

Model Development Phase Template

| | |
|---------------|--|
| Date | 24 April 2024 |
| Team ID | Team-738169 |
| Project Title | Rainfall Prediction Using Machine Learning |
| Maximum Marks | 6 Marks |

Model Selection Report

In the forthcoming Model Selection Report, various models will be outlined, detailing their descriptions, hyperparameters, and performance metrics, including Accuracy or F1 Score. This comprehensive report will provide insights into the chosen models and their effectiveness.

Model Selection Report:

| Model | Description | Hyperparameters | Performance Metric (e.g., Accuracy, F1 Score) |
|---------------------|--|-----------------|---|
| Logistic Regression | Logistic regression is a statistical model used for binary classification tasks, where the outcome variable is categorical with two possible classes (e.g., yes/no, true/false, 0/1). It's called "logistic" | - | Accuracy Score = 77 % |
| Random Forest | Ensemble of decision trees; robust, handles complex relationships, reduces overfitting, and provides feature importance for loan approval prediction. | Yes | Accuracy Score = 83 % |

| | | | |
|---------------|---|-----|-----------------------|
| Decision Tree | Simple tree structure; interpretable, captures non-linear relationships, suitable for initial insights into loan approval patterns. | - | Accuracy Score = 80 % |
| KNN | Classifies based on nearest neighbors; adapts well to data patterns, effective for local variations in loan approval criteria. | - | Accuracy Score = 76 % |
| SVM | SVM aims to find the optimal hyperplane that best separates data points belonging to different classes in a high-dimensional space. | - | Accuracy Score = 75 % |
| XGBoost | XGBoost, short for eXtreme Gradient Boosting, is a highly scalable and accurate supervised learning algorithm known for its effectiveness in various machine learning tasks, particularly in structured/tabular data scenarios. | Yes | Accuracy Score = 84 % |