

# GASES DE EFECTO INVERNADERO Y SU RELACION CON LA TEMPERATURA MUNDIAL

## GRUPO 2 :

- Lautaro Villafañe ([ivilla1357@gmail.com](mailto:ivilla1357@gmail.com)) – Ing. Químico.
- Diego Murature ([diegofmurature@gmail.com](mailto:diegofmurature@gmail.com)) – Ing. En Telecomunicaciones.
- Julio Mansilla ([julio.tingo@gmail.com](mailto:julio.tingo@gmail.com)) – Geólogo.

## OBJETIVOS (¿QUE QUEREMOS HACER?)

### Objetivo general

El objetivo de este trabajo es utilizar Python, a través de Google Colab, para poder analizar las variables que influyen en el cambio climático y en las variaciones de temperatura de superficie medias mundiales. Además, realizar un modelo que prediga las temperaturas a partir de las emisiones y concentraciones de gases de efecto invernadero emitidas por la población e industrias.

### Objetivos particulares

- Analizar las tendencias de las emisiones de gases a lo largo de los años por países.
- Determinar cuáles son los países que mas colaboran con las emisiones de gases.
- Cuales son las concentraciones de gases que se encuentran en la atmosfera terrestre.
- De los países que más emisiones de gases liberan, cuales son sus fuentes de energía y de que tipo son (renovables o no renovables).
- Como es el consumo de combustibles fósiles en el mundo a lo largo de los años.
- Cuales variables son las que mas contribuyen a las anomalías de temperaturas mundiales.
- Realizar un Análisis de Regresión Lineal Múltiple, utilizando variables como las concentraciones de gases de efecto invernadero en la atmosfera para predecir las anomalías de T (°C) en superficie.
- Utilizar un método de aprendizaje automático de modelos supervisados como Random Forest para predecir las anomalías de temperatura global.

## DESCRIPCIÓN DE LA PROBLEMÁTICA (¿POR QUÉ ES IMPORTANTE ESTE PROBLEMA?)

Aumento de las Emisiones de Gases de Efecto Invernadero y sus consecuencias en el Cambio Climático

En la última década, las emisiones de dióxido de carbono (CO<sub>2</sub>) , metano (CH<sub>4</sub>), y óxidos de nitrógeno (NO) han alcanzado niveles alarmantes, resultado en un aumento significativo en la temperatura media global. Este incremento no solo afecta el clima, sino que también presenta riesgos para la salud pública, la biodiversidad y la seguridad alimentaria. A medida que los países

industrializados y en desarrollo contribuyen a estas emisiones, se observa una falta de consenso y acción global efectiva para mitigar el impacto del cambio climático. Por ello, es fundamental comprender las variables que inciden en este fenómeno para poder mitigar sus efectos y proponer soluciones sostenibles.

### Impactos claves de esta problemática

1. **Desigualdad en las Emisiones:** Algunos países son responsables de una gran parte de las emisiones globales, mientras que otros, a pesar de ser menos contaminantes, sufren desproporcionadamente los efectos del cambio climático, como desastres naturales, sequías y aumento del nivel del mar.
2. **Fuentes de Energía No Renovables:** Muchos de los países que más contribuyen a las emisiones dependen en gran medida de fuentes de energía no renovables, lo que agrava la situación. Esto plantea una necesidad urgente de transición hacia fuentes de energía más sostenibles y limpias.
3. **Incertidumbre en las Predicciones Climáticas:** La falta de modelos precisos que relacionen las concentraciones de gases de efecto invernadero y sus respectivas anomalías en temperaturas dificulta la formulación de políticas eficaces para combatir el cambio climático. Además, el sistema climático es complejo, y las interacciones entre sus componentes son difíciles de modelar con precisión.
4. **Cambio en los Patrones Climáticos:** Las anomalías de temperatura que resultan de altas concentraciones de GEI (gases de efecto invernadero) están relacionadas con cambios peligrosos en los patrones climáticos, lo que puede llevar a problemas en la agricultura, acceso al agua y sanidad pública.

Por lo tanto, las investigaciones son cruciales para entender las tendencias y variables que contribuyen a estas problemáticas. Utilizando herramientas como Python y modelos de aprendizaje automático, se puede profundizar en el análisis de datos sobre emisiones, concentraciones de gases y sus impactos en la temperatura. Se debe entender que este problema no solo es un desafío científico, sino también una responsabilidad ética y social.

### FUENTES DE LOS DATOS

Los datos utilizados para este estudio provienen de fuentes confiables y reconocidas a nivel mundial, entre ellas:

- **NASA:** Para datos de temperatura global y registros históricos climáticos.
- **NOAA** (National Oceanic and Atmospheric Administration): Para información sobre concentraciones de gases de efecto invernadero.
- **Our World in Data:** Para estadísticas sobre emisiones de gases por país y tipo de energía.
- **Banco Mundial:** Para información socioeconómica relacionada con el consumo de energía y las emisiones industriales

## METODOLOGÍA (PASOS DEL ANÁLISIS)

1. **Recolección de datos:** Se recopilaron conjuntos de datos provenientes de las fuentes mencionadas.
2. **Limpieza y preprocesamiento:** Se eliminaron valores atípicos, datos faltantes y se ajustaron los formatos para un análisis uniforme.
3. **Análisis exploratorio:** Se realizaron visualizaciones para identificar tendencias, patrones y correlaciones entre las variables.
4. **Selección de características:** Se utilizó esta herramienta para investigar las variables más significativas en el comportamiento de la variable target.
5. **Análisis de regresión lineal múltiple:** Se implementó este modelo para evaluar la relación entre las concentraciones de GEI y las anomalías de temperatura. Se realiza una validación simple y una cruzada. Se toma un 70% de los datos para entrenamiento y 30% de test set.
6. **Se plantearon los análisis de supuestos:** Para la regresión lineal múltiple, realizando el diagnostico del modelo para la linealidad, homocedasticidad, normalidad y la autocorrelación (independencia de los errores).
7. **Modelo de predicción:** Se utilizó el algoritmo Random Forest para predecir las anomalías de temperatura en función de las emisiones y otras variables relevantes. Se utilizan hiperparámetros como  $n\_estimators = 50$ ,  $max\_depth = 5$ ,  $min\_samples\_split = 30$ .
8. **Regresión ordinaria de mínimos cuadrados con Statsmodels (OLS):** Se obtienen coeficientes para evaluar las relaciones entre variables y comprobar los supuestos del modelo. Se obtienen valores como el p-value,  $R^2$ , Durbin-Watson, etc.
9. **Evaluación del modelo:** Se midieron métricas como el error cuadrático medio (MSE), el error absoluto medio (MAE), el Error cuadrado medio de raíz (RMSE) y el coeficiente de determinación ( $R^2$ ) para validar el desempeño del modelo.

## RESULTADOS

- ✓ Se identificaron los países que más contribuyen a las emisiones de gases de efecto invernadero, siendo China, Estados Unidos e India los principales emisores.
- ✓ Se demostró que el aumento de la concentración de CO<sub>2</sub>, CH<sub>4</sub> y N<sub>2</sub>O en la atmósfera está fuertemente correlacionado con el incremento de la temperatura media global.
- ✓ La creación de matrices de correlación (Pearson y Spearman) mostraron las principales correlaciones positivas y negativas entre las distintas variables. Siento esto una fuente importante para relacionar otras variables y realizar distintos análisis.

- ✓ El modelo de regresión lineal múltiple obtuvo un  $R^2$  del 0.83054, mostrando una alta capacidad explicativa de las anomalías de temperatura, es decir, el 83.05% de la variabilidad de la temperatura puede ser explicada por las concentraciones de estos gases. Además, el RMSE da 0.1198 lo sugiere que, en promedio, las predicciones del modelo tienen un error de aproximadamente 0.119 grados Celsius.
- ✓ Los análisis de los supuestos para la regresión lineal múltiple dan como resultados: Normalidad: los residuos están aproximadamente distribuidos de manera normal (según histograma y Gráfico Q-Q).
- ✓ El modelo Random Forest presentó una mayor precisión en la predicción de temperaturas futuras debido a su capacidad para manejar datos no lineales. Los resultados parecen indicar que el modelo es consistente entre los conjuntos de entrenamiento y testeo. No hay una diferencia significativa entre las métricas, lo que indica que el modelo tiene un buen equilibrio y no hay overfitting. Tanto el MSE como el MAE y la RMSE son bajos, lo que implica que el modelo hace predicciones precisas. El alto valor de  $R^2$  (Entrenamiento: 0.95073 Testeo: 0.9354) dado indica que el modelo se ajusta bien a los datos.
- ✓ Además, para el modelo Random Forest, se obtiene un p-value : 0.000 en  $x_1$ . Este valor p es muy bajo, lo que sugiere que el coeficiente de  $x_1$  es altamente significativo. Por lo que hay evidencia suficiente para afirmar que  $x_1$  tiene un efecto significativo en la variable dependiente  $T^{\circ}C$ .
- ✓ El valor de Durbin-Watson: 2.389. Por lo que hay una ligera autocorrelación negativa en los residuos, pero NO es tan alta ya que si el valor estadístico está entre 1.5 y 2.5 se considera normal. No hay un problema grave de autocorrelación en los residuos.

## BIBLIOGRAFÍA

- ❖ NASA, "Global Climate Change: Vital Signs of the Planet". <https://climate.nasa.gov/>
- ❖ NOAA, "Greenhouse Gas Index". <https://www.noaa.gov/>
- ❖ Our World in Data, "CO2 and Greenhouse Gas Emissions". <https://ourworldindata.org/>
- ❖ Banco Mundial, "Data Catalog". <https://datacatalog.worldbank.org/>