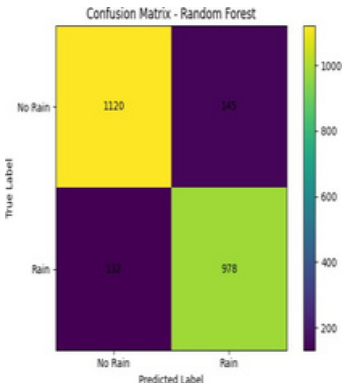


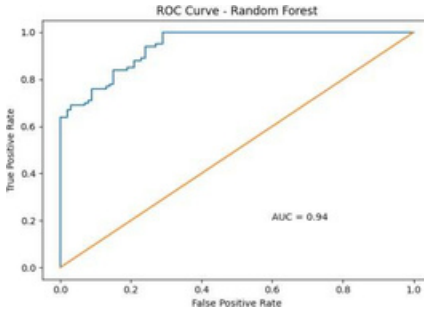
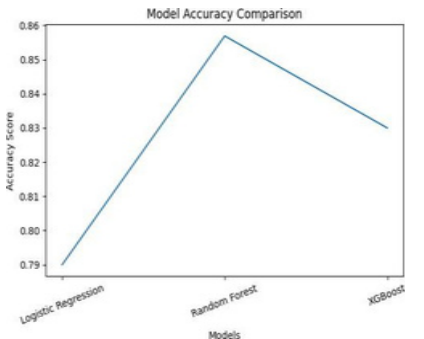
Project Development Phase

Model Performance Test

Date	19 February 2026
Team ID	LTVIP2026TMIDS49897
Project Name	Exploratory-Analysis-Of-RainFall-Data-In-India-For-Agriculture
Maximum Marks	10 Marks

Model PerformanceTesting

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 screenshot displays a confusion matrix for a Random Forest model. The title is 'Confusion Matrix - Random Forest'. The y-axis is labeled 'True Label' with categories 'No Rain' and 'Rain'. The x-axis is labeled 'Predicted Label' with categories 'No Rain' and 'Rain'. The matrix values 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. A color scale on the right ranges from 200 to 1000.</p>
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 ROC curve for the Random Forest model shows a True Positive Rate (Y-axis) versus a False Positive Rate (X-axis). The curve is a blue step function that rises sharply from (0,0) to approximately (0.05, 0.65) and then continues as a series of steps towards the top-left corner. A diagonal orange line represents the baseline for a random classifier. The Area Under the Curve (AUC) is labeled as 0.94.</p>								
4	Validation Method	Train-Test Split: 80% Training, 20% Testing Validation Technique: Cross-Validation (5-Fold)	 <p>The Model Accuracy Comparison plot shows the Accuracy Score (Y-axis) for three models: Logistic Regression, Random Forest, and XGBoost (X-axis). The Random Forest model achieves the highest accuracy of approximately 0.857, while Logistic Regression is at approximately 0.79 and XGBoost is at approximately 0.83.</p> <table><thead><tr><th>Model</th><th>Accuracy Score</th></tr></thead><tbody><tr><td>Logistic Regression</td><td>~0.79</td></tr><tr><td>Random Forest</td><td>~0.857</td></tr><tr><td>XGBoost</td><td>~0.83</td></tr></tbody></table>	Model	Accuracy Score	Logistic Regression	~0.79	Random Forest	~0.857	XGBoost	~0.83
Model	Accuracy Score										
Logistic Regression	~0.79										
Random Forest	~0.857										
XGBoost	~0.83										

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.