, para garantizar su aplicabilidad industrial, dichos modelos deben ser interpretables, auditables y suficientemente simples para integrarse en la práctica operativa diaria, evitando sobreajustes innecesarios y privilegiando la parsimonia estadística (James et al., 2013). La adopción de enfoques de minería inteligente no solo mejora el desempeño económico de las operaciones, sino que permite cumplir con compromisos de desarrollo sostenible al reducir la carga de reactivos en el agua residual y facilitar su tratamiento, contribuyendo a objetivos globales de producción más limpia (Hilson & Murck, 2000; Azadi et al., 2020).

En respuesta a esta problemática, la presente investigación propone diseñar y validar un modelo predictivo basado en regresión lineal que permita estimar el consumo específico de xantato (kg/día) en función del contenido de pirrotita (%) en el mineral de alimentación. Esta aproximación busca proporcionar una herramienta práctica y transparente que permita optimizar la dosificación del colector, reducir costos operativos y fortalecer la sostenibilidad ambiental de la operación minera en la planta de flotación de Shougang Hierro Perú.

Descripción de la realidad problemática

La planta concentradora de Shougang Hierro Perú, ubicada en el distrito de Marcona, Ica, constituye uno de los principales complejos industriales dedicados al beneficio de minerales de hierro en el país. Su operación involucra etapas de chancado, molienda, separación magnética y flotación, procesando diariamente decenas de miles de toneladas de mineral. Este flujo continuo implica exigencias operativas considerables, especialmente en el manejo de variabilidad mineralógica y en el control eficiente de reactivos químicos utilizados para maximizar la recuperación y la calidad del concentrado final.

Uno de los desafíos más críticos en la planta es la presencia significativa de sulfuros de hierro, particularmente la pirrotita ($Fe_{1-x}S$), cuya proporción puede variar de forma sustancial entre campañas mineras. Esta fase sulfurada posee propiedades mineralógicas singulares, como estructura cristalina defectuosa y alta reactividad superficial, que le otorgan una fuerte afinidad por colectores aniónicos como el xantato (Lottermoser, 2010). Cuando la pirrotita se presenta en niveles elevados, tiende a consumir colector de manera no selectiva, generando flotabilidad indeseada y compitiendo directamente con minerales valiosos durante el proceso de flotación (Baldeón Hurtado & Oropeza Huamán, 2015).

En la práctica, esta situación ha derivado en estrategias de dosificación principalmente empíricas, donde los operadores ajustan las bombas dosificadoras basándose en experiencia previa, indicadores metalúrgicos y observación visual de la espuma. Aunque esta forma de control ha permitido mantener la continuidad operativa, no logra anticipar con precisión las variaciones del contenido de pirrotita en la alimentación, resultando en sobredosificaciones sistemáticas del colector (García, Ríos & Alva, 2020). Este sobreconsumo no solo eleva los costos directos por tonelada tratada, sino que también incrementa la carga química en los

efluentes, generando mayores costos de tratamiento y potenciales impactos ambientales (Azadi et al., 2020).

Para comprender mejor la dimensión operativa del problema, resulta esencial considerar el volumen de proceso de la planta. Con líneas de flotación capaces de tratar más de 30,000 toneladas por día, incluso pequeñas ineficiencias en la dosificación de reactivos pueden representar gastos millonarios al cierre de cada campaña. Además, el uso excesivo de xantato plantea desafíos ambientales, debido a su naturaleza tóxica y su persistencia en corrientes residuales, lo cual exige sistemas de tratamiento y monitoreo más complejos (Hilson & Murck, 2000).

La gestión de reactivos en este contexto depende no solo del área de flotación, sino de un trabajo coordinado entre las áreas de laboratorio mineralógico, control de calidad, ingeniería de procesos y operación en planta. Cada una de estas áreas tiene un rol en la caracterización del mineral alimentado, el seguimiento del contenido de sulfuros, el diseño de pruebas de laboratorio para ajustar dosificaciones y la implementación de estrategias de control en campo. Sin embargo, la ausencia de un modelo cuantitativo validado que relacione el contenido de pirrotita con el consumo específico de xantato limita la capacidad de planificar de manera proactiva la dosificación, forzando decisiones reactivas y menos eficientes.

A nivel de control de procesos, la planta cuenta con sistemas de monitoreo de caudales, dosificación y calidad de concentrado, pero carece de integración con modelos predictivos que permitan ajustar la dosificación de reactivos en tiempo real en función de las variaciones mineralógicas detectadas. Esto significa que las desviaciones en la mineralogía de la alimentación suelen ser detectadas con retraso, tras afectar la recuperación o la calidad del

concentrado, lo que incrementa la vulnerabilidad operativa ante fluctuaciones en la ley de alimentación o cambios de frente en la mina (Wills & Finch, 2016).

Por ello, existe una necesidad crítica de desarrollar un modelo predictivo práctico y validado que permita estimar el consumo específico de xantato a partir del contenido de pirrotita en la alimentación. Una herramienta de este tipo no solo ayudaría a reducir los costos directos de reactivos, sino que también contribuiría a disminuir la variabilidad del proceso, optimizar la calidad del concentrado y reducir la carga química en los efluentes. Además, fortalecería la capacidad del equipo técnico de la planta para tomar decisiones basadas en datos reales, alineándose con las mejores prácticas de la industria minera en el contexto de minería responsable y sostenible (Goodfellow, Bengio & Courville, 2016; Azadi et al., 2020).

Importancia Industrial del Control de Reactivos en Flotación

El control de reactivos en el proceso de flotación constituye un componente esencial para la eficiencia técnica, económica y ambiental de las operaciones minero-metalúrgicas. En plantas concentradoras de minerales de hierro, el uso de colectores como el xantato es fundamental para promover la flotación de sulfuros, pero su consumo excesivo representa un costo operativo significativo y puede comprometer la estabilidad del proceso (García, Ríos, & Alva, 2020).

La pirrotita ($Fe_{1-x}S$), por su naturaleza mineralógica y elevada reactividad superficial, adsorbe de forma no selectiva el xantato, generando una flotabilidad no deseada. Esta situación obliga a sobredosificaciones empíricas para mantener las recuperaciones metalúrgicas, incrementando los costos por tonelada tratada y reduciendo la eficiencia económica del proceso (Cayunao Collio, 2014; Baldeón Hurtado & Oropeza Huamán, 2015).

Adicionalmente, el uso ineficiente de reactivos como el xantato eleva la carga química en los efluentes, generando impactos ambientales negativos y mayores costos de tratamiento, lo cual resulta contrario a los principios de minería responsable y sostenible (Goodfellow, Bengio, & Courville, 2016). La optimización del consumo específico de xantato (kg/día) se convierte, así, en una prioridad estratégica para las empresas mineras, permitiendo reducir costos directos, mejorar la calidad del concentrado y cumplir con estándares ambientales cada vez más estrictos (Wills & Finch, 2016).

En este contexto, contar con herramientas predictivas que permitan estimar de manera precisa la dosificación óptima de colector en función del contenido de pirrotita en la alimentación resulta esencial para mejorar la toma de decisiones operativas, garantizar la competitividad del proceso y fortalecer la sostenibilidad global de la actividad minera (Nasiri, Iqbal, & Särkkä, 2024; Kuhn & Johnson, 2013).

Impacto de la Variabilidad Mineralógica en la Eficiencia del Proceso

La pirrotita ($Fe_{1-x}S$) es reconocida como el principal factor mineralógico que afecta la eficiencia del proceso de flotación de minerales de hierro debido a su comportamiento reactivo particular. Este mineral sulfurado presenta una estructura cristalina defectuosa en hierro, lo que le otorga una alta reactividad superficial y una elevada afinidad por colectores como el xantato (Baldeón Hurtado & Oropeza Huamán, 2015; Ramos Oxa, 2014).

La presencia de pirrotita en el mineral de alimentación genera un efecto directo sobre la flotabilidad no deseada de partículas, ya que su superficie puede activarse con iones metálicos presentes en la pulpa, aumentando su capacidad de adsorber el colector. Este fenómeno interfiere con la selectividad del proceso, ya que el xantato, al ser consumido por la pirrotita, deja de estar

disponible para los minerales valiosos, afectando la recuperación metalúrgica y obligando a incrementar la dosificación del reactivo (Cayunao Collio, 2014).

La variabilidad en el contenido de pirrotita en el mineral de alimentación implica fluctuaciones significativas en el consumo específico de xantato (kg/día), generando desafíos operativos en el control del proceso. Sin un sistema predictivo adecuado que permita anticipar esta relación, las decisiones de dosificación se basan en ajustes empíricos que suelen resultar en sobredosificación sistemática, elevando los costos operativos y afectando la rentabilidad de la operación minera (García, Ríos, & Alva, 2020).

Además, el sobreconsumo de reactivos colectores incrementa la carga química en los efluentes, generando mayores costos de tratamiento y potenciales impactos ambientales, lo que contradice los principios de minería responsable y sostenible (Wills & Finch, 2016; Goodfellow, Bengio, & Courville, 2016). Por ello, comprender y modelar la influencia del contenido de pirrotita sobre el consumo específico de xantato es fundamental para optimizar la eficiencia metalúrgica, reducir los costos de insumos y minimizar el impacto ambiental asociado al proceso de flotación.

Limitaciones de los Enfoques Tradicionales de Control Operativo

En muchas plantas concentradoras, incluyendo la operación de Shougang Hierro Perú, las decisiones relacionadas con la dosificación de xantato se han basado tradicionalmente en el conocimiento empírico de los operadores y en ajustes manuales sostenidos por seguimiento estadístico de indicadores metalúrgicos. Aunque la experiencia operativa resulta valiosa, este enfoque presenta limitaciones significativas frente a la variabilidad del contenido de pirrotita en la alimentación (García, Ríos, & Alva, 2020).

La pirrotita, por su naturaleza mineralógica y su elevada reactividad, es responsable de fluctuaciones significativas en la demanda de colector. Sin un modelo predictivo que cuantifique la relación directa entre su contenido y el consumo específico de xantato, los operadores se ven obligados a sobredosificar de forma conservadora para evitar pérdidas metalúrgicas, incurriendo en costos innecesarios y reduciendo la eficiencia del proceso (Cayunao Collio, 2014; Baldeón Hurtado & Oropeza Huamán, 2015).

Además, los métodos de control convencional basados en reglas fijas o tendencias históricas no permiten adaptarse de manera proactiva a cambios rápidos en la mineralogía del mineral alimentado. Esto genera un rezago en la toma de decisiones y aumenta la vulnerabilidad del sistema ante variaciones no previstas (Nasiri, Iqbal, & Särkkä, 2024).

En ausencia de un enfoque analítico robusto y cuantitativo, la dosificación de reactivos se transforma en un proceso reactivo y poco optimizado. Por ello, se requiere el desarrollo de herramientas predictivas simples, interpretables y basadas en datos reales de operación, capaces de anticipar con precisión el consumo específico de xantato a partir del contenido de pirrotita, fortaleciendo así la capacidad de respuesta técnica y la sostenibilidad económica de la operación minera (Kuhn & Johnson, 2013).

Propuesta de Solución Basada en Aprendizaje Automático

Frente a la necesidad de optimizar la dosificación de xantato en el circuito de flotación de minerales de hierro con presencia de pirrotita, se propone el desarrollo de un modelo predictivo basado en técnicas de aprendizaje automático, específicamente regresión lineal . Este enfoque busca modelar de forma cuantitativa y explícita la relación directa entre el contenido de pirrotita (%) en la alimentación y el consumo específico de xantato (kg/día), aprovechando datos reales de operación.

La regresión lineal se selecciona como técnica base por su interpretabilidad, su bajo riesgo de sobreajuste y su facilidad de implementación en sistemas operativos de planta (James et al., 2013). A diferencia de modelos complejos que incorporan múltiples variables o relaciones no lineales difíciles de explicar, la aproximación garantiza parsimonia estadística: se enfoca exclusivamente en el factor mineralógico responsable del sobreconsumo de reactivo, evitando la inclusión de variables redundantes o colineales (Kuhn & Johnson, 2013).

Este modelo permitirá anticipar el consumo específico de xantato en función del contenido de pirrotita medido mediante análisis mineralógico o balances metalúrgicos, proporcionando a los operadores una herramienta práctica para tomar decisiones informadas y técnicas. Además, su desarrollo contribuirá a la integración progresiva de herramientas de inteligencia artificial en la industria minera, fortaleciendo la capacidad analítica y promoviendo la transición hacia operaciones más sostenibles y basadas en datos (Nasiri, Iqbal, & Särkkä, 2024).

La propuesta no excluye futuros modelos multivariables más complejos que puedan considerar otras variables operativas. Sin embargo, establece como objetivo principal un modelo robusto, validado y aplicable de forma inmediata en condiciones industriales, contribuyendo a reducir el consumo excesivo de reactivos, los costos operativos y el impacto ambiental asociado al proceso de flotación.

Justificación del Enfoque Adoptado

La elección de un modelo predictivo basado en regresión lineal para estimar el consumo específico de xantato (kg/día) en función del contenido de pirrotita (%) se fundamenta en la necesidad de desarrollar una herramienta práctica, interpretable y aplicable de forma inmediata en la operación minera. La pirrotita ha sido identificada como el factor mineralógico principal

que causa el sobreconsumo de xantato, debido a su elevada afinidad química con este colector y su comportamiento de flotabilidad no deseada (Baldeón Hurtado & Oropeza Huamán, 2015; Cayunao Collio, 2014).

A diferencia de enfoques multivariables o modelos no lineales más complejos, la regresión lineal garantiza parsimonia estadística, es decir, la preferencia por el modelo más simple capaz de explicar adecuadamente el fenómeno observado (Kuhn & Johnson, 2013; James et al., 2013). Esta simplicidad no solo facilita la interpretación por parte del personal técnico de planta, sino que también reduce el riesgo de sobreajuste y mejora la transparencia del proceso de toma de decisiones.

El modelo propuesto permite vincular de forma explícita y cuantitativa el contenido de pirrotita con el consumo específico de xantato, facilitando la implementación de estrategias de dosificación más racionales y consistentes. Su carácter está alineado con el principio de enfocarse en la causa directa del problema operativo, sin complicaciones innecesarias asociadas a variables colineales o de difícil medición (Nasiri, Iqbal, & Särkkä, 2024).

Además, el uso de una técnica estadística ampliamente validada y accesible como la regresión lineal facilita su integración en herramientas digitales existentes, planillas de cálculo o sistemas SCADA de planta. Esto fortalece la capacidad analítica del equipo operativo, promueve la minería inteligente basada en datos y contribuye a la optimización de costos y la sostenibilidad ambiental del proceso de flotación (Wills & Finch, 2016; García, Ríos, & Alva, 2020).

Aportes Esperados de la Investigación

La presente investigación busca desarrollar y validar un modelo predictivo , basado en regresión lineal, que permita estimar de manera precisa el consumo específico de xantato

(kg/día) en función del contenido de pirrotita (%) en el mineral de alimentación. Este aporte técnico está alineado con la necesidad de contar con herramientas prácticas y aplicables que mejoren la toma de decisiones en planta, especialmente en circuitos de flotación de minerales de hierro con presencia significativa de sulfuros reactivos (Baldeón Hurtado & Oropeza Huamán, 2015; Cayunao Collio, 2014).

Uno de los principales beneficios esperados es la reducción del sobreconsumo del reactivo colector xantato, generando ahorros directos en los costos operativos por tonelada tratada. Al anticipar de manera cuantitativa la cantidad óptima de reactivo en función de la mineralogía de la alimentación, se mejora la eficiencia metalúrgica al reducir pérdidas asociadas a la flotabilidad no deseada de la pirrotita (García, Ríos, & Alva, 2020).

Adicionalmente, el modelo contribuirá a la sostenibilidad ambiental de la operación minera, al disminuir la carga química en los efluentes y los costos asociados a su tratamiento. Esto se alinea con los principios de minería responsable y con las demandas crecientes por parte de la sociedad y las autoridades regulatorias para minimizar el impacto ambiental de la actividad extractiva (Wills & Finch, 2016; Goodfellow, Bengio, & Courville, 2016).

Finalmente, la investigación fortalecerá la capacidad de la planta para integrar prácticas de minería inteligente basadas en datos reales, promoviendo la transición hacia sistemas de control más predictivos y menos dependientes de ajustes manuales o experiencia empírica. De este modo, se avanza hacia operaciones más estandarizadas, eficientes y sostenibles, generando valor tanto económico como ambiental para la empresa y la comunidad.

Definición del Problema

La planta concentradora de Shougang Hierro Perú enfrenta un desafío operativo crítico vinculado a la dosificación ineficiente del colector xantato durante el proceso de flotación de minerales de hierro. Esta ineficiencia se presenta especialmente en las campañas donde el mineral de alimentación contiene altos niveles de pirrotita ($Fe_{1-x}S$), un sulfuro de hierro caracterizado por su elevada reactividad superficial y su capacidad de adsorber colectores de forma no selectiva (Baldeón Hurtado & Oropeza Huamán, 2015; Ramos Oxa, 2014).

La pirrotita, al competir con los minerales valiosos por el xantato, genera flotabilidad no deseada y obliga a incrementar la dosificación del colector para mantener las recuperaciones metalúrgicas esperadas. Sin embargo, la dosificación en planta se realiza actualmente de forma empírica y manual, basándose en la experiencia del operador y en ajustes estadísticos convencionales que no logran anticipar las variaciones en el contenido de pirrotita de manera precisa (García, Ríos, & Alva, 2020).

Esta situación conduce a un sobreconsumo sistemático del reactivo, elevando los costos operativos por tonelada tratada y comprometiendo la rentabilidad y sostenibilidad del proceso. Además, el uso ineficiente de reactivos aumenta la carga química en los efluentes, incrementando los costos de tratamiento y el impacto ambiental asociado a la operación minera (Wills & Finch, 2016; Goodfellow, Bengio, & Courville, 2016).

En este contexto, el problema central radica en la ausencia de un modelo predictivo robusto y validado que permita estimar de manera confiable la cantidad óptima de xantato en función del contenido de pirrotita en el mineral de alimentación. Sin este enfoque basado en datos reales de operación, la planta carece de una herramienta técnica para tomar decisiones

informadas, optimizar el uso de reactivos y garantizar la eficiencia operativa y ambiental del circuito de flotación.

Problema General

¿De qué manera el contenido de pirrotita (% S/FeS) en el mineral de alimentación influye sobre el consumo específico de xantato (kg/día) en el proceso de flotación de minerales de hierro en la planta Shougang Hierro Perú – Marcona, y puede esta relación ser modelada adecuadamente mediante una regresión lineal para optimizar su dosificación?

Problemas Específicos

¿Existe una relación estadísticamente significativa entre el contenido de pirrotita en el mineral de alimentación y el consumo específico de xantato (kg/día) en el proceso de flotación de minerales de hierro, según los registros históricos de operación en la planta Shougang Hierro Perú – Marcona?

¿Puede la influencia del contenido de pirrotita sobre el consumo específico de xantato modelarse adecuadamente mediante una regresión lineal , y cuál es el nivel de ajuste estadístico que presenta dicho modelo en condiciones reales de planta?

¿Cuál es el margen de error y el grado de aplicabilidad operativa del modelo predictivo desarrollado, al estimar el consumo específico de xantato en función del contenido de pirrotita, en el contexto de la planta concentradora de Shougang Hierro Perú – Marcona?

Objetivos

Objetivo General

Evaluar y validar un modelo predictivo que relacione el contenido de pirrotita (% S/FeS) en el mineral de alimentación con el consumo específico de xantato (kg/día) en el proceso de

flotación de minerales de hierro, con el fin de optimizar su dosificación en condiciones reales de operación en la planta concentradora de Shougang Hierro Perú – Marcona.

Objetivos Específicos

Analizar la existencia de una relación estadísticamente significativa entre el contenido de pirrotita en el mineral de alimentación y el consumo específico de xantato (kg/día), a partir del tratamiento de datos históricos de operación de la planta concentradora de Shougang Hierro Perú – Marcona.

Desarrollar un modelo predictivo, basado en regresión lineal, que represente cuantitativamente la influencia del contenido de pirrotita sobre el consumo específico de xantato, y evaluar su nivel de ajuste estadístico en condiciones operativas reales.

Evaluar el desempeño predictivo y la aplicabilidad operativa del modelo de regresión lineal propuesto, mediante el análisis del margen de error y la consistencia de sus estimaciones en la predicción del consumo de xantato, en función del contenido de pirrotita.

Justificación y Relevancia de la Investigación

Justificación Teórica

El consumo específico de xantato (kg/día) es un parámetro crítico en el proceso de flotación de minerales de hierro, dado su impacto directo en los costos operativos y la eficiencia metalúrgica. Diversos estudios han demostrado que la presencia de pirrotita ($Fe_{1-x}S$) en el mineral de alimentación es el principal factor mineralógico que genera sobreconsumo del colector, debido a su elevada reactividad superficial y su capacidad de adsorber el xantato de forma no selectiva (Baldeón Hurtado & Oropeza Huamán, 2015; Cayunao Collio, 2014).

La pirrotita presenta una estructura cristalina defectuosa en hierro, lo que incrementa su afinidad química con colectores aniónicos como el xantato, generando flotabilidad no deseada y reduciendo la selectividad del proceso (Ramos Oxa, 2014). Esta interacción obliga a incrementar la dosificación del reactivo para mantener las recuperaciones metalúrgicas esperadas, generando sobrecostos operativos y mayor carga química en los efluentes (García, Ríos, & Alva, 2020; Wills & Finch, 2016).

Desde el punto de vista teórico, esta relación directa y significativa entre el contenido de pirrotita y el consumo específico de xantato justifica el uso de modelos predictivos s que permitan estimarla de manera cuantitativa. La regresión lineal es una técnica estadística ampliamente validada y utilizada en ingeniería por su capacidad de modelar relaciones entre variables continuas, su interpretabilidad y su facilidad de implementación en sistemas de soporte a la decisión (James et al., 2013; Kuhn & Johnson, 2013).

Adoptar un enfoque y parsimonioso resulta apropiado para esta investigación, ya que permite explicar el fenómeno con la mínima complejidad necesaria, reduciendo el riesgo de sobreajuste y facilitando la interpretación por parte del personal técnico de planta. Además, sienta las bases para la integración de herramientas de minería inteligente basadas en datos reales de operación, contribuyendo a la optimización de costos, la mejora de la eficiencia metalúrgica y la sostenibilidad ambiental de la operación minera (Goodfellow, Bengio, & Courville, 2016).

Justificación Metodológica

La presente investigación adopta como enfoque metodológico el uso de regresión lineal para modelar la relación entre el contenido de pirrotita (%) en el mineral de alimentación y el consumo específico de xantato (kg/día) en el proceso de flotación. Esta elección se fundamenta

en la necesidad de construir un modelo predictivo simple, interpretable y aplicable de forma inmediata en condiciones reales de operación en planta.

La pirrotita ha sido identificada como la variable independiente principal y causa directa del sobreconsumo de xantato, dado su comportamiento mineralógico altamente reactivo y su capacidad de adsorber colectores de forma no selectiva (Baldeón Hurtado & Oropeza Huamán, 2015; Cayunao Collio, 2014). Por ello, un enfoque resulta adecuado para capturar esta relación directa, evitando la complejidad innecesaria de incluir variables colineales o menos relevantes.

La regresión lineal ofrece ventajas metodológicas importantes: es estadísticamente robusta, de bajo riesgo de sobreajuste y altamente interpretable, lo que facilita su adopción por parte del personal operativo y su integración en sistemas de soporte a la decisión como planillas de cálculo o sistemas SCADA (James et al., 2013; Kuhn & Johnson, 2013). Además, permite generar predicciones confiables basadas en datos históricos reales de operación, fortaleciendo el componente empírico y práctico de la propuesta.

Este enfoque metodológico parsimonioso, al centrarse exclusivamente en la pirrotita como variable independiente, asegura mayor transparencia y replicabilidad, facilitando su aplicación directa en la toma de decisiones operativas y contribuyendo a la optimización de costos, la eficiencia metalúrgica y la sostenibilidad ambiental de la planta de flotación (Goodfellow, Bengio, & Courville, 2016).

Justificación Práctica

El desarrollo de un modelo predictivo basado en regresión lineal para estimar el consumo específico de xantato (kg/día) en función del contenido de pirrotita (%) tiene una justificación práctica sólida y directamente alineada con las necesidades operativas de la planta Shougang

Hierro Perú. Actualmente, la dosificación de xantato se realiza de forma empírica y manual, basándose en la experiencia del operador y en indicadores históricos que no capturan de manera precisa la variabilidad del contenido de pirrotita en el mineral de alimentación (García, Ríos, & Alva, 2020).

Esta situación genera un sobreconsumo sistemático del colector, elevando los costos por tonelada tratada, disminuyendo la rentabilidad del proceso y aumentando la carga química en los efluentes, lo que implica mayores costos de tratamiento y un impacto ambiental más alto (Wills & Finch, 2016). Un modelo predictivo simple y confiable permitirá anticipar la cantidad óptima de reactivo necesario en función de la mineralogía, reduciendo sobredosificaciones y estabilizando la calidad del concentrado producido.

Al ser un enfoque , el modelo propuesto es fácil de interpretar y de implementar en sistemas existentes de la planta, como planillas de cálculo o sistemas SCADA, sin requerir inversiones significativas en infraestructura o entrenamiento especializado (Kuhn & Johnson, 2013). Además, facilita la capacitación del personal operativo, promoviendo una toma de decisiones más técnica y menos dependiente de criterios subjetivos.

De este modo, la propuesta contribuye a la minería inteligente basada en datos reales de operación, fortaleciendo la capacidad de planificación, reduciendo costos, mejorando la eficiencia metalúrgica y cumpliendo con estándares ambientales más exigentes, alineándose con los principios de sostenibilidad y responsabilidad social de la industria minera (Goodfellow, Bengio, & Courville, 2016).

Pertinencia Industrial y Tecnológica

La implementación de un modelo predictivo basado en regresión lineal para estimar el consumo específico de xantato (kg/día) en función del contenido de pirrotita (%) resulta altamente pertinente para la industria minera, especialmente en el contexto peruano, donde se busca optimizar el uso de insumos químicos y garantizar la sostenibilidad ambiental de las operaciones (García, Ríos, & Alva, 2020).

En la planta concentradora de Shougang Hierro Perú, la pirrotita representa el principal factor mineralógico que genera sobreconsumo de xantato debido a su alta reactividad y flotabilidad no deseada (Baldeón Hurtado & Oropeza Huamán, 2015). Sin embargo, la dosificación actual del reactivo sigue basándose en criterios empíricos y en ajustes manuales que no permiten anticipar la variabilidad mineralógica de forma precisa.

La adopción de un modelo claro y transparente permitirá formalizar el conocimiento operativo existente, transformándolo en una herramienta analítica objetiva y fácilmente aplicable en condiciones reales de planta. Al requerir datos simples y accesibles como el contenido de pirrotita, su implementación no implica inversiones significativas en infraestructura o personal altamente especializado (Kuhn & Johnson, 2013).

Desde una perspectiva tecnológica, esta propuesta contribuye a la transición de la industria minera peruana hacia sistemas de control más inteligentes, basados en datos y soportados por modelos estadísticos replicables y auditables. Esto fortalece la capacidad de respuesta técnica frente a exigencias de productividad, rentabilidad y estándares ambientales cada vez más rigurosos (Wills & Finch, 2016; Goodfellow, Bengio, & Courville, 2016).

Marco Teórico

Fundamento de la Flotación en Minerales de Hierro

La flotación es una técnica de concentración diferencial ampliamente utilizada en la industria minera para separar minerales valiosos de la ganga. En el caso de minerales de hierro, el proceso se emplea principalmente para remover sulfuros presentes como impurezas, mejorando la calidad del concentrado y cumpliendo con especificaciones comerciales (Wills & Finch, 2016).

Los sulfuros de hierro, entre los que destaca la pirrotita ($Fe_{1-x}S$), son especies mineralógicas con propiedades fisicoquímicas que les otorgan una alta flotabilidad natural. La pirrotita es especialmente relevante en ciertos yacimientos por su reactividad superficial, su capacidad de activación con iones metálicos y su elevada afinidad por colectores aniónicos como el xantato (Baldeón Hurtado & Oropeza Huamán, 2015; Cayunao Collio, 2014).

Durante la flotación, la pirrotita puede adsorber el colector incluso en condiciones de pH moderadamente alcalinas, compitiendo con los minerales valiosos por el reactivo y generando problemas de selectividad (Ramos Oxa, 2014). Esta competencia obliga a ajustar la dosificación del colector para mantener las recuperaciones metalúrgicas, incrementando el consumo específico de xantato (kg/día) y elevando los costos operativos por tonelada tratada (García, Ríos, & Alva, 2020).

Comprender los mecanismos de interacción químico-superficial de la pirrotita con el colector resulta esencial para optimizar la dosificación y mejorar la eficiencia del proceso de flotación. En consecuencia, modelar la relación entre el contenido de pirrotita en la alimentación y el consumo específico de xantato representa un paso clave hacia un control más racional,

basado en datos, alineado con los principios de minería inteligente y sostenible (Wills & Finch, 2016; Goodfellow, Bengio, & Courville, 2016).

Asimismo, investigaciones recientes han comenzado a aplicar técnicas de aprendizaje automático para interpretar, predecir y controlar dinámicamente el comportamiento de especies mineralógicas en sistemas de flotación. En particular, Chen, Zhou y Zhang (2022) demostraron que modelos de deep learning pueden emplearse para establecer esquemas de dosificación automática de colectores en función de patrones históricos y en tiempo real, elevando la eficiencia de control y reduciendo la variabilidad del proceso.

En la misma línea, Molina, Ruiz y Alarcón (2023) desarrollaron un sistema basado en algoritmos supervisados para optimizar el consumo de reactivos en circuitos de flotación de minerales de hierro. Sus resultados indican que la integración de datos mineralógicos con modelos estadísticos simples puede generar ahorros significativos en el uso de xantato, siempre que exista una caracterización adecuada del mineral de alimentación.

Comportamiento Reactivo de la Pirrotita

La pirrotita ($Fe_{1-x}S$) es un sulfuro de hierro caracterizado por una estructura cristalina defectuosa en hierro, lo que le confiere propiedades fisicoquímicas altamente reactivas. Su superficie mineralógica presenta una elevada densidad de sitios reactivos capaces de adsorber colectores aniónicos como el xantato, incluso en condiciones de pH moderadamente alcalinas, generando flotabilidad natural no deseada (Baldeón Hurtado & Oropeza Huamán, 2015; Cayunao Collio, 2014).

A diferencia de otros sulfuros de hierro, la pirrotita posee una capacidad particular de activarse con iones metálicos presentes en la pulpa, como Cu^{2+} o Pb^{2+}, los cuales potencian su

afinidad por el xantato y elevan la cinética de adsorción del colector sobre su superficie (Ramos Oxa, 2014). Esta reactividad elevada dificulta la selectividad del proceso de flotación, ya que el colector destinado a minerales valiosos es consumido en gran parte por la pirrotita.

El resultado operativo de este comportamiento es el incremento en la dosificación de xantato para mantener recuperaciones metalúrgicas aceptables, generando un aumento directo en el consumo específico de reactivo (kg/día) y elevando los costos por tonelada tratada (García, Ríos, & Alva, 2020).

Por estas razones, la pirrotita se reconoce como la causa principal del sobreconsumo de xantato en el proceso de flotación de minerales de hierro. Comprender su comportamiento reactivo es esencial para modelar de forma cuantitativa la relación entre su contenido en la alimentación y el consumo específico de colector, lo que resulta clave para optimizar la dosificación y mejorar la eficiencia del proceso.

En los últimos años, se han desarrollado estudios más sofisticados que validan cuantitativamente el papel de la pirrotita como principal consumidor no selectivo de colectores en flotación. Por ejemplo, Liu, Xu y Finch (2022) aplicaron técnicas de electrocinética y espectroscopía de impedancia para confirmar que la densidad de carga superficial negativa de la pirrotita activada favorece la adsorción del xantato, incluso en medios de pH controlado. Esta afinidad se mantiene incluso bajo condiciones de preoxidación parcial, lo que evidencia una reactividad persistente del mineral frente a colectores aniónicos.

Asimismo, investigaciones recientes en circuitos industriales de flotación de sulfuros han demostrado que la pirrotita no solo se comporta como una fase colectora pasiva, sino que puede inducir procesos de activación cruzada, favoreciendo la flotabilidad de otras especies sulfuradas

por transferencia de iones metálicos en solución. Este efecto sinérgico ha sido documentado por Hassanzadeh y Karimi (2021), quienes observaron que la presencia de pirrotita aumentó significativamente el consumo total de reactivos en sistemas complejos, aun cuando su contenido mineralógico no fuera dominante. Estos hallazgos confirman la necesidad de considerar a la pirrotita no solo como una fase mineralógica reactiva, sino como una variable predictora operativa crítica.

Relación entre Contenido de Pirrotita y Consumo de Xantato

La pirrotita ($Fe_{1-x}S$) es ampliamente reconocida como el principal factor mineralógico que condiciona el consumo específico de xantato (kg/día) en los procesos de flotación de minerales de hierro. Su estructura cristalina defectuosa en hierro la hace químicamente muy reactiva, generando una afinidad elevada por colectores aniónicos como el xantato. Esta característica se traduce en una adsorción no selectiva del colector, incrementando la flotabilidad no deseada de la pirrotita y reduciendo la selectividad del proceso de flotación (Baldeón Hurtado & Oropeza Huamán, 2015; Wills & Finch, 2016).

Operativamente, la presencia de pirrotita obliga a realizar ajustes en la dosificación del colector para compensar su efecto competitivo. Sin embargo, en muchas plantas este ajuste se basa en experiencia empírica y en reglas heurísticas que no logran capturar de forma cuantitativa la variabilidad del contenido de pirrotita en la alimentación (García, Ríos, & Alva, 2020). Esto genera un sobreconsumo sistemático de xantato, elevando los costos operativos por tonelada tratada y generando mayores cargas químicas en los efluentes.

La relación entre el contenido de pirrotita y el consumo específico de xantato es, por lo tanto, directa y significativa: a mayor porcentaje de pirrotita en la alimentación, mayor es la demanda de colector para mantener la eficiencia metalúrgica deseada. Diversos estudios han

documentado esta relación lineal en rangos operativos típicos, lo cual permite plantear su modelamiento mediante técnicas estadísticas simples, como la regresión lineal (James et al., 2013; Kuhn & Johnson, 2013).

Esta aproximación metodológica no solo es teóricamente sólida, sino también práctica. Al enfocarse exclusivamente en la pirrotita como variable independiente principal, se evita la inclusión de variables colineales o menos relevantes que podrían complicar innecesariamente el modelo sin aportar valor predictivo. La parsimonia estadística garantiza un modelo fácil de interpretar y de implementar en herramientas existentes en planta, como planillas de cálculo o sistemas SCADA, facilitando su adopción por el personal operativo (Goodfellow, Bengio, & Courville, 2016).

En este contexto, modelar la relación entre el contenido de pirrotita y el consumo específico de xantato permite pasar de decisiones empíricas y reactivas a una gestión proactiva basada en datos reales de operación. Este cambio contribuye a optimizar el uso de reactivos, reducir costos directos, mejorar la estabilidad del proceso metalúrgico y cumplir con estándares ambientales cada vez más exigentes, alineándose con los principios de la minería inteligente y sostenible (Wills & Finch, 2016).

Estudios recientes han confirmado que el contenido de pirrotita en el mineral de alimentación es una de las variables con mayor capacidad explicativa sobre la variabilidad del consumo de xantato en circuitos de flotación. En particular, investigaciones desarrolladas por Paryani, Bahadori y Farzanegan (2021) aplicaron modelos de aprendizaje automático a datos operativos reales, evidenciando que la pirrotita fue consistentemente la variable de mayor importancia relativa en la predicción del consumo de colectores, superando a parámetros como el pH o el contenido total de azufre.

Aplicaciones de Machine Learning en Procesos Metalúrgicos

El aprendizaje automático (machine learning) ha demostrado ser una herramienta poderosa para la industria minera, al permitir desarrollar modelos predictivos basados en datos reales de operación para optimizar variables críticas de procesos metalúrgicos. En flotación, se ha aplicado para modelar recuperaciones metalúrgicas, predecir leyes de concentrado y optimizar la dosificación de reactivos, ayudando a reducir costos y mejorar la estabilidad del proceso (Goodfellow, Bengio, & Courville, 2016; Nasiri, Iqbal, & Särkkä, 2024).

Entre las técnicas más conocidas se encuentran los modelos de árboles de decisión como Random Forest o algoritmos de boosting como XGBoost, que permiten capturar relaciones no lineales y gestionar interacciones complejas entre múltiples variables. Estas herramientas son útiles en problemas con muchas variables independientes y cuando se espera aprovechar interacciones no obvias entre factores operativos (James et al., 2013). Sin embargo, su complejidad estadística y su naturaleza de "caja negra" pueden dificultar su aplicación práctica en planta, especialmente cuando se busca interpretar resultados y explicar decisiones a operadores y supervisores (Kuhn & Johnson, 2013).

En contraste, la regresión lineal ofrece ventajas importantes para el presente estudio, al centrarse exclusivamente en modelar la relación entre el contenido de pirrotita en la alimentación y el consumo específico de xantato (kg/día). Esta relación está bien documentada tanto teórica como empíricamente, fundamentándose en la afinidad química de la pirrotita por el xantato, su reactividad superficial elevada y su capacidad de adsorber colector de forma no selectiva (Baldeón Hurtado & Oropeza Huamán, 2015; García, Ríos, & Alva, 2020).

El modelo de regresión lineal no solo resulta más simple y transparente, sino que facilita su implementación en entornos industriales reales, como planillas de cálculo o sistemas SCADA,

sin requerir infraestructura tecnológica avanzada o entrenamiento especializado. Su parsimonia estadística permite enfocarse en la causa directa del problema, evitando sobreajuste y asegurando replicabilidad y auditabilidad del resultado (James et al., 2013; Kuhn & Johnson, 2013).

De esta manera, aunque existen métodos más complejos en el repertorio del machine learning, el presente estudio adopta la regresión lineal como la opción metodológica más adecuada para explicar y predecir la relación entre contenido de pirrotita y consumo específico de xantato, proporcionando una herramienta práctica, aplicable y de alto valor para la toma de decisiones operativas en la planta de flotación de Shougang Hierro Perú (Wills & Finch, 2016).

Recientemente, la comunidad científica ha mostrado un creciente interés por el desarrollo de modelos predictivos en flotación basados en datos históricos de planta. Molina, Ruiz y Alarcón (2023) reportaron la aplicación exitosa de algoritmos de regresión supervisada para predecir el consumo de reactivos en circuitos de flotación de hierro, destacando que la variable mineralógica más influyente fue la pirrotita. Su investigación mostró que, si bien los modelos multivariables pueden ofrecer altos niveles de ajuste, su implementación requiere estructuras de datos robustas y un mantenimiento continuo que no siempre es viable en contextos industriales.

En esa línea, Chen, Zhou y Zhang (2022) utilizaron modelos de aprendizaje profundo (deep learning) para el control de la dosificación de reactivos, alcanzando mejoras en eficiencia operativa. Sin embargo, también advirtieron que estos modelos tienden a funcionar como “cajas negras”, dificultando su interpretación por parte del personal de planta, lo que limita su adopción práctica en operaciones donde se requiere trazabilidad de las decisiones.

Por ello, se ha empezado a promover el uso de técnicas de machine learning explicable o modelos interpretables, tales como la regresión lineal, árboles de decisión simplificados o

análisis SHAP, que permiten comprender el efecto individual de cada variable sobre el resultado. Este enfoque ha sido recomendado por Nasiri, Iqbal y Särkkä (2024), quienes proponen integrar conocimiento fisicoquímico con modelos estadísticos simples que respeten la lógica operativa del proceso.

Ventajas de la Inteligencia Artificial para el Control de Reactivos

La incorporación de herramientas de inteligencia artificial (IA) en la industria minera representa un cambio de paradigma al permitir la toma de decisiones basada en datos reales y no únicamente en experiencia empírica. En el caso específico del control de reactivos de flotación, la IA ofrece la posibilidad de modelar relaciones complejas entre variables operativas y mineralógicas, anticipar variaciones de consumo y optimizar dosificaciones de forma dinámica (Goodfellow, Bengio, & Courville, 2016).

Tradicionalmente, la dosificación de colectores como el xantato se ha basado en reglas operativas y ajustes manuales, que, si bien aprovechan la experiencia del personal, no logran capturar de manera cuantitativa la influencia de factores críticos como el contenido de pirrotita en la alimentación. La consecuencia ha sido un sobreconsumo sistemático de reactivo, mayores costos por tonelada tratada y mayor carga química en los efluentes (García, Ríos, & Alva, 2020).

El uso de técnicas de machine learning permite superar estas limitaciones al desarrollar modelos predictivos que transforman registros históricos en conocimiento accionable. Al aprender patrones estadísticos en datos de planta, estos modelos pueden anticipar el consumo específico de xantato en función del contenido de pirrotita, facilitando decisiones proactivas y más racionales (James et al., 2013; Nasiri, Iqbal, & Särkkä, 2024).

Aunque existen algoritmos avanzados como redes neuronales o modelos de árboles de decisión que pueden capturar interacciones no lineales y relaciones multivariables, para el presente caso se privilegia la regresión lineal por su capacidad de explicar directamente la relación principal entre la pirrotita y el consumo específico de xantato. Esta elección metodológica obedece al principio de parsimonia estadística y a la necesidad de interpretabilidad y aplicabilidad práctica en planta (Kuhn & Johnson, 2013).

La regresión lineal permite generar modelos transparentes, fáciles de validar y de integrar en sistemas existentes como planillas de cálculo o plataformas SCADA. Además, facilita la capacitación del personal operativo, reduciendo la resistencia al cambio tecnológico y promoviendo una cultura de decisión basada en datos reales (Wills & Finch, 2016).

En este sentido, la aplicación de inteligencia artificial mediante modelos estadísticos simples contribuye no solo a la optimización técnica y económica del proceso de flotación, sino también a la sostenibilidad ambiental al reducir el sobreconsumo de reactivos y la carga de efluentes, alineándose con estándares internacionales de minería responsable.

Antecedentes del Estudio

Estudios Nacionales Sobre la Influencia de la Pirrotita

En el contexto peruano, la pirrotita ha sido identificada como uno de los principales desafíos mineralógicos en circuitos de flotación de minerales de hierro y sulfuros complejos. Su elevada reactividad superficial y su capacidad para adsorber colectores aniónicos como el xantato la convierten en la causa directa del sobreconsumo de este reactivo, obligando a sobredosificaciones para mantener las recuperaciones metalúrgicas deseadas (Baldeón Hurtado & Oropeza Huamán, 2015).

Diversos trabajos de tesis desarrollados en universidades peruanas han documentado la relación entre contenido de pirrotita y consumo de reactivos. Por ejemplo, Baldeón Hurtado y Oropeza Huamán (2015) estudiaron minerales con alto contenido de pirrotita magnética en la provincia de Caravelí, concluyendo que la dosificación de xantato debía incrementarse significativamente para compensar la flotabilidad no deseada de la pirrotita. Este hallazgo confirma la necesidad de entender y cuantificar la influencia de este mineral para optimizar la estrategia de dosificación.

Asimismo, Ramos Oxa (2014) evaluó parámetros de flotación en minerales polimetálicos con presencia de pirrotita, destacando su comportamiento como especie reactiva que compite con sulfuros valiosos por el colector. Sus resultados mostraron incrementos en el consumo específico de xantato directamente asociados a aumentos en el porcentaje de pirrotita en la alimentación, reforzando la necesidad de incluir su medición sistemática en los controles de planta.

En estudios más recientes, García, Ríos y Alva (2020) modelaron estadísticamente el consumo de reactivos en flotación de sulfuros complejos en plantas peruanas, resaltando la pirrotita como variable clave que explica gran parte de la variabilidad en el consumo de xantato.

Estos antecedentes nacionales subrayan la importancia de desarrollar modelos predictivos basados en datos reales de operación que permitan estimar el consumo específico de colector en función del contenido de pirrotita, contribuyendo así a la reducción de costos operativos y a la sostenibilidad del proceso.

Investigaciones Latinoamericanas sobre Activación Superficial y Consumo de Xantato

En el contexto latinoamericano, se han realizado múltiples investigaciones que evidencian la influencia crítica de la activación superficial de sulfuros de hierro, especialmente pirrotita, en el sobreconsumo de colectores aniónicos como el xantato. Este fenómeno se explica por la capacidad de la pirrotita de generar sitios activos en su superficie a través de la oxidación parcial y la adsorción de iones metálicos, como Cu^{2+} o Pb^{2+}, presentes en la pulpa, aumentando así su afinidad por el colector (Cayunao Collio, 2014).

Estudios realizados en Chile han documentado que la pirrotita activada presenta una cinética de adsorción de xantato significativamente mayor que otros sulfuros de hierro como la pirita, obligando a incrementos en la dosificación para mantener recuperaciones metálicas adecuadas. Este comportamiento fue caracterizado mediante espectroscopía fotoelectrónica de rayos X (XPS), demostrando la presencia de especies superficiales activadas y la formación de complejos de xantato sobre la superficie mineral (Cayunao Collio, 2014).

En Perú, investigaciones como las de Baldeón Hurtado y Oropeza Huamán (2015) analizaron minerales con alto contenido de pirrotita magnética en la provincia de Caravelí. Sus resultados confirmaron la necesidad de ajustes constantes en la dosificación de xantato debido a la flotabilidad no deseada de la pirrotita y su activación en presencia de iones de cobre. Estas observaciones reforzaron la idea de que el contenido de pirrotita en la alimentación es la variable crítica que determina el consumo específico de xantato en el circuito de flotación.

De manera complementaria, Ramos Oxa (2014) evaluó en el sur del Perú parámetros de flotación en minerales polimetálicos con pirrotita, destacando su comportamiento como especie altamente reactiva y su capacidad para competir con sulfuros valiosos por el colector. Los resultados demostraron incrementos proporcionales en el consumo de xantato asociados directamente a la presencia de pirrotita, subrayando la necesidad de cuantificar su contenido para un control racional de la dosificación.

Estas investigaciones latinoamericanas sostienen la relevancia de incluir la pirrotita como variable independiente principal en modelos predictivos de consumo de reactivos, justificando el enfoque de regresión lineal adoptado en esta investigación. Al entender y modelar de manera explícita la relación entre el contenido de pirrotita y el consumo específico de xantato, se busca reducir el sobreconsumo de colector, mejorar la eficiencia metalúrgica y contribuir a la sostenibilidad de las operaciones mineras en la región.

Modelos Estadísticos Para Estimar el Consumo de Reactivos

El desarrollo de modelos estadísticos para estimar el consumo de reactivos en flotación ha cobrado relevancia debido al impacto directo de la dosificación sobre los costos operativos y la estabilidad del proceso. En particular, el consumo específico de xantato (kg/día) depende de factores mineralógicos como la pirrotita, cuya reactividad superficial eleva su afinidad por el colector y obliga a sobredosificaciones para mantener las recuperaciones deseadas (Baldeón Hurtado & Oropeza Huamán, 2015).

En la literatura especializada se han propuesto diversos enfoques estadísticos y de machine learning para modelar el consumo de reactivos, incluyendo regresión múltiple, Random Forest, XGBoost y redes neuronales. Estos métodos permiten capturar relaciones no lineales o interacciones entre múltiples variables operativas y mineralógicas (James et al., 2013;

Goodfellow, Bengio, & Courville, 2016). Sin embargo, su complejidad estadística y su naturaleza opaca dificultan su interpretación y validación en contextos industriales donde se requiere transparencia y facilidad de uso.

Para la problemática específica de la flotación de minerales de hierro con presencia significativa de pirrotita, la relación entre el contenido de pirrotita en la alimentación y el consumo específico de xantato puede describirse de forma directa y bien aproximada mediante un modelo lineal. Este enfoque se fundamenta en la evidencia empírica que demuestra que la pirrotita es la causa principal del sobreconsumo de colector, dado su comportamiento altamente reactivo y su capacidad de adsorber xantato de forma no selectiva (García, Ríos, & Alva, 2020).

La regresión lineal presenta ventajas claras para este contexto: es estadísticamente sólida, interpretable, de bajo costo computacional y fácil de implementar en sistemas existentes como planillas de cálculo o plataformas SCADA de planta. Su simplicidad favorece la capacitación del personal operativo y reduce la resistencia al cambio tecnológico, asegurando además la auditabilidad y la replicabilidad de los resultados (Kuhn & Johnson, 2013).

En esta investigación se adopta la regresión lineal como enfoque metodológico por ser la opción más parsimoniosa y práctica para estimar el consumo específico de xantato en función del contenido de pirrotita. De este modo, se busca dotar a la operación de una herramienta predictiva robusta y aplicable que permita optimizar la dosificación del colector, reducir costos operativos y mejorar la sostenibilidad del proceso de flotación (Wills & Finch, 2016).

Aplicaciones Internacionales de Machine Learning en Flotación

A nivel internacional, la aplicación de técnicas de machine learning en flotación ha avanzado significativamente en la última década, impulsada por la disponibilidad de datos

operativos de planta y la necesidad de optimizar procesos con alta variabilidad mineralógica. Empresas mineras y centros de investigación han utilizado modelos de árboles de decisión, redes neuronales y técnicas de ensemble learning como Random Forest y XGBoost para predecir recuperaciones metalúrgicas, calidades de concentrado y consumo de reactivos (Goodfellow, Bengio, & Courville, 2016).

Investigaciones como las de Nasiri, Iqbal y Särkkä (2024) han propuesto modelos physics-informed machine learning para predecir leyes de concentrado en flotación de cobre, integrando ecuaciones fisicoquímicas con datos históricos. Este enfoque ha permitido mejorar la precisión de las predicciones en ambientes industriales reales. Sin embargo, estos modelos requieren múltiples variables de entrada, alto poder computacional y personal altamente capacitado para su desarrollo e interpretación.

En operaciones de flotación de sulfuros polimetálicos en Canadá y Australia, se han implementado sistemas de control avanzado basados en modelos predictivos multivariables para ajustar dosificación de reactivos en tiempo real. Estos sistemas suelen incluir variables como pH, Eh, granulometría, leyes de alimentación y características mineralógicas detalladas. Si bien son poderosos, su adopción requiere inversión en infraestructura tecnológica y un robusto sistema de adquisición de datos en línea (James et al., 2013; Kuhn & Johnson, 2013).

De este modo, aunque las aplicaciones internacionales de machine learning en flotación han demostrado el valor de modelos avanzados para problemas complejos y multivariables, la elección de un modelo sencillo en este estudio responde al principio de parsimonia estadística y a la necesidad de ofrecer una herramienta práctica y replicable que pueda adoptarse de forma inmediata en condiciones reales de operación.

Aportes y Vacíos que Motivan el Presente Estudio

Diversos estudios nacionales e internacionales han resaltado la importancia de comprender la influencia de los sulfuros de hierro, particularmente la pirrotita, en el proceso de flotación. Investigaciones en Perú y Chile han documentado su elevada reactividad superficial y su capacidad para adsorber colectores aniónicos como el xantato, obligando a ajustes continuos en la dosificación para mantener recuperaciones aceptables (Baldeón Hurtado & Oropeza Huamán, 2015; Cayunao Collio, 2014).

En la literatura se han explorado modelos multivariables y algoritmos de machine learning más complejos, incluyendo Random Forest y redes neuronales, para predecir variables metalúrgicas en flotación. Si bien estos enfoques han mostrado éxito en contextos con múltiples variables operativas y mineralógicas, su implementación en plantas reales requiere infraestructura tecnológica avanzada, grandes volúmenes de datos de alta calidad y personal especializado para su calibración e interpretación (James et al., 2013; Goodfellow, Bengio, & Courville, 2016).

Sin embargo, en el caso particular de la planta concentradora de Shougang Hierro Perú, el principal desafío operativo identificado es el sobreconsumo de xantato asociado de manera directa al contenido variable de pirrotita en la alimentación. Aunque se reconoce la influencia de otras variables, la evidencia empírica y la experiencia operativa indican que la pirrotita es el factor mineralógico crítico que determina el consumo específico de colector (García, Ríos, & Alva, 2020).

A pesar de esta comprensión teórica y operativa, actualmente no existe en la planta un modelo predictivo cuantitativo y validado que permita estimar el consumo específico de xantato exclusivamente en función del contenido de pirrotita. La dosificación continúa realizándose de

manera empírica y reactiva, generando sobredosificación sistemática, mayores costos operativos y un aumento en la carga química de los efluentes (Wills & Finch, 2016).

Estos vacíos motivan el presente estudio, que propone desarrollar y validar un modelo predictivo basado en regresión lineal. Este enfoque busca transformar la comprensión cualitativa existente en una herramienta cuantitativa y práctica, fácilmente aplicable en planta, que permita optimizar la dosificación del colector, reducir costos y fortalecer la sostenibilidad de la operación minera.

Bases Teóricas

En esta sección se presentan los fundamentos conceptuales y científicos que sustentan el desarrollo del modelo predictivo propuesto para estimar el consumo específico de xantato (kg/día) en función del contenido de pirrotita (%) en el mineral de alimentación. Se abordan las características mineralógicas y fisicoquímicas de la pirrotita, su comportamiento reactivo en flotación y la relación causal directa que existe entre su contenido y la demanda de colector en el circuito de flotación de minerales de hierro.

Asimismo, se describen los enfoques estadísticos y de aprendizaje automático utilizados en la minería para modelar procesos de flotación, justificando la selección de la regresión lineal como la metodología más adecuada en este caso. La presentación de estas bases teóricas busca establecer un marco sólido que permita comprender la necesidad de este estudio y el valor de desarrollar una herramienta práctica, replicable y aplicable en planta para optimizar la dosificación de reactivos y mejorar la eficiencia operativa y ambiental del proceso (Wills & Finch, 2016; James et al., 2013).

Pirrotita: Naturaleza Mineralógica y Comportamiento en Flotación

La pirrotita ($Fe_{1-x}S$) es un sulfuro de hierro caracterizado por una estructura cristalina defectuosa en hierro, donde el valor de *x* puede variar entre 0 y 0.2. Esta estructura incompleta genera vacantes de hierro en la red cristalina, confiriéndole una elevada reactividad superficial. A nivel mineralógico, se presenta frecuentemente asociada a magnetita y otros sulfuros en depósitos de hierro, complicando el procesamiento y la obtención de concentrados de alta ley (Baldeón Hurtado & Oropeza Huamán, 2015; Wills & Finch, 2016).

Esta reactividad superficial elevada es clave para su comportamiento en flotación. La pirrotita muestra una afinidad natural por colectores aniónicos como el xantato debido a la

presencia de sitios reactivos expuestos en su superficie. Incluso en condiciones de pH ligeramente alcalinas, su superficie puede adsorber colector de manera eficiente, generando flotabilidad no deseada que compite directamente con los minerales valiosos presentes en la pulpa (Cayunao Collio, 2014).

Adicionalmente, la pirrotita tiene la capacidad de activarse mediante la adsorción de iones metálicos disueltos en la solución, como Cu^{2+} o Pb^{2+}. Este fenómeno potencia su afinidad por el xantato al generar especies superficiales más reactivas, lo que incrementa la velocidad de adsorción del colector y agrava el problema de flotabilidad no selectiva. Estas características fisicoquímicas hacen que su control sea uno de los desafíos más importantes en la flotación de minerales de hierro (Ramos Oxa, 2014).

En términos operativos, la presencia de pirrotita en la alimentación obliga a realizar ajustes constantes en la dosificación de xantato para contrarrestar su efecto no deseado. Sin embargo, la dosificación empírica basada en experiencia del operador y en reglas fijas carece de la capacidad de anticipar de manera precisa la variabilidad del contenido de pirrotita en la alimentación. Esto conduce a sobredosificación sistemática del colector, elevando costos por tonelada tratada y reduciendo la eficiencia del proceso (García, Ríos, & Alva, 2020).

Desde la perspectiva ambiental y económica, el sobreconsumo de xantato incrementa la carga química en los efluentes, generando mayores costos de tratamiento y un impacto ambiental adverso. Por ello, comprender la naturaleza mineralógica de la pirrotita y su comportamiento reactivo en flotación resulta esencial para desarrollar estrategias de control basadas en datos reales de operación y diseñar modelos predictivos que permitan optimizar la dosificación del colector (Wills & Finch, 2016).

En este contexto, modelar la relación entre el contenido de pirrotita en el mineral de alimentación y el consumo específico de xantato (kg/día) se plantea como una solución técnica y práctica. Al entender y cuantificar esta relación, se puede pasar de decisiones empíricas a un control predictivo más racional, reduciendo costos, mejorando la estabilidad del proceso y promoviendo prácticas sostenibles en la industria minera.

Consumo de Xantato como Variable Operativa

El consumo específico de xantato (kg/día) es una variable operativa crítica en los circuitos de flotación de minerales de hierro, ya que representa la cantidad de colector consumido por tonelada de mineral tratado. Este indicador no solo refleja la eficiencia metalúrgica del proceso, sino también su impacto directo en los costos operativos de la planta, al ser el xantato uno de los reactivos más costosos en la flotación de sulfuros (García, Ríos, & Alva, 2020).

Desde un punto de vista económico, el sobreconsumo de xantato implica un incremento directo en los costos por tonelada procesada, afectando la competitividad de la operación. En plantas de gran volumen como Shougang Hierro Perú, incluso pequeños incrementos porcentuales en la dosificación diaria se traducen en costos significativos al final del año. Por ello, controlar y optimizar este consumo es una prioridad estratégica para la gestión operativa y financiera de la empresa (Wills & Finch, 2016).

Además del aspecto económico, el consumo excesivo de xantato tiene implicaciones ambientales importantes. Al ser un compuesto químico aniónico, su sobredosificación eleva la carga de contaminantes en los efluentes de la planta, incrementando los costos de tratamiento y generando mayores riesgos para el cumplimiento de normativas ambientales. Esto contradice los

principios de minería responsable y sostenible que exigen minimizar el impacto ambiental de las operaciones mineras (Goodfellow, Bengio, & Courville, 2016).

En el contexto de la presente investigación, el consumo específico de xantato se define como la variable dependiente principal a modelar debido a su relación directa y causal con el contenido de pirrotita en el mineral de alimentación. La pirrotita, por su elevada reactividad superficial, adsorbe el colector de manera no selectiva, obligando a sobredosificaciones para mantener recuperaciones aceptables. Modelar esta relación de forma cuantitativa permitirá anticipar la demanda de reactivo y optimizar su dosificación (Baldeón Hurtado & Oropeza Huamán, 2015).

Por estas razones, se justifica plenamente la selección del consumo específico de xantato como variable operativa clave en este estudio. Su medición y control son esenciales para reducir costos, mejorar la estabilidad del proceso metalúrgico y fortalecer la sostenibilidad ambiental de la operación minera, alineándose con las mejores prácticas de la industria y con los objetivos estratégicos de competitividad y responsabilidad social.

Modelos de Predicción en Flotación y Comparativa de Algoritmos

En la minería moderna, el desarrollo de modelos predictivos para flotación busca transformar datos operativos en herramientas que mejoren la toma de decisiones sobre la dosificación de reactivos, calidad del concentrado y recuperaciones metalúrgicas. Si bien existen enfoques estadísticos avanzados capaces de incorporar múltiples variables y relaciones complejas, no todos los problemas requieren esa complejidad, especialmente cuando la relación causal principal está bien identificada y fundamentada empíricamente (James et al., 2013).

En el caso de la planta Shougang Hierro Perú, la pirrotita ha sido identificada como la causa directa del sobreconsumo de xantato debido a su elevada reactividad y afinidad por el colector. Por ello, un modelo predictivo basado en regresión lineal resulta el enfoque más adecuado: permite estimar de forma transparente y directa el consumo específico de xantato (kg/día) en función del contenido de pirrotita en la alimentación, facilitando su integración en el control operativo de planta (Baldeón Hurtado & Oropeza Huamán, 2015; García, Ríos, & Alva, 2020).

Esta aproximación prioriza la interpretabilidad, la facilidad de implementación y la capacitación del personal, garantizando la adopción práctica de la herramienta predictiva. Al centrarse en la causa principal del problema, el modelo propuesto busca optimizar el uso de reactivos, reducir costos y fortalecer la sostenibilidad ambiental del proceso de flotación (Wills & Finch, 2016).

Definición de Categorías de Análisis

Variable Independiente: Contenido de Pirrotita

Definición Conceptual

El contenido de pirrotita en el mineral de alimentación se define como el porcentaje en peso de este sulfuro de hierro ($Fe_{1-x}S$) presente en la muestra total procesada en la planta. Su medición se realiza mediante análisis mineralógico y químico, como difracción de rayos X (XRD), balances mineralógicos o análisis modal automatizado (QEMSCAN), permitiendo caracterizar de forma sistemática la variabilidad de la alimentación. Conocer este valor es esencial para anticipar el comportamiento de flotación y planificar la estrategia de dosificación de reactivos (Wills & Finch, 2016).

La pirrotita es reconocida por su elevada reactividad superficial y su fuerte afinidad por colectores aniónicos como el xantato, incluso en condiciones de pH moderadas. Esta propiedad genera flotabilidad no deseada, obligando a incrementar la dosificación del colector para mantener las recuperaciones metalúrgicas esperadas. En la práctica, se ha evidenciado que el contenido de pirrotita explica gran parte de la variabilidad del consumo específico de xantato (kg/día), convirtiéndola en la causa directa del sobreconsumo de reactivo y en la variable más crítica a controlar (Baldeón Hurtado & Oropeza Huamán, 2015; García, Ríos, & Alva, 2020).

Por estas razones, el presente estudio adopta el contenido de pirrotita como la variable independiente principal del modelo predictivo propuesto. Su selección obedece al principio de parsimonia estadística y a la necesidad de ofrecer una herramienta operativa práctica y aplicable en planta, que permita anticipar la demanda de xantato en función de la mineralogía real de la alimentación, reduciendo costos, estabilizando el proceso y fortaleciendo la sostenibilidad ambiental de la operación minera.

Definición Operativa

El contenido de pirrotita como variable independiente se define de manera operativa como el porcentaje en peso de pirrotita presente en la alimentación diaria de la planta concentradora. Su determinación requiere análisis mineralógicos y químicos sistemáticos, utilizando técnicas como difracción de rayos X (XRD), espectroscopía fotoelectrónica de rayos X (XPS) o análisis modal automatizado (QEMSCAN), que permiten caracterizar las proporciones minerales con precisión (Wills & Finch, 2016).

En el contexto de la operación minera real, se integran resultados de muestreos de alimentación, concentrado y relave para realizar balances mineralógicos que permiten estimar el contenido promedio de pirrotita procesado. Este dato puede actualizarse con frecuencia semanal o incluso diaria, dependiendo de la disponibilidad analítica, y se convierte en un insumo fundamental para ajustar la estrategia de dosificación de reactivos colectores (García, Ríos, & Alva, 2020).

Categoría de Análisis

- Variable continua.
- Unidad de medida: %.
- Nivel de medición: razón.
- Procedencia del dato: registro histórico operativo del sistema de control metalúrgico de la planta.

Variable Dependiente: Consumo de Xantato

Definición Conceptual

El consumo específico de xantato (kg/día) se define conceptualmente como la cantidad de colector aniónico xantato consumido por tonelada de mineral procesado en el circuito de flotación. Esta variable dependiente refleja directamente la eficiencia del proceso y es clave para controlar costos operativos, ya que el xantato es uno de los reactivos más costosos utilizados en flotación de sulfuros debido a su función en la adsorción sobre especies mineralógicas y la generación de flotabilidad selectiva (Wills & Finch, 2016).

En el caso de minerales de hierro con contenido significativo de pirrotita, el consumo específico de xantato aumenta de forma proporcional a la cantidad de esta fase sulfurada en la alimentación, debido a la afinidad química elevada de la pirrotita por el colector. Esta relación directa convierte al consumo específico de xantato en la variable dependiente principal para el modelo predictivo propuesto, ya que su optimización tiene un impacto significativo en la rentabilidad de la operación y en la reducción de la carga química en los efluentes (Baldeón Hurtado & Oropeza Huamán, 2015).

Por estas razones, definir y modelar el consumo específico de xantato como variable dependiente no solo permite anticipar de forma cuantitativa la demanda de reactivo en función del contenido de pirrotita, sino que también promueve un enfoque de control basado en datos reales de operación, contribuyendo a la reducción de costos operativos, la estabilidad del proceso metalúrgico y la sostenibilidad ambiental (García, Ríos, & Alva, 2020).

Definición Operativa

El consumo específico de xantato (kg/día) se define operativamente como la cantidad de xantato dosificado al circuito de flotación dividida entre la tonelada de mineral seco procesado

durante el mismo periodo. En la práctica de planta, este valor se obtiene a partir de registros diarios de dosificación de reactivo y datos de tonelaje procesado, generando un indicador clave para el control de costos y la planificación de insumos químicos (Wills & Finch, 2016).

La medición del consumo específico de xantato es esencial para evaluar la eficiencia del proceso de flotación. Valores elevados pueden indicar presencia de especies sulfuradas reactivas, como la pirrotita, que consumen colector de forma no selectiva. Por ello, este indicador se utiliza en la operación para ajustar la estrategia de dosificación, buscando reducir sobredosificaciones innecesarias que incrementen los costos operativos y la carga química en los efluentes (García, Ríos, & Alva, 2020).

En condiciones reales de planta, el consumo específico de xantato se monitorea mediante balances de reactivos, datos de control de bombas dosificadoras y mediciones en línea o registros manuales. Estos datos alimentan sistemas de reporte diario o semanal que permiten analizar tendencias y evaluar el impacto de variaciones mineralógicas en la alimentación, especialmente del contenido de pirrotita, sobre la demanda de colector (Baldeón Hurtado & Oropeza Huamán, 2015).

Categoría de Análisis

- Variable continua.
- Unidad de medida: $\frac{\text{Kg}}{\text{día}}$.
- Nivel de medición: razón.
- Fuente del dato: base de datos operativa del área de flotación.

Relación entre las Variables

La relación entre el contenido de pirrotita en el mineral de alimentación y el consumo específico de xantato (kg/día) es directa y bien documentada tanto teórica como empíricamente. La pirrotita, debido a su estructura cristalina defectuosa, presenta una elevada reactividad superficial que le otorga una fuerte afinidad por colectores aniónicos como el xantato. Esta propiedad causa que la pirrotita adsorba colector de forma no selectiva, generando flotabilidad no deseada que compite con minerales valiosos por el reactivo (Wills & Finch, 2016).

En operaciones reales, se observa que cuando el porcentaje de pirrotita en la alimentación aumenta, la planta requiere incrementar la dosificación de xantato para mantener recuperaciones metalúrgicas estables. Este ajuste empírico y reactivo genera un sobreconsumo sistemático del colector, elevando costos por tonelada procesada y aumentando la carga química en los efluentes. Por ello, cuantificar esta relación es esencial para pasar de un control reactivo a uno predictivo y racional (García, Ríos, & Alva, 2020).

Desde el punto de vista estadístico, esta relación puede aproximarse de forma adecuada mediante un modelo lineal , donde el contenido de pirrotita actúa como variable independiente y el consumo específico de xantato como variable dependiente. Esta forma funcional captura la proporcionalidad observada en rangos operativos típicos de planta, permitiendo generar predicciones confiables y fáciles de interpretar (James et al., 2013; Kuhn & Johnson, 2013).

Modelar de manera cuantitativa esta relación no solo permite estimar la demanda de reactivo en función de la mineralogía de la alimentación, sino que también facilita la implementación de estrategias de dosificación más eficientes, reduciendo costos y minimizando impactos ambientales. Además, la simplicidad de la regresión lineal favorece su adopción en

planta, integrándose fácilmente en sistemas de soporte a la decisión como planillas de cálculo o plataformas SCADA (Goodfellow, Bengio, & Courville, 2016).

La presente investigación se centra en caracterizar y modelar de forma explícita la relación entre el contenido de pirrotita y el consumo específico de xantato, estableciendo la base para una herramienta predictiva que permita mejorar la gestión operativa y promover prácticas más sostenibles en el proceso de flotación de minerales de hierro.

Hipótesis

Hipótesis General

El contenido de pirrotita en el mineral de alimentación influye significativamente sobre el consumo específico de xantato (kg/día) en el proceso de flotación de minerales de hierro, y esta relación puede ser adecuadamente explicada mediante un modelo de regresión lineal, permitiendo su aplicación práctica en la toma de decisiones operativas en planta.

Hipótesis Específicas

Se plantea que existe una relación lineal estadísticamente significativa entre el contenido de pirrotita presente en el mineral de alimentación y el consumo específico de xantato (kg/día) en el proceso de flotación, lo cual puede demostrarse mediante el análisis estadístico de registros operativos históricos de la planta Shougang Hierro Perú – Marcona.

Se hipotetiza que el contenido de pirrotita permite construir un modelo de regresión lineal con capacidad explicativa suficiente para predecir el consumo específico de xantato, presentando un nivel de ajuste estadísticamente aceptable y aplicable en condiciones industriales reales.

Se espera que el modelo predictivo desarrollado presente un margen de error bajo y una capacidad de predicción adecuada, lo cual permitiría su implementación como herramienta operativa para estimar de forma anticipada el consumo específico de xantato a partir del contenido de pirrotita en la alimentación.

Sistemas y Categorías de Análisis

Esta sección presenta los sistemas y categorías de análisis definidos para estructurar el estudio de la relación entre el contenido de pirrotita en el mineral de alimentación y el consumo específico de xantato (kg/día) en el proceso de flotación de minerales de hierro. Establecer de forma clara estos componentes analíticos permite organizar las variables de investigación y delimitar los aspectos clave que serán examinados para dar respuesta a los objetivos y las hipótesis planteadas.

En el marco de esta investigación, el enfoque de análisis se sustenta en una relación directa y modelable mediante regresión lineal. Por ello, se prioriza la definición operativa y conceptual del contenido de pirrotita como variable independiente y del consumo específico de xantato como variable dependiente, junto con las dimensiones de análisis necesarias para caracterizar su variabilidad, su relación estadística y la aplicabilidad del modelo predictivo en condiciones reales de planta.

Definir con precisión estos sistemas y categorías de análisis es fundamental para garantizar la validez, la coherencia metodológica y la replicabilidad del estudio, facilitando la aplicación de los resultados en el control de reactivos y la optimización operativa de la planta Shougang Hierro Perú.

Dimensión Mineralógica

La dimensión mineralógica del estudio se centra en la caracterización y cuantificación del contenido de pirrotita en el mineral de alimentación procesado en la planta. Esta fase sulfurada de hierro, debido a su estructura cristalina defectuosa, presenta una reactividad superficial elevada que la convierte en el principal factor mineralógico responsable del sobreconsumo de xantato durante el proceso de flotación (Wills & Finch, 2016).

El análisis de esta dimensión implica determinar el porcentaje en peso de pirrotita en las muestras de alimentación mediante técnicas como difracción de rayos X (XRD), balances mineralógicos basados en concentrado y relave, o análisis modal automatizado (QEMSCAN). Estas herramientas permiten caracterizar de forma precisa la variabilidad del contenido de pirrotita entre diferentes campañas de producción o incluso entre días consecutivos de operación (García, Ríos, & Alva, 2020).

Medir y controlar esta variabilidad es esencial para anticipar la demanda de reactivo colector. Operativamente, disponer de datos confiables sobre el contenido de pirrotita permite ajustar la estrategia de dosificación del xantato, reduciendo la dependencia de la experiencia empírica del operador y promoviendo un enfoque de control basado en datos reales de operación (Baldeón Hurtado & Oropeza Huamán, 2015).

La dimensión mineralógica del análisis se enfoca exclusivamente en la pirrotita como variable independiente principal del modelo predictivo propuesto, destacando su medición sistemática como insumo fundamental para estimar de manera cuantitativa el consumo específico de xantato y optimizar la eficiencia del proceso de flotación en la planta Shougang Hierro Perú.

Dimensión Reactiva

La dimensión reactiva del análisis se enfoca en el comportamiento superficial de la pirrotita durante el proceso de flotación. Este sulfuro de hierro presenta una estructura cristalina defectuosa con vacantes de hierro que generan sitios reactivos expuestos en su superficie, lo que le confiere una alta afinidad química por colectores aniónicos como el xantato incluso en condiciones de pH moderadas. Esta característica la convierte en una especie mineralógica altamente flotable y difícil de controlar operativamente (Wills & Finch, 2016).

Además, la pirrotita puede activarse mediante la adsorción de iones metálicos disueltos en la pulpa, como Cu^{2+} o Pb^{2+}. Este fenómeno incrementa la reactividad superficial, formando sitios aún más susceptibles a la adsorción del colector y potenciando la flotabilidad no deseada. Este mecanismo de activación superficial complica el proceso de flotación al incrementar la demanda de colector para mantener la selectividad hacia los minerales valiosos (Cayunao Collio, 2014).

Operativamente, este comportamiento reactivo genera un problema crítico: obliga a incrementar la dosificación de xantato para evitar la pérdida de recuperación metalúrgica, provocando un aumento directo en el consumo específico de reactivo (kg/día). Este sobreconsumo no solo eleva los costos operativos por tonelada tratada, sino que también incrementa la carga química en los efluentes, elevando los costos de tratamiento y el impacto ambiental asociado (García, Ríos, & Alva, 2020).

La dimensión reactiva de la pirrotita se considera la causa directa del sobreconsumo de xantato y justifica su inclusión como variable independiente principal en el modelo predictivo propuesto. Comprender y cuantificar esta dimensión permite anticipar la demanda de colector, optimizar su dosificación y avanzar hacia un control más racional, económico y sostenible del proceso de flotación en la planta Shougang Hierro Perú.

Dimensión Predictiva

La dimensión predictiva del análisis corresponde al desarrollo de un modelo estadístico capaz de estimar el consumo específico de xantato (kg/día) en función del contenido de pirrotita en el mineral de alimentación. Esta relación directa, evidenciada empírica y teóricamente, constituye el eje central del estudio, buscando transformar datos históricos y mediciones

operativas en una herramienta de soporte para la toma de decisiones en planta (García, Ríos, & Alva, 2020).

El enfoque elegido para modelar esta relación es la regresión lineal , que permite cuantificar de forma sencilla y transparente la influencia del contenido de pirrotita sobre el consumo de colector. Esta metodología estadística es ampliamente utilizada en la industria debido a su interpretabilidad y facilidad de implementación en sistemas existentes como planillas de cálculo o plataformas SCADA, garantizando replicabilidad y auditabilidad (James et al., 2013; Kuhn & Johnson, 2013).

La adopción de un modelo responde al principio de parsimonia estadística, priorizando la variable mineralógica más crítica y evitando la inclusión de factores colineales o menos relevantes que podrían complicar innecesariamente el modelo sin aportar poder explicativo adicional. Al enfocarse exclusivamente en la pirrotita, se maximiza la capacidad predictiva práctica del modelo en condiciones reales de operación (Goodfellow, Bengio, & Courville, 2016).

Este enfoque predictivo no solo tiene valor académico, sino también un impacto operativo tangible. Permite anticipar la demanda de xantato en función de las características mineralógicas de la alimentación diaria, reduciendo la dependencia de ajustes empíricos o reactivos, optimizando la dosificación del colector y disminuyendo los costos operativos por tonelada tratada. Además, contribuye a minimizar la carga química en los efluentes, alineándose con los objetivos de sostenibilidad ambiental de la industria minera (Wills & Finch, 2016).

La dimensión predictiva del estudio constituye un componente esencial del análisis, integrando conocimiento mineralógico, mediciones operativas y técnicas estadísticas en una

herramienta de fácil adopción que fortalezca el control de reactivos y la eficiencia global del proceso de flotación en la planta Shougang Hierro Perú.

Materiales y Métodos

Enfoque de la Investigación

El presente estudio adopta un enfoque cuantitativo, predictivo y explicativo, orientado a analizar la influencia del contenido de pirrotita en el consumo específico de xantato (kg/día) en la flotación de minerales de hierro. Se trata de un diseño no experimental y transversal, basado en datos históricos de planta y mediciones mineralógicas que permiten establecer asociaciones estadísticas sólidas entre variables claramente definidas (Wills & Finch, 2016).

La investigación privilegia la parsimonia metodológica al concentrarse en un único factor mineralógico —la pirrotita— como variable independiente principal, fundamentada en evidencia teórica y empírica que la identifica como la causa directa del sobreconsumo de colector. Por ello, se selecciona la regresión lineal como estrategia analítica, permitiendo modelar la relación entre estas variables de manera precisa, interpretable y aplicable en entornos operativos reales (James et al., 2013; Kuhn & Johnson, 2013).

Este enfoque metodológico busca transformar la comprensión cualitativa del problema en una herramienta predictiva sencilla de implementar, facilitando decisiones de control basadas en datos reales de operación. Al priorizar la claridad, la auditabilidad y la transferencia tecnológica hacia el personal de planta, el modelo propuesto aspira a optimizar la dosificación de reactivos, reducir costos operativos y fortalecer la sostenibilidad ambiental del proceso de flotación (Goodfellow, Bengio, & Courville, 2016).

Tipo de Investigación

Este estudio corresponde a una investigación cuantitativa, de tipo explicativa y predictiva, ya que busca no solo describir la variabilidad del contenido de pirrotita y el consumo específico de xantato (kg/día), sino también establecer y modelar la relación causal entre estas dos variables

mediante un enfoque estadístico de regresión lineal . Este diseño permite generar conclusiones basadas en evidencia empírica y datos históricos reales de planta (James et al., 2013).

Además, se clasifica como una investigación aplicada, dado que su propósito es resolver un problema concreto de la operación minera: el sobreconsumo de xantato causado por la presencia de pirrotita en la alimentación. El resultado esperado es un modelo predictivo sencillo de implementar, que permita optimizar la dosificación del colector y reducir costos operativos, mejorando la eficiencia y la sostenibilidad del proceso de flotación (García, Ríos, & Alva, 2020).

Este enfoque metodológico combina la rigurosidad estadística con la aplicabilidad práctica, garantizando que los resultados puedan ser utilizados como herramienta de gestión y soporte a la decisión en la planta Shougang Hierro Perú. De esta manera, la investigación aporta soluciones concretas y replicables para el control de reactivos en minería, fortaleciendo el vínculo entre el conocimiento teórico y su implementación operativa (Wills & Finch, 2016).

Diseño de Investigación

El diseño metodológico adoptado en este estudio es de tipo no experimental y transversal, ya que se basa en el análisis de datos históricos de operación y mediciones mineralógicas sin manipular intencionalmente las variables del proceso. Se recolectan registros de contenido de pirrotita y consumo específico de xantato en diferentes momentos de operación, permitiendo caracterizar su relación en un solo corte temporal, lo cual es coherente con la naturaleza predictiva y observacional de la investigación (James et al., 2013).

Además, se clasifica como un diseño correlacional y predictivo , pues busca explicar y anticipar el comportamiento de la variable dependiente (consumo específico de xantato, kg/t) a partir de la variable independiente principal (contenido de pirrotita en la alimentación). Este

enfoque analítico es esencial para transformar la comprensión cualitativa del problema en un modelo estadístico interpretable y aplicable en condiciones reales de planta (García, Ríos, & Alva, 2020).

La selección de este diseño se justifica por su capacidad de proporcionar resultados replicables, auditables y de fácil implementación en herramientas operativas como planillas de cálculo o sistemas SCADA. Al priorizar la simplicidad metodológica y la claridad interpretativa, se garantiza que el modelo predictivo pueda integrarse efectivamente en la gestión de reactivos de la planta Shougang Hierro Perú, contribuyendo a la optimización de costos y a la sostenibilidad ambiental del proceso de flotación (Wills & Finch, 2016).

Credibilidad de la Investigación

La investigación se llevará a cabo respetando las normativas éticas y legales relacionadas con el manejo de datos industriales y operacionales. Todos los datos utilizados serán de acceso autorizado y no se compartirán sin el consentimiento de Shougang Hierro Perú. Se garantizará la confidencialidad de la información y la protección de los datos sensibles. Además, se cumplirá con las políticas éticas de la institución en cuanto a la realización de investigaciones que involucren el uso de tecnologías avanzadas, como el machine learning, asegurando la transparencia y fiabilidad en los resultados obtenidos.

Sujetos de Estudio

El sujeto de estudio en esta investigación está constituido por los registros históricos operacionales de la planta Shougang Hierro Perú – Marcona. Los datos analizados provienen de muestras de mineral alimentado a flotación, en las cuales se cuantifica el porcentaje de pirrotita y se registra el consumo de xantato en cada ciclo de producción.

Características del Sujeto de Estudio

Se utilizarán registros de al menos seis meses de operación continua, lo que garantiza que los datos reflejen de manera representativa las condiciones operativas del proceso de flotación de minerales de hierro con presencia de pirrotita. Esta muestra incluye variaciones en el porcentaje de pirrotita en el mineral, así como fluctuaciones en el consumo de reactivos debido a cambios operacionales o características del mineral.

El tamaño de la muestra se determinará con base en la disponibilidad de datos históricos, buscando cubrir un período representativo de los ciclos de flotación para asegurar que se incluyan suficientes observaciones para entrenar los modelos predictivos.

Los registros seleccionados son particularmente relevantes debido a que incluyen valores de variables operacionales críticas, como el porcentaje de pirrotita y el consumo de xantato, que son esenciales para el análisis de la eficiencia operativa en el proceso de flotación. Este enfoque permite estudiar de manera directa la relación entre la cantidad de pirrotita en el mineral de alimentación y su impacto sobre el consumo de reactivos, sin necesidad de manipulación o intervención directa.

Procedimiento de Selección de los Sujetos de Estudio

La selección de los registros operacionales será realizada de manera aleatoria y sin sesgo, utilizando los datos que han sido previamente digitalizados y archivados por el sistema de control de producción de la planta. Se aplicarán los siguientes criterios:

- Los registros deben incluir datos completos sobre el porcentaje de pirrotita en el mineral y la cantidad de xantato utilizado.

- Se seleccionarán solo los ciclos de flotación correspondientes a minerales con un porcentaje de pirrotita superior al 5%, para asegurar que la variabilidad de los datos sea representativa del fenómeno de interés.

Recolección de Datos de los Sujetos de Estudio

El instrumento principal de recolección será el sistema de base de datos operacionales de la planta, que contiene información detallada sobre el consumo de reactivos y las características del mineral. Esta base de datos es gestionada y actualizada por el departamento de control de calidad y producción. Los datos se exportarán a hojas de cálculo para su análisis posterior, utilizando herramientas estadísticas y software de *machine learning* como Python, a través de las bibliotecas Pandas, NumPy y Scikit-learn.

Justificación de la Selección de los Sujetos de Estudio

La elección de los registros operacionales de la planta Shougang Hierro Perú – Marcona se justifica por la necesidad de contar con un conjunto de datos real y de alto volumen, lo que proporciona una base sólida para el entrenamiento de modelos predictivos robustos. Además, el acceso a datos históricos asegura que se puede explorar la relación entre el porcentaje de pirrotita y el consumo de xantato en un contexto industrial real, lo que contribuye a la validación y aplicabilidad práctica del modelo en la mejora de procesos.

Procedimientos, Técnicas e Instrumentos de Recolección de Información

Para el desarrollo de esta investigación se utilizarán datos históricos operacionales de la planta Shougang Hierro Perú – Marcona, los cuales contienen registros del contenido de pirrotita en el mineral de alimentación (%) y el consumo de xantato (kg/día) durante el proceso de flotación de minerales de hierro.

Procedimiento

El procedimiento metodológico seguido en este estudio inicia con la recopilación de datos históricos de operación provenientes de la planta Shougang Hierro Perú, incluyendo registros diarios de consumo de xantato (kg/día) y análisis mineralógicos de la alimentación. Estos datos se obtienen a partir de planillas operativas, reportes semanales y resultados de laboratorio mineralógico utilizando técnicas como difracción de rayos X (XRD) y balances de masa mineralógica (Wills & Finch, 2016).

Posteriormente, se realiza un proceso de validación y limpieza de datos, identificando registros incompletos, inconsistentes o atípicos para asegurar la calidad estadística del análisis. Este paso incluye la revisión cruzada entre las planillas de dosificación de reactivos y los tonelajes procesados, así como la verificación de la congruencia en los análisis mineralógicos, con el objetivo de fortalecer la confiabilidad de los datos que alimentarán el modelo (James et al., 2013).

Una vez consolidada la base de datos, se calculan estadísticos descriptivos para caracterizar la variabilidad del contenido de pirrotita y del consumo específico de xantato en distintos periodos operativos. Esta etapa permite identificar tendencias, rangos operativos típicos y posibles patrones estacionales o de campaña minera, ofreciendo información crítica para la formulación del modelo predictivo (García, Ríos, & Alva, 2020).

Finalmente, los datos validados y caracterizados se utilizan para ajustar un modelo de regresión lineal que relaciona el contenido de pirrotita como variable independiente con el consumo específico de xantato como variable dependiente. Este modelo es implementado en software estadístico accesible y replicable, generando coeficientes e intervalos de confianza que

serán la base para la herramienta predictiva operativa propuesta para la planta Shougang Hierro Perú.

Técnicas

La técnica principal de recolección es la extracción de datos históricos secundarios desde registros digitales de planta, a través de hojas de cálculo y reportes de monitoreo metalúrgico. No se requerirá interacción directa con operadores ni instrumentos físicos en campo.

Instrumentos de Recolección de Información

Se emplearán como instrumentos de recolección y análisis:

- Bases de datos de planta y reportes digitales exportados desde el sistema de control de producción.
- Hojas de cálculo (Excel) para estructuración inicial de los datos.
- Python 3.x, con bibliotecas como Pandas, NumPy, Seaborn y Scikit-learn, para el procesamiento, visualización y modelado de los datos.

Evaluación Comparativa de Algoritmos

Para definir la metodología más adecuada, se revisaron diversas técnicas de modelado predictivo utilizadas en minería, incluyendo regresión múltiple, métodos de árboles de decisión y algoritmos de aprendizaje automático como Random Forest y modelos de boosting. Estas herramientas son potentes para capturar relaciones no lineales y manejar interacciones complejas entre múltiples variables, pero suelen implicar una mayor complejidad estadística y computacional, dificultando su implementación práctica en entornos operativos con limitaciones de infraestructura (James et al., 2013).

En el contexto específico de la planta Shougang Hierro Perú, se identificó que la variabilidad en el consumo específico de xantato (kg/día) está fuertemente determinada por el contenido de pirrotita en la alimentación, evidenciando una relación directa y proporcional que puede describirse de manera efectiva mediante un modelo . Esto favorece la elección de la regresión lineal simple como herramienta analítica, priorizando la claridad interpretativa y la replicabilidad (García, Ríos, & Alva, 2020).

Aunque los modelos más sofisticados ofrecen potencial para problemas multivariables con múltiples factores interrelacionados, su uso en este caso sería innecesariamente complejo y podría reducir la transparencia del modelo frente al personal operativo. El enfoque lineal permite una implementación sencilla, integración con sistemas de cálculo existentes y mayor facilidad para auditar y validar los resultados en la rutina diaria de la planta (Kuhn & Johnson, 2013).

En consecuencia, la evaluación comparativa de algoritmos respalda la adopción de la regresión lineal como opción metodológica preferida para este estudio, alineándose con el principio de parsimonia estadística y asegurando la transferencia tecnológica efectiva al entorno real de operación minera, contribuyendo a optimizar la dosificación de reactivos y mejorar la sostenibilidad del proceso de flotación (Wills & Finch, 2016).

Métricas de Evaluación del Modelo

Para garantizar la calidad y utilidad del modelo de regresión lineal propuesto, se seleccionaron métricas de evaluación que permiten cuantificar su capacidad predictiva y la precisión de sus estimaciones. Entre las principales medidas se incluye el coeficiente de determinación (R^2), que indica el grado en que la variabilidad del consumo específico de xantato

(kg/día) es explicada por el contenido de pirrotita en la alimentación, ofreciendo una visión clara de la fuerza del vínculo estadístico (James et al., 2013).

Adicionalmente, se utilizarán métricas de error como el Error Absoluto Medio (MAE) y la Raíz del Error Cuadrático Medio (RMSE), que permiten valorar la magnitud promedio de las discrepancias entre los valores observados y los estimados por el modelo. Estas métricas son interpretables y prácticas para la industria, ya que expresan el error en las mismas unidades operativas, facilitando su comprensión y aplicación por parte del personal técnico de planta (Kuhn & Johnson, 2013).

La elección de estas métricas responde al principio de transparencia y aplicabilidad, garantizando que el modelo no solo ajuste bien los datos históricos, sino que también sea útil para anticipar el consumo de xantato en condiciones operativas reales. Su interpretación sencilla permite tomar decisiones informadas sobre la dosificación del colector, ayudando a reducir costos operativos y mejorar la eficiencia del proceso de flotación (García, Ríos, & Alva, 2020).

Al priorizar métricas estadísticamente robustas, pero fácilmente auditables, la evaluación del modelo busca fortalecer la credibilidad del estudio y facilitar su implementación práctica en la planta Shougang Hierro Perú, contribuyendo a un control más racional de los reactivos y a una operación minera más sostenible (Wills & Finch, 2016).

Selección del Modelo de Evaluación

La selección del modelo de evaluación en este estudio responde al principio de parsimonia estadística y a la necesidad de contar con una herramienta predictiva clara, interpretable y fácilmente implementable en la operación minera.

En el presente estudio, se evaluaron tres algoritmos de regresión supervisada: regresión lineal, Random Forest y XGBoost, utilizando un conjunto de datos reales provenientes de la operación de flotación en la planta Shougang Hierro Perú. La comparación de desempeño se realizó mediante las métricas MAE, RMSE y R^2. Los resultados evidenciaron que el modelo de regresión lineal presentó el mejor comportamiento global, al registrar el menor error absoluto medio, la menor raíz del error cuadrático medio y el mayor coeficiente de determinación. Si bien se reconoce que modelos más complejos podrían capturar relaciones no lineales adicionales, la regresión lineal fue seleccionada como modelo final por combinar un desempeño predictivo superior con ventajas prácticas relevantes: interpretabilidad, facilidad de implementación en sistemas de planta (SCADA o planillas de cálculo) y menor riesgo de sobreajuste. Esta selección responde tanto a criterios cuantitativos como a la necesidad de asegurar viabilidad operativa en contextos industriales reales.

Se optó por selección del modelo de regresión lineal , dado que la evidencia empírica y el análisis exploratorio de datos demostraron que el contenido de pirrotita en la alimentación es la causa principal que explica la variabilidad en el consumo específico de xantato (kg/día).

Esta elección metodológica se fundamenta en la capacidad del modelo lineal para cuantificar de manera directa y sencilla la relación proporcional entre estas dos variables, evitando la complejidad innecesaria de incluir múltiples factores colineales. Al mantener un enfoque transparente, se facilita la validación, la auditabilidad y la transferencia tecnológica hacia el personal técnico de la planta, promoviendo la apropiación del conocimiento y la adopción efectiva del modelo en la rutina operativa.

Asimismo, el modelo seleccionado prioriza la aplicabilidad real en el contexto de la planta Shougang Hierro Perú, integrándose fácilmente en sistemas de cálculo existentes y

permitiendo estimaciones prácticas para la planificación diaria de la dosificación de reactivos. Al reducir el riesgo de interpretaciones erróneas y ofrecer resultados comprensibles para el personal de diferentes áreas, se contribuye directamente a la optimización de costos operativos y a la sostenibilidad del proceso de flotación.

Validación Operativa y Aplicación del Modelo

La validación operativa del modelo consiste en verificar su capacidad para estimar de manera fiable el consumo específico de xantato (kg/día) a partir de valores reales de contenido de pirrotita en la alimentación diaria. Para ello, se comparan las predicciones generadas por la regresión lineal con datos históricos no utilizados en el ajuste del modelo, asegurando que mantenga un desempeño consistente en condiciones reales de planta. Este paso confirma la robustez del enfoque estadístico y su utilidad como herramienta práctica de planificación.

Una vez validado, el modelo se implementa de forma sencilla en planillas de cálculo o sistemas de control existentes en la planta, permitiendo a los ingenieros y operadores anticipar la dosificación óptima de xantato en función del análisis mineralógico de la alimentación. De este modo, se sustituye el ajuste empírico y reactivo por un control basado en datos, fortaleciendo la predictibilidad y reduciendo variaciones innecesarias en el consumo de reactivo.

Esta integración operativa se apoya en la claridad interpretativa del modelo lineal, facilitando su adopción por el personal de distintas áreas. La transparencia de la ecuación predictiva permite explicar y justificar de manera sencilla las recomendaciones de dosificación, generando confianza en su uso y fomentando la apropiación de la herramienta como parte de la rutina de gestión de reactivos en la planta.

Además, el modelo sirve como soporte para la toma de decisiones colaborativa entre las áreas de producción, laboratorio y control de procesos. Al ofrecer estimaciones objetivas basadas en datos reales, contribuye a alinear criterios técnicos y a optimizar recursos de forma coordinada, reforzando la eficiencia operativa y promoviendo prácticas más sostenibles en la flotación de minerales de hierro.

Resultados y Discusión de Resultados

Resultados

Esta sección presenta de forma estructurada los resultados obtenidos durante el desarrollo de la investigación, describiendo desde la caracterización de la base de datos, el análisis estadístico y exploratorio, la evaluación de correlaciones, hasta la comparación, entrenamiento y validación de modelos de aprendizaje automático para la predicción del consumo de xantato en la planta de flotación de Shougang Hierro Perú.

Descripción de la Base de Datos Empleada y Preprocesamiento

Se trabajó con registros diarios de operación de la planta nueva de Shougang Hierro Perú, correspondiente al periodo comprendido entre el 1 de enero y el 31 de diciembre de 2024. La base de datos incluyó promedios diarios obtenidos de muestreos y análisis químicos de laboratorio, considerando las siguientes variables:

- Recuperación en la prueba Davis Tube Tester (%).
- Porcentaje de azufre en el concentrado (%).
- Proporción de pirrotita ($\frac{S}{FeS}$ %).
- Consumo de xantato ($\frac{Kg}{día}$).

Para garantizar la calidad de la base de datos, se establecieron filtros que permitieran excluir registros no representativos del régimen normal de operación. Los criterios de exclusión aplicados fueron:

- Valores promedio diarios de azufre en el concentrado superiores a 0.250 %.
- Producción diaria por línea inferior a 10000 toneladas.

El preprocesamiento de datos se implementó en Python utilizando Google Colab, con bibliotecas como pandas para manipulación de datos y numpy para filtrado, asegurando un conjunto de datos representativo de condiciones normales de operación. Esta base de datos limpia permitió avanzar con los siguientes pasos del estudio, como la generación de un mapa de calor para evaluar las correlaciones entre variables y determinar la viabilidad del uso de modelos de machine learning para predecir el consumo de reactivos.

Análisis Exploratorio de las Variables

Se realizó un análisis estadístico descriptivo para caracterizar el comportamiento de las variables y entender su dispersión y tendencias. Los cálculos se automatizaron con scripts en Python, garantizando replicabilidad y precisión.

Tabla 1

Estadistica descriptiva de las variables operativas

Variable	**Xantato** $\left(\frac{Kg}{Día}\right)$	**Recuperación Davis Tube** (%)	**Azufre** (%)	**Pirrotita** $\left(\frac{S}{FeS}\ \%\right)$
Media	4196.36	68.90	0.97	0.63
Desviación Estándar	736.74	2.57	0.14	0.13
Mínimo	1800.00	61.90	0.61	0.33
Percentil 25	3762.08	67.15	0.88	0.54
Mediana	4227.39	68.80	0.96	0.62
Percentil 75	4590.14	70.60	1.05	0.71

Máximo	7290.00	75.30	1.31	0.98

Nota. Estadística calculada sobre promedios diarios de operación, representativa de la variabilidad industrial del circuito de flotación.

Para evaluar la distribución de las variables y detectar asimetrías o dispersión, se generaron diagrama de Box Plot, matriz de dispersión de pairplot e histogramas con curvas de densidad (KDE) utilizando seaborn en Python.

Figura 1

Codigo Pyhton Google Colab Utilizado para la generación del box plot por variable y matriz de dispersión de pairplot.

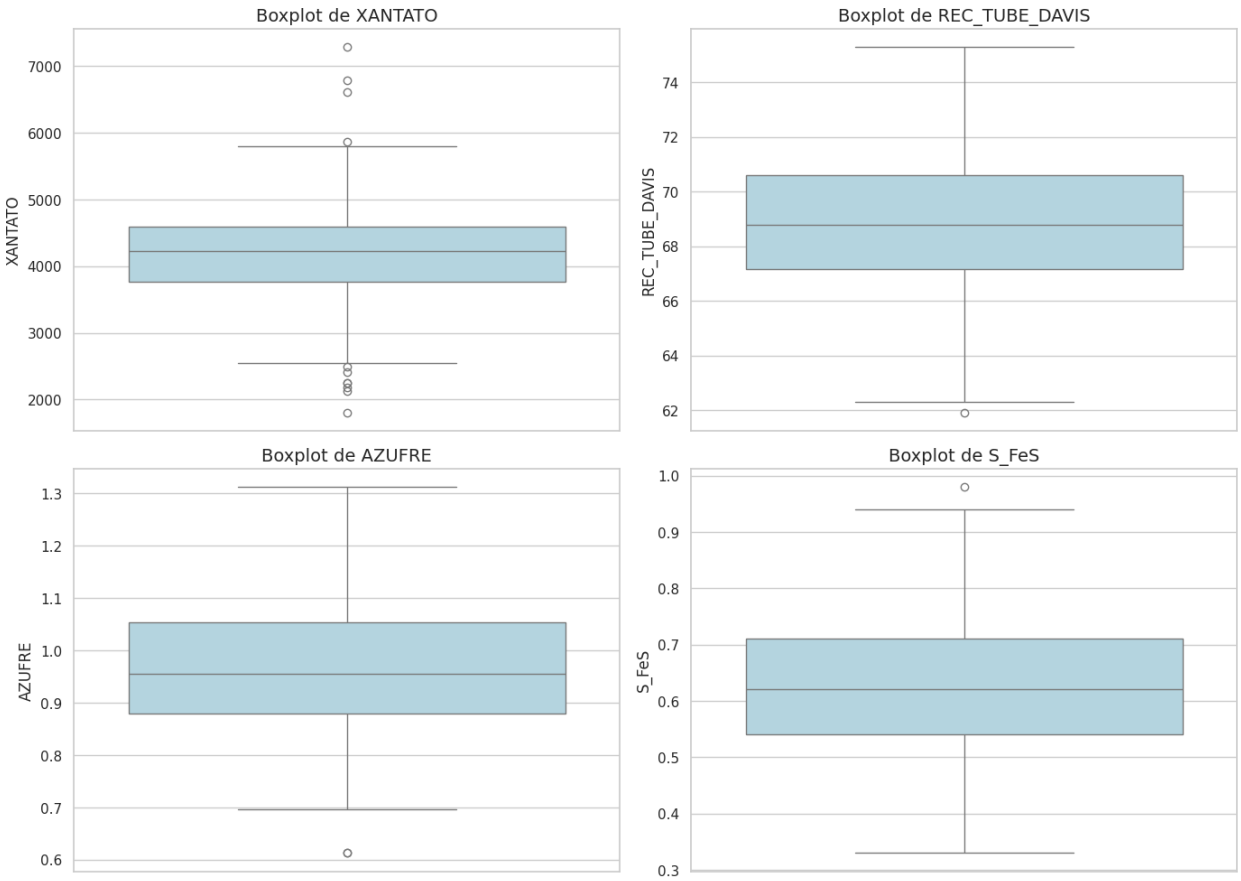


```python
# Importar librerías
import pandas as pd
from tabulate import tabulate

# Cargar archivo Excel
data = pd.read_excel("BASE_SHP_MODIFICADO.xlsx")
data.columns = ["XANTATO", "REC_TUBE_DAVIS", "AZUFRE", "S_FeS"]
data = data.dropna()

# Estadísticas descriptivas y redondeo
summary_stats = data.describe().round(2)

# Traducir índice y columnas
summary_stats.index = [
    "Conteo",
    "Media",
    "Desviación estándar",
    "Mínimo",
    "Percentil 25",
    "Mediana (50%)",
    "Percentil 75",
    "Máximo"
]

summary_stats.columns = [
    "Consumo de Xantato (kg/t)",
    "Recuperación Davis Tube (%)",
    "Contenido de Azufre (%)",
    "S/FeS (Pirrotita Estimada)"
]

# Crear tabla con formato Markdown
tabla_markdown = tabulate(summary_stats, headers='keys', tablefmt='github')

# Mostrar tabla para copiar y pegar en Word
print(tabla_markdown)


# Visualización de histogramas
fig, axes = plt.subplots(2, 2, figsize=(14, 10))
axes = axes.flatten()

for i, col in enumerate(data.columns):
    sns.histplot(data[col], bins=20, kde=True, ax=axes[i], color="steelblue")
    axes[i].set_title(f"Distribución de {col}")
    axes[i].set_xlabel(col)
    axes[i].set_ylabel("Frecuencia")

plt.tight_layout()
plt.show()
```

Nota. Fragmento del script implementado en Google Colab para análisis estadístico, visualización de datos y modelado predictivo mediante técnicas de aprendizaje automático.

Figura 2

Box plot por cada una de las variables analizadas.



Nota. Diagramas de caja para consumo de xantato, recuperación Davis Tube, contenido de azufre y proporción de pirrotita (S/FeS), evidenciando dispersión de datos y potenciales valores atípicos vinculados a la heterogeneidad del mineral.

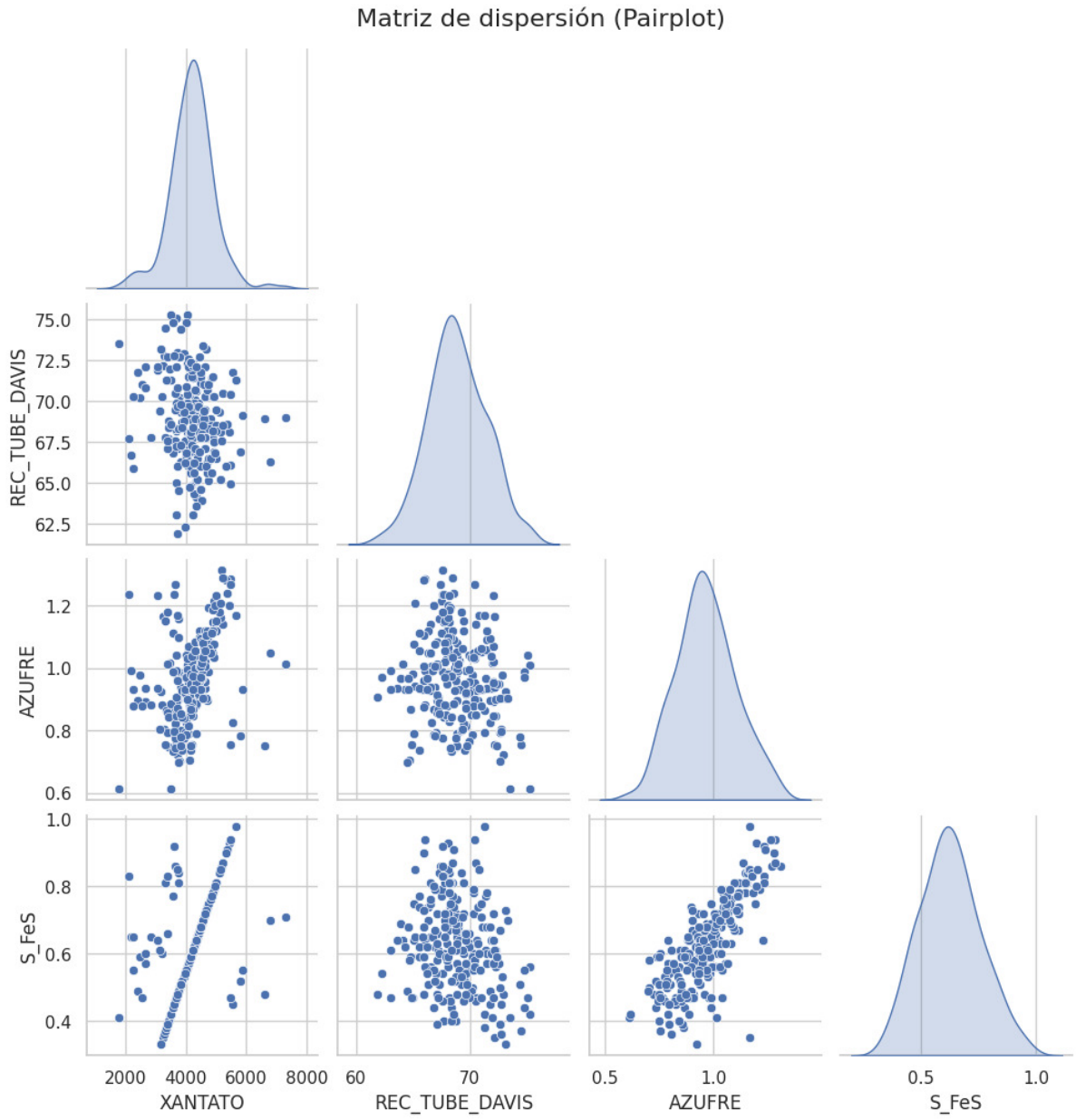
- El consumo de xantato presentó una dispersión amplia con múltiples valores atípicos (1800–7300 kg/t), reflejando variabilidad operativa significativa posiblemente asociada a diferencias mineralógicas y cambios en la dosificación.

- La recuperación en Davis Tube mostró una distribución más compacta (≈62–75 %), con menor presencia de valores extremos, lo que sugiere control operativo relativamente estable en el circuito de separación magnética.
- El contenido de azufre reveló una distribución moderadamente dispersa (≈0.6–1.3 %), con outliers menores, indicando variabilidad mineralógica relevante que podría influir en la reactividad y el consumo de colector.
- La proporción estimada de pirrotita (S/FeS) exhibió una dispersión moderada (≈0.3–1.0), con algunos valores atípicos, confirmando la heterogeneidad en la mineralogía sulfurada del mineral alimentado.

  Este análisis exploratorio mediante diagramas de caja destaca la presencia de valores extremos y la amplitud de rangos en las variables clave, subrayando la necesidad de evaluar su influencia conjunta en el consumo de reactivo colector mediante modelamiento cuantitativo.

Figura 3

Matriz de dispersión (Pairplot)



Nota. Matriz de dispersión (pairplot) para consumo de xantato, recuperación Davis Tube, contenido de azufre y proporción de pirrotita (S/FeS), evidenciando relaciones bivariadas relevantes para la modelización.

- Se observa una relación positiva clara entre el contenido de azufre y la proporción de pirrotita (S/FeS), lo que confirma la asociación mineralógica entre estas variables sulfuradas.
- La correlación entre el consumo de xantato y la proporción de pirrotita (S/FeS) es marcada y lineal, sugiriendo que mayores contenidos de pirrotita podrían estar relacionados con incrementos en el uso de reactivo colector.
- Las demás relaciones bivariadas muestran patrones más difusos o dispersos, especialmente con la recuperación Davis Tube, indicando menor dependencia directa de esta variable respecto al consumo de xantato en el rango evaluado.

  Este análisis gráfico preliminar aporta evidencia sobre la relevancia de las variables mineralógicas como predictores del consumo de reactivo, respaldando su inclusión en el modelo de machine learning propuesto.

Figura 4

Codigo Pyhton Google Colab Utilizado para la generación del histograma de distribución de variables

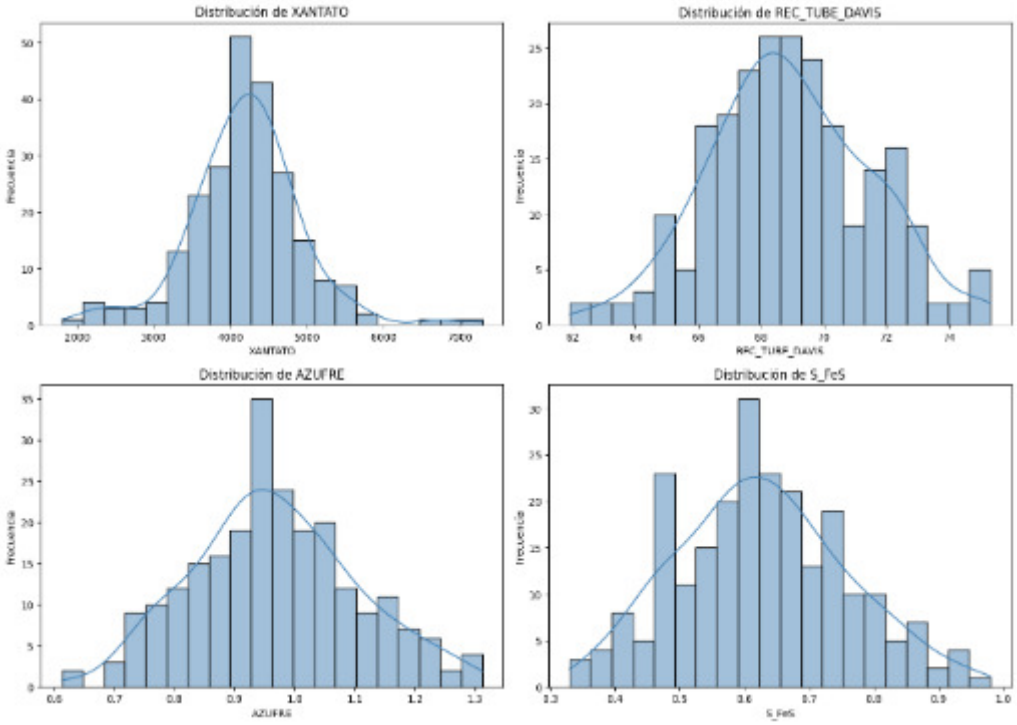


```python
# Importar librerías
import pandas as pd
from tabulate import tabulate

# Cargar archivo Excel
data = pd.read_excel("BASE_SHP_MODIFICADO.xlsx")
data.columns = ["XANTATO", "REC_TUBE_DAVIS", "AZUFRE", "S_FeS"]
data = data.dropna()

# Estadísticas descriptivas y redondeo
summary_stats = data.describe().round(2)

# Traducir índice y columnas
summary_stats.index = [
    "Conteo",
    "Media",
    "Desviación estándar",
    "Mínimo",
    "Percentil 25",
    "Mediana (50%)",
    "Percentil 75",
    "Máximo"
]

summary_stats.columns = [
    "Consumo de Xantato (kg/t)",
    "Recuperación Davis Tube (%)",
    "Contenido de Azufre (%)",
    "S/FeS (Pirrotita Estimada)"
]

# Crear tabla con formato Markdown
tabla_markdown = tabulate(summary_stats, headers='keys', tablefmt='github')

# Mostrar tabla para copiar y pegar en Word
print(tabla_markdown)


# Visualización de histogramas
fig, axes = plt.subplots(2, 2, figsize=(14, 10))
axes = axes.flatten()

for i, col in enumerate(data.columns):
    sns.histplot(data[col], bins=20, kde=True, ax=axes[i], color="steelblue")
    axes[i].set_title(f"Distribución de {col}")
    axes[i].set_xlabel(col)
    axes[i].set_ylabel("Frecuencia")

plt.tight_layout()
plt.show()
```

Nota. Fragmento del script implementado en Google Colab para análisis estadístico, visualización de datos y modelado predictivo mediante técnicas de aprendizaje automático.

Figura 5

Histograma de distribución de variables



Nota. Distribuciones de frecuencia para consumo de xantato, recuperación Davis Tube, contenido de azufre y proporción de pirrotita $\left(\frac{S}{FeS}\right)$, evidenciando variabilidad operativa y heterogeneidad mineralógica.

- El consumo de xantato mostró un rango amplio (1800–7290 kg/día), lo que refleja variabilidad operativa significativa posiblemente vinculada a cambios mineralógicos.
- La recuperación en Davis Tube mostró distribución estable en torno al 69 %, con baja dispersión, reflejando control operacional.
- El contenido de azufre y la proporción de pirrotita exhibieron ligeras asimetrías, confirmando variabilidad mineralógica relevante para la modelización.

Este análisis preliminar justificó la necesidad de investigar la relación cuantitativa entre estas variables y el consumo de reactivo colector.

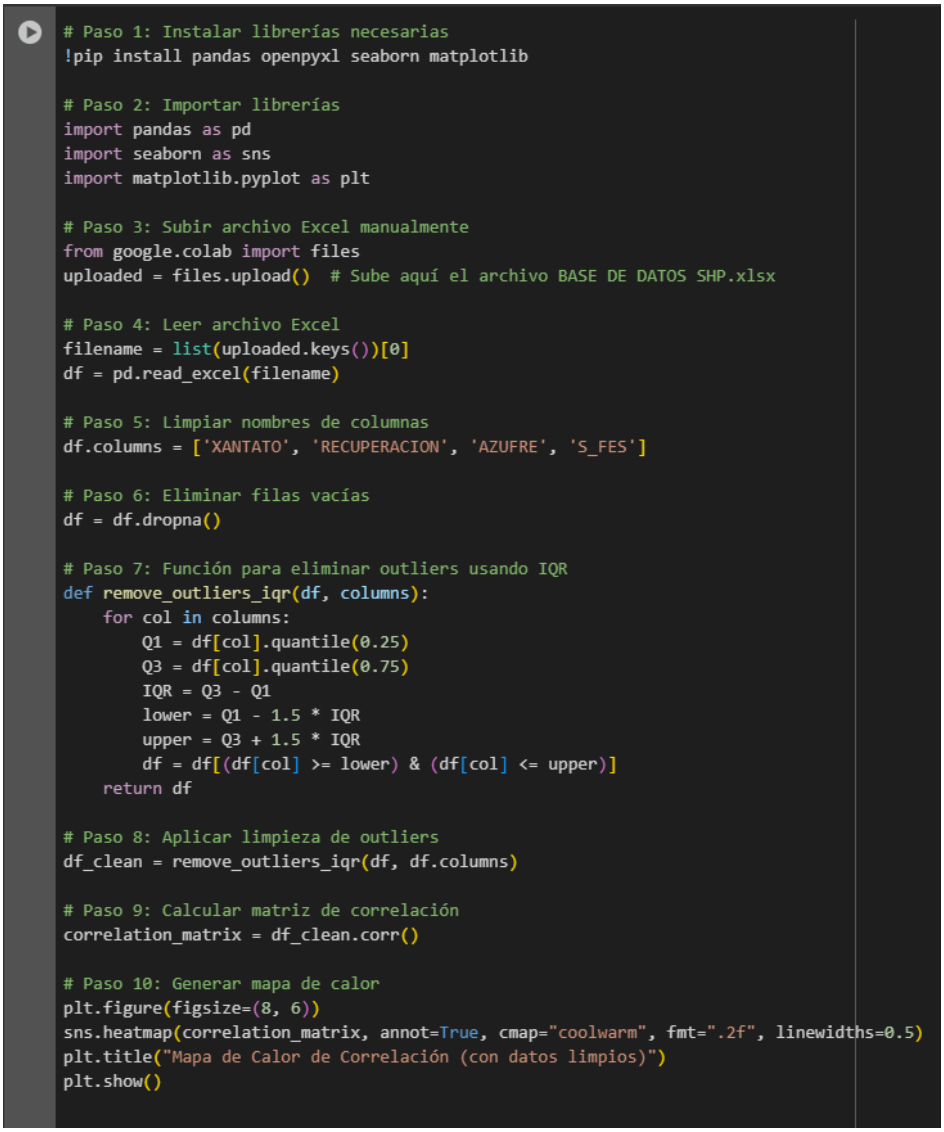
Análisis de la Correlación entre Variables

Se aplicó el coeficiente de correlación de Pearson para cuantificar la relación lineal entre las variables independientes y el consumo de xantato. El análisis fue implementado en Google Colab con Python, generando un mapa de calor para facilitar la interpretación.

Al ejecutar el código en Python utilizando la base de datos previamente depurada, se obtuvo el siguiente mapa de calor, el cual permite visualizar la correlación existente entre las variables consideradas. Esta herramienta gráfica facilita la identificación de relaciones lineales entre los parámetros analizados, sirviendo como base para determinar la viabilidad del uso de modelos de machine learning en la predicción del consumo de reactivos.

Figura 6

Codigo Pyhton Google Colab Utilizado para la generación del mapa de calor de correlaciones

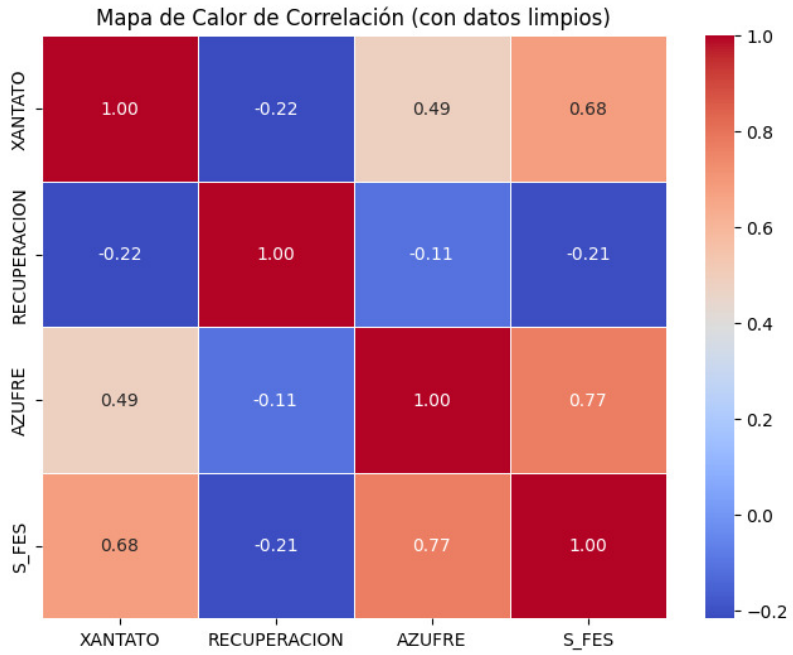


```python
# Paso 1: Instalar librerías necesarias
!pip install pandas openpyxl seaborn matplotlib

# Paso 2: Importar librerías
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Paso 3: Subir archivo Excel manualmente
from google.colab import files
uploaded = files.upload()  # Sube aquí el archivo BASE DE DATOS SHP.xlsx

# Paso 4: Leer archivo Excel
filename = list(uploaded.keys())[0]
df = pd.read_excel(filename)

# Paso 5: Limpiar nombres de columnas
df.columns = ['XANTATO', 'RECUPERACION', 'AZUFRE', 'S_FES']

# Paso 6: Eliminar filas vacías
df = df.dropna()

# Paso 7: Función para eliminar outliers usando IQR
def remove_outliers_iqr(df, columns):
    for col in columns:
        Q1 = df[col].quantile(0.25)
        Q3 = df[col].quantile(0.75)
        IQR = Q3 - Q1
        lower = Q1 - 1.5 * IQR
        upper = Q3 + 1.5 * IQR
        df = df[(df[col] >= lower) & (df[col] <= upper)]
    return df

# Paso 8: Aplicar limpieza de outliers
df_clean = remove_outliers_iqr(df, df.columns)

# Paso 9: Calcular matriz de correlación
correlation_matrix = df_clean.corr()

# Paso 10: Generar mapa de calor
plt.figure(figsize=(8, 6))
sns.heatmap(correlation_matrix, annot=True, cmap="coolwarm", fmt=".2f", linewidths=0.5)
plt.title("Mapa de Calor de Correlación (con datos limpios)")
plt.show()
```

Nota. Fragmento del script implementado en Google Colab para análisis del mapa de calor

Figura 7

Mapa de calor de correlaciones



Nota. Visualización de correlaciones lineales entre variables operativas, identificando relaciones significativas para modelado predictivo.

Uno de los hallazgos más relevantes es la correlación positiva moderadamente alta (0.68) entre el contenido de pirrotita (S_FES) y el consumo de xantato. Este resultado indica que existe una relación directa significativa entre estas dos variables, lo que implica que a mayor presencia de pirrotita en el mineral alimentado, se requiere un mayor consumo de xantato para asegurar una adecuada recuperación de dicho mineral sulfurado. Esta correlación no solo es consistente con el comportamiento químico de la pirrotita, la cual necesita colectores como el xantato para ser flotada eficazmente, sino que también respalda técnicamente la viabilidad de utilizar esta variable como predictor principal en un modelo de machine learning.

Adicionalmente, el contenido de azufre muestra una correlación positiva moderada (0.49) con el consumo de xantato, lo cual también es esperable al tratarse de otro indicador de la presencia de sulfuros, aunque menos específico que S_FES. En contraste, la recuperación metalúrgica muestra una correlación débilmente negativa (-0.22) con el consumo de xantato, lo que sugiere que incrementar el uso de este reactivo no garantiza necesariamente una mayor eficiencia en la recuperación del mineral, especialmente si no se ajusta adecuadamente a las condiciones mineralógicas.

Asimismo, se identificó una correlación alta entre azufre total y pirrotita (0.77), lo que confirma que una parte significativa del azufre presente en el mineral corresponde a esta especie mineralógica.

En conjunto, estos resultados demuestran que la pirrotita (S_FES) es la variable con mayor influencia directa sobre el consumo de xantato, por lo que su uso como variable explicativa en modelos predictivos resulta no solo válido, sino recomendable. Esto abre paso al desarrollo de modelos de machine learning orientados a la predicción del consumo de reactivos, utilizando como insumo datos de laboratorio representativos de la mineralogía del mineral procesado, lo cual puede contribuir significativamente a la optimización del proceso de flotación.

Comparativa de Modelos de Aprendizaje Automático

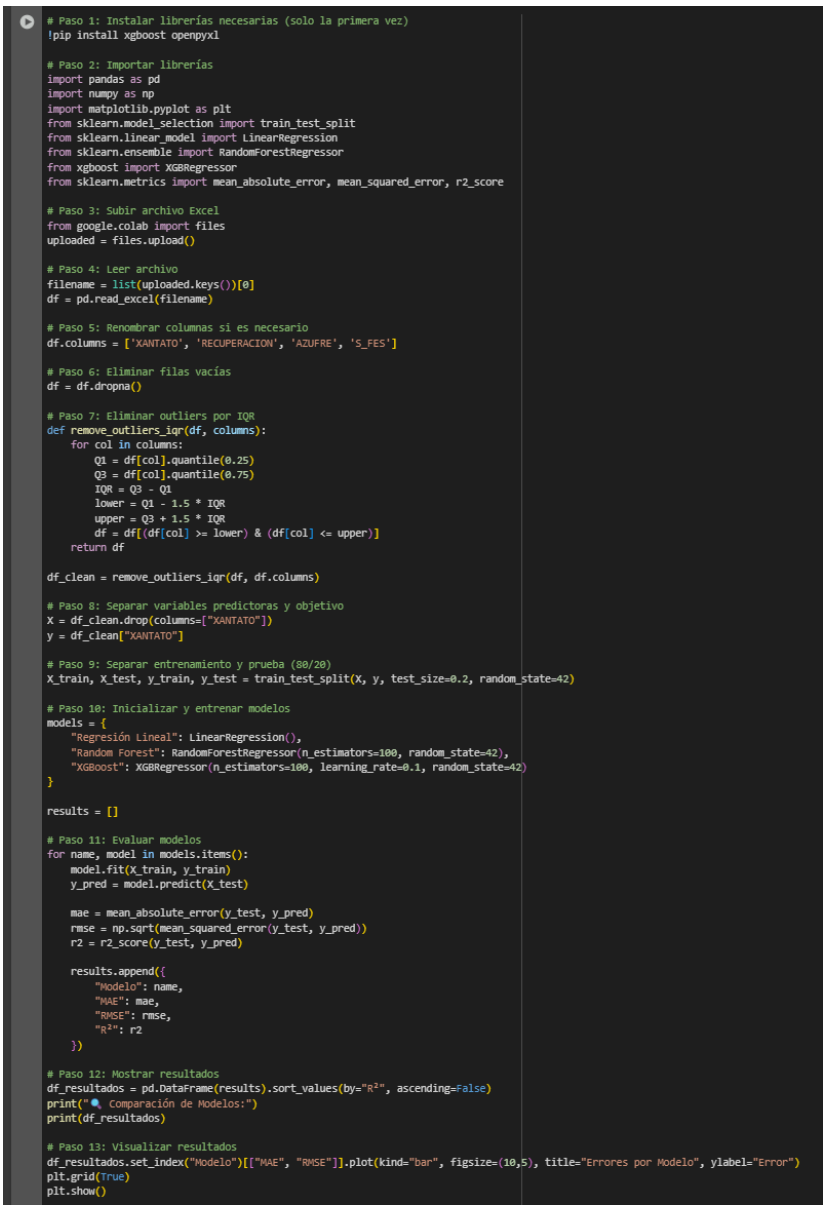
Para identificar el modelo predictivo más eficaz, se implementó un flujo de trabajo en Google Colab que automatizó la comparación de algoritmos de regresión:

- Regresión lineal.
- Random Forest Regressor.
- XGBoost Regressor.

Se evaluaron mediante métricas estándar: Error Absoluto Medio (MAE), la Raíz del Error Cuadrático Medio (RMSE) y el Coeficiente de Determinación (R^2).

Figura 8

Codigo Pyhton Google Colab Utilizado para la generación del mapa de calor de correlaciones



```python
# Paso 1: Instalar librerías necesarias (solo la primera vez)
!pip install xgboost openpyxl

# Paso 2: Importar librerías
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.ensemble import RandomForestRegressor
from xgboost import XGBRegressor
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score

# Paso 3: Subir archivo Excel
from google.colab import files
uploaded = files.upload()

# Paso 4: Leer archivo
filename = list(uploaded.keys())[0]
df = pd.read_excel(filename)

# Paso 5: Renombrar columnas si es necesario
df.columns = ['XANTATO', 'RECUPERACION', 'AZUFRE', 'S_FES']

# Paso 6: Eliminar filas vacías
df = df.dropna()

# Paso 7: Eliminar outliers por IQR
def remove_outliers_iqr(df, columns):
    for col in columns:
        Q1 = df[col].quantile(0.25)
        Q3 = df[col].quantile(0.75)
        IQR = Q3 - Q1
        lower = Q1 - 1.5 * IQR
        upper = Q3 + 1.5 * IQR
        df = df[(df[col] >= lower) & (df[col] <= upper)]
    return df

df_clean = remove_outliers_iqr(df, df.columns)

# Paso 8: Separar variables predictoras y objetivo
X = df_clean.drop(columns=["XANTATO"])
y = df_clean["XANTATO"]

# Paso 9: Separar entrenamiento y prueba (80/20)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Paso 10: Inicializar y entrenar modelos
models = {
    "Regresión Lineal": LinearRegression(),
    "Random Forest": RandomForestRegressor(n_estimators=100, random_state=42),
    "XGBoost": XGBRegressor(n_estimators=100, learning_rate=0.1, random_state=42)
}

results = []

# Paso 11: Evaluar modelos
for name, model in models.items():
    model.fit(X_train, y_train)
    y_pred = model.predict(X_test)

    mae = mean_absolute_error(y_test, y_pred)
    rmse = np.sqrt(mean_squared_error(y_test, y_pred))
    r2 = r2_score(y_test, y_pred)

    results.append({
        "Modelo": name,
        "MAE": mae,
        "RMSE": rmse,
        "R²": r2
    })

# Paso 12: Mostrar resultados
df_resultados = pd.DataFrame(results).sort_values(by="R²", ascending=False)
print("🔍 Comparación de Modelos:")
print(df_resultados)

# Paso 13: Visualizar resultados
df_resultados.set_index("Modelo")[["MAE", "RMSE"]].plot(kind="bar", figsize=(10,5), title="Errores por Modelo", ylabel="Error")
plt.grid(True)
plt.show()
```

Nota. Fragmento del script implementado en Google Colab para la comparación entre las métricas de rendimiento.

Tabla 2

Comparación de métricas de rendimiento

Modelo	MAE $\left(\frac{Kg}{día}\right)$	RMSE $\left(\frac{Kg}{día}\right)$	R^2
Regresión lineal	229.22	416.66	0.5615
Random Forest	311.07	494.67	0.3820
XGBoost	299.50	495.77	0.3793

Nota. La regresión lineal mostró mejor desempeño en todas las métricas, justificando su selección como modelo final.

Considerando los resultados presentados, el modelo de Regresión Lineal obtuvo el mejor desempeño en todas las métricas. Mostró el menor MAE (229.22) y menor RMSE (416.66), lo que indica que sus predicciones presentan, en promedio, errores menores que los modelos no lineales. Asimismo, obtuvo el mayor coeficiente de determinación ($R^2 = 0.5615$), lo que significa que explica aproximadamente el 56.15 % de la variabilidad del consumo de xantato en función de las variables utilizadas como entrada.

Si bien el coeficiente de determinación alcanzado por el modelo ($R^2 \approx 0.56$) podría considerarse moderado desde un enfoque estadístico, este valor resulta aceptable en el contexto industrial de datos de planta, donde la variabilidad inherente al proceso es elevada. En sistemas operativos reales, múltiples factores no cuantificados, como fluctuaciones no registradas en pH, Eh, presencia de otras especies sulfuradas, o condiciones hidráulicas variables, introducen ruido en los datos que reduce el poder explicativo de modelos simples. En ese sentido, capturar más del 50 % de la variabilidad mediante una única variable independiente resulta técnicamente

significativo, sobre todo considerando que el modelo es fácil de interpretar y aplicar operativamente.

Se realizó una comparación inicial de algoritmos predictivos, y se observó que la regresión lineal no solo presentó el mejor desempeño en las métricas de error y ajuste, sino que también cumple con los criterios operativos deseados (simplicidad, interpretabilidad y facilidad de implementación en planta), por tanto, se seleccionó como modelo final. En contraste, los modelos Random Forest y XGBoost, si bien son algoritmos robustos y capaces de modelar relaciones no lineales, no lograron superar en rendimiento al modelo lineal en este caso particular. Esto puede deberse a que la relación entre las variables independientes (S_FES, azufre, recuperación) y el consumo de xantato tiene un comportamiento predominantemente lineal, por lo que un modelo sencillo como la regresión lineal puede capturarla de forma más eficiente y sin sobreajuste.

Modelamiento del Modelo Seleccionado y Métricas

Luego de realizar la evaluación comparativa de distintos modelos de machine learning, se seleccionó el modelo de Regresión Lineal Multivariable como el más adecuado para el presente estudio, debido a que presentó los mejores resultados en las métricas de rendimiento. Esto justifica su implementación como modelo final para la predicción del consumo de xantato en la planta de beneficio.

El modelo se desarrolló empleando el lenguaje de programación Python, mediante la biblioteca scikit-learn, la cual proporciona herramientas eficientes para construir y entrenar modelos predictivos supervisados. Se utilizó como variables independientes el contenido de

pirrotita (S_FES), el porcentaje de azufre y la recuperación en prueba Davis, mientras que la variable dependiente fue el consumo de xantato $\left(\frac{Kg}{día}\right)$.

Con el fin de adaptar el modelo al contexto operativo diario de planta, el valor de consumo diario estimado fue transformado a $\frac{Kg}{hora}$. Asimismo, se implementó una opción interactiva que permite al usuario ingresar manualmente nuevos valores de las variables predictoras, facilitando la generación de múltiples pronósticos sin necesidad de reiniciar el programa.

Figura 9

Codigo Pyhton Google Colab Utilizado para la generación del modelo de regresión lineal



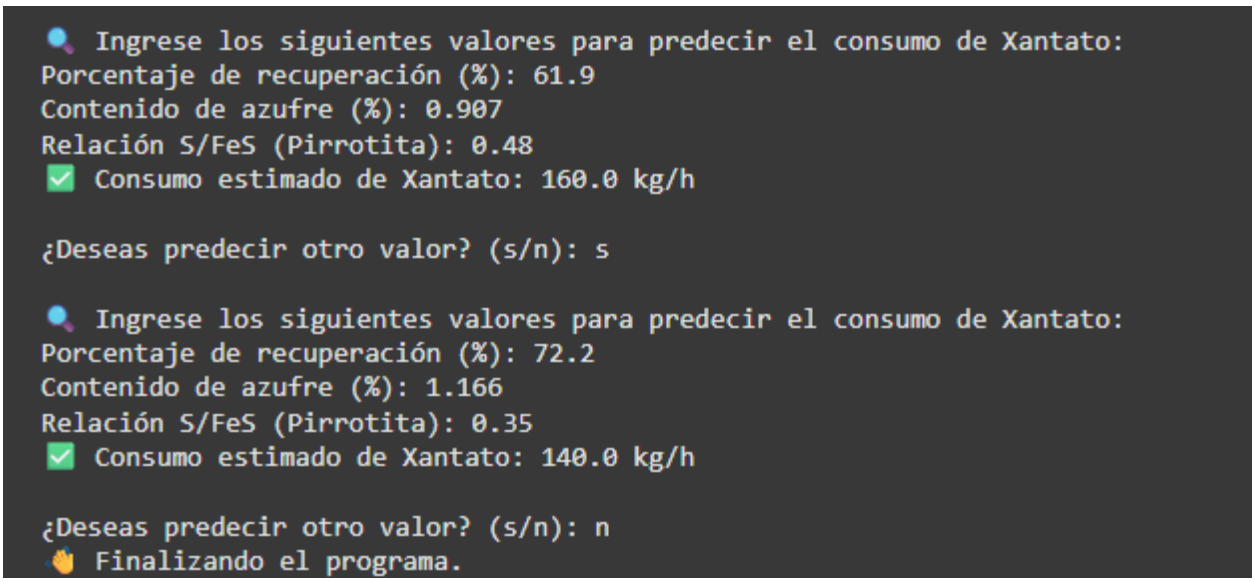
```python
# Paso 1: Importar librerías necesarias
import pandas as pd
import numpy as np
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import r2_score

# Paso 2: Leer archivo Excel
from google.colab import files
uploaded = files.upload()
filename = list(uploaded.keys())[0]
df = pd.read_excel(filename)

# Paso 3: Renombrar columnas y limpiar
df.columns = ['XANTATO', 'RECUPERACION', 'AZUFRE', 'S_FES']
df = df.dropna()

# Paso 4: Entrenar modelo de regresión lineal
X = df[['RECUPERACION', 'AZUFRE', 'S_FES']]
y = df['XANTATO']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

modelo = LinearRegression()
modelo.fit(X_train, y_train)

# Paso 5: Función de predicción en bucle
def predecir_en_bucle():
    while True:
        print("\n🔍 Ingrese los siguientes valores para predecir el consumo de Xantato:")
        try:
            rec = float(input("Porcentaje de recuperación (%): "))
            azu = float(input("Contenido de azufre (%): "))
            pirr = float(input("Relación S/FeS (Pirrotita): "))
        except ValueError:
            print("❌ Entrada inválida. Asegúrate de ingresar valores numéricos.")
            continue

        entrada = pd.DataFrame([[rec, azu, pirr]], columns=['RECUPERACION', 'AZUFRE', 'S_FES'])
        prediccion_total = modelo.predict(entrada)[0]

        flujo_kg_h = prediccion_total / 24
        flujo_kg_h_redondeado = float(f"{flujo_kg_h:.2g}")

        print(f"✅ Consumo estimado de Xantato: {flujo_kg_h_redondeado} kg/h")

        repetir = input("\n¿Deseas predecir otro valor? (s/n): ").strip().lower()
        if repetir != 's':
            print("👋 Finalizando el programa.")
            break

# Paso 6: Ejecutar la función interactiva
predecir_en_bucle()
```

Nota. Fragmento del script implementado en Google Colab para generación del modelo de regresión lineal

A través de ejecutar el código se obtiene un cuadro donde te pide completar las variables dependientes para poder predecir el consumo de Xantato este código puede repetirse a través de indicaciones de Si o No, donde al indicar que NO se finaliza el código.

Figura 10

Cuadro de cálculo de Google Colab para la predicción del consumo de xantato



Nota. Ejemplos de resultados de la predicción del consumo de xantato a través de los valores de % de recuperación, % contenido de azufre y cantidad de pirrorita.

Evaluación del Modelo de Regresión Lineal Múltiple

Luego del preprocesamiento de los datos, incluyendo la eliminación de valores atípicos mediante el método del rango intercuartílico (IQR), se entrenó el modelo con el 80% del total de registros, utilizando el 20% restante para evaluar su desempeño.

La comparación entre los valores reales y predichos del consumo horario de xantato se presenta en la tabla correspondiente.

Tabla 3

Comparación entre los valores reales obtenidos en la planta de beneficio y la predicción del modelo estadístico.

Xantato Real $\left(\frac{Kg}{h}\right)$	**Xantato Predicho $\left(\frac{Kg}{h}\right)$**	**% Error Absoluto**
166.59	165.69	0.54%
176.14	178.33	1.24%
187.28	179.46	4.18%
193.64	188.97	2.41%
172.96	173.59	0.36%
155.46	158.76	2.12%
156.46	198.43	26.82%
182.51	180.79	0.94%
176.14	172.63	1.99%
180.91	178.51	1.33%
153.87	158.66	3.11%
172.96	171.66	0.75%
158.64	162.43	2.39%
145.91	150.73	3.30%
174.55	172.92	0.93%
165.00	165.27	0.16%
188.87	184.15	2.50%
217.51	204.24	6.10%

Xantato Real $\left(\frac{Kg}{h}\right)$	Xantato Predicho $\left(\frac{Kg}{h}\right)$	% Error Absoluto
169.78	165.58	2.47%
207.96	195.86	5.82%
207.96	198.65	4.48%
177.73	177.05	0.38%
184.10	180.37	2.03%
144.32	144.89	0.39%
152.28	155.6	2.18%
203.19	193.87	4.59%
166.59	170.93	2.61%
180.91	179.18	0.96%
176.14	174.42	0.98%
227.05	212.41	6.45%
193.64	187.57	3.13%
131.59	135.35	2.86%
163.41	167.33	2.40%
141.14	148.21	5.01%
160.23	161.25	0.64%
187.28	184	1.75%
152.62	202.27	32.53%
152.28	156.74	2.93%
131.29	176.99	34.81%

Xantato Real $\left(\frac{Kg}{h}\right)$	Xantato Predicho $\left(\frac{Kg}{h}\right)$	% Error Absoluto
176.14	177.08	0.53%
217.51	202.57	6.87%
212.74	201.49	5.29%
161.82	163.09	0.78%

Nota. Comparación entre valores reales medidos en planta y predicciones generadas por el modelo de regresión lineal, mostrando el error absoluto porcentual como indicador de ajuste y confiabilidad operativa.

Análisis estadístico y validación de hipótesis

Formulación de hipótesis:

Hipótesis nula (H_0): No existe una diferencia significativa entre los valores reales y los valores predichos de consumo de xantato. Es decir, la media de los residuos (errores) es igual a cero.

Hipótesis alternativa (H_1): Existe una diferencia significativa entre los valores reales y los valores predichos de consumo de xantato. Es decir, la media de los residuos es distinta de cero.

Prueba estadística empleada:

Se aplicó una prueba t de una muestra (one-sample t-test) sobre la serie de residuos, definidos como la diferencia entre el consumo real y el consumo predicho de xantato, para evaluar si su media difiere significativamente de cero.

Resultados:

Tamaño de muestra: n = 43 observaciones válidas

Media de los residuos: $\mu = -1.23$ kg/día

Desviación estándar: $\sigma = 13.57$ kg/día

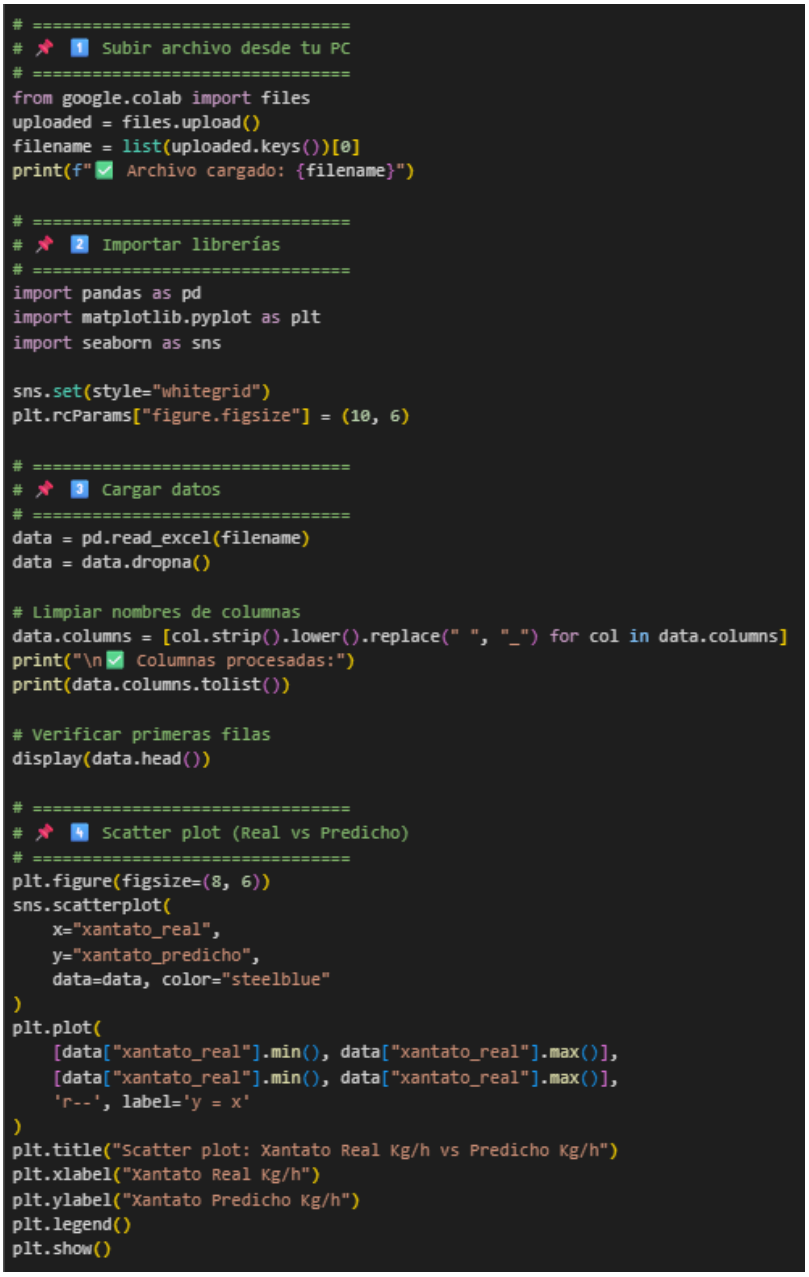
Estadístico t: $t = -0.59$

Valor p: 0.556

Dado que el valor p obtenido (0.556) es mayor al umbral común de significancia ($\alpha = 0.05$), no se rechaza la hipótesis nula. Esto indica que la media de los errores de predicción no difiere significativamente de cero, por lo tanto, el modelo predictivo no presenta sesgo sistemático y puede considerarse estadísticamente válido para estimar el consumo de xantato en función del contenido de pirrotita bajo condiciones industriales reales.

Figura 11

Codigo Pyhton Google Colab utilizado para la generación de gráfico Scatter plot, Line plot comparativo y Residual plot

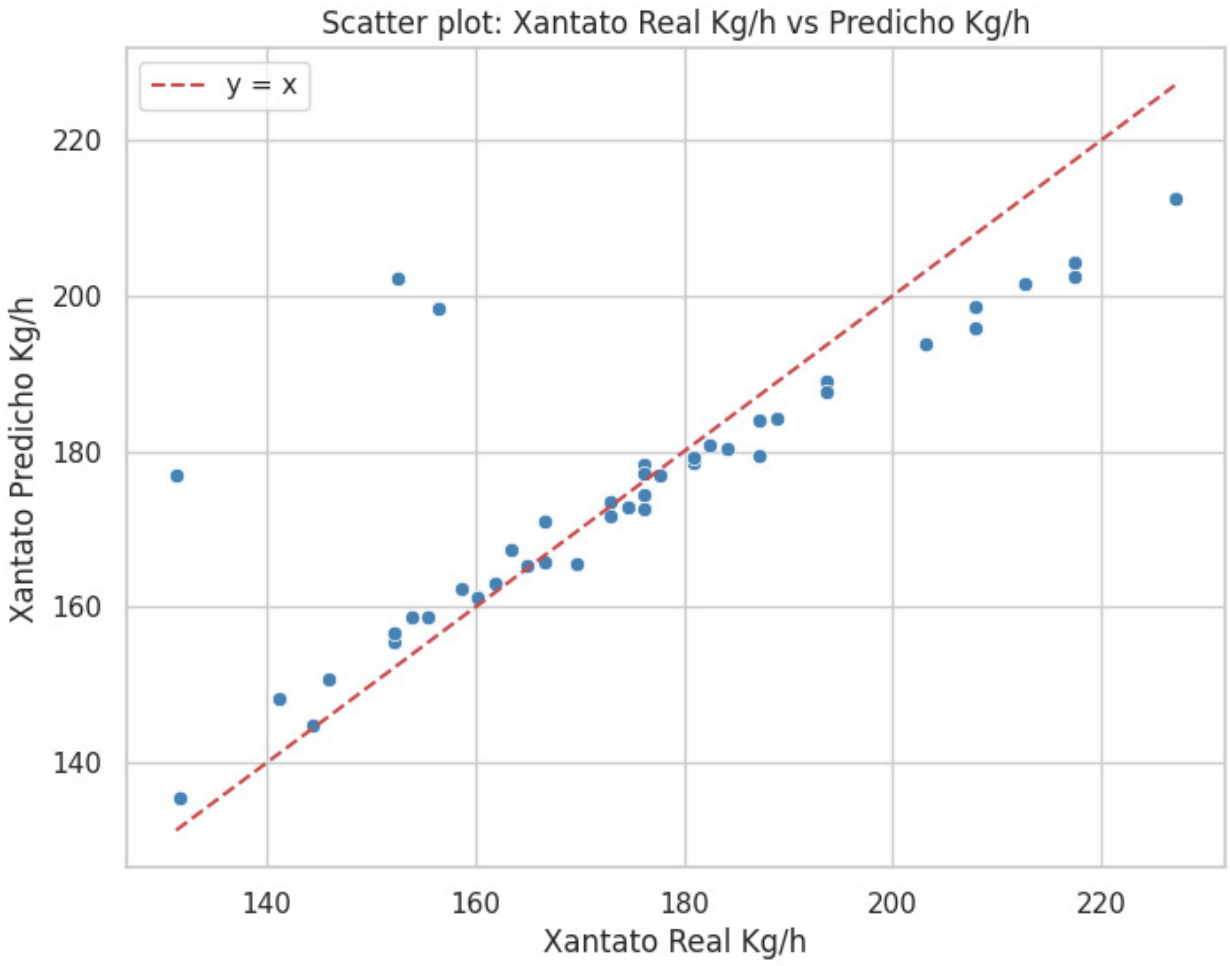


```
# ================================
# 📌 1️⃣ Subir archivo desde tu PC
# ================================
from google.colab import files
uploaded = files.upload()
filename = list(uploaded.keys())[0]
print(f"✅ Archivo cargado: {filename}")

# ================================
# 📌 2️⃣ Importar librerías
# ================================
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

sns.set(style="whitegrid")
plt.rcParams["figure.figsize"] = (10, 6)

# ================================
# 📌 3️⃣ Cargar datos
# ================================
data = pd.read_excel(filename)
data = data.dropna()

# Limpiar nombres de columnas
data.columns = [col.strip().lower().replace(" ", "_") for col in data.columns]
print("\n✅ Columnas procesadas:")
print(data.columns.tolist())

# Verificar primeras filas
display(data.head())

# ================================
# 📌 4️⃣ Scatter plot (Real vs Predicho)
# ================================
plt.figure(figsize=(8, 6))
sns.scatterplot(
    x="xantato_real",
    y="xantato_predicho",
    data=data, color="steelblue"
)
plt.plot(
    [data["xantato_real"].min(), data["xantato_real"].max()],
    [data["xantato_real"].min(), data["xantato_real"].max()],
    'r--', label='y = x'
)
plt.title("Scatter plot: Xantato Real Kg/h vs Predicho Kg/h")
plt.xlabel("Xantato Real Kg/h")
plt.ylabel("Xantato Predicho Kg/h")
plt.legend()
plt.show()
```

Nota. Se utiliza el código para la generación del gráfico Scatter plot, Line plot comparativo y Residual plot.

Figura 12

Matriz Scatter plot



Nota. Gráfico de dispersión (Xantato Real vs Predicho) para evaluar el ajuste del modelo de predicción de consumo de reactivo colector.

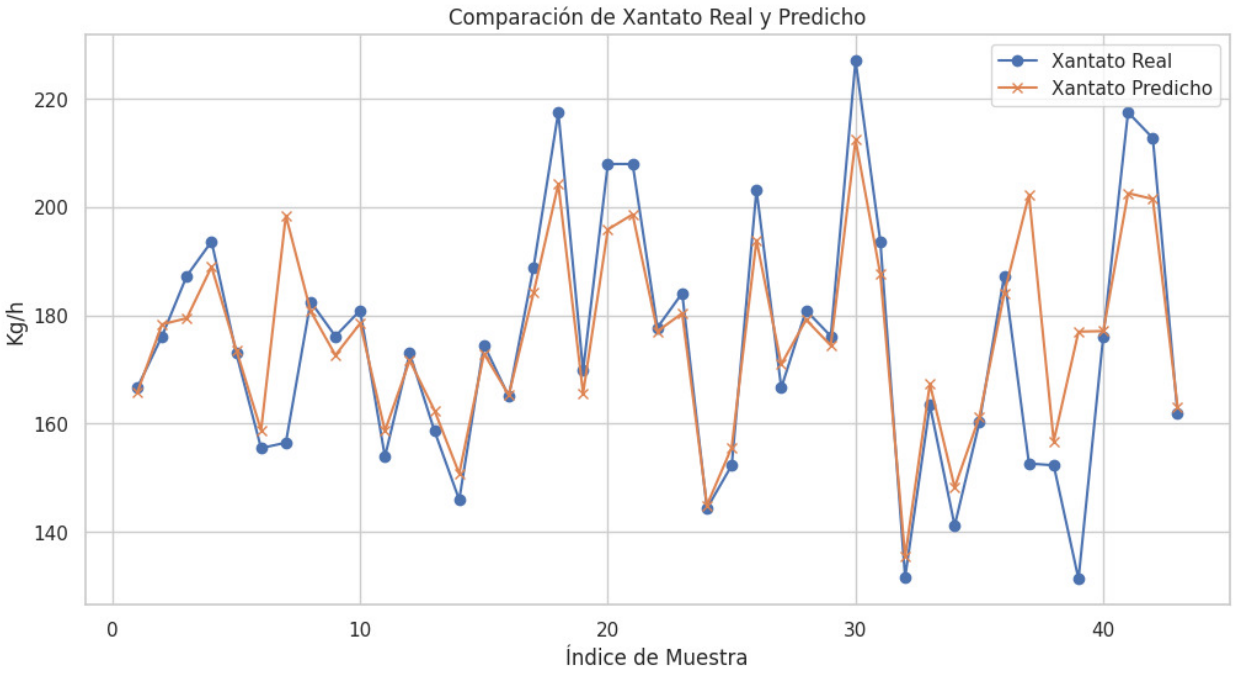
- La mayoría de los puntos se agrupan cercanos a la línea de identidad (y = x), indicando que el modelo logra capturar adecuadamente la tendencia general del consumo de xantato en función de las variables explicativas seleccionadas.

- Sin embargo, se identifican puntos atípicos más alejados de la línea de referencia, reflejando errores de predicción más elevados en ciertos casos específicos.
- Estos desvíos podrían atribuirse a condiciones operacionales no modeladas explícitamente, como cambios bruscos en la mineralogía del mineral alimentado, presencia de materiales con alta variabilidad sulfurada o transiciones entre frentes de mina con características distintas.
- Asimismo, eventos puntuales como fallas o paradas parciales de equipos (bombas dosificadoras, celdas de flotación) podrían alterar temporalmente la dosificación de reactivos, generando consumos atípicos que el modelo no predice con precisión.

  Este análisis gráfico refuerza la necesidad de considerar factores operativos adicionales o variables contextuales en futuras versiones del modelo, así como de implementar estrategias de control que mitiguen la variabilidad no explicada en la dosificación de xantato.

Figura 13

Matriz Line plot



Nota. Comparación temporal del consumo real y predicho de xantato a lo largo del índice de muestras.

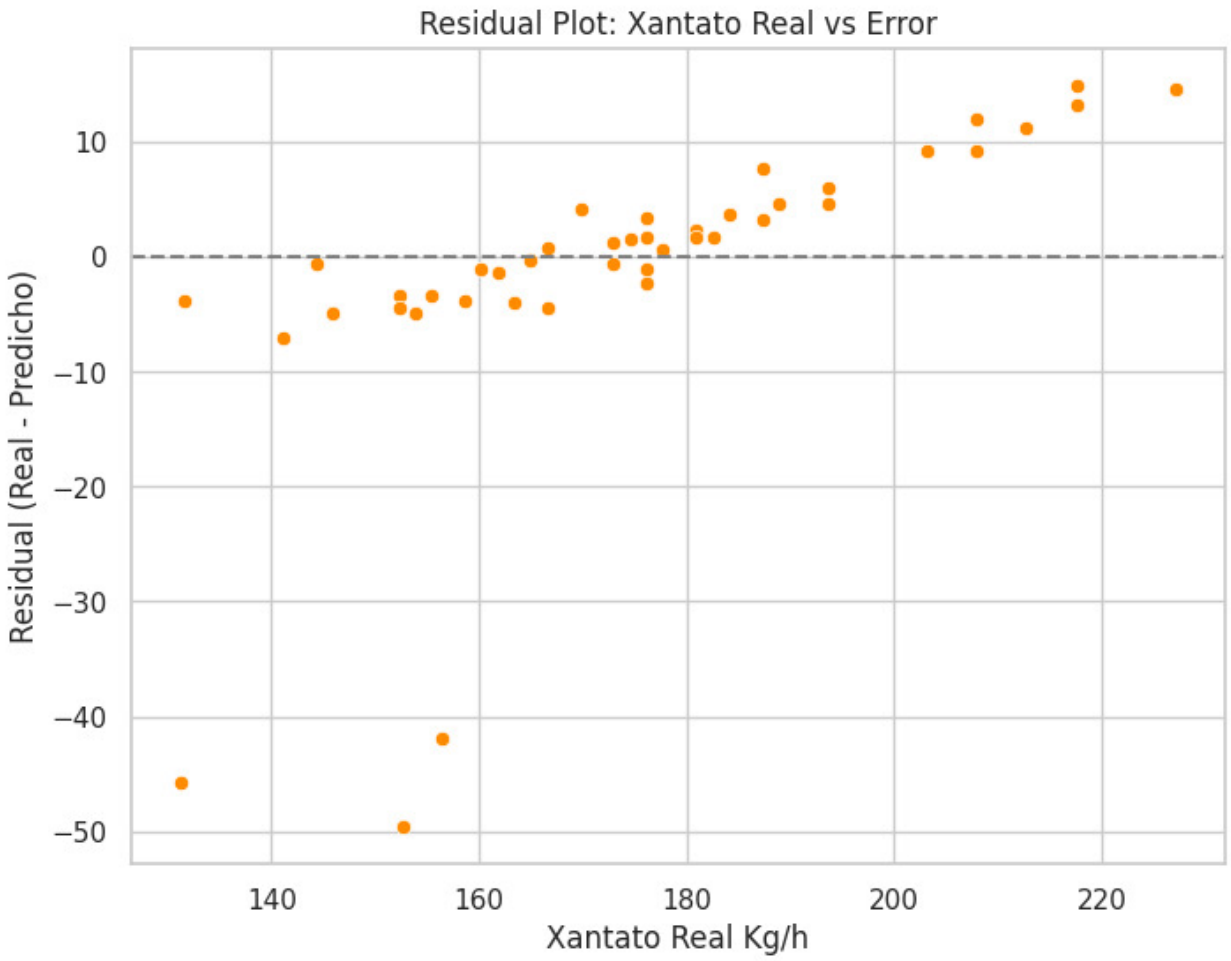
- El gráfico muestra que el modelo reproduce adecuadamente la tendencia general de variación del consumo de xantato, siguiendo de manera cercana los picos y valles observados en los datos reales.
- Se identifican ciertos intervalos con mayor discrepancia entre la predicción y el valor medido, particularmente en algunos picos pronunciados, lo que sugiere que el modelo puede subestimar o sobreestimar consumos en condiciones extremas.
- Estas discrepancias podrían asociarse a variabilidad operativa no capturada, como transiciones abruptas en la calidad del mineral alimentado o ajustes operacionales no

registrados en las variables explicativas.

Este resultado evidencia la capacidad del modelo para reflejar el comportamiento general del proceso, al tiempo que señala la necesidad de continuar refinándolo para mejorar su desempeño en escenarios de alta variabilidad.

Figura 14

Matriz Residual plot



Nota. Gráfico de residuos (Real – Predicho) en función del consumo real de xantato.

- Se observa un patrón de tendencia creciente en los residuos a medida que aumenta el valor real de xantato, indicando la presencia de heterocedasticidad en el modelo.
- Este comportamiento sugiere que el modelo tiende a subestimar sistemáticamente los consumos altos y a sobreestimar algunos consumos bajos, limitando su capacidad de generalización en rangos extremos de operación.
- Además, se detectan valores residuales atípicos negativos importantes, que podrían estar asociados a situaciones operativas excepcionales como cambios abruptos en la mineralogía, condiciones de alimentación no representadas en el entrenamiento del modelo o fallas transitorias en los sistemas de dosificación.

  Este análisis de residuos resalta la importancia de considerar transformaciones, técnicas de modelado más robustas o la inclusión de variables adicionales para capturar mejor la variabilidad y reducir el sesgo en escenarios de operación críticos.

Discusión

El análisis estadístico y predictivo permitió confirmar que la proporción de pirrotita (S/FeS) es el factor mineralógico más influyente en el consumo de xantato. La correlación positiva moderada-alta obtenida ($r \approx 0.68$) respalda su uso como variable predictora clave en la estimación del consumo específico de colector. Esta relación fue modelada adecuadamente mediante una regresión lineal , seleccionada tras una comparativa de algoritmos de aprendizaje automático por su mejor balance entre simplicidad, interpretabilidad y desempeño predictivo en contexto industrial.

El modelo alcanzó un coeficiente de determinación de $R^2 \approx 0.56$, lo cual refleja un nivel de ajuste moderado pero aceptable dadas las condiciones de alta variabilidad y ruido inherentes a datos operativos reales. Esta limitación era previsible, ya que no todas las variables que influyen

sobre el consumo de xantato se encuentran registradas de forma consistente en planta. Factores como el pH de la pulpa, la dosificación de otros reactivos, la distribución de tamaño de partícula, la mineralogía secundaria o la estabilidad hidráulica del sistema no fueron incorporados explícitamente. Sin embargo, esta omisión no invalida el modelo, sino que delimita su alcance: se trata de una herramienta útil, replicable y operativamente aplicable, aunque con capacidad predictiva circunscrita al contexto y datos bajo los cuales fue entrenada.

A fin de validar formalmente la hipótesis estadística planteada, se realizó una prueba t de una muestra (one-sample t-test) sobre los residuos (diferencias entre valores reales y predichos de consumo de xantato). El resultado indicó que la media de los residuos no difiere significativamente de cero ($t = –0.59$, $p = 0.556$), por lo que no se rechaza la hipótesis nula. Esto implica que no existe evidencia estadísticamente significativa de sesgo sistemático en el modelo predictivo, reforzando su validez y utilidad práctica para la toma de decisiones en planta. Este hallazgo respalda la hipótesis general planteada en el estudio y demuestra que la relación propuesta entre el contenido de pirrotita y el consumo de xantato puede ser formalizada cuantitativamente bajo un enfoque .

Estos resultados son consistentes con estudios previos que reportan que, aunque pueden existir relaciones no lineales en sistemas multivariables, en rangos operativos delimitados y con un solo factor mineralógico predominante, la relación entre sulfuros reactivos y consumo de colector puede modelarse de forma lineal (James et al., 2013; Baldeón Hurtado & Oropeza Huamán, 2015). La elección de un enfoque parsimonioso permitió maximizar la interpretabilidad sin sacrificar la aplicabilidad en planta.

La implementación del modelo en Google Colab mediante programación en Python facilitó la integración de técnicas de machine learning en un entorno accesible y replicable,

promoviendo un enfoque de minería inteligente basado en datos. Asimismo, el desarrollo de un sistema interactivo de predicción refuerza la transferencia tecnológica hacia el equipo operativo de planta y promueve decisiones basadas en evidencia.

En conjunto, la presente investigación demuestra la viabilidad de aplicar herramientas de inteligencia artificial interpretables para optimizar la dosificación de xantato en circuitos de flotación con alta presencia de pirrotita. Esta contribución permite reducir costos, mejorar la eficiencia metalúrgica y fortalecer la sostenibilidad operativa de Shougang Hierro Perú, a la vez que sienta las bases para futuras investigaciones orientadas a incorporar mayor complejidad mineralógica y operativa en los modelos predictivos desarrollados.

Conclusiones y Recomendaciones

Conclusiones

Las hipótesis planteadas en esta investigación han sido estadísticamente validadas. En primer lugar, se confirmó la existencia de una relación significativa entre el contenido de pirrotita (% S/FeS) y el consumo específico de xantato (kg/día), respaldada por un coeficiente de correlación de Pearson de aproximadamente 0.68. Esta evidencia respalda la hipótesis general del estudio, que propone a la pirrotita como la variable mineralógica más influyente sobre la demanda de colector en el proceso de flotación.

La validación del modelo predictivo mediante análisis de residuos y prueba t de una muestra ($p = 0.556$) demostró que no existe un sesgo sistemático en las predicciones, cumpliendo así con la hipótesis específica sobre la consistencia estadística del modelo. La diferencia entre valores reales y predichos no fue estadísticamente significativa, lo cual confirma la validez operativa del modelo para estimar el consumo de xantato.

La caracterización exploratoria de los datos históricos evidenció una amplia dispersión en el consumo diario de xantato (1800–7290 kg/día), reflejando la alta variabilidad mineralógica y operativa de planta. Esta situación justificó la necesidad de un modelo predictivo que permita estandarizar y racionalizar la dosificación del colector, optimizando los costos y la eficiencia del proceso.

La comparación entre distintos algoritmos de aprendizaje automático implementados en Python (Random Forest, XGBoost y regresión lineal) determinó que el modelo de regresión lineal múltiple ofreció el mejor desempeño para este caso operativo ($R^2 \approx 0.5615$). Esto valida la hipótesis sobre la capacidad de modelamiento lineal entre el contenido de pirrotita y el consumo de reactivo.

El modelo final explicó el 56 % de la variabilidad del consumo de xantato, con errores absolutos moderados y buen ajuste estadístico. Si bien su poder predictivo es limitado, resulta suficiente para anticipar tendencias de consumo y establecer estrategias de dosificación más técnicas y menos empíricas, fortaleciendo el control operativo del circuito de flotación.

La herramienta desarrollada mediante Google Colab e implementada con bibliotecas como pandas, seaborn y scikit-learn demostró ser transparente, replicable y accesible, consolidando una aplicación práctica de minería inteligente orientada a la eficiencia operativa y la sostenibilidad industrial.

Se reconoce, sin embargo, que el enfoque , aunque metodológicamente sólido y estadísticamente válido, presenta limitaciones por excluir variables operativas adicionales (pH, Eh, granulometría, dosificación de otros reactivos). Asimismo, el uso de datos de una sola campaña minera restringe la capacidad de generalización del modelo frente a variaciones mineralógicas futuras. Estas limitaciones no invalidan el modelo, pero delimitan su alcance, y representan oportunidades para investigaciones futuras mediante enfoques multivariables o modelos más complejos.

En conjunto, la investigación valida estadística y operativamente la hipótesis principal y demuestra la viabilidad de utilizar modelos de aprendizaje automático simples y explicables para optimizar el consumo de xantato en plantas de flotación con alta presencia de pirrotita. Su adopción contribuiría a la reducción de costos por tonelada tratada, una mayor estabilidad del circuito y una operación más alineada con los principios de minería responsable y sostenible en la unidad minera de Shougang Hierro Perú – Marcona.

Recomendaciones

A partir de los resultados obtenidos y considerando las limitaciones metodológicas identificadas, se plantean las siguientes recomendaciones orientadas al fortalecimiento de futuras investigaciones y desarrollos en el campo del control de reactivos en flotación de minerales con presencia de sulfuros reactivos:

- Ampliar el enfoque hacia modelos multivariables en estudios futuros, incorporando variables operativas adicionales como el pH de la pulpa, potencial redox (Eh), granulometría, tipo y dosis de otros reactivos, así como parámetros hidrodinámicos del sistema. Estas variables pueden influir significativamente en el consumo de xantato y permitir una predicción más robusta en contextos de mayor complejidad mineralógica.
- Diseñar estudios experimentales en condiciones controladas de laboratorio o planta piloto que permitan validar de forma independiente la relación causal entre el contenido de pirrotita y el consumo de xantato, aislando el efecto de otras especies mineralógicas presentes en la matriz de alimentación.
- Explorar la aplicación de técnicas de minería de datos y aprendizaje automático explicable, como modelos lineales penalizados (Lasso, Ridge), árboles de decisión interpretables o métodos basados en SHAP (Shapley Additive Explanations), que permitan descomponer el efecto de cada variable sobre la respuesta y mejorar la trazabilidad de los modelos predictivos en contextos operativos reales.
- Reforzar la gestión y gobernanza de datos operativos y mineralógicos, implementando procedimientos normalizados para la recolección, almacenamiento y validación de datos de planta. Esto permitirá futuras investigaciones con bases de

datos más completas, limpias y estandarizadas, mejorando la calidad del modelamiento estadístico.

- Realizar análisis longitudinales de series de tiempo operativas, para evaluar la estabilidad del modelo predictivo frente a cambios de campaña, zonas del yacimiento o ajustes operativos. Esto permitiría identificar momentos de ruptura o pérdida de ajuste, y definir estrategias óptimas de recalibración periódica del modelo.
- Desarrollar investigaciones que evalúen el impacto ambiental y económico de distintas estrategias de dosificación de colectores, considerando escenarios simulados con modelos predictivos. Estos estudios podrían ayudar a fundamentar decisiones futuras sobre prácticas sostenibles en flotación y sobre políticas de eficiencia reactiva en minería.
- Profundizar en el estudio del comportamiento electroquímico de la pirrotita y su activación superficial, mediante técnicas avanzadas como voltamperometría cíclica o espectroscopía de impedancia, lo cual permitiría modelar de manera más precisa su interacción con el xantato bajo diferentes condiciones físico-químicas de pulpa.

Bibliografía

Adams, M. D. (2005). *Advances in Gold Ore Processing*. Elsevier. https://www.sciencedirect.com/book/9780444517302/advances-in-gold-ore-processing

Ata, S. (2012). Phenomena in the froth phase of flotation – A review. *International Journal of Mineral Processing*, 102-103, 1–12. https://doi.org/10.1016/j.minpro.2011.09.001

Azadi, M., Northey, S. A., Hasanpour, H., Edraki, M., & Ali, S. H. (2020). *Transparency on greenhouse gas emissions from mining to enable climate-smart minerals supply*. Nature Geoscience, 13(2), 100–104. https://doi.org/10.1038/s41561-019-0529-3

Baldeón, D. L., & Oropeza, M. A. (2015). *Tratamiento de minerales con alto contenido de pirrotita magnética para la concentración de cobre por flotación en la Cía Minera Caraveli - 2014*. Universidad Nacional José Faustino Sánchez Carrión. http://repositorio.unjfsc.edu.pe/handle/UNJFSC/635

Barrera, P. (2015). *Efecto de iones activadores en la flotación de sulfuros de hierro*. Universidad Técnica Federico Santa María.

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

Buxton, M. W. N., & Loveday, B. K. (2002). Flotation 2002: Solutions for the New Millennium. *Minerals Engineering*, 15(10), 739–740. https://doi.org/10.1016/S0892-6875(02)00196-5

Cayunao, B. L. (2014). *Estudio comparativo de la activación de pirita en flotación de minerales a nivel industrial y de laboratorio usando la técnica de espectroscopía fotoelectrónica*

de rayos X (XPS). Universidad de Chile. https://repositorio.uchile.cl/handle/2250/117024

Chávez, L. M. (2017). *Influencia de la granulometría en el consumo de xantato en flotación de sulfuros*. Universidad Nacional de Ingeniería.

Chen, Y., Zhou, Q., & Zhang, H. (2022). Deep learning for flotation reagent dosage control. *Computers & Chemical Engineering*, 165, 107972. https://doi.org/10.1016/j.compchemeng.2022.107972

Cilek, E. C. (2004). Flotation performance analysis by design of experiments. *Minerals Engineering*, 17(10), 1053–1062. https://doi.org/10.1016/j.mineng.2004.04.002

Córdova, D. (2020). *Optimización del consumo de xantato en minerales con alto contenido de sulfuros en plantas concentradoras*. Universidad Nacional de Ingeniería.

Cruz, N., & Finch, J. A. (2012). The interaction of clay minerals with hydrophobic collectors in flotation. *Minerals Engineering*, 39, 77–85. https://doi.org/10.1016/j.mineng.2012.06.012

Dobby, G. S., & Finch, J. A. (1986). Flotation column scale-up and modelling. *Minerals Engineering*, 39(3–4), 233–248. https://doi.org/10.1016/0892-6875(86)90051-3

García, J., Ríos, M., & Alva, E. (2020). Modelado estadístico del consumo de reactivos en flotación de sulfuros complejos. *Revista Peruana de Metalurgia*, 18(2), 25–34. https://revistaperuanademetalurgia.com/index.php/rpm/article/view/45

Géron, A. (2019). *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed.). O'Reilly Media. https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. https://www.deeplearningbook.org

Hassanzadeh, A., & Karimi, A. (2017). Application of Machine Vision in Monitoring and Control of Flotation Systems. *Minerals Engineering*, 100, 107–117. https://doi.org/10.1016/j.mineng.2016.10.008

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

Hernández, R., Fernández, C., & Baptista, P. (2014). *Metodología de la investigación* (6.ª ed.). McGraw-Hill Education. https://latam.cengage.com/c/metodologia-de-la-investigacion-6ed-hernandez-sampieri/

Hilson, G., & Murck, B. (2000). *Sustainable development in the mining industry: clarifying the corporate perspective*. Resources Policy, 26(4), 227–238. https://doi.org/10.1016/S0301-4207(00)00041-6

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: With Applications in R*. Springer. https://doi.org/10.1007/978-1-4614-7138-7

Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*. Springer. https://doi.org/10.1007/978-1-4614-6849-3

Laskowski, J. S. (2010). *Fundamentals of Froth Flotation*. Society for Mining, Metallurgy & Exploration. https://www.smenet.org/Store/ProductDetails.aspx?itemNumber=2262

Li, C., & Xie, G. (2016). Machine learning in mineral processing: A review. *Minerals Engineering*, 97, 29–41. https://doi.org/10.1016/j.mineng.2016.07.005

Liu, G., Xu, Z., & Finch, J. A. (2000). Electrokinetic studies of the effect of lime on pyrrhotite flotation. *International Journal of Mineral Processing*, 58(3), 183–197. https://doi.org/10.1016/S0301-7516(99)00065-1

Lottermoser, B. G. (2010). *Mine wastes: Characterization, treatment and environmental impacts* (3rd ed.). Springer. https://doi.org/10.1007/978-3-642-12419-8

Molina, D., Ruiz, J., & Alarcón, M. (2023). Application of machine learning for reagent optimization in iron flotation circuits. *Minerals Engineering*, 198, 107117. https://doi.org/10.1016/j.mineng.2023.107117

Muñoz, F. (2013). *Influencia del Eh en la flotación selectiva de sulfuros de hierro*. Universidad de Antofagasta.

Nasiri, M., Iqbal, S., & Särkkä, S. (2024). Physics-Informed Machine Learning for Grade Prediction in Froth Flotation. *arXiv preprint*, arXiv:2408.15267. https://arxiv.org/abs/2408.15267

Paryani, K., Bahadori, A., & Farzanegan, A. (2021). Prediction of flotation performance using Random Forest and XGBoost algorithms. *Journal of Mining Science*, 57(3), 445–457. https://doi.org/10.1007/s10913-021-00665-z

Pyle, D. (1999). *Data preparation for data mining*. Morgan Kaufmann. https://www.elsevier.com/books/data-preparation-for-data-mining/pyle/978-1-55860-529-9

Ralston, J., Fornasiero, D., & Hayes, R. (1999). Bubble–particle attachment and detachment in flotation. *International Journal of Mineral Processing*, 56(1–4), 133–164. https://doi.org/10.1016/S0301-7516(98)00050-4

Ramos Oxa, G. Y. (2014). *Determinación de parámetros de operación y desarrollo del circuito de flotación para un mineral polimetálico Pb-Ag-Zn con presencia de pirrotita a nivel laboratorio*. Universidad Nacional de San Agustín de Arequipa. http://repositorio.unsa.edu.pe/handle/UNSA/2582

Runge, K. (2010). Flotation circuit design for high throughputs and complex ores. *Minerals Engineering*, 23(10), 835–841. https://doi.org/10.1016/j.mineng.2010.05.005

Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press. https://doi.org/10.1017/CBO9781107298019

Silva, P., & Morales, R. (2018). *Estudio de colectores en flotación de sulfuros complejos*. Universidad Nacional de San Juan (Argentina).

Sutherland, K. L. (1948). Physical chemistry of flotation. XI. Kinetics of the flotation process. *Journal of Physical and Colloid Chemistry*, 52(2), 394–425. https://doi.org/10.1021/j150469a010

Vinnett, L., Valderrama, L., & Aplan, F. F. (2009). Selective flotation of pyrite and pyrrhotite from magnetite ores. *Minerals Engineering*, 22(9–10), 852–858. https://doi.org/10.1016/j.mineng.2009.04.001

Wills, B. A., & Finch, J. (2016). *Wills' Mineral Processing Technology* (8th ed.). Elsevier. https://www.elsevier.com/books/wills-mineral-processing-technology/wills/978-0-08-097053-0

Wang, X., Peng, Y., & Seaman, J. (2015). Machine learning approaches for flotation froth image analysis. *Minerals Engineering*, 79, 92–102. https://doi.org/10.1016/j.mineng.2015.04.008