



PREDICTING TOMORROW'S RAINFALL USING MACHINE LEARNING

A CLASSIFICATION APPROACH WITH WEATHER DATA
GROUP 15

PROBLEM STATEMENT

Objective: To develop a classification model that predicts whether it will rain tomorrow based on historical weather data.

Significance: Accurate rainfall predictions are crucial for agriculture, disaster management, and daily planning.

Challenge: Traditional forecasting methods may lack precision, especially for localized predictions.



INTRODUCTION



- Weather Forecasting: The process of predicting atmospheric conditions using data and models.
- Machine Learning in Meteorology: Leveraging algorithms to identify patterns in historical weather data for predictive insights.
- Classification Task: Binary classification—predicting 'Rain' or 'No Rain' for the next day.

DATASET OVERVIEW

Source: Historical weather data from [<https://www.kaggle.com/datasets/jsphyg/weather-dataset-rattle-package>].

Features: Temperature, humidity, wind speed, atmospheric pressure, cloud cover, etc.

Target Variable: RainTomorrow (Yes/No)

Preprocessing: Handling missing values, encoding categorical variables, feature scaling.



METHODOLOGY OVERVIEW

Data Collection: Gathered historical weather data.

Data Preprocessing: Cleaned and prepared data for modeling.

Model Selection: Evaluated multiple classification algorithms.

Training & Testing: Split data into training and testing sets.

Evaluation: Assessed models using performance metrics.



CLASSIFICATION ALGORITHMS USED



Logistic Regression: Statistical model for binary outcomes.



Random Forest: Ensemble of decision trees for improved accuracy.



K-Nearest Neighbors (KNN): Classifies based on proximity to neighbors.

Decision Tree: Model that splits data based on feature values.

Support Vector Machine (SVM): Finds optimal hyperplane for classification.

Ensemble Methods: Combines multiple models for better performance.

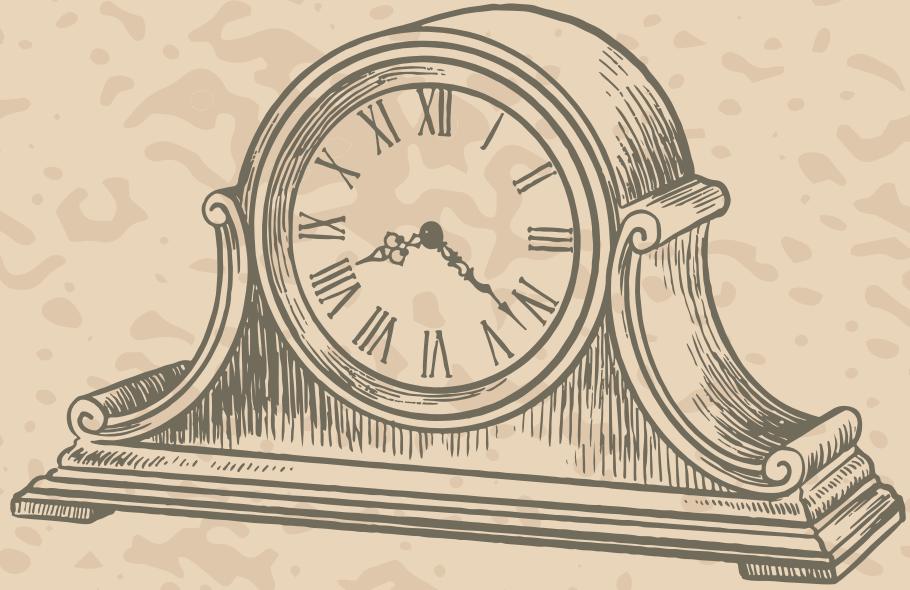
MODEL TRAINING & VALIDATION

Data Split: 80% training, 20% testing.

Cross-Validation: Used k-fold cross-validation for robustness.

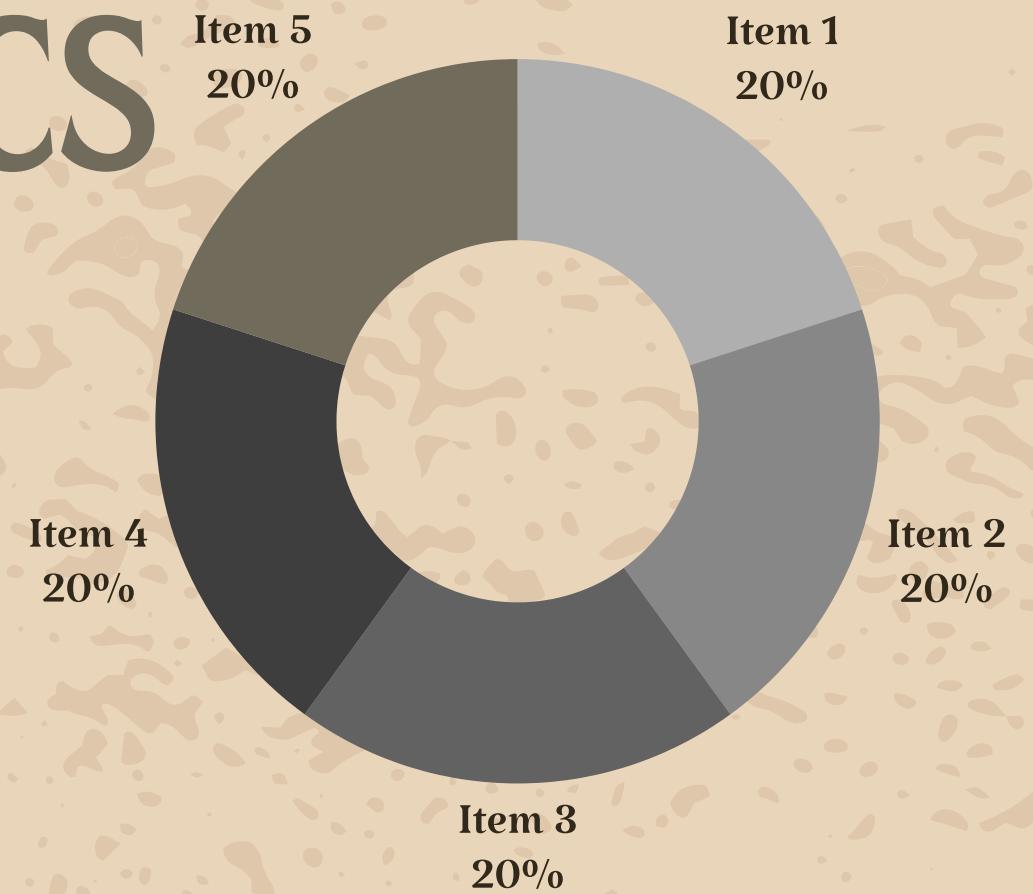
Hyperparameter Tuning: Optimized model parameters using grid search.

Tools: Implemented using Python libraries like scikit-learn.

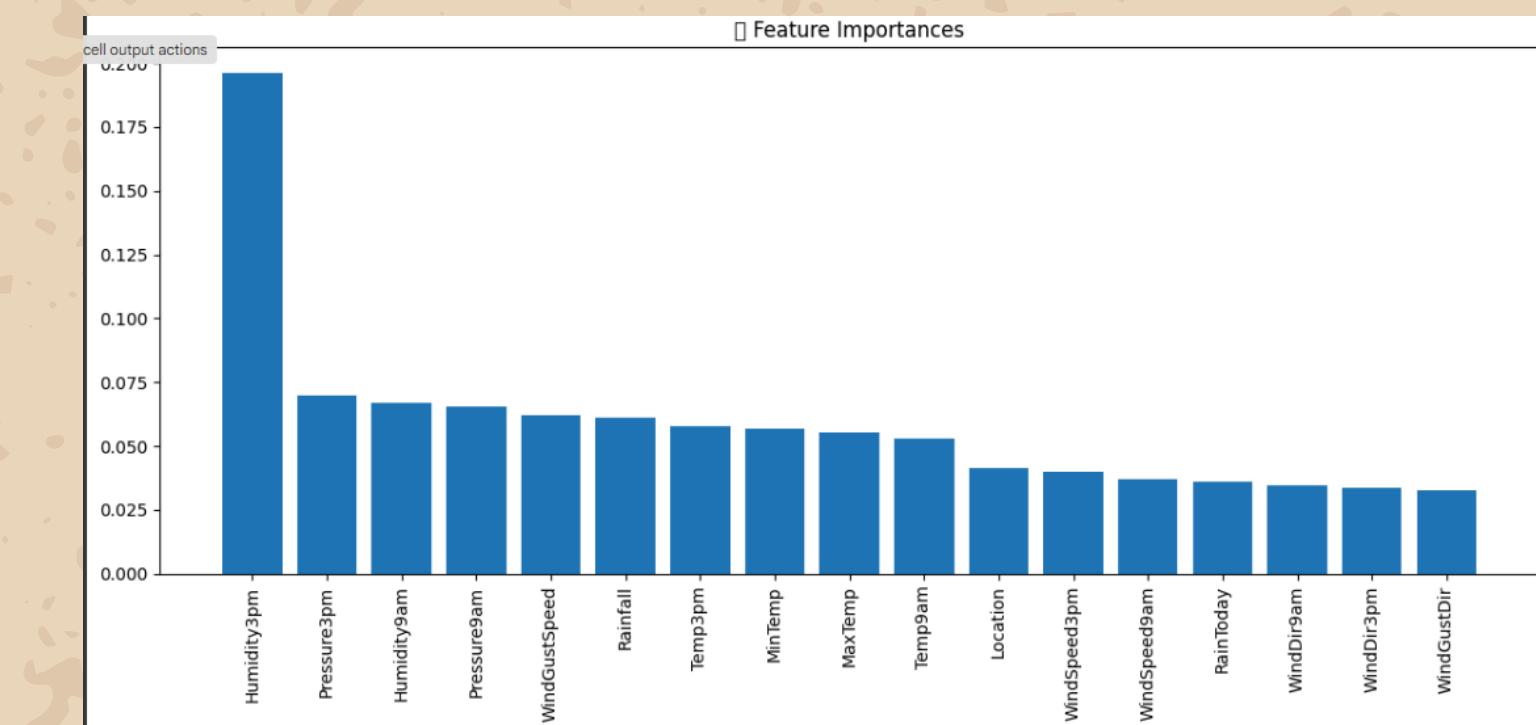
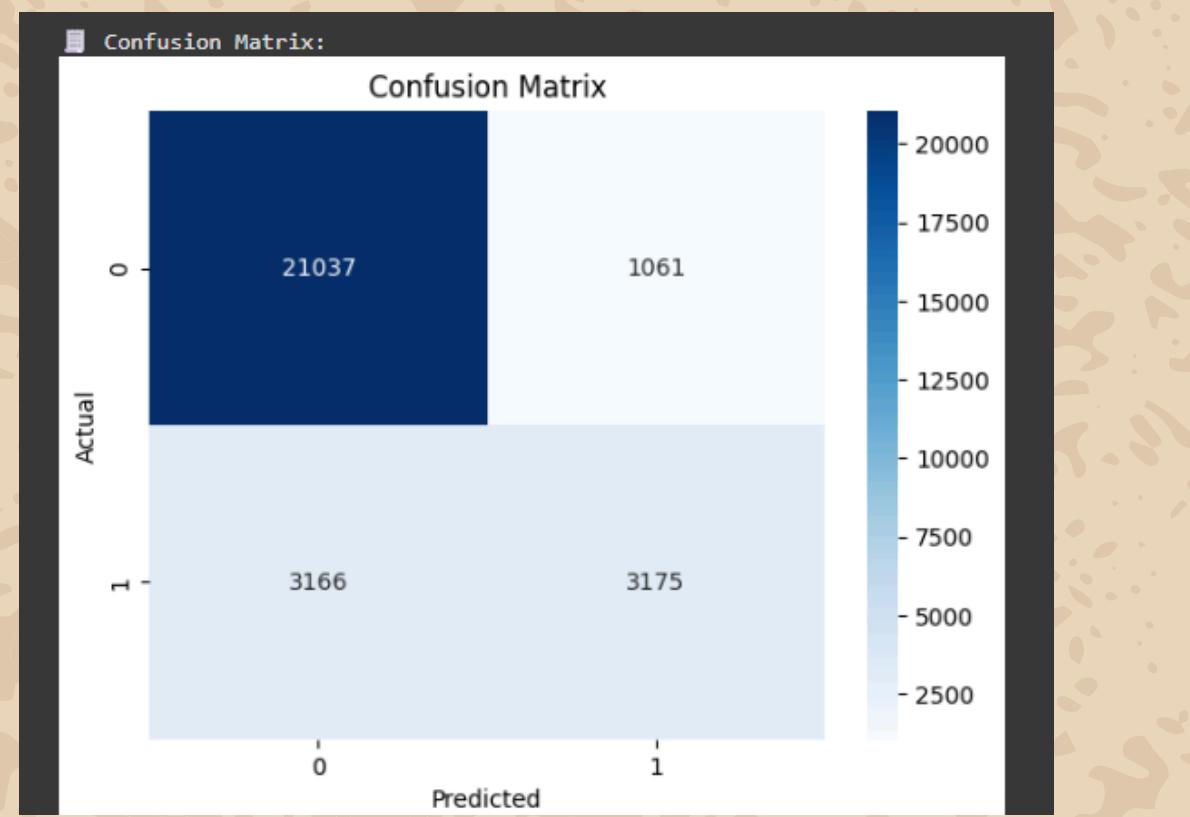
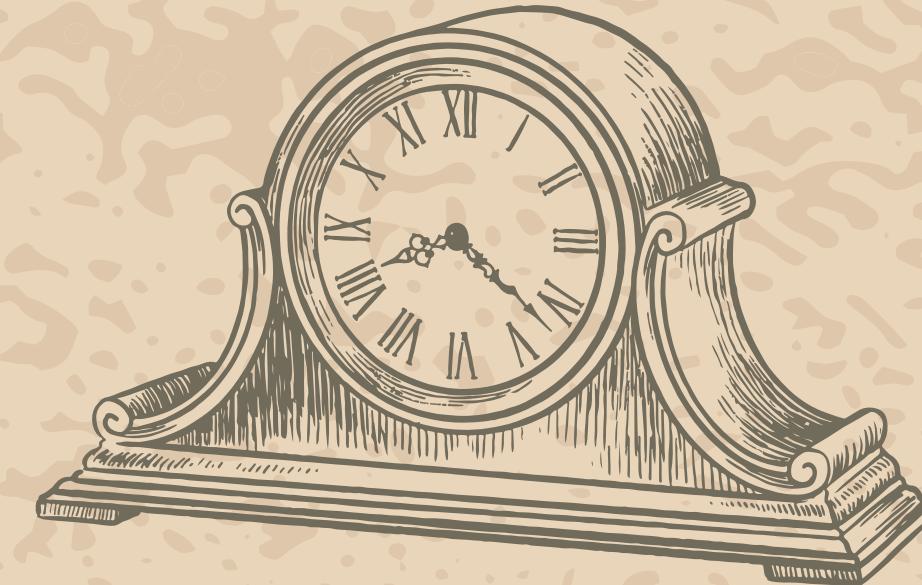
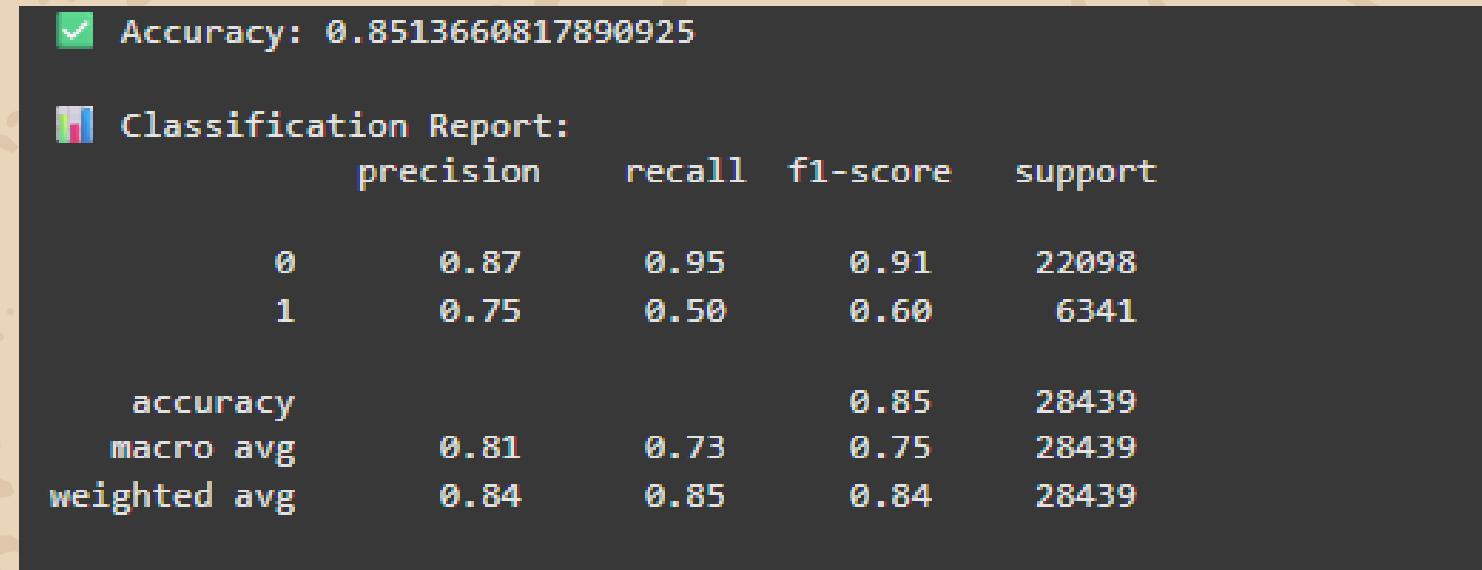


PERFORMANCE METRICS

- Accuracy: Proportion of correct predictions.
- Precision: True positives divided by total predicted positives.
- Recall (Sensitivity): True positives divided by total actual positives.
- F1 Score: Harmonic mean of precision and recall.
- Confusion Matrix: Displays true vs. predicted classifications.



RESULTS SUMMARY



CONFUSION MATRIX EXAMPLE

True Positives (TP): Correctly predicted rain.

True Negatives (TN): Correctly predicted no rain.

False Positives (FP): Incorrectly predicted rain.

False Negatives (FN): Missed predictions of rain.



DISCUSSION

Best Performing Model: Ensemble Classifier with highest accuracy.

Trade-offs: Higher recall but lower precision indicates more false positives.

Model Selection: Depends on application—e.g., for agriculture, higher recall may be preferred to avoid missing rain predictions.



CONCLUSION

Achievement: Developed a machine learning model to predict rainfall with reasonable accuracy.

Implications: Can aid in planning and decision-making in weather-dependent sectors.

Future Work: Incorporate more features, explore deep learning models, and test on real-time data.



REFERENCES

"Weather Forecasting Using Machine Learning Techniques: Rainfall and Temperature Analysis," Preprints.org, 2024.

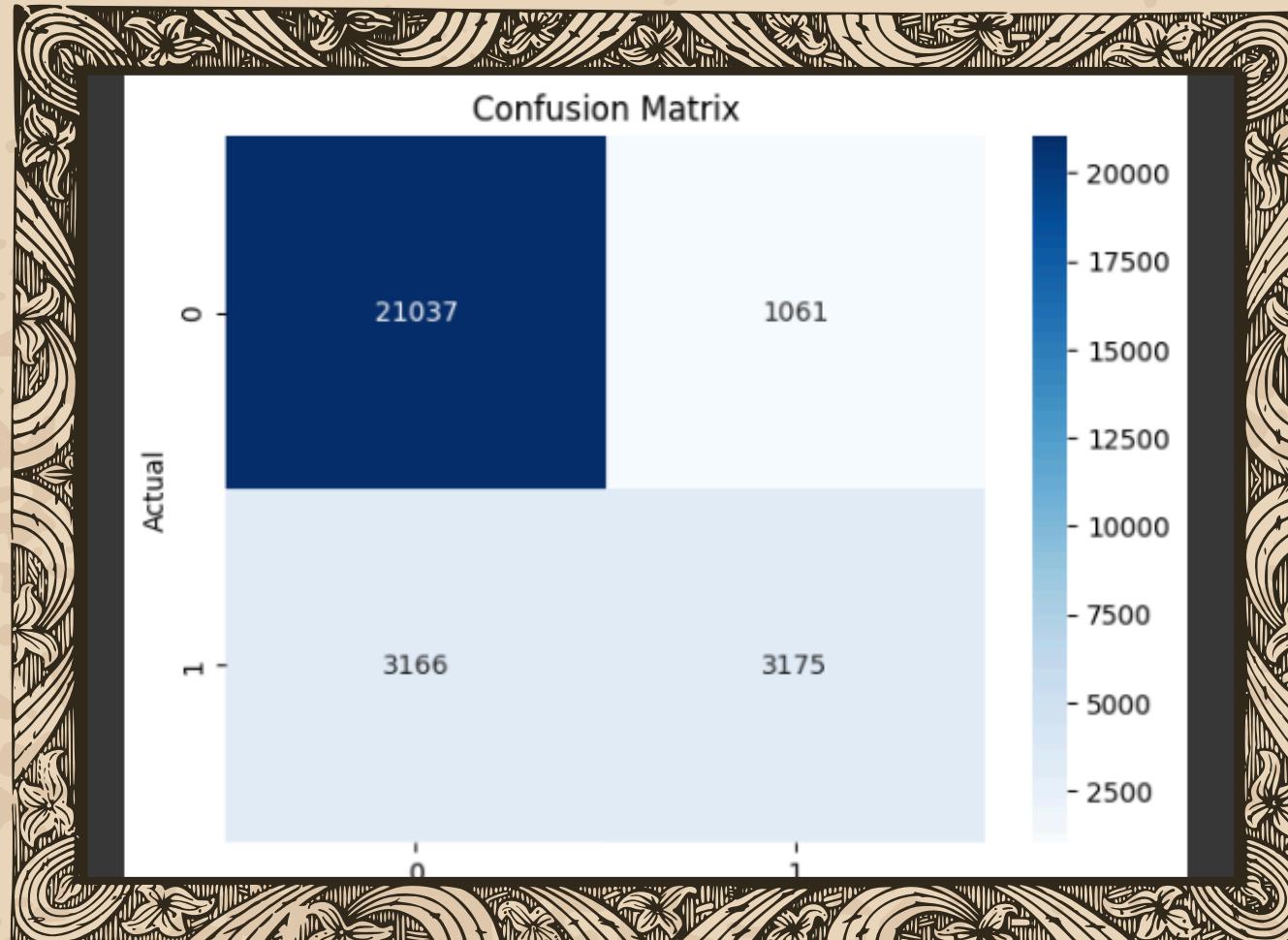
"Machine Learning-based Algorithms for Weather Forecasting," ResearchGate, 2022.

Wikipedia articles on F-score, Precision and Recall, Confusion Matrix, Brier Score.

Kagglelink:
<https://www.kaggle.com/datasets/jsphyg/weather-dataset-rattle-package>



DOCUMENTATION



User today's weather data to predict if it will rain tomorrow.

- Temperature (°C): 25
- Temperature (°C): 38
- Humidity (mm): 3
- Wind Speed (km/h): 40
- Wind Speed at 9am (km/h): 15
- Wind Speed at 3pm (km/h): 19
- Cloudy at 9am (%): 22
- Cloudy at 3pm (%): 26
- Pressure at 9am (hPa): 17
- Pressure at 3pm (hPa): 12
- Temperature at 9am (°C): 32
- Temperature at 3pm (°C): 31

Will rain today? (Yes/No): No

Prediction: It will ☀ not rain tomorrow.

THANK YOU