

Predicción de contenido de Pd y Pt utilizando modelos multivariados de Aprendizaje Automático y su Integración con Geoestadística en la Generación de nuevos targets de Exploración en un Yacimiento de Sulfuros de Ni-Cu-PGEs.
(Geología y Exploraciones - Tecnologías digitales y de IA)

Neil Harrison Berrospi Rios¹, Daniel Abinadab Basilio Rojas² y Carlos Josué Mendoza Kuong³

¹ Autor: Universidad Nacional de Ingeniería, Lima, Lima, Perú (neilberrospir@uni.pe)

² Coautor 1: Atticus Peru SAC, Lima, Lima, Perú (daniel.basilio@atticusgeo.pe)

³ Coautor 2: Beijing Shougang, Lima, Lima, Perú (cmendozak@uni.pe)

RESUMEN

El yacimiento en estudio corresponde a un sistema de sulfuros magmáticos de Ni, Cu y elementos del grupo del platino (PGE), donde cerca del 90 % de la base de datos geoquímica carece de información sobre Pd y Pt, debido a limitaciones operativas en campañas históricas y la imposibilidad de reanalizar núcleos deteriorados. Frente a esta situación, se propone una metodología basada en aprendizaje automático multivariado para estimar los contenidos de Pd y Pt a partir de datos existentes de Ni, Cu y Co, sin necesidad de realizar nuevos análisis de laboratorio.

Como primer paso, se desarrollaron modelos de regresión con el fin de evaluar el poder predictivo de las variables Ni, Cu y Co sobre los contenidos de Pd y Pt. Tras validar su solidez mediante métricas como MAE y R^2 , se avanzó hacia un modelo de clasificación multiclase, más robusto para la toma de decisiones operativas. Este modelo permitió categorizar las leyes de Pd y Pt en clases baja, media y alta, y fue evaluado con métricas ajustadas por desbalance como precisión y Kappa de Cohen. La consistencia de los resultados reflejó patrones coherentes con el contexto geológico del depósito.

Los modelos fueron aplicados sobre sondeos históricos y sobre el modelo de bloques geoestadístico, permitiendo la generación de inferencias en zonas sin datos primarios. Esto facilitó la identificación de dos nuevas zonas con alto potencial de exploración con alta ley de Pd + Pt, y representó un ahorro estimado de 19,710 USD en análisis de laboratorio, al evitar el reensayo de núcleos deteriorados. La metodología demuestra ser replicable, escalable y alineada con principios de eficiencia y sostenibilidad en exploración minera, ofreciendo un marco aplicable a otros elementos críticos y a proyectos con baja densidad de muestreo.

1. Introducción

El yacimiento en estudio corresponde a un sistema de sulfuros magmáticos de níquel (Ni), cobre (Cu) y elementos del grupo del platino (PGE), emplazado en la Subprovincia Abitibi, al margen sur del Cratón Superior (Jackson & Fyon, 1992). La secuencia estratigráfica está dominada por unidades del Ensamble Kidd-Munro, típicamente asociado a flujos komatíticos vinculados a eventos magmáticos generados por plumas del manto. La litología del depósito está compuesta mayoritariamente por rocas volcánicas ultramáficas, como komatitas, dunitas, peridotitas y piroxenita, e incluye también andesitas y riolitas de composición intermedia a félsica (Scott Jobin-Bevans et al., 2024). La mineralización se presenta como capas irregulares de sulfuros masivos, transicionando hacia cuerpos con textura reticular (net-textured) y sulfuros diseminados, mientras que en algunos sectores se observan vetillas mineralizadas cortando la roca caja. Los principales minerales sulfurados observados, en orden decreciente de abundancia visual, son pirrotita, pentlandita, millerita y calcopirita, responsables del contenido metálico de interés (Scott Jobin-Bevans et al., 2025).

Las campañas geoquímicas desarrolladas se fundamentaron mayoritariamente en el análisis de núcleos provenientes de perforaciones históricas, que superan ampliamente en número (aprox. 20 a 1) a los datos generados por perforaciones recientes. Factores como la optimización del presupuesto, limitaciones operativas y el contexto del precio de los metales durante la ejecución de los programas de exploración, derivaron en que los elementos del grupo del platino (Pd y Pt) no fueran incluidos de forma sistemática en los análisis. Como resultado, cerca del 90 % de la base de datos, principalmente aquella asociada a perforaciones históricas carece de información sobre estos metales. A ello se suma la imposibilidad

de reanalizar dichos núcleos debido a la pérdida o deterioro de la mayoría de las cajas. Ante esta situación, se plantea la necesidad de aplicar metodologías de prospección más ágiles, costo-eficientes y apoyadas en el aprovechamiento estratégico de datos heredados.

En este contexto, los datos geoquímicos se integran con algoritmos de inteligencia artificial para desarrollar una solución incipiente, orientada a modernizar de forma inferencial la información heredada sin necesidad de realizar nuevos análisis. Tal como lo plantean Bourdeau et al. (2023), esta metodología se basa en modelos de aprendizaje automático entrenados con datos combinados de campañas de perforación recientes e históricas. La cual surge como respuesta a uno de los desafíos aún vigentes en la exploración: la generación de datos continúa dependiendo en gran medida de prácticas operativamente costosas y poco ágiles.

En el presente trabajo, se implementó un enfoque inferencial dentro del flujo de trabajo exploratorio, orientado a predecir los contenidos de Pd y Pt a partir de los datos geoquímicos existentes de Ni, Cu y Co. Esta inferencia fue complementada mediante técnicas geoestadísticas, evaluando el comportamiento espacial de Pd y Pt en el modelo de bloques, con el objetivo de identificar zonas con alto potencial de valor agregado para el proyecto. A diferencia de los métodos tradicionales, que dependen de la generación de datos primarios a través de nuevas campañas, esta metodología permite orientar la toma de decisiones estratégicas, como la selección de áreas puntuales para perforar o que núcleos de sondajes reanalizar, mediante la creación de datos derivados a partir de la información ya disponible.

2. Objetivos

- Presentar una metodología que utilice algoritmos de aprendizaje automático multivariado para estimar contenidos de Pd y Pt a partir de datos geoquímicos sencillos (Ni, Cu y Co).
- Cuantificar el impacto económico y operacional mediante el ahorro en análisis de laboratorio y la mejora en la selección de muestras.
- Demostrar el valor añadido de combinar técnicas de ciencia de datos con los flujos de trabajo geoestadísticos para la identificación de nuevas zonas mineralizadas.
- Establecer un marco metodológico escalable y extensible para su aplicación a

otros elementos críticos y a proyectos con baja densidad de muestreo.

3. Compilación de Datos y Desarrollo del Trabajo

Se utilizó la totalidad de los datos geoquímicos disponibles con análisis de paladio (Pd) y platino (Pt), provenientes principalmente de la campaña de perforación de 2021, seguida —en orden decreciente de cantidad de muestras— por las campañas de 2007, 2005 y 2004 (ver Tabla 1). Estos datos conformaron el conjunto de entrenamiento, validación y prueba para el desarrollo del modelo de predicción.

Adicionalmente, se empleó una base de datos histórica que, si bien carece en su mayoría de análisis para Pd y Pt, contiene información geoquímica relevante de níquel (Ni), cobre (Cu) y cobalto (Co). Esta base fue utilizada como insumo para que el modelo realice las predicciones de Pd y Pt en aquellos intervalos donde dichos elementos no fueron muestreados.

Tabla 1. Cantidad de análisis geoquímicos disponibles para Pd y Pt por año de campaña de perforación.

Campañas de perforacion						
Año	Total de Sondajes	Total de Muestras	Muestras con Data de Pd		Muestras con Data de Pt	
			Cantidad	%	Cantidad	%
1955 to 1974	97	1969	12	0.6%	12	0.6%
1989 to 1999	11	623	31	4.1%	31	4.1%
2001	9	149	-	-	-	-
2004	63	3487	114	2.6%	114	2.6%
2005	95	3954	301	6.0%	301	6.0%
2007	5	60	56	92.7%	60	100.0%
2021	19	509	508	99.7%	508	99.7%
Total	299	10751	1022	9.5%	1026	9.5%

El desarrollo del trabajo se estructuró en tres fases:

- **Fase 1:** Desarrollo del modelo de regresión
- **Fase 2:** Desarrollo del modelo de clasificación
- **Fase 3:** Aplicación del modelo a datos históricos y al modelo de bloques

3.1. Desarrollo del modelo de regresión

Esta primera fase tuvo como objetivo evaluar la capacidad predictiva de Ni, Cu y Co sobre Pd y Pt a través de modelos de regresión multivariada. Se utilizaron métricas de desempeño como el error absoluto medio (MAE) y el coeficiente de determinación (R²) para validar el modelo.

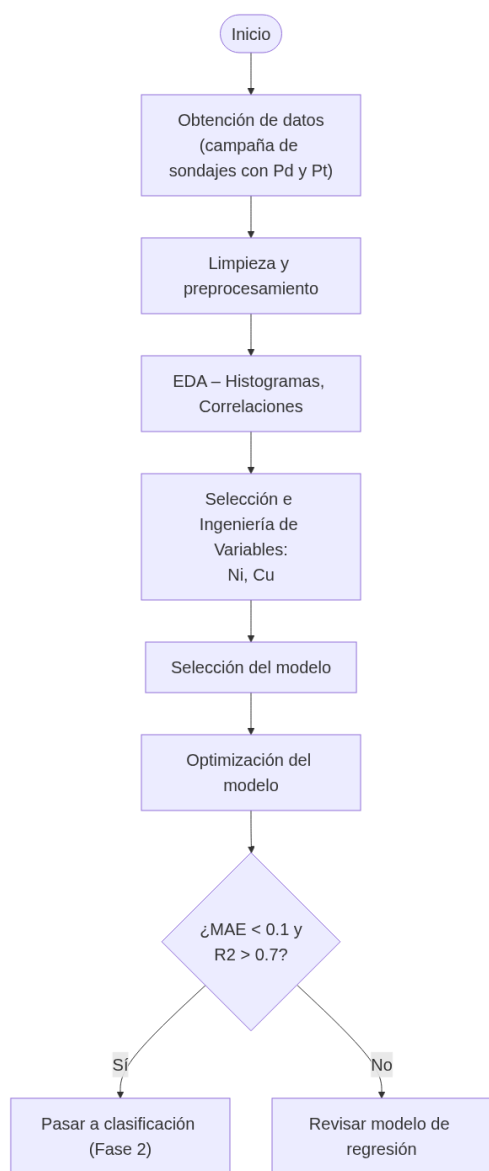


Figura 1. Diagrama de flujo de la Fase 1: desarrollo del modelo de regresión.

3.1.1. Obtención de datos

Se utilizaron alrededor de 1,000 registros analíticos que incluían análisis geoquímicos multielementales de Ni, Cu, Co, Pd y Pt.

3.1.2. Limpieza y preprocesamiento

Se eliminaron valores nulos, negativos y posibles errores de medición.

3.1.3. Análisis exploratorio de datos (EDA)

Se aplicaron técnicas estadísticas y visuales para entender la distribución de las variables, identificar patrones y evaluar relaciones entre predictores y variables objetivo. Las correlaciones de Pearson mostraron asociaciones lineales moderadas entre Ni, Cu y Pd/Pt, mientras que el coeficiente de

Spearman reveló correlaciones no lineales más fuertes, indicando la pertinencia de utilizar modelos capaces de capturar relaciones complejas y no lineales entre las variables.

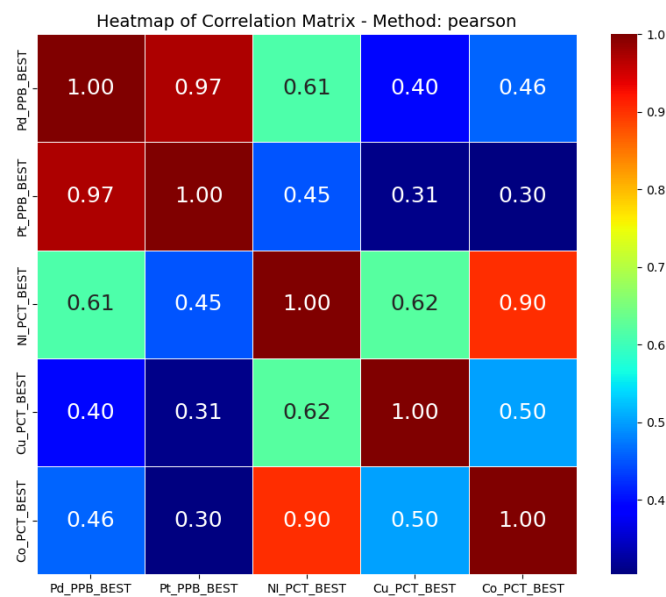


Figura 2. Matriz de correlación de Pearson entre Ni, Cu, Co, Pd y Pt.

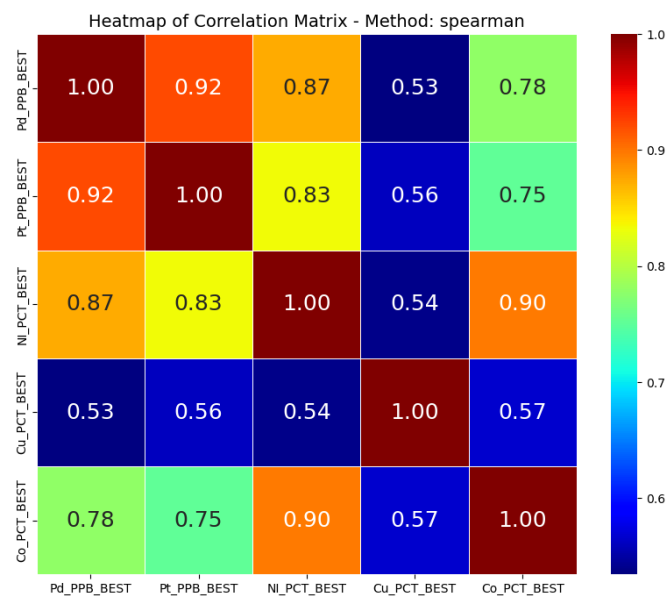


Figura 3. Matriz de correlación de Spearman entre Ni, Cu, Co, Pd y Pt.

3.1.4. Selección e Ingeniería de Variables

Se seleccionaron las variables Ni y Cu como predictores principales, descartando Co debido a su alta colinealidad con Ni, según lo evidenciado en el análisis de multicolinealidad. Esta decisión se fundamentó en el principio de parsimonia, buscando reducir la redundancia sin afectar la capacidad explicativa del modelo. Adicionalmente, las variables objetivo fueron transformadas a unidades de ppm (partes por millón) para garantizar

la coherencia dimensional y facilitar la interpretación de los resultados.

3.1.5. Selección del modelo

En esta etapa se procedió al entrenamiento de diversos algoritmos de regresión, con el objetivo de identificar el modelo con mejor desempeño predictivo para los valores de ley de Pd y Pt. Inicialmente, el conjunto de datos fue dividido en dos subconjuntos: 80% para entrenamiento y 20% para testeo. Para obtener métricas más representativas y reducir la varianza asociada a una sola partición, se implementó un esquema de validación cruzada k-fold.

Se evaluaron diversos algoritmos de regresión no lineales, entre ellos:

- Árboles de decisión de regresión (Decision Tree Regressor)
- Vecinos más cercanos de regresión (K-Nearest Neighbors Regressor)
- Bosque aleatorio de regresión (Random Forest Regressor)
- XGBoost Regressor
- Máquinas de vectores de soporte de regresión (Support Vector Regressor)

Cada modelo fue entrenado y evaluado utilizando como métricas principales el Error Absoluto Medio (MAE) y el coeficiente de determinación (R^2). Los resultados demostraron que el modelo Random Forest Regressor superó al resto en términos de precisión, alcanzando los valores de MAE más bajos tanto para el Pd como para el Pt. Este rendimiento, sumado a su estabilidad y bajo riesgo de sobreajuste, motivó su elección para las etapas posteriores de refinamiento.

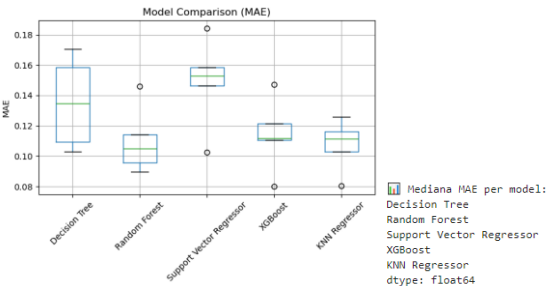


Figura 4. Diagrama de caja del MAE por modelo para Pd y resumen de medianas.

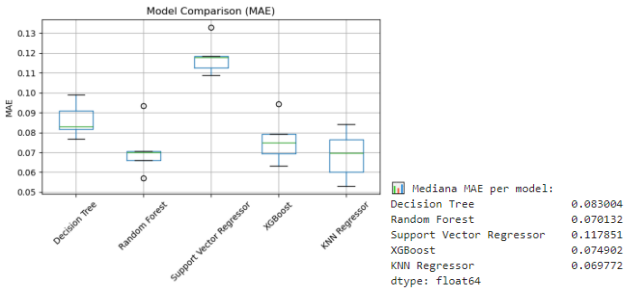


Figura 5. Diagrama de caja del MAE por modelo para Pt y resumen de medianas.

3.1.6. Optimización del modelo

Se procedió al ajuste de hiperparámetros del modelo Random Forest utilizando búsqueda en malla (grid search). Los modelos optimizados alcanzaron métricas satisfactorias, como se presenta a continuación:

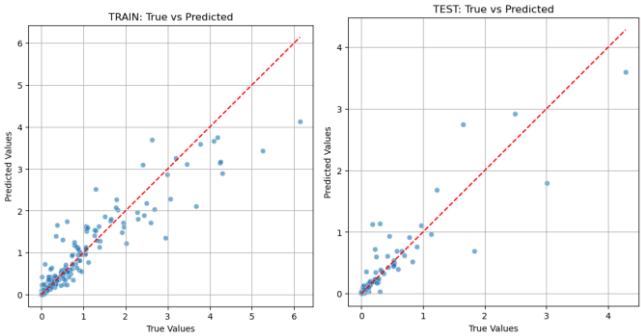


Figura 6. Diagrama de dispersión de valores reales versus predichos para Pd (entrenamiento y testeo).

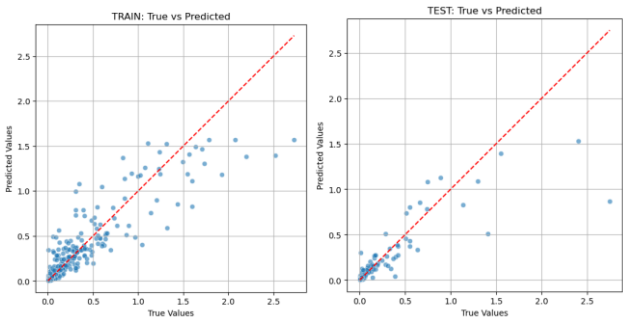


Figura 7. Diagrama de dispersión de valores reales versus predichos para Pt (entrenamiento y testeo).

Tabla 2. Resumen de métricas de desempeño de los modelos optimizados de regresión para el Pd y Pt.

Metal	Entrenamiento		Testeo	
	MAE (ppm)	R2	MAE (ppm)	R2
Pd	0.0658	0.8945	0.065	0.8289
Pt	0.0477	0.8353	0.0541	0.7326

Estos resultados confirmaron la capacidad

predictiva de Ni y Cu sobre Pd y Pt, lo cual habilitó el avance hacia el desarrollo de un modelo de clasificación en la siguiente fase.

3.2. Desarrollo del modelo de clasificación

En esta fase se desarrolló un modelo de clasificación supervisada, considerando que predecir clases de ley puede ser más robusto operativamente que estimar valores exactos.

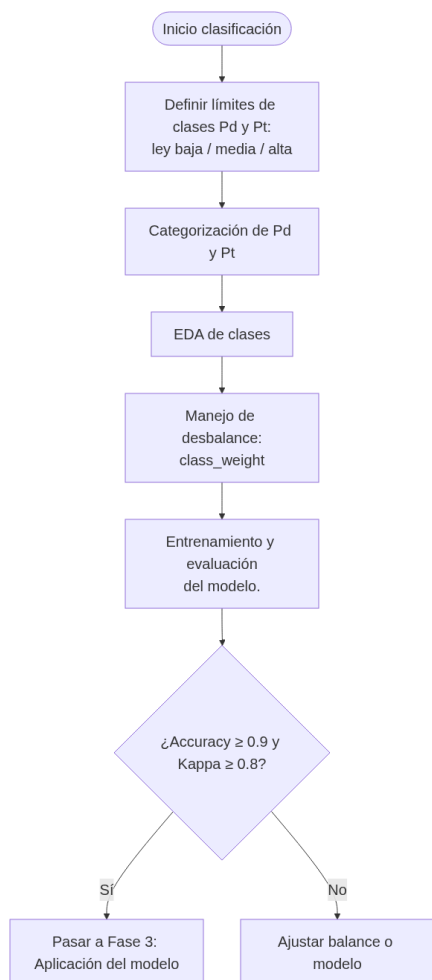


Figura 8. Diagrama de flujo de la Fase 2: desarrollo del modelo de clasificación.

3.2.1. Definición de límites de clases de Pd y Pt

Se establecieron umbrales operativos de ley para Pd y Pt en función de los percentiles y condiciones del yacimiento:

- Baja ley (Low grade): 0 – 0.1 g/t
- Media ley (Medium grade): 0.1 – 0.5 g/t
- Alta ley (High grade): > 0.5 g/t

3.2.2. Categorización de las Variables Objetivo

Se creó una nueva variable categórica para cada metal (Pd y Pt), asignando la clase correspondiente

según el rango de ley. Esto transformó el problema de regresión en uno de clasificación multiclase.

3.2.3. Análisis exploratorio de clases

Se evidenció un desbalance significativo en la distribución de clases, con predominancia de muestras de baja ley.

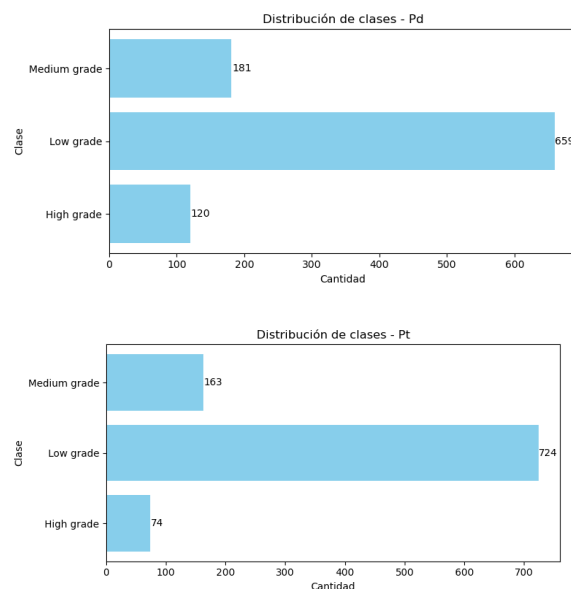


Figura 9. Distribución de clases para Pd y Pt en el conjunto de datos.

3.2.4. Manejo del desbalance de clases

Para evitar la generación artificial de datos y mantener la integridad geológica, se optó por un enfoque de ponderación por clase (class weighting) durante el entrenamiento. Dado el desbalance, se empleó Cohen's Kappa como métrica robusta de desempeño complementaria a la exactitud (accuracy).

3.2.5. Entrenamiento y Evaluación del Modelo

Se utilizó nuevamente el algoritmo Random Forest, ajustado para clasificación multiclase con ponderación. Las métricas de desempeño en entrenamiento y testeo se muestran a continuación:

TRAIN CLASSIFICATION REPORT:				
	precision	recall	f1-score	support
High grade	0.93	0.94	0.93	96
Low grade	1.00	0.96	0.98	527
Medium grade	0.85	0.95	0.90	145
accuracy			0.96	768
macro avg	0.93	0.95	0.94	768
weighted avg	0.96	0.96	0.96	768

TEST CLASSIFICATION REPORT:				
	precision	recall	f1-score	support
High grade	0.79	0.92	0.85	24
Low grade	0.98	0.96	0.97	132
Medium grade	0.82	0.78	0.80	36
accuracy			0.92	192
macro avg	0.86	0.89	0.87	192
weighted avg	0.92	0.92	0.92	192

Figura 10. Reporte de clasificación para Pd (entrenamiento y testeo).

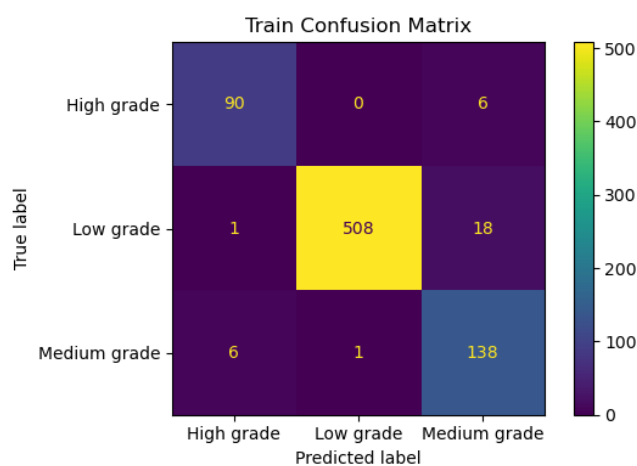


Figura 11. Matriz de confusión en la etapa de entrenamiento para el Pd.

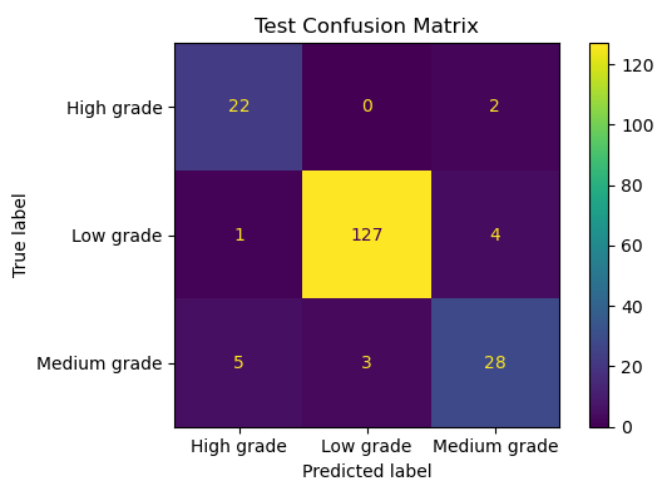


Figura 12. Matriz de confusión en la etapa de testeo para el Pd.

TRAIN CLASSIFICATION REPORT:				
	precision	recall	f1-score	support
High grade	0.81	0.97	0.88	59
Low grade	1.00	0.96	0.98	579
Medium grade	0.83	0.92	0.87	130
accuracy			0.95	768
macro avg	0.88	0.95	0.91	768
weighted avg	0.96	0.95	0.95	768

TEST CLASSIFICATION REPORT:				
	precision	recall	f1-score	support
High grade	0.71	0.80	0.75	15
Low grade	0.99	0.96	0.97	145
Medium grade	0.74	0.79	0.76	33
accuracy			0.92	193
macro avg	0.81	0.85	0.83	193
weighted avg	0.92	0.92	0.92	193

Figura 13. Reporte de clasificación para Pt (entrenamiento y testeo).

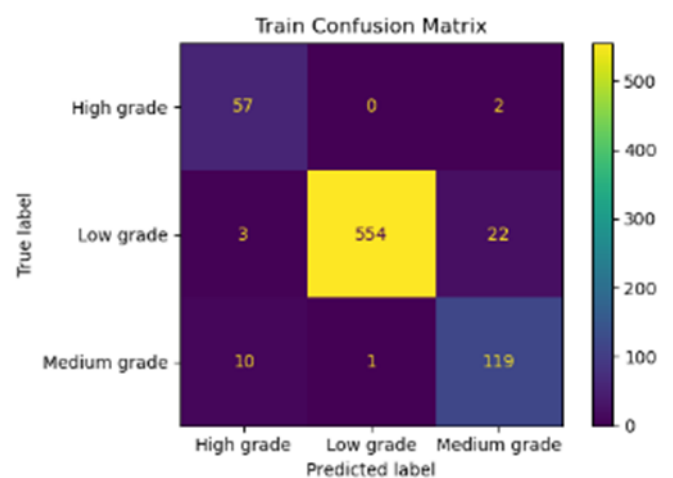


Figura 14. Matriz de confusión en la etapa de entrenamiento para el Pt.

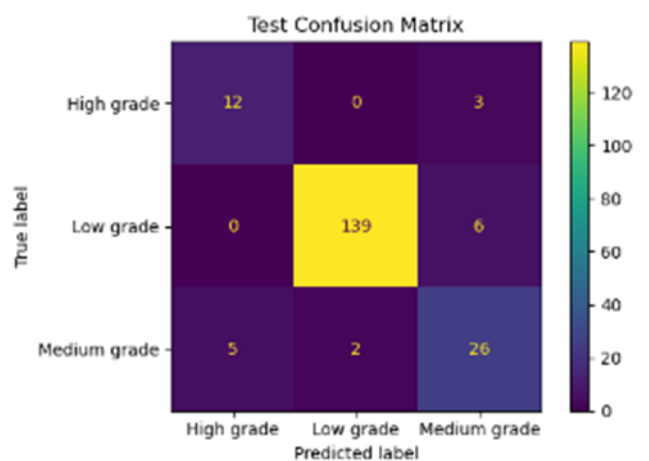


Figura 15. Matriz de confusión en la etapa de entrenamiento para el Pt.

Tabla 3. Resumen de métricas de desempeño de los modelos de clasificación para el Pd y Pt.

Metal	Entrenamiento	Testeo	Cohen's Kappa
	Exactitud	Exactitud	
Pd	0.96	0.92	0.84
Pt	0.95	0.92	0.80

3.3. Fase 3: Aplicación del Modelo

3.3.1. Predicción en sondajes históricos

Se aplicó el modelo de clasificación sobre los registros históricos de perforación que no contaban con análisis de Pd y Pt. Los resultados fueron exportados en formato .csv e integrados en el software Leapfrog Geo para su visualización y análisis espacial.

3.3.2. Predicción en modelo de bloques y propuesta de targets de exploración

El modelo entrenado se aplicó al modelo de bloques geoestadístico existente, el cual contenía valores interpolados de Ni y Cu. A partir de los resultados de predicción, se generaron nuevas columnas para Pd y Pt, que luego fueron integradas en una única variable compuesta (Pd + Pt) con el fin de representar el contenido conjunto de elementos del grupo del platino (PGE).

Para esta nueva variable, se redefinieron los rangos de ley sumando los umbrales previamente establecidos para cada metal, quedando clasificados de la siguiente manera:

- Baja ley (Low grade): 0 – 0.2 g/t
- Media ley (Medium grade): 0.2 – 1 g/t
- Alta ley (High grade): > 1 g/t

Se aplicó un filtro para seleccionar los bloques clasificados como de alta ley en la variable Pd + Pt. Las zonas resultantes fueron interpretadas como potenciales targets de exploración para PGE y propuestas como áreas prioritarias para una futura campaña de perforación.

Los resultados obtenidos a partir de la aplicación del modelo tanto en sondajes históricos como en el modelo de bloques geoestadístico se presentan y analizan en detalle en la Sección 4. Adicionalmente, se incluyen imágenes complementarias en la sección de Anexos para respaldar visualmente los hallazgos obtenidos.

4. Presentación y discusión de resultados

Los modelos desarrollados fueron aplicados tanto a datos históricos de sondajes como al modelo de bloques estimado para los elementos de Ni, Cu y

Co del yacimiento. Esta sección presenta los resultados obtenidos, su validación geológica, y una discusión de sus implicancias técnicas y operativas.

4.1. Impacto práctico y validación técnica

Los modelos lograron altos niveles de precisión y consistencia:

- El modelo de clasificación alcanzó un 92% de accuracy tanto para el Pd como el Pt, con valores de Cohen's Kappa superiores a 0.80, lo que indica un acuerdo sustancial incluso en presencia de desbalance de clases.
- El modelo de regresión demostró buen ajuste con R^2 de 0.82 para Pd y 0.73 para Pt, y bajos valores de MAE.
- La validación visual de los resultados en Leapfrog mostró coherencia geoespacial con unidades mineralizadas conocidas, reforzando la confiabilidad técnica del enfoque.

4.2. Resultados de la Aplicación del Modelo a sondajes históricos

Se aplicó el modelo de clasificación a un total de 9729 muestras históricas que no contaban con análisis de Pd ni Pt. Los resultados muestran que más del 95% de las muestras históricas corresponden a zonas de baja ley con valores menores a 0.2 ppm de Pd+Pt, un 4.5% corresponden a zonas de media ley con valores de entre 0.2 a 1 ppm de Pd+Pt y solo un 0.5% de las muestras se encuentran en la zona de mayor interés con leyes mayores a 1 ppm de Pd+Pt (Tabla 4).

Tabla 4. Distribución de clases predichas en los sondajes históricos del grupo Pd + Pt.

Clase	Muestras históricas	
	Cantidad	Porcentaje (%)
High grade	30	0.31
Low grade	9275	95.19
Medium grade	439	4.51
Total	9744	

La distribución espacial resultante del modelo de clasificación (Figuras 16 y 17) revela una relación coherente entre las categorías de Pd y Pt y las leyes de Ni, donde los intervalos identificados como media a alta ley de PGE coinciden con valores

elevados de Ni (>1%), mientras que los clasificados como baja ley se asocian con concentraciones más moderadas, entre 0.01% y 0.8%, reafirmando la consistencia geoquímica del sistema.

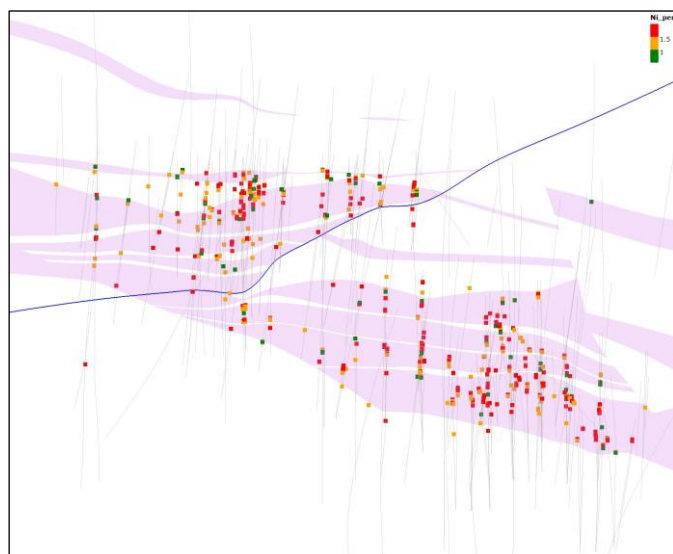


Figura 16. Distribución de leyes de Ni en tramos sin análisis de Pd-Pt, pero clasificados como mineralización media a alta en PGE según el modelo.

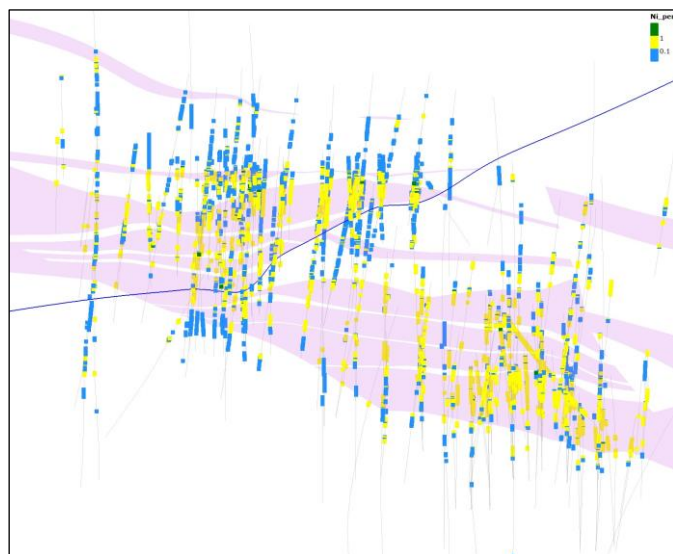


Figure 17. Distribución de leyes de Ni en tramos sin análisis de Pd-Pt, pero clasificados como mineralización baja en PGE según el modelo.

Este patrón se sustenta en la asociación del contenido Pd y Pt en lentes de sulfuros masivos, dominados por pirrotita y pentlandita, y en menor proporción calcopirita. Sin embargo, aunque la presencia de valores altos de PGE suele coincidir con altos contenidos de Ni, la relación no es bidireccional, es decir no toda zona con Ni >1% muestra contenido elevados de Pd-Pt, como se evidencia en la Figura 18.

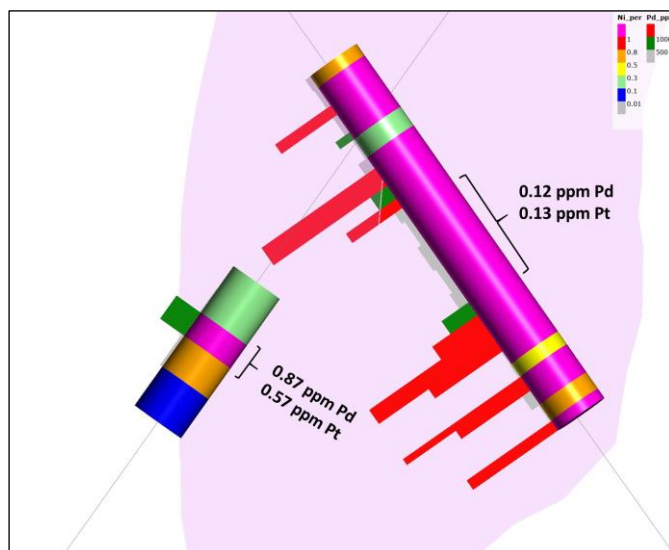


Figura 18. En la imagen a la izquierda, sondaje de 2021; a la derecha, de 2004. Ambos exhiben el contenido de Ni como cilindros y los valores de PGE en barras superpuestas.

Allí radica el valor del modelo de clasificación: permite priorizar estratégicamente el re-muestreo de núcleos con alta probabilidad de contener altos valores de Pd-Pt, superando decisiones basadas solo en el contenido del Ni o presencia visual de sulfuros masivos.

En un escenario hipotético donde fuese posible reanalizar el 100 % de los núcleos históricos preservados, se habrían identificado visualmente un total de 687 muestras potencialmente relevantes, basándose en la presencia de sulfuros masivos y altos contenidos de Ni. Asumiendo un costo conservador de 30 USD por muestra para el análisis de Pd + Pt (ver Tabla 5), el gasto total estimado ascendería a 20,610 USD. Sin embargo, mediante la aplicación del modelo predictivo propuesto, solo 30 muestras fueron clasificadas como de alta ley en Pd + Pt, lo que reduce significativamente el número de muestras a enviar a laboratorio. Esto representa un ahorro operativo aproximado de 19,710 USD, al evitar el análisis innecesario de 657 muestras sin potencial económico relevante.

Tabla 5. Benchmarking de precios para análisis de Pd y Pt según laboratorio y metodología aplicada.

Laboratorio	Análisis	Precio USD
Hazen Research	Pt o Pd (Fire Assay + ICP-OES)	50
ALS Global	Pt + Pd + Au (30 g Fire Assay)	28.55
Intertek Minerals	Pt + Pd	338
Ontario Geo Labs (IMP-101)	Pt + Pd (Lead Fire Assay + ICP-MS)	28.89
Ontario Geo Labs (IMP-200)	Pt + Pd (NiS Fire Assay + ICP-MS)	216.77

4.3. Resultados de la Aplicación al Modelo de Bloques Geoestadístico

Se utilizó el modelo de clasificación previamente entrenado para inferir la categoría de ley de Pd y Pt en el modelo de bloques 3D del yacimiento. Los predictores utilizados fueron los valores krigeados de Ni y Cu, ya disponibles en el modelo de bloques estimado.

Posteriormente, se aplicó un filtro para seleccionar bloques clasificados simultáneamente como alta ley tanto en Pd como en Pt. El resultado fue una delimitación de zonas de interés con potencial mineral no explorado.

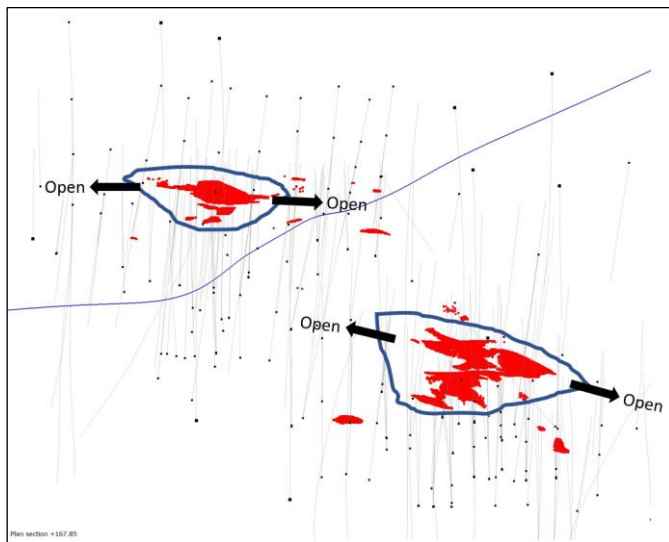


Figura 19. Vista en planta que evidencia dos zonas con alto potencial en PGE, delimitadas por bloques clasificados como alta ley y abiertas a continuidad a lo largo del rumbo de la estructura principal.

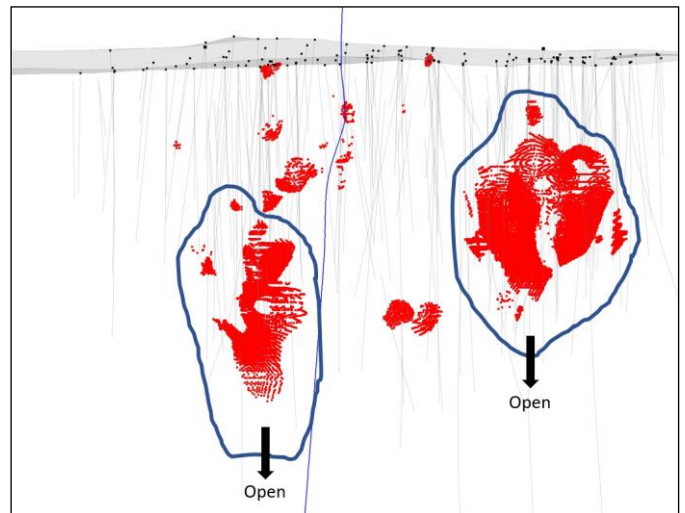


Figura 20. Sección longitudinal mirando al norte que evidencia dos zonas con alto potencial en PGE, delimitadas por bloques clasificados como alta ley y abiertas a continuidad en profundidad.

Las zonas identificadas mediante el modelo coinciden espacialmente con regiones donde se ha evidenciado mineralización asociada a sulfuros masivos, lo cual refuerza la validez geológica del enfoque. Además, el modelo permitió detectar nuevos sectores en profundidad que no habían sido previamente priorizados por el modelo geoestadístico original.

Esta capacidad de extender el conocimiento del yacimiento más allá del área perforada convierte al modelo en una herramienta de targeting exploratorio, capaz de proponer campañas de perforación más focalizadas, con base en inferencia matemática y validación geológica.

4.4. Desempeño estadístico del modelo: ¿por qué funcionó?

El modelo de regresión logró niveles de ajuste notables ($MAE = 0.065$ y $R^2 = 0.83$ para Pd y $MAE = 0.054$ y $R^2 = 0.73$ para Pt), lo que evidencia su alta capacidad para capturar relaciones complejas entre los contenidos de Ni, Cu, Co y los PGEs, relaciones que son típicamente no lineales y multivariadas, y que fueron modeladas eficazmente mediante algoritmos no paramétricos como Random Forest y XGBoost.

Este buen desempeño se sustenta en un rasgo distintivo de los sistemas magmáticos de sulfuros Ni-Cu-PGE: la persistencia de una mineralogía portadora de PGE similar en dominios tanto de alta como de baja ley. Estudios como el de Hanley (2007) confirman que esta homogeneidad mineralógica se debe por el hecho de que ambos grupos de muestras comparten un mismo origen

geológico y una evolución genética común, lo que respalda la validez del modelo predictivo incluso en contextos de mineralización menos intensa.

Además, Hanley (2007) destaca que minerales como michenerita (PdBiTe), sperrylite (PtAs_2) y pentlandita ($(\text{Fe,Ni})_9\text{S}_8$), esta última con Pd incorporado en solución sólida sustituyendo parcialmente al Ni o Fe, se encuentran tanto en mineralización de alta como de baja ley. Esta observación sugiere que la presencia de Pd y Pt no está necesariamente restringida a zonas de alta ley, sino que puede distribuirse en diferentes niveles de enriquecimiento, lo que amplía el espectro de interés exploratorio dentro del sistema.

4.5. Limitaciones y Consideraciones

A pesar de los resultados promisorios, es importante reconocer algunas limitaciones:

- El modelo fue entrenado sobre una única campaña moderna, lo que puede restringir su capacidad de generalización en otras unidades litológicas o zonas estructuralmente distintas.
- La calidad de las predicciones depende de la precisión de los predictores (Ni y Cu), cuya interpolación en el modelo de bloques también introduce incertidumbre.
- El desbalance natural de clases, característico en la distribución de PGEs, limita la sensibilidad del modelo ante casos raros, aunque se ha mitigado con técnicas de ponderación.

5. Conclusiones

Este trabajo presentó una metodología basada en algoritmos de aprendizaje automático multivariado para estimar los contenidos de Pd y Pt a partir de datos geoquímicos simples como Ni, Cu y Co, comúnmente disponibles en campañas de exploración temprana. Los resultados obtenidos validan la solidez del enfoque, tanto en términos de precisión predictiva como de aplicabilidad práctica.

En la etapa de regresión, los modelos entrenados (Random Forest Regressor) alcanzaron desempeños satisfactorios, con valores de MAE de 0.065 para Pd y 0.054 para Pt en la fase de testeo, y coeficientes de determinación (R^2) de 0.83 y 0.73, respectivamente. Estos resultados respaldan la existencia de relaciones no lineales aprovechables entre los metales base y los elementos del grupo del platino. Esta constatación resulta fundamental para los procesos de exploración y estimación de recursos, ya que pone en evidencia la ventaja de

haber aplicado la metodología propuesta. A diferencia de los enfoques tradicionales, que suelen apoyarse en modelos lineales y simplificados, esta estrategia permitió capturar y explicar de forma coherente el comportamiento de las variables Pd y Pt en función de su relación con Ni, Cu y Co. Tal comportamiento refleja el control ejercido por el evento principal de mineralización, el cual no habría sido posible describir adecuadamente mediante técnicas estadísticas convencionales.

Al reformular el problema como clasificación multiclase, los modelos lograron exactitudes superiores al 90 % y valores de Kappa de Cohen de 0.84 para Pd y 0.80 para Pt en los datos de testeo, evidenciando un alto grado de acuerdo ajustado por el desbalance de clases. Esto reafirma la robustez metodológica de la combinación entre regresión y clasificación supervisada.

Desde una perspectiva operativa, la implementación de esta metodología permite reducir significativamente los costos asociados al análisis de laboratorio. En particular, se evitó enviar a ensayo un número considerable de muestras históricas que carecían de información sobre Pd y Pt, generando un ahorro estimado de 19,710 USD.

Asimismo, la integración del modelo de aprendizaje automático con el modelo de bloques geoestadístico permitió identificar dos nuevas zonas con potencial de mineralización de PGE, a partir de la clasificación combinada Pd + Pt. Estas zonas fueron propuestas como targets de exploración para futuras campañas de perforación, contribuyendo directamente a la eficiencia de los flujos de trabajo en exploración avanzada.

Finalmente, la metodología demostrada presenta un alto potencial de escalabilidad. Su diseño modular y dependiente de variables geoquímicas de fácil acceso la convierte en una herramienta transferible a otros metales de interés económico, como el oro (Au), y a contextos geológicos con baja densidad de muestreo. Esto refuerza su valor como contribución práctica y sostenible tanto para el sector minero como para el desarrollo de nuevas técnicas en exploración basada en datos.

6. Anexos

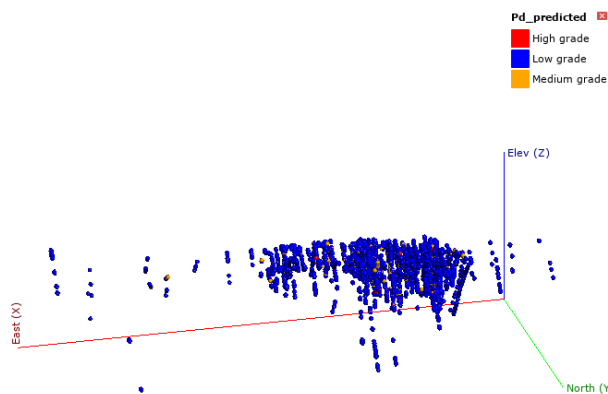


Figura 21. Vista en 3D de la clasificación inferida de leyes de Pd en sondajes históricos.

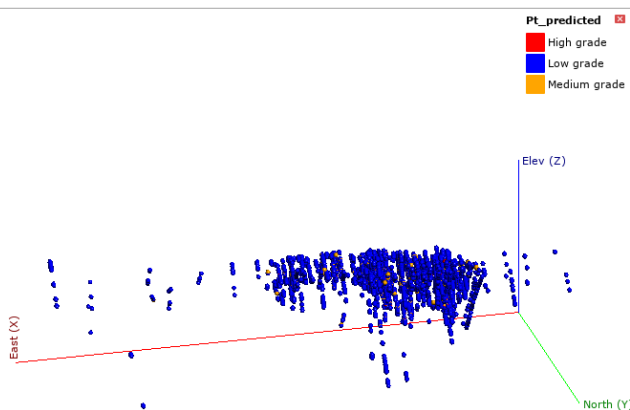


Figura 22. Vista en 3D de la clasificación inferida de leyes de Pt en sondajes históricos.

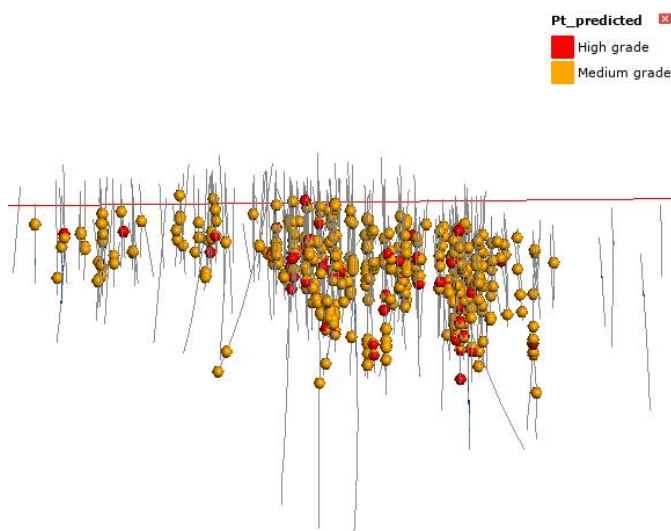


Figura 23. Vista en 3D de las clases predichas como alta y media ley de Pt en sondajes históricos.

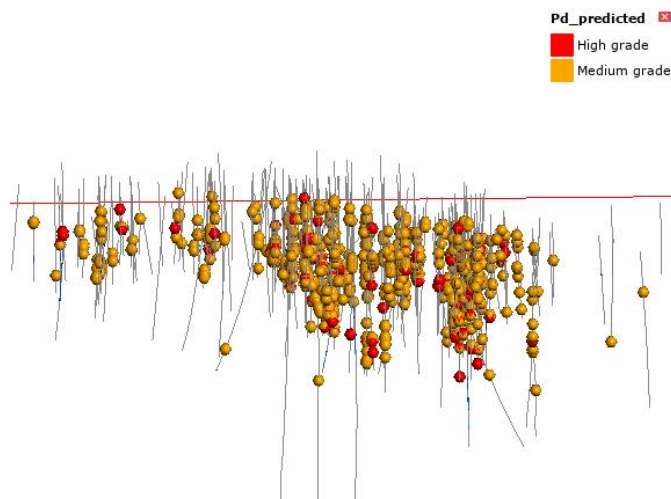


Figura 24. Vista en 3D de las clases predichas como alta y media ley de Pd en sondajes históricos.

7. Referencias bibliográficas

ALS Global. (2025). *Geochemistry Fee Schedule – USD*. <https://www.alsglobal.com/-/media/ALSGlobal/Resources-Grid/Fee-Schedule-Geochemistry/2025/ALS-Geochemistry-Fee-Schedule-USD-2025.pdf>

Bourdeau, J. E., Zhang, S. E., Lawley, C. J. M., Parsa, M., Nwaila, G. T., & Ghorbani, Y. (2023). *Predictive Geochemical Exploration: Inferential Generation of Modern Geochemical Data, Anomaly Detection and Application to Northern Manitoba*. *Natural Resources Research*, v. 32(6), p. 2355–2386. <https://doi.org/10.1007/s11053-023-10273-6>

Hanley, J. J. (2007). *The role of arsenic-rich melts and mineral phases in the development of high-grade Pt-Pd mineralization within komatiite-associated magmatic Ni-Cu sulfide horizons at Dundonald Beach South, Abitibi subprovince, Ontario, Canada*. *Economic Geology*, v. 102, p. 305–317. <https://doi.org/10.2113/gsecongeo.102.2.305>

Hazen Research, Inc. (2024). *Assaying and Laboratory Services*. https://www.hazenresearch.com/sites/default/files/hazen_fee_schedule.pdf

Intertek Group plc. (2024). *Analytical Services Pricing – Minerals Division*. <https://www.intertek.com/siteassets/minerals/intertek-lsi-2024-analytical-fees-usd.pdf>

Jackson, S.L. and Fyon, J.A. (1992) *The Western Abitibi Subprovince in Ontario*. In *Thurston, P. C., et*

al. (eds.). Geology of Ontario, Ontario Geological Survey Special Volume 4, Part 1, pp. 405–483

Scott Jobin-Bevans, Simon J.A. Mortimer, & John Siriunas (2024). *Technical Report and Mineral Resource Estimates Alexo-Dundonald Nickel Sulphide Project*. Nickel and Technologies, 71-80. https://class1nickel.com/class_1/wp-content/uploads/2024/07/Caracle_Class1-Alexo-North-TR_MRE-July8_2024-F1.pdf

Scott Jobin-Bevans, Simon J.A. Mortimer, & John Siriunas. (2025). *Technical Report and Mineral Resource Estimates for the Alexo-Dundonald Nickel Sulphide Project: Including Updated Dundonald North Mineral Resource Estimate*. Nickel and Technologies. <https://newsfile.futunn.com/public/NN-PersistNoticeAttachment/7781/20250513/SEDAR PLUS/CSA SEDAR PLUS NOTICE RECORD ID 2111477.pdf>

Ontario Geological Survey. (2024). *Geo Labs Brochure – Services and Analytical Fees*. https://www.geologyontario.mndm.gov.on.ca/mines/ogs/geo_labs/mines-2024-geo-labs-brochure-en-2024-04-16.pdf

8. Reseña Profesional

Neil Harrison Berrospi Rios es egresado de la Universidad Nacional de Ingeniería (UNI) de la

carrera de Ingeniería Geológica y especialista en Inteligencia Artificial por la Pontificia Universidad Católica de Valparaíso. Ha trabajado en el área de Inteligencia Artificial en Consorcio Minero Horizonte (CMH) por más de 2 años haciendo mapas prospectivos aplicando Inteligencia Artificial, actualmente es científico de datos en el área de Innovación de Nexa Resources.

Daniel Abinadab Basilio Rojas es egresado de la Universidad Nacional de Ingeniería (UNI). Cuenta con más de cinco años de experiencia en Atticus Perú SAC, desempeñándose en las áreas de base de datos, modelamiento geológico y estimación de recursos. Durante este tiempo, ha participado en el desarrollo de modelos geológicos integrados en diversos tipos de depósitos, incluidos sistemas epitermales, vetiformes, magmáticos, VMS, de elementos de tierras raras (REE), entre otros, brindando soporte técnico a proyectos internacionales.

Carlos Josue Mendoza Kuong es egresado de la Universidad Nacional de Ingeniería de la carrera de Ingeniería de minas, con estudios de análisis de datos y experiencia en la validación de data para reportabilidad de sistemas de despacho, experiencia ganada en Antamina y Beijing Shougang.

AUTORIZACIÓN DE PARTICIPACIÓN

Yo, Simon James Atticus Mortimer con cargo de Gerente General de la empresa Atticus Peru SAC; autorizo que el trabajo titulado "Predicción de contenido de Pd y Pt utilizando modelos multivariados de Aprendizaje Automático y su Integración con Geoestadística en la Generación de nuevos targets de Exploración en un Yacimiento de Sulfuros de Ni-Cu-PGEs" presentado por el autor Neil Harrison Berrospi Rios y coautores Daniel Abinadab Basilio Rojas y Carlos Josue Mendoza Kuong sea presentado en el concurso del Premio Nacional de Minería del evento PERUMIN 37 Convención Minera en las fechas del 22 al 26 de setiembre del 2025 en la ciudad de Arequipa.



Firma
DNI/Pasaporte: 47580628
Fecha: 04/07/2025