

Proyecto de Ciencia de Datos

Autor: Agustín Garibaldi

Predicción y Monitoreo del Consumo Energético en una Unidad de Destilación

Introducción

Este proyecto tiene como objetivo principal predecir el consumo energético horario en una unidad de destilación de una refinería de petróleo. Para ello, se aplican técnicas de Ciencia de Datos que incluyen:

- Análisis exploratorio de datos
- Preprocesamiento y transformación de variables
- Entrenamiento de modelos de regresión (Lineal y Random Forest)
- Evaluación comparativa de desempeño
- Visualización de resultados

Como complemento, se desarrolló un dashboard interactivo en Looker Studio para el monitoreo de variables clave del proceso. El panel fue diseñado con una estructura responsive, permitiendo su visualización fluida tanto en computadoras de escritorio como en celulares o tablets, lo que facilita la consulta desde cualquier dispositivo en planta o en operación remota.

Dataset Simulado

Para este análisis se trabajó con un dataset sintético de 720 registros horarios, que simula condiciones reales de operación. Cada fila representa una hora de funcionamiento continuo de la unidad. Las variables incluidas son:

- fecha_hora: marca temporal de cada observación
- tipo_crudo: categoría del crudo procesado (liviano, medio, pesado)
- caudal_crudo_m3h: volumen de crudo procesado por hora (en m³/h)
- temperatura_entrada_C: temperatura del crudo al ingresar (°C)
- temperatura_salida_C: temperatura del crudo a la salida (°C)
- presion_bar: presión interna de operación de la columna de destilación (en bar)
- consumo_energia_kwh: energía eléctrica consumida durante esa hora (variable objetivo)

Herramientas utilizadas

- Python (Pandas, Seaborn, Scikit-Learn, Matplotlib) para procesamiento, visualización y modelado.
- Looker Studio (Google Data Studio) para construir un panel de control accesible, interactivo y responsive.
- MS Word para la documentación técnica del proyecto.