**Climate Change Modeling Project - Final Report**

**Objective:**

The goal of this project was to analyze over 500 user comments from NASA’s Facebook page related to climate change and use **sentiment analysis** to predict user engagement (likes) using a **Random Forest Regressor** model. The project aimed to understand how sentiment impacts public engagement with climate change-related posts.

**Steps Taken:**

**1. Understanding the Problem:**

* The problem was defined as predicting user engagement (measured by likesCount) based on the sentiment expressed in user comments. This required analyzing the comments, processing text data, and building a predictive model.

**2. Dataset Preparation:**

* The dataset contained over 500 user comments with columns like date, likesCount, profileName, commentsCount, and text.
* **Data Cleaning**: Missing values were handled, and the text column was processed to remove non-alphanumeric characters and convert all text to lowercase.

**3. Data Exploration and Preprocessing:**

* The data was explored using basic statistics, and sentiment analysis was performed on the text column to classify comments as **Positive**, **Negative**, or **Neutral**.

**4. Sentiment Analysis:**

* The **TextBlob** library was used to analyze the sentiment of each comment. The comments were categorized into three sentiment labels:
  + Positive
  + Negative
  + Neutral

**5. Model Training:**

* The **Random Forest Regressor** model was trained using sentiment features (sentiment\_Positive, sentiment\_Neutral) to predict user engagement (likesCount).
* The model was trained on 80% of the data, and 20% was reserved for testing.

**6. Saving the Model:**

* The trained model was saved using **joblib**, ensuring that it could be easily reloaded and used for future predictions without retraining.

**7. Scenario Analysis:**

* Three scenarios were simulated with different sentiment configurations:
  + **Scenario 1**: High Positive Sentiment
  + **Scenario 2**: High Neutral Sentiment
  + **Scenario 3**: Balanced Sentiment
* Predictions for likesCount were made for each scenario to understand how sentiment influences engagement.

**8. Visualization:**

* A bar chart was created to visualize the predicted likesCount for each sentiment scenario. The chart showed:
  + High Positive Sentiment led to the highest engagement.
  + High Neutral Sentiment led to the lowest engagement.
  + Balanced Sentiment had a moderate level of engagement.

**Results:**

* **Scenario 1** (High Positive Sentiment): **6.83 likes**
* **Scenario 2** (High Neutral Sentiment): **2.65 likes**
* **Scenario 3** (Balanced Sentiment): **5.60 likes**

The analysis revealed that **positive sentiment** tends to drive higher engagement, while **neutral sentiment** results in lower engagement.

**Conclusion:**

* The project successfully predicted user engagement based on sentiment using a **Random Forest Regressor**.
* It demonstrated that sentiment plays a significant role in determining public engagement with climate change content on social media.
* The model could be further improved by adding more features (e.g., commentsCount, date) and tuning the model for better accuracy.