CLIMATE CHANGE MODELING USING NASA CLIMATE DATA

Aarya Shukla

UNIFIED MENTOR PROJECT

CLIMATE CHANGE MODELING USING NASA CLIMATE DATA

Abstract

This project focuses on climate change modeling using real-world data obtained from NASA's climate datasets. With the growing concern around global warming, rising sea levels, and changing precipitation patterns, this analysis attempts to uncover key trends in global climate parameters. The objective is to process, visualize, and model the data to gain insights into long-term climate behavior and support future climate projections.

1. Introduction

Climate change is one of the most pressing challenges facing humanity today. The increasing levels of greenhouse gases, global temperature anomalies, and other climate indicators have made it imperative for scientists and policymakers to rely on data-driven modeling to anticipate the future state of our planet. This project utilizes historical climate data from NASA to analyze trends and build predictive models. The analysis aims to understand temperature variations over time and explore potential future changes using machine learning techniques.

2. Dataset Description

- **Dataset Name**: climate_nasa.csv
- **Source**: NASA Climate Data Portal

- **Key Columns**:
    - Year: The year of observation
    - Month: The month of observation
    - Mean: Mean monthly temperature anomaly
    - High: Highest temperature anomaly
    - Low: Lowest temperature anomaly

This data represents **global monthly average temperature anomalies** over a long time span, which serves as a useful indicator for observing warming trends.

3. Methodology

The notebook follows these main steps:

- **Loading and Cleaning the Data**:
    - Checked for missing values and ensured correct data types.
    - Combined Year and Month into a Datetime column.
    - Sorted data chronologically.
- **Exploratory Data Analysis (EDA)**:
    - Used line plots and rolling averages to visualize trends.
    - Explored seasonality and anomalies in temperature patterns.
- **Feature Engineering**:
    - Extracted features such as Year, Month, Quarter.
    - Created lagged features and rolling mean values for temporal modeling.

- **Modeling**:
  - Implemented a **Linear Regression** model to forecast temperature anomaly.
  - Split the dataset into training and testing sets using a time-series approach.
  - Evaluated model performance using metrics like MAE and RMSE.

4. Exploratory Data Analysis (EDA)

Some key insights from EDA:

- **Temperature Anomaly Increasing**: The mean global temperature anomaly shows a **clear upward trend**, especially post-1980.
- **Monthly Seasonality**: There is a recurring seasonal pattern indicating periodic highs and lows.
- **Rolling Averages**: A 12-month rolling mean smooths the data and reveals the long-term warming trend.

Visualizations Included:

- Line plot of monthly mean temperature anomaly over time
- Rolling average plots
- Boxplots showing anomaly distribution per month

5. Modeling and Predictions

- **Model Used**: Scikit-learn's LinearRegression
- **Features**: Year, Month (as numerical), rolling means
- **Target**: Mean temperature anomaly

- **Train-Test Split**: 80-20 based on chronological order

Model Evaluation:

- **Mean Absolute Error (MAE)**: ~0.06
- **Root Mean Square Error (RMSE)**: ~0.09

The model successfully captured the rising temperature trend and provided reasonable short-term forecasts. However, it may not capture nonlinear effects or complex interactions, which are common in climate data.

6. Results and Discussion

- **Forecast Visualization**: Forecasts were plotted against actual temperature anomalies. The predicted curve followed the actual trend closely for near-future values.
- **Interpretation**: Linear regression is a good first approach, but climate dynamics often require more complex models like ARIMA, LSTM, or ensemble methods.

7. Conclusion

This project demonstrates how global temperature anomaly data can be processed, analyzed, and modeled to understand climate change trends. Key takeaways:

- The Earth is experiencing a **consistent upward shift in average temperatures**.
- **Machine learning techniques** can help forecast short-term changes.
- This modeling lays the foundation for more sophisticated climate simulation tools.

8. Future Work

- Incorporate more features like CO_2 levels, ocean temperature, ice coverage, etc.
- Try advanced models such as **LSTM**, **Random Forest**, or **XGBoost**.
- Use **climate simulation models (CMIP6)** for more robust long-term forecasts.

9. References

- NASA Global Climate Change: https://climate.nasa.gov
- IPCC Reports
- Scikit-learn Documentation
- Python Libraries: Pandas, Matplotlib, Seaborn, Scikit-learn