Model Development Phase Template

Date	20 July 2024
Team ID	SWTID1721319573
Project Title	Blueberry Yield Prediction
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)
Linear Regression	Linear Regression is a simple model that assumes a linear relationship between the independent variables and the dependent variable.	• fit_intercept: True • normalize: False (deprecated in newer versions; normalization is handled by StandardScaler if needed) • copy_X: True	Linear Regression: MAE: 97.318 MSE: 16219.955 RMSE: 127.358 R-Square: 0.992 Accuracy: 99.18%

Random Forest Regressor	Random Forest is an ensemble learning method that constructs multiple decision trees during training and outputs the mean prediction of the individual trees.	• n_estimators: 100 (default) • criterion: 'squared_error' (default) • max_depth: None (default, trees are expanded until all leaves are pure or until all leaves contain less than min_samples_split samples) • min_samples_split: 2 (default) • min_samples_leaf: 1 (default) • max_features: 'auto' (default)	Random Forest Regressor: MAE: 116.868 MSE: 22604.657 RMSE: 150.348 R-Square: 0.989 Accuracy: 98.86%
Decision Tree Regressor	Decision Tree Regressor creates a model in the form of a tree structure, where each node represents a	• criterion: 'squared_error' (default) • splitter: 'best' (default)	Decision Tree Regressor: MAE: 155.595 MSE: 42571.611 RMSE: 206.329 R-Square: 0.978 Accuracy: 97.84%

	decision based on the features, and the leaves represent the predicted values.	• max_depth : None (default) • min_samples_split: 2 (default) • min_samples_leaf: 1 (default)	
XGBoost Regressor	XGBoost is an advanced gradient boosting method that optimizes the performance of boosting algorithms and is known for its accuracy and efficiency.	• n_estimators: 100 (default) • learning_rate: 0.3 (default) • max_depth : 6 (default) • subsample: 1 (default) • colsample_bytree: 1 (default)	XGBoost Regressor: MAE: 111.927 MSE: 20632.639 RMSE: 143.641 R-Square: 0.990 Accuracy: 98.96%