

## Model Development Phase

Date	12 July 2024
Team ID	SWTID1720077079
Project Title	Wild Blueberry Yield Prediction
Maximum Marks	6 Marks

### Model Selection Report:

Model	Description	Hyperparameters	Performance Metric
Linear Regression	Linear Regression is a simple and interpretable model, but it assumes a linear relationship between the features and the target variable	fit_intercept=True copy_X=True n_jobs=None positive=False	Mean Absolute Error: 351.5273933689664 Root Mean Squared Error: 463.7929580320785 R2: 88.81392550043651
Decision Tree	It constructs a tree-like model of decisions and their possible consequences.	random_state=42 max_depth=5	Mean Absolute Error: 421.6096623777866 Root Mean Squared Error: 539.5911930066827 R2: 82.80694144829309
Random Forest	Random Forest is a collection of individual Decision Tree models, where each tree is trained on a random subset	n_estimators=100 random_state=42 max_depth=5	Mean Absolute Error: 382.0077129253888 Root Mean Squared Error: 499.75198453244883 R2: 84.93700199371773

	of the training data and a random subset of the features.		
XGBoost	The XGBoost model is a gradient boosting algorithm used for regression tasks, with the objective function set to 'reg:squarederror' to optimize the squared error loss, which is appropriate for regression problems.	max_depth=5 n_estimators=100 learning_rate=0.1	Mean Absolute Error: 180.54001388799836 Root Mean Squared Error: 260.16663946930686 R2: 96.60866093301176
SVM Regression	SVM Regression is a supervised learning algorithm used for solving regression problems. It works by finding the best-fitting hyperplane in a high-dimensional feature space that minimizes the error between the predicted and actual target values.	C=100, epsilon=0.001 gamma='auto' kernel='linear'	Mean Absolute Error: 455.29076862926655 Root Mean Squared Error: 586.6050431338977 R2: 81.513327311015