

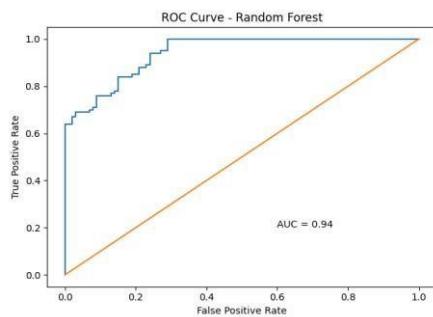
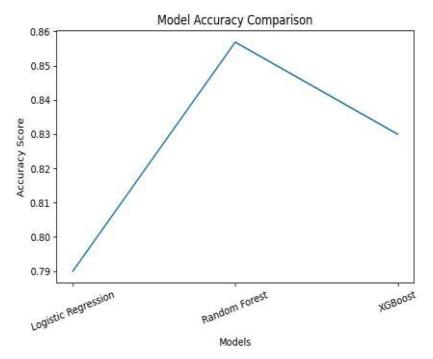
## Project Development Phase

### Model Performance Test

Date	17 February 2026
Team ID	LTVIP2026TMIDS48850
Project Name	Exploratory analysis of Rainfall data in India 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 figure is a heatmap titled "Confusion Matrix - Random Forest". It has "True Label" on the y-axis and "Predicted Label" on the x-axis, both with categories "No Rain" and "Rain". The matrix values are: True Label   Predicted Label No Rain   [1120, 145] Rain   [132, 978]. A color scale bar on the right ranges from dark purple (~200) to bright yellow (~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	
4	Validation Method	Train-Test Split: 80% Training, 20% Testing Validation Technique: CrossValidation (5-Fold)	

### 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.