

Primera Entrega del Proyecto

Por:

Edwin David Duque Grajales cc. 1017255650 Ing Ambiental
Peter Daniel Garrido Rodriguez cc. 1085952515 Ing. Industrial
Jheison Andres Benavides Rincon cc.1015466242 Ing. Civil

Materia:

Introducción a la Inteligencia Artificial para las Ingenierías

Profesor:

Raúl Ramos Pollan



Universidad de Antioquia
Facultad de Ingeniería
Medellín 2023

Descripción del Problema

El objetivo de este desafío es crear modelos de aprendizaje automático que utilicen datos de emisiones de fuente abierta (de observaciones del satélite Sentinel-5P) para predecir las emisiones de carbono.

Dataset

Se utilizará un dataset de kaggle que corresponde a la siguiente competición: [Predict CO2 Emissions in Rwanda | Kaggle](#), que tiene 24.354 filas y 76 columnas.

En esta competencia seleccionaron aproximadamente 497 ubicaciones únicas de múltiples áreas de Ruanda, distribuidas entre tierras agrícolas, ciudades y plantas de energía. Los datos de esta competición están divididos por tiempo; Los años 2019 - 2021 están incluidos en los datos de entrenamiento y su tarea es predecir los datos de emisiones de CO2 desde 2022 hasta noviembre.

Se extrajeron siete características principales semanalmente de Sentinel-5P desde enero de 2019 hasta noviembre de 2022. Cada característica (dióxido de azufre, monóxido de carbono, etc.) contiene subcaracterísticas como column number density, qué es la densidad de la columna vertical a nivel del suelo, calculada mediante la técnica DOAS. Puede leer más sobre cada característica en los enlaces a continuación, incluido cómo se miden y las definiciones de las variables. Se le proporcionan los valores de estas características en el conjunto de prueba y su objetivo es predecir las emisiones de CO2 utilizando información de tiempo además de estas características.

- Dióxido de azufre - COPENICUS/S5P/NRTI/L3_SO2
- Monóxido de carbono - COPENICUS/S5P/NRTI/L3_CO
- Dióxido de nitrógeno - COPENICUS/S5P/NRTI/L3_NO2
- Formaldehído - COPENICUS/S5P/NRTI/L3_HCHO
- Índice de aerosoles UV - COPENICUS/S5P/NRTI/L3_AER_AI
- Ozono - COPENICUS/S5P/NRTI/L3_O3
- Nube - COPENICUS/S5P/OFFL/L3_CLOUD

Métricas

La métrica de evaluación principal para el modelo será error absoluto medio (MAE) el cual se calcula mediante la siguiente expresión:

$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - x_i|}{n} = \frac{\sum_{i=1}^n |e_i|}{n}.$$

Donde el MAE se calcula como la suma de los errores absolutos dividida por el tamaño de la muestra.

El error absoluto medio sirve para cuantificar la precisión de una técnica de predicción comparando por ejemplo los valores predichos frente a los observados, el tiempo real frente al tiempo previsto, o una técnica de medición frente a otra técnica alternativa de medición. Esta métrica nos permitirá ver el nivel de error que se obtienen de los datos calculados y los reales y determinar en qué aspectos del modelo requieren ajustes o soluciones.

Primer Criterio

Lo esperado de la modelación de este proyecto es obtener valores significativos con un nivel de confianza del 60% que puedan predecir las emisiones de gases de carbono con referencia a la fuente abierta de observaciones del satélite Sentinel-5P. Además de procesar la información con la técnica DOAS teniendo en cuenta las variables de Dióxido de azufre, Monóxido de carbono, Dióxido de nitrógeno, Formaldehído, Índice de aerosoles UV, Ozono y Nubes. Es importante mencionar que los criterios están sujetos a cambios mediante se está desarrollando el modelo.

Bibliografía

- Differential Optical Absorption Spectroscopy, 2011. Recuperado el 22 de septiembre de 2023, de http://chrome-extension://efaidnbnmnibpcajpcglclefindmkaj/http://repositorio.gestiondelriesgo.gov.co/bitstream/20.500.11762/20564/3/Fenomeno_nino-2016.pdf
- *Predict CO2 emissions in Rwanda*. (s/f). Kaggle.com. Recuperado el 26 de septiembre de 2023, de <https://www.kaggle.com/competitions/playground-series-s3e20/data>
- Wikipedia contributors. (s/f). *Error absoluto medio*. Wikipedia, The Free Encyclopedia. https://es.wikipedia.org/w/index.php?title=Error_absoluto_medio&oldid=146234007