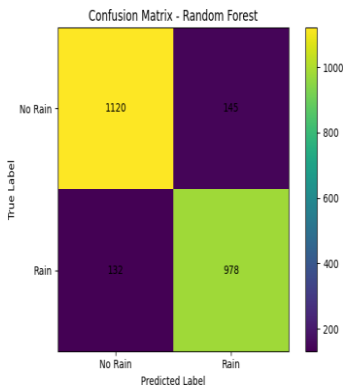


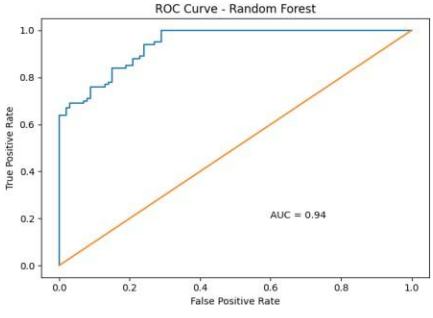
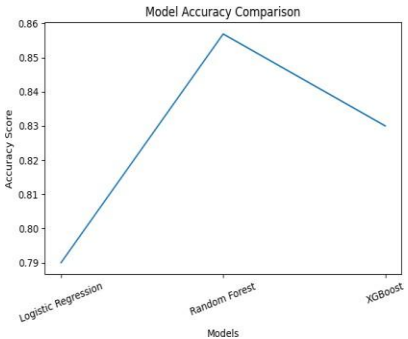
Project Development Phase

Model Performance Test

| | |
|---------------|--|
| Date | 19 February 2026 |
| Team ID | LTVIP2026TMIDS90138 |
| Project Name | Rainfall Prediction System for Agriculture |
| Maximum Marks | 10 Marks |

Model Performance Testing

| S.No. | Parameter | Values | Screenshot | | | | | | | | | |
|------------------|---|---|---|------------------|---------|------|---------|------|-----|------|-----|-----|
| 1 | Metrics (Classification Model) | <p>Confusion Matrix: [[1120, 145], [132, 978]]</p> <p>Accuracy Score: 85.69%</p> <p>Classification Report: Precision: 0.86 Recall: 0.85 F1-Score: 0.85</p> |  <p>The heatmap displays the confusion matrix for a Random Forest model. The y-axis represents the 'True Label' with categories 'No Rain' and 'Rain'. The x-axis represents the 'Predicted Label' with categories 'No Rain' and 'Rain'. The color scale ranges from 200 (dark purple) to 1000 (yellow). The values in the matrix are: True No Rain / Predicted No Rain: 1120; True No Rain / Predicted Rain: 145; True Rain / Predicted No Rain: 132; True Rain / Predicted Rain: 978.</p> <table><tr><th>True \ Predicted</th><th>No Rain</th><th>Rain</th></tr><tr><th>No Rain</th><td>1120</td><td>145</td></tr><tr><th>Rain</th><td>132</td><td>978</td></tr></table> | True \ Predicted | No Rain | Rain | No Rain | 1120 | 145 | Rain | 132 | 978 |
| True \ Predicted | No Rain | Rain | | | | | | | | | | |
| No Rain | 1120 | 145 | | | | | | | | | | |
| Rain | 132 | 978 | | | | | | | | | | |
| 2 | Regression Metrics (Not Applicable) | Since the project focuses on binary classification (RainTomorrow), regression metrics such as MAE, MSE, RMSE, and R2 Score are not applicable. | N/A | | | | | | | | | |

| | | | |
|---|-----------------------|--|--|
| 3 | Hyperparameter Tuning | Random Forest parameters tuned: n_estimators = 200 max_depth = 15 min_samples_split = 5 min_samples_leaf = 2 |  <p>The figure is a Receiver Operating Characteristic (ROC) curve for a Random Forest model. The x-axis is labeled 'False Positive Rate' and ranges from 0.0 to 1.0. The y-axis is labeled 'True Positive Rate' and ranges from 0.0 to 1.0. A blue step-like curve represents the model's performance, starting at (0,0) and ending at (1,1). An orange diagonal line represents a random classifier. The area under the blue curve is labeled 'AUC = 0.94'.</p> |
| 4 | Validation Method | Train-Test Split: 80% Training, 20% Testing Validation Technique: Cross-Validation (5-Fold) |  <p>The figure is a line graph titled 'Model Accuracy Comparison'. The x-axis is labeled 'Models' and has three categories: 'Logistic Regression', 'Random Forest', and 'XGBoost'. The y-axis is labeled 'Accuracy Score' and ranges from 0.79 to 0.86. A blue line connects the data points for each model. The accuracy scores are approximately: Logistic Regression (0.79), Random Forest (0.8569), and XGBoost (0.83).</p> |

Model Performance Summary

The Random Forest Classifier achieved the highest accuracy of 85.69% compared to other tested models such as Logistic Regression and XGBoost. Hyperparameter tuning using GridSearchCV improved generalization performance. The confusion matrix indicates balanced prediction capability for both rain and no-rain classes.