**Rainfall Prediction Model**

In this project I explored over the historical weather API data from date 1st Jan 2019 to 10th Feb 2023 and build the model to predict the rainfall. To build this model I have used regression models in google colab.

**Steps Involved in this Project:**

**Loading and Discovering Data:**

Here, I have downloaded the data from <https://open-meteo.com/> in the format of the CSV (Comma Separated Values) file and then uploaded that dataset into the drive, then mounted the dataset and loaded it into the google colab.

**Pre-processing the Data:**

Data cleaning is an important step in the data analytics process in which you either remove or update information that is incomplete or improperly formatted. Null values Treatment by different methods. We have our dataset in hand which is raw and unfiltered. This step involves cleaning our data first by eliminating the columns which are not needed for our analysis. We have around 1502 rows × 9 columns in our dataset. Since, there were no null values, we have left it untouched.

**Exploratory Data Analysis:**

Exploratory Data Analysis (EDA) is an approach to analysing and summarizing data sets in order to extract insights and understand the data better. It involves using a variety of techniques and visualizations to get a sense of the structure, patterns, and relationships within the data.

**Feature Engineering:**

This section refers to manipulation — addition, deletion, combination, mutation — of your data set to improve machine learning model training, leading to better performance and greater accuracy.

**Train Test Split:**

In this section I split the data into a training set and a testing set. The training set is used for training the model, and the testing set is used to test your model. This allows us to train your models on the training set, and then test their accuracy on the unseen testing set.

**Models:**

* Lasso Regression
* Random Forest (With GridSearchCV)

**Accuracy:**

* Lasso Regression: 53.08 %
* Random Forest: 82.36%

**Model Architecture:**

The Lasso regression model used in the previous section attempts to find a linear relationship between the input features and the target variable. The Random Forest Regressor model, on the other hand, builds a multiple decision tree to make predictions. It splits the input data into subsets based on the values of the input features and recursively splits the subsets until a stopping criterion is met. The final predictions are made by averaging the rainfall values of the training examples that reach the same leaf node.

**Conclusion:**

The Random Forest Regressor model outperformed the Lasso regression model in predicting daily rainfall. Further experimentation with different models and input features may be necessary to improve the model's performance even further.