Model Development Phase Template

Date	15 April 2024
Team ID	Team-738164
Project Title	Rainfall Prediction Using Machine Learning
Maximum Marks	6 Marks

Model Selection Report:

Model	Description	Hyperparameters	Performance Metric (e.g., Accuracy, F1 Score)
Logistic Regression	A linear model used for classification problems. It predicts the probability of a sample belonging to a specific class.	C (Regularization strength), fit_intercept (Whether to fit an intercept term), max_iter (Maximum number of iterations)	Accuracy: 79% AUC: 0.86 (Baseline Scores)
Random Forest	An ensemble learning method that combines multiple decision trees. It is robust to overfitting and can handle high-dimensional data.	n_estimators (Number of decision trees), criterion (Splitting function - gini or entropy), max_depth (Maximum depth of each tree), min_samples_split (Minimum samples required for a split), min_samples_leaf (Minimum samples allowed in a leaf node)	Accuracy: 85% AUC: 0.88 (Baseline Scores)

Decision Tree	A tree-based model that makes predictions by splitting the data based on features. It is interpretable and can handle both categorical and numerical features.	criterion (Splitting function - gini or entropy), max_depth (Maximum depth of the tree), min_samples_split (Minimum samples required for a split), min_samples_leaf (Minimum samples allowed in a leaf node)	Accuracy: 79% AUC: 0.70 (Baseline Scores)
XGBoost	An advanced implementation of gradient boosting that uses decision trees as base learners. It is known for its speed, accuracy, and ability to handle complex data.	n_estimators (Number of decision trees), max_depth (Maximum depth of each tree), learning_rate (Step size in the boosting process)	Accuracy: 86% AUC: 0.89 (Baseline Scores)