

PROYECTO INTRODUCCIÓN A LA INTELIGENCIA ARTIFICIAL.

Raúl Felipe Berrio - Esteban machado - Diego Andrés Rodríguez Galeano.

(1) Descripción del Problema Predictivo a Resolver:

El problema predictivo se enfoca en la monitorización precisa de las emisiones de dióxido de carbono (CO₂), una parte fundamental en la lucha contra el cambio climático. La capacidad de obtener lecturas precisas de carbono permite a los investigadores y gobiernos comprender las fuentes y patrones de emisión de CO₂. Aunque Europa y América del Norte cuentan con sistemas extensos para monitorear las emisiones de carbono en tierra, en África hay una falta de tales sistemas disponibles.

El objetivo de este desafío es crear modelos de machine learning utilizando datos de emisiones de CO₂ de código abierto provenientes de observaciones satelitales Sentinel-5P para predecir las futuras emisiones de carbono. Estas soluciones pueden ayudar a los gobiernos y otras partes interesadas a estimar los niveles de emisiones de carbono en toda África.

(2) Dataset Utilizado: <https://www.kaggle.com/competitions/playground-series-s3e20/data>

Utilizaremos el conjunto de datos GRACED proporcionado por Carbon Monitor (con 76 columnas y más de 79000 datos), basado en observaciones satelitales Sentinel-5P. Este conjunto de datos contiene mediciones de concentración de CO₂ y otros gases relacionados en diferentes ubicaciones de África, así como variables temporales y geoespaciales relevantes. El conjunto de datos contiene siete características principales que se extrajeron semanalmente de las observaciones de Sentinel-5P desde enero de 2019 hasta noviembre de 2022. Estas características son:

- Sulfuro de Dioxido (Sulphur Dioxide): Medido por COPERNICUS/S5P/NRTI/L3_SO2.
- Monóxido de Carbono (Carbon Monoxide): Medido por COPERNICUS/S5P/NRTI/L3_CO.
- Dióxido de Nitrógeno (Nitrogen Dioxide): Medido por COPERNICUS/S5P/NRTI/L3_NO2.
- Formaldehído (Formaldehyde): Medido por COPERNICUS/S5P/NRTI/L3_HCHO.
- Índice de Aerosol UV (UV Aerosol Index): Medido por COPERNICUS/S5P/NRTI/L3_AER_AI.
- Ozono (Ozone): Medido por COPERNICUS/S5P/NRTI/L3_O3.
- Nubes (Cloud): Medido por COPERNICUS/S5P/OFFL/L3_CLOUD.

(3) Métricas de Desempeño Requeridas:

Métrica de Machine Learning: La métrica de desempeño para evaluar los modelos será el Error Cuadrático Medio (RMSE, por sus siglas en inglés). El RMSE mide la raíz cuadrada de la diferencia entre los valores predichos y los valores reales, lo que proporciona una medida de cuán cerca están las predicciones de los valores reales de emisiones de CO₂.

(4) Desempeño Deseable en Producción:

El desempeño deseable en producción sería lograr un valor bajo de RMSE, lo que indicaría que el modelo es capaz de hacer predicciones precisas de las emisiones futuras de CO₂ en África. Dado que el RMSE es una métrica de error, un valor bajo significa que las predicciones se ajustan bien a los datos reales y son útiles para estimar las emisiones de CO₂ en regiones donde no se puede realizar seguimiento en el terreno. La precisión y la capacidad de generalización del modelo serán esenciales para su utilidad en la estimación de emisiones de carbono en toda África.