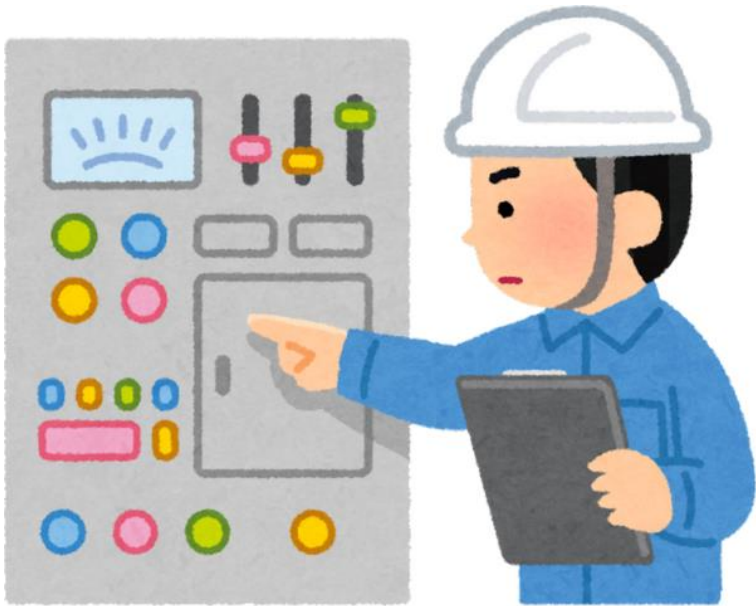


# Optimization of Iron Concentrate Percentage Using Machine Learning / Optimización del % de Hierro en Concentrado usando Machine Learning

Proyecto aplicado en contexto industrial

Fabio Lopez - October 2025



# Project Overview / Visión General del Proyecto

- **Business Context / Contexto del Negocio**

The flotation process in mineral beneficiation plants plays a crucial role in optimizing metal recovery and product purity. Efficient control of process variables can lead to increased iron concentrate quality and reduced silica content — two key metrics for production value.

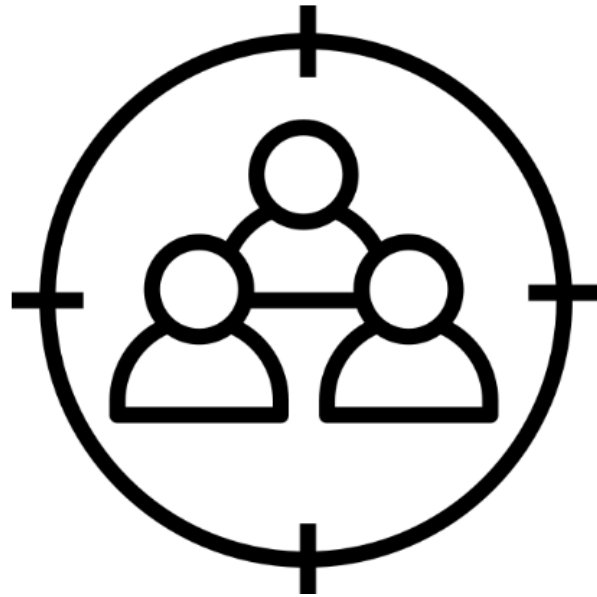
El proceso de flotación en plantas de beneficio mineral es fundamental para optimizar la recuperación del metal y la pureza del producto. Un control eficiente de las variables del proceso puede mejorar la calidad del concentrado de hierro y reducir el contenido de sílice, dos métricas clave en la rentabilidad de la producción.

# Project Objective / Objetivo del Proyecto

- To develop a **machine learning-based solution** capable of:
  - **Predicting % Iron Concentrate** from controllable process variables.
  - **Detecting anomalies** in production that may indicate process deviations.
  - **Optimizing input parameters** to maximize concentrate quality.
- 
- Desarrollar una solución basada en **machine learning** capaz de:
  - **Predecir el porcentaje de concentrado de hierro** a partir de variables del proceso controlables.
  - **Detectar anomalías** que puedan indicar desviaciones en la producción.
  - **Optimizar parámetros de entrada** para maximizar la calidad del concentrado.

# Target Audience / Audiencia objetivo:

Data science recruiters, production engineers, and operations managers in the mining and metallurgical sectors.



# Dataset Overview / Descripción del Conjunto de Datos

- **Dataset Overview**
- **Source:** Real industrial data from a flotation plant (available online).
- **File:** MiningProcess\_Flotation\_Plant\_Database.csv
- **Total Records:** 736,282
- **Time Range:** March 10, 2017 onward (hourly records)
- **Features:**
  - **Inputs:**
    - % Iron Feed
    - % Silica Feed
    - Starch Flow
    - Amina Flow
    - Ore Pulp pH
    - Ore Pulp Density
    - Flotation Column 01 Air Flow
  - **Targets:**
    - % Iron Concentrate
    - % Silica Concentrate



- **Descripción del Dataset**
- **Fuente:** Datos industriales reales de una planta de flotación (disponible en línea).
- **Archivo:** MiningProcess\_Flotation\_Plant\_Database.csv
- **Total de Registros:** 736,282
- **Rango de Tiempo:** Desde el 10 de marzo de 2017 (registros por hora)
  - **Variables: Entradas:**
    - % Hierro en la Alimentación
    - % Sílice en la Alimentación
    - Flujo de Almidón
    - Flujo de Amina
    - pH de la Pulpa de Mineral
    - Densidad de la Pulpa
    - Flujo de Aire de la Columna de Flotación 01
  - **Salidas:**
    - % Hierro en el Concentrado
    - % Sílice en el Concentrado



# Data Cleaning & Exploratory Data Analysis (EDA)/ Limpieza de Datos y Análisis Exploratorio

## **Data Cleaning:**

- Converted commas to decimal points for numeric consistency.
- Removed rows with missing or invalid data.
- Selected relevant features and target variables.
- Ensured all data types were correctly set for modeling.

## **Limpieza de Datos:**

- Se reemplazaron comas por puntos para uniformidad numérica.
- Se eliminaron filas con datos faltantes o inválidos.
- Se seleccionaron variables relevantes para el modelo.
- Se aseguró que todos los tipos de datos estuvieran correctos para el modelado.

- **Exploratory Data Analysis (EDA)**

- Examined distributions and ranges of features.
- Identified correlations between inputs and target (% Iron Concentrate).
- Detected and handled outliers.
- Visualized key variables for insights (histograms, scatter plots).

- **Análisis Exploratorio de Datos (EDA)**

- Se examinaron distribuciones y rangos de las variables.
- Se identificaron correlaciones entre entradas y el objetivo (% Hierro en Concentrado).
- Se detectaron y gestionaron valores atípicos.
- Se visualizaron variables clave para obtener insights (histogramas, gráficos de dispersión).



# Predictive Model & Results/Modelo Predictivo y Resultados

- **Model Used: Random Forest Regressor**
  - Captures nonlinear relationships between variables.
  - Robust to overfitting with proper tuning.
- 
- **Modelo Utilizado: Random Forest Regressor**
  - Captura relaciones no lineales entre variables.
  - Robusto contra el sobreajuste con ajuste adecuado.

- **Performance Metrics:**

- Mean Squared Error (MSE): 0.0532 (low error)
- $R^2$  Score: 0.9574 (explains 95.74% of variance)

- **Métricas de Desempeño:**

- Error Cuadrático Medio (MSE): 0.0532 (error bajo)
- $R^2$  Score: 0.9574 (explica 95.74% de la varianza)

# Interpretation/Interpretación:

- Model predicts % Iron Concentrate accurately.
  - High  $R^2$  indicates strong explanatory power.
  - Suitable for process optimization.
- 
- El modelo predice con precisión el % de Hierro en Concentrado.
  - $R^2$  alto indica fuerte poder explicativo.
  - Adecuado para la optimización del proceso.

# Process Optimization Results/Resultados de la Optimización del Proceso

- **Optimization Goal:**

- Maximize % Iron Concentrate by adjusting process parameters.

- **Objetivo de la Optimización:**

- Maximizar el % de Hierro en Concentrado ajustando los parámetros del proceso

- **Optimal Parameters Found:**

- Ore Pulp pH: **10.50**
- Starch Flow: **5036** units
- Amina Flow: **346** units
- Air Flow: **360** units
- % Iron Feed: **57.98%**
- % Silica Feed: **22.76%**
- Pulp Density: **1.76**

- **Parámetros Óptimos Encontrados:**

- pH Pulpa de Mineral: **10.50**
- Flujo de Almidón: **5036** unidades
- Flujo de Amina: **346** unidades
- Flujo de Aire: **360** unidades
- % Alimentación de Hierro: **57.98%**
- % Alimentación de Sílice: **22.76%**
- Densidad de Pulpa: **1.76**

**Result:**

Maximum predicted % Iron Concentrate: **66.52%**

**Resultado:**

% Máximo predicho de Hierro en Concentrado: **66.52%**

# Conclusions & Next Steps/Conclusiones y Próximos Pasos

- **Conclusions:**

- Machine Learning model accurately predicts % Iron Concentrate with high  $R^2$  (0.96).
- Optimization identified key process parameters to maximize iron recovery.
- Implementing recommended settings can improve concentrate quality and process efficiency.

- **Conclusiones:**

- El modelo de Machine Learning predice con precisión el % de Hierro en Concentrado con un  $R^2$  alto (0.96).
- La optimización identificó parámetros clave del proceso para maximizar la recuperación de hierro.
- Implementar los ajustes recomendados puede mejorar la calidad del concentrado y la eficiencia del proceso.

- **Next Steps:**

- Validate optimized parameters with pilot or plant trials.
- Integrate predictive model into process control systems for real-time adjustments.
- Continue data collection to refine and improve the model.
- Explore applying ML to other flotation variables or downstream processes.

- **Próximos Pasos:**

- Validar los parámetros optimizados con pruebas piloto o en planta.
- Integrar el modelo predictivo en los sistemas de control para ajustes en tiempo real.
- Continuar con la recolección de datos para refinar y mejorar el modelo.
- Explorar la aplicación de ML en otras variables de flotación o procesos posteriores.

# Thank You & Final Remarks/Agradecimientos y Comentarios Finales

**Thank you for your attention!**

- This project reflects my ability to apply data analytics and machine learning techniques to solve real-world industrial problems.  
I'm ready to contribute as a **Data Analyst**, and I'm fully committed to continue learning and adapting to any technical or business need.

Let's talk about how I can add value to your team!

**¡Gracias por su atención!**

- Este proyecto demuestra mi capacidad para aplicar análisis de datos y técnicas de machine learning a problemas reales del entorno industrial.  
Estoy listo para aportar como **Analista de Datos**, y tengo toda la disposición para seguir aprendiendo y adaptarme a cualquier necesidad técnica o del negocio.

¡Conversemos sobre cómo puedo aportar valor a su equipo!

