WEATHER TREND FORECASTING REPORT

PREPARED BY:

PAVANKUMAR T

PM ACCELERATOR MISSION:

- Our mission is to break down financial barriers and achieve educational fairness.

- With the goal of establishing 200 schools worldwide over the next 20 years.

- We aim to empower more kids for a better future in their life and career, simultaneously fostering a diverse landscape in the tech industry.

WEATHER TREND FORECASTING USING MACHINE LEARNING AND TIME SERIES MODELS

Project Overview:

1. Objective: Forecasting weather patterns using Machine Learning and time series models.

2. Data Source: GlobalWeatherRepository.csv(taken from kaggle).

3. Key Analysis:

   * Basic Analysis:

      Data Cleaning & Preprocessing, EDA, Basic model building.

    * Advanced Analysis:

      Advanced EDA, Forecasting with multiple models and unique analyses.

DATA CLEANING AND PREPROCESSING

- **Data Cleaning:**

  Handled missing values.

  Handled outliers using Inter Quartile Range (IQR).

- **Normalization:**

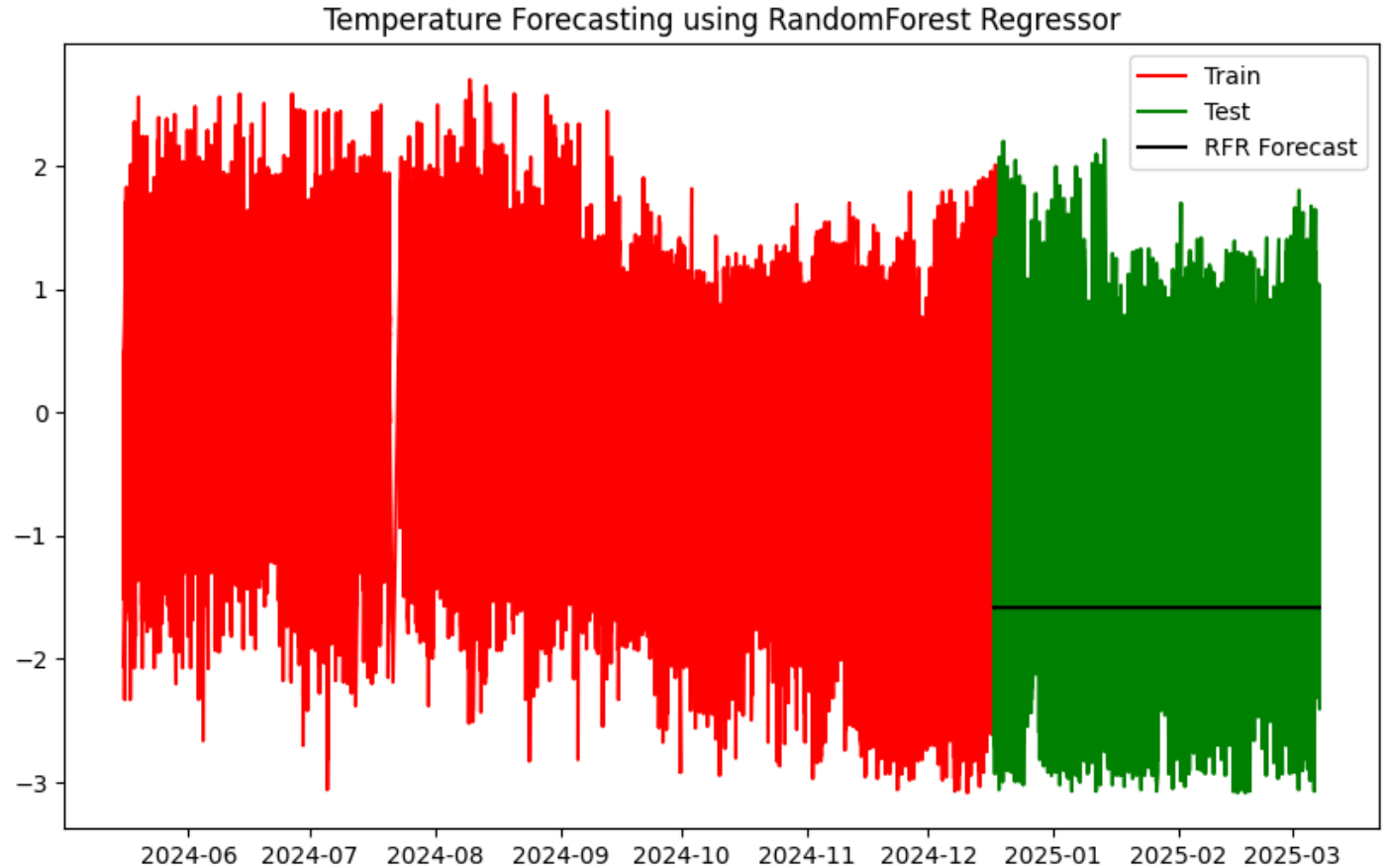 Normalized the numerical data using StandardScaler(preprocessing).

EXPLORATORY DATA ANALYSIS (EDA)

- ■ **Descriptive Statistics:**

    Displayed descriptive statistics of the data.

- ■ **Visualization:**

    Visualized data for understanding data distribution and trends

Temperature Forecasting using RandomForest Regressor

FORECASTING MODELS

- **RandomForest Regressor:**

  Model trained and made predictions.

  Evaluation of RFR model is done using MSE and RMSE

  For RFR model the mse = 2.4606 and rmse = 1.5686

- **XGBoost Regressor:**

  Model trainedand made predictions.

  Evaluation of XGBR model is done using MSE and RMSE
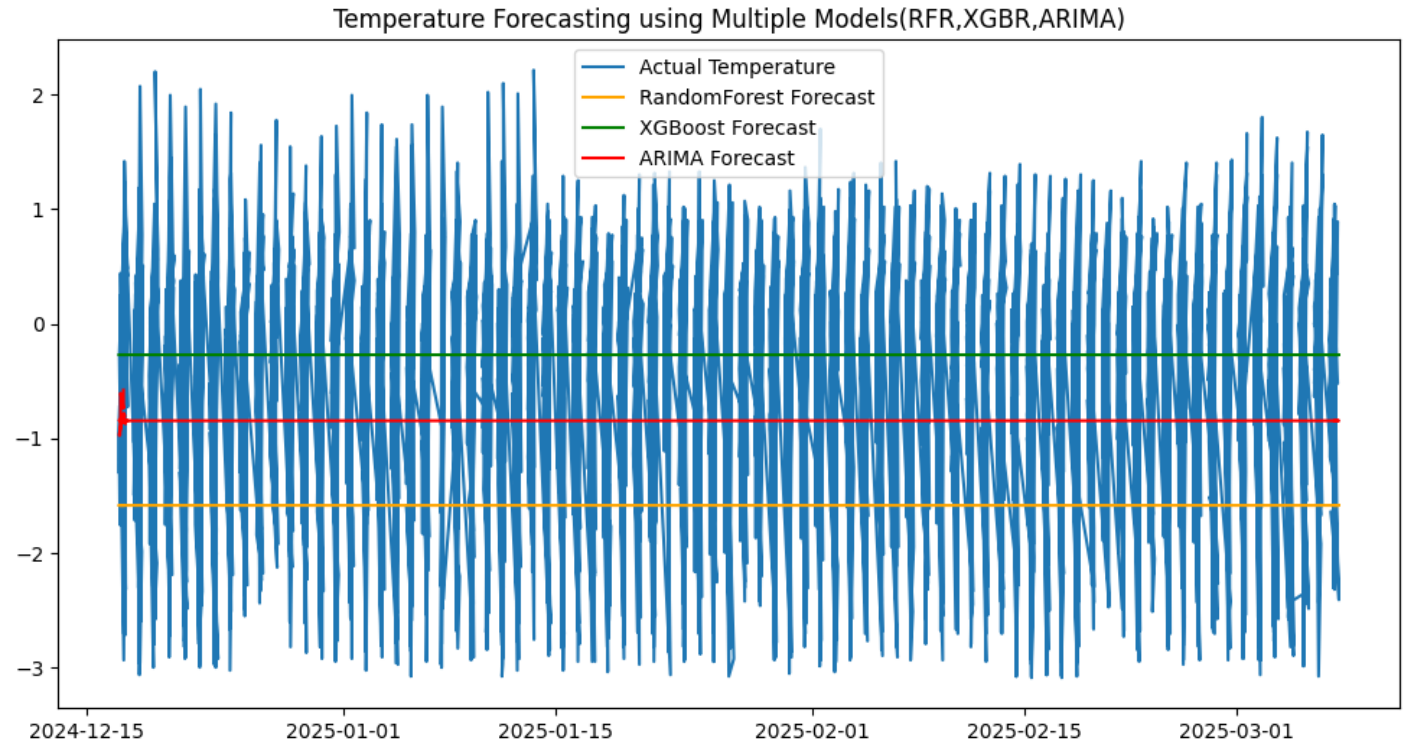
  For XGBR mode, mse = 1.3419 and rmse = 1.1584

- **ARIMA Model:**

  Model trained and made predictions.

  Evaluation of ARIMA model is done using MSE and RMSE.

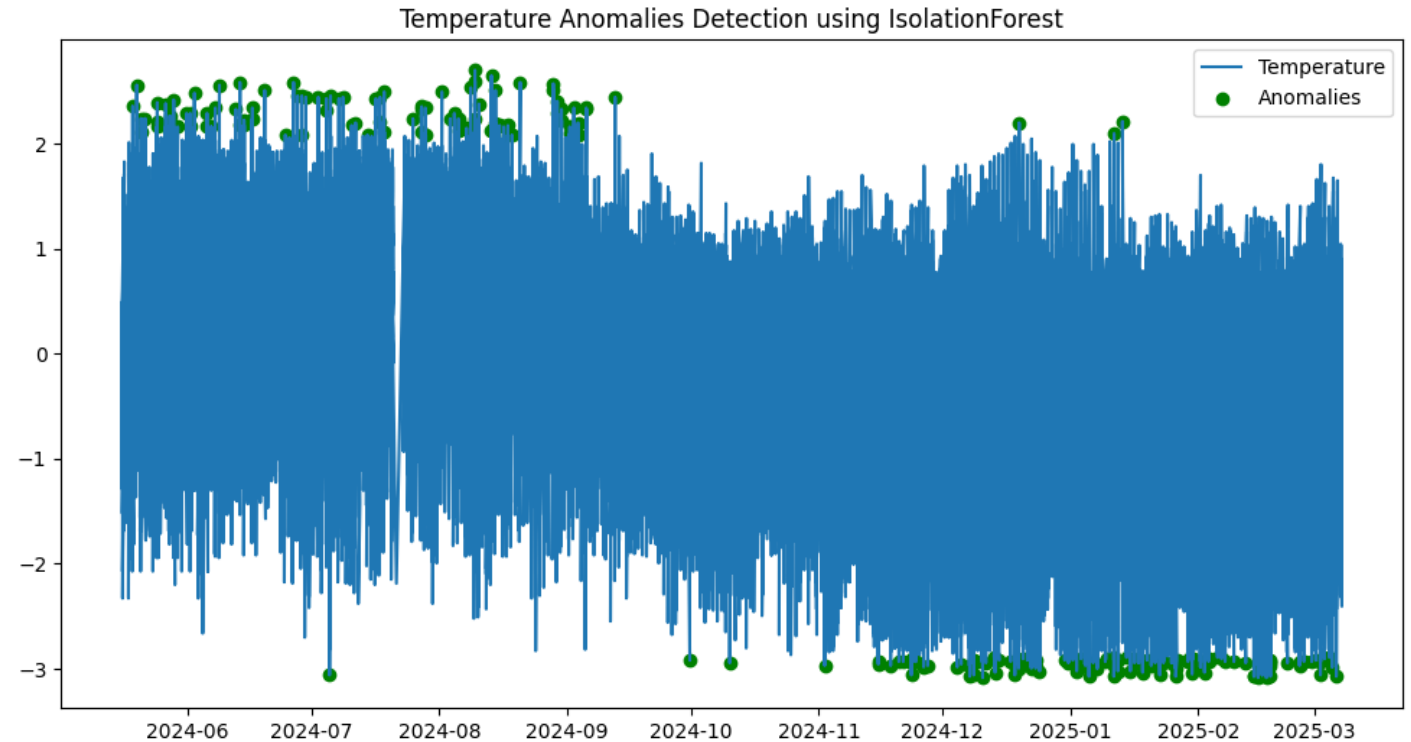 For ARIMA model, mse = 1.8909 and rmse = 1.3455.

Temperature Forecasting using Multiple Models(RFR,XGBR,ARIMA)

ANOMALY DETECTION

- Isolation Forest:

    -> Using Isolation Forest technique detected some anomalies

    -> Visualization of anomalies.

Temperature Anomalies Detection using IsolationForest

ENSEMBLE MODELS

- Combining Models:

  -> Combined XGBoost and ARIMA for better accuracy of ensemble of models.

  -> Ensemble of models performance evaluation is done using MSE and RMSE.

  -> For Ensemble of model, mse = 1.2918 and rmse = 1.1633

  ->Visualization of ensemble forecasts.

Temperature Forecasting using Ensemble of Models
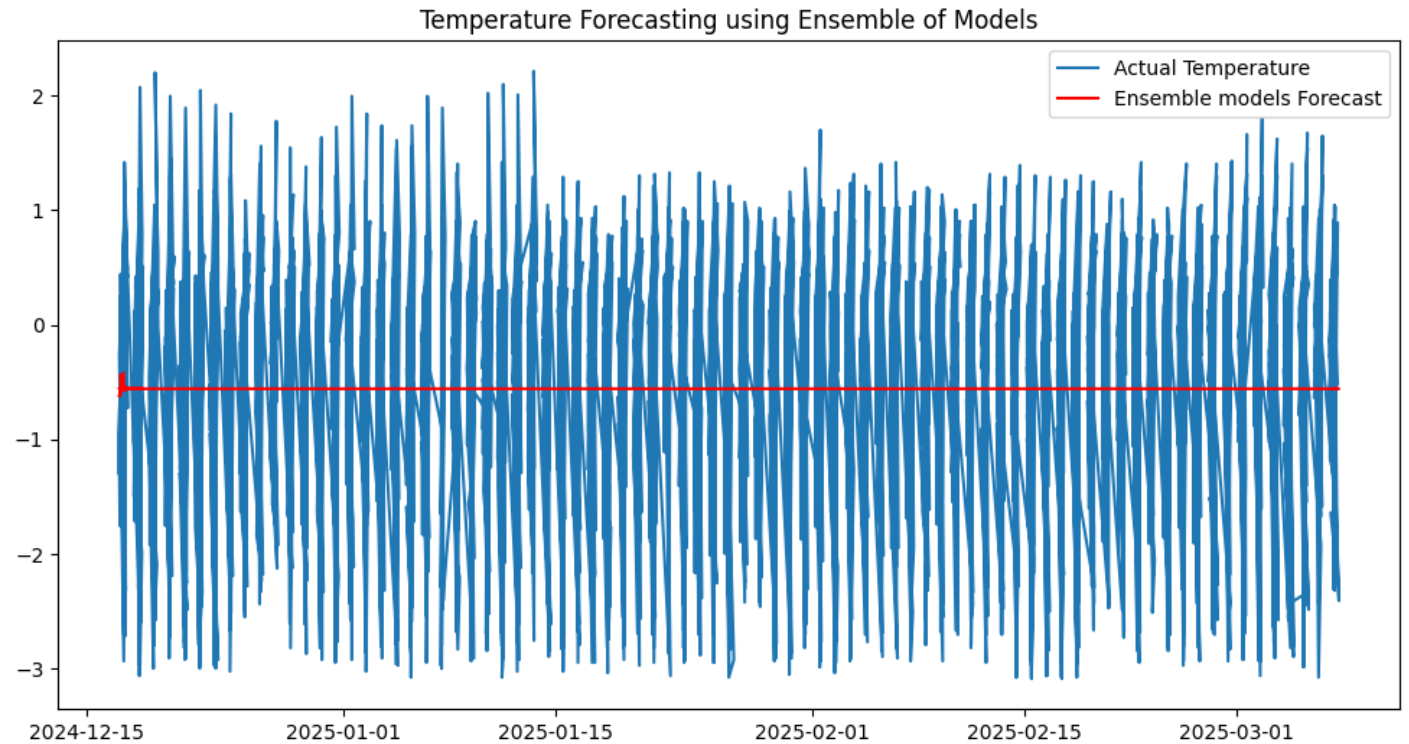
UNIQUE ANALYSES

- Climate Analysis:

- Studied long-term climate patterns and variations in different regions(countries).

- Environmental Impact:

- Analyzed air quality and correlation with various weather parameters.

- Spatial Analysis:

- Generated Geographical patterns of temperature using folium and HeatMap.

- Visualized the average temperature by country.

- Feature Importance:

- Displaying feature importance from RandomForest model

SPATIAL ANALYSIS VISUALIZATION USING MAP

GEOGRAPHICAL PATTERNS

- Geographical patterns of countries based on average temperature.
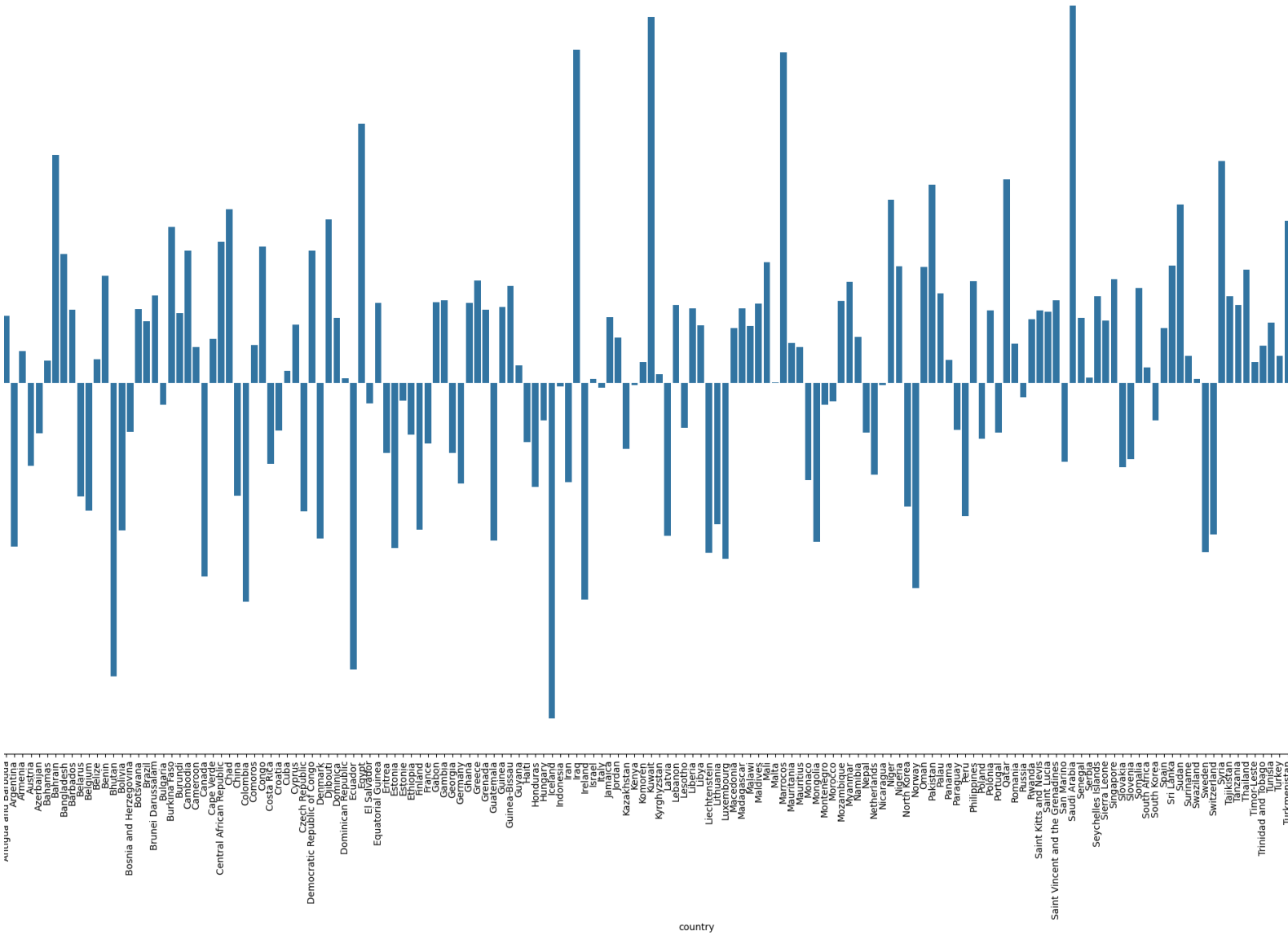
CONCLUSION AND INSIGHTS

- Outliers in data are detected and handled using IQR.

- visualized some anomalies and handled them using IsolationForest.

- XGBoost Regressor model forecasted the best weather patterns than other models.

- Ensemble of models also performed well and ready to use for real time weather forecast.

- Geo spatial analysis is done and displayed the results in map.

- Saudi Arabia and Iceland has highest average temperature and Dominican Republic has lowest average temperature.

THANK YOU