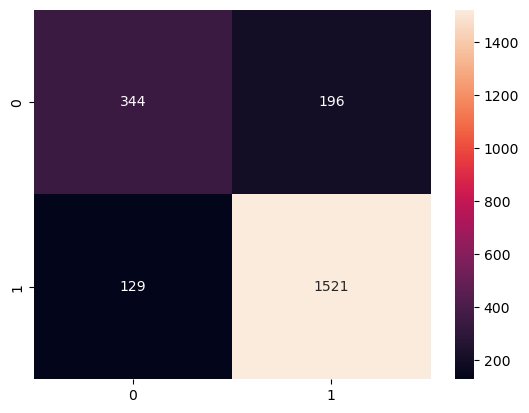
**Data Modeling Report:**

I trained four different models but selected two to compare here ..

#### **Rainfall Prediction using XGBoost**



#### **Overview**

#### This project involves training an XGBoost classification model to predict whether it will rain based on various weather-related features. The dataset includes temperature, humidity, wind speed, and other atmospheric variables.

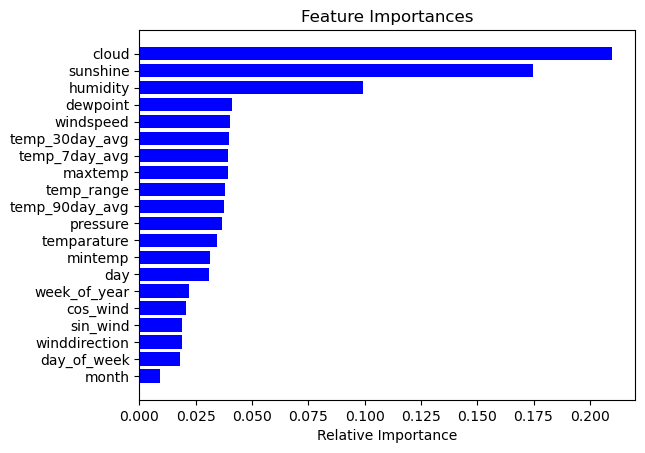
#### **Model Performance**

#### **Confusion Matrix Analysis:**

* + True Positives: 1521
  + False Positives: 196
  + True Negatives: 344
  + False Negatives: 129

**ROC Curve:** The model achieved an **AUC score of 0.78**, indicating good predictive performance.

**Feature Importance:** The most influential features in the model include **day, pressure, maximum temperature, wind speed, and sunshine levels**.



**Findings and Conclusion**

The model demonstrates reasonable accuracy in predicting rainfall. Further improvements could involve hyperparameter tuning, feature engineering, or incorporating additional weather data sources to enhance predictions.

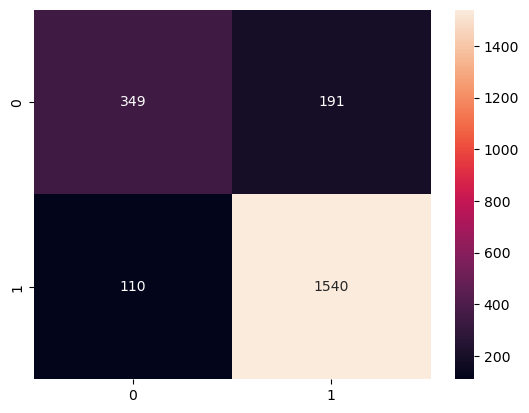
* Rainfall Prediction using Random Forest

Overview

This project involves training a Random Forest classification model to predict whether it will rain based on various weather-related features. The dataset includes temperature, humidity, wind speed, and other atmospheric variables.

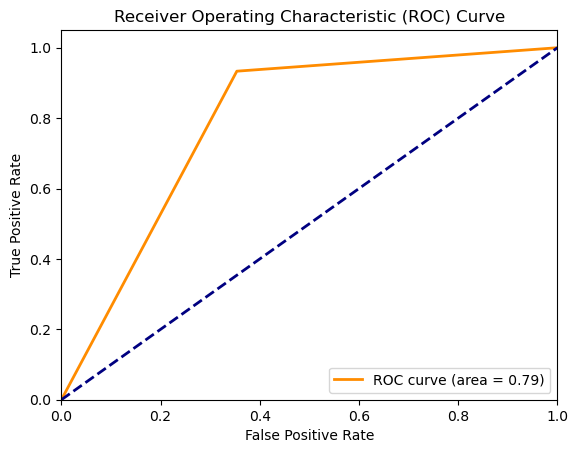
Model Performance

Confusion Matrix Analysis:

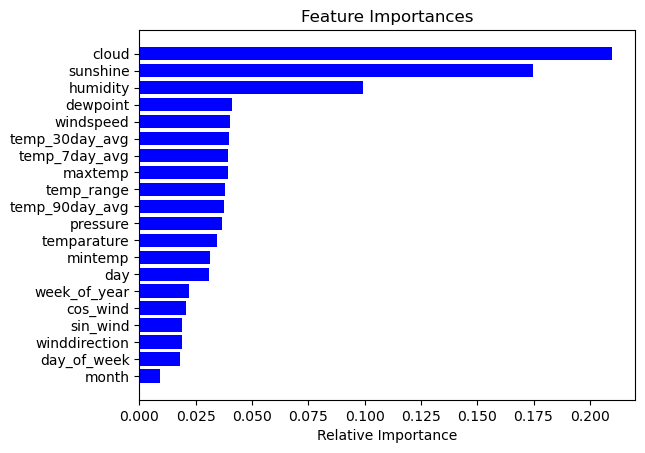


* True Positives: 1540
* False Positives: 191
* True Negatives: 349
* False Negatives: 110

ROC Curve: The model achieved an AUC score of 0.79, showing good predictive capability.



Feature Importance: The most influential features in the model include cloud cover, sunshine, humidity, and dew point.



Findings and Conclusion

The model performs well, with slightly improved performance compared to XGBoost. Cloud cover and sunshine appear to be key indicators of rainfall.