TEAM MEMBERS
1. Sameer Ahamed Rizwan Basha – 202381922
2. Peter Utomakili – 201477924

PROJECT OVERVIEW

With a primary focus on greenhouse gas emissions, particularly carbon dioxide ($CO_2$). These emissions result from a variety of human activities, including the use of fossil fuels for energy, transportation, manufacturing, and deforestation. Rising global temperatures, more frequent and severe extreme weather events, the melting of the polar ice caps, an increase in sea level, and disruptions to ecosystems and biodiversity are some of the negative repercussions that follow. To ensure a healthy and habitable planet for both the present and future generations, addressing the emissions problem is a global necessity that requires cooperative efforts to reduce and ultimately eradicate these harmful emissions.

WHAT INTERESTING THING DO YOU WANT TO TRY TO MODEL USING DATA?

A regression model would be a good option to forecast $CO_2$ emissions from a dataset gathered in Rwanda. We have planned to train different models with our dataset and use a model that best represents our data. A few of the models we have planned to use are Linear regression (Lasso), Random Forest, XGB Boost, Decision Trees, Support Vector Machines (SVM), etc. You can use these models to generate quantitative forecasts based on the dataset and understand how various factors affect $CO_2$ emissions.

EXPLAIN WHY IT WOULD BE USEFUL TO SOMEONE.

There are various reasons why using a regression model to forecast $CO_2$ emissions is important. Firstly, it is an effective means in the struggle against climate change on a worldwide scale. Researchers and policymakers can decide how to mitigate the environmental impact of $CO_2$ emissions in Rwanda by having a better understanding of the elements that drive these emissions. Secondly, there isn't quality data that we can get from the subcontinent. Therefore, we can use the predictive model from Rwanda as a generalization to all of Africa. It equips decision-makers to develop efficient policies and allocate resources targeted at lowering emissions and attaining sustainability development goals.

DESCRIPTION OF THE DATA SOURCES YOU WILL USE TO TRAIN YOUR MODEL

We obtained the dataset from [Kaggle.com | CO2 Emissions in Rwanda](#)

Our project is predicting $CO_2$ emissions using regression modeling. The factors we've identified encompass a wide range of variables, both atmospheric and pollutant-related, which can significantly influence $CO_2$ emissions. Relevance of each factor is discussed briefly below:

1. **Sulfur Dioxide (SO2)**: $SO_2$ is a precursor to sulfuric acid formation and can contribute to air pollution. Understanding its levels can shed light on industrial emissions and their environmental impact.
2. **Carbon Monoxide (CO)**: CO is a product of incomplete combustion and is a key indicator of combustion-related emissions. Monitoring CO levels helps assess the efficiency of energy and transportation systems.
3. **Nitrogen Dioxide (NO2)**: $NO_2$ is a major component of smog and contributes to respiratory issues. It's an important pollutant to track due to its links with emissions from vehicles and industrial sources.
4. **Formaldehyde**: Formaldehyde is a volatile organic compound (VOC) and is a crucial air quality parameter. It is linked to various health concerns and is used to gauge indoor and outdoor pollution sources.
5. **Aerosol Content**: Aerosols affect climate and air quality. Measuring aerosol content is valuable for understanding their role in emissions and atmospheric processes.
6. **Ozone**: Ozone levels are relevant as they are both a greenhouse gas and an air pollutant. Its presence in the atmosphere is a key factor in air quality and climate change.
7. **Atmospheric Features**: Monitoring atmospheric features such as cloud cover can provide insights into how weather patterns impact emissions. Clouds can affect the amount of solar radiation reaching the Earth's surface, which, in turn, influences $CO_2$ concentrations.