Weather Forecasting and Analysis Project Report

# 1. Introduction

This project focuses on weather data analysis and temperature prediction using real-world meteorological data. The aim was to understand the patterns across cities and years, build models to forecast temperatures, and identify anomalies.

# 2. Data Collection and Preprocessing

- Historical weather data was collected and filtered.

- Unnecessary columns were dropped, and missing values were imputed using median strategy.

- Feature engineering was performed to derive useful features like month, hour, and weather conditions.

# 3. City-wise Analysis (2020)

- Data was grouped city-wise for the year 2020.

- Monthly averages of temperature, humidity, and wind speed were plotted for each city.

- These plots help visualize seasonal variations and compare city-specific patterns.

# 4. Year-wise Trend Analysis (2017–2020)

- Temperature trends were compared year-wise using line plots.

- Humidity levels were visualized using boxplots to show distributions.

- Wind speed was also compared across years.

- This helped in identifying consistent patterns or deviations in recent years.

# 5. Feature Importance

- A Random Forest model was trained for temperature prediction.

- Feature importance was extracted and plotted to understand the influence of different variables.

- Features like humidity and wind speed emerged as significant predictors.

# 6. Anomaly Detection

- Isolation Forest algorithm was applied to detect anomalies in the data.

- Scatter plots helped visualize outliers in the temperature-humidity relationship.

- This ensured data quality and highlighted abnormal conditions.

# 7. Temperature Prediction Models

- Three models were used: Linear Regression, Random Forest, and KNN.

- Model predictions were plotted against actual values to compare performance.

- Predictions were visualized with both sample indices and actual datetime for clarity.

# 8. Visualization and Web Integration

- Matplotlib plots were generated for all analyses.

- These plots were saved as images and embedded into a web interface.

- City-wise and year-wise comparison charts were also displayed dynamically on the webpage.

# 9. Conclusion

- The project successfully analyzed weather data across cities and years.

- Models built for prediction were able to reasonably forecast temperatures.

- Visualization offered insights into seasonal and regional differences.

- The web-based dashboard makes the analysis accessible and interpretable.