Model Development Phase Template

Date	13 July 2024
Team ID	SWTID1720151584
Project Title	E-Commerce Shipping 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)
Support vector machine	Support Vector Machines (SVMs) classify data by finding the optimal hyperplane that maximizes the margin between classes, using support vectors (key data points) to define this boundary.	gamma='auto',C=5,kernel='rbf'	Accuracy: 0.6612737572687 F1 Score: 0.6518436904729
Logistic Regression	Logistic regression classifies data by estimating probabilities using a logistic function, ideal for binary classification tasks.	Default parameters used	Accuracy: 0.6467509640927 F1 Score: 0.6931470546290

Decision Tree Classifier	A Decision Tree classifier uses a tree-like model of decisions, splitting data into branches based on feature values to classify outcomes at leaf nodes.	criterion='entropy', random_state=0	Accuracy: 0.66570932 F1 Score: 0.71465442
K-neighbors Classifier	The K-Neighbors classifier predicts a data point's class based on the majority class of its K nearest neighbors, using distance metrics for proximity.	Default parameters used	Accuracy: 0.647323943 F1 Score: 0.705149700
Naive Bayes Classifier	Naive Bayes classifier predicts classes based on Bayes' theorem, assuming feature independence, and is effective for large datasets with categorical input features.	Default parameters used	Accuracy: 0.649859154 F1 Score: 0.603664302
XG boost Classifier	XGBoost classifier is an optimized gradient boosting algorithm, combining multiple weak learners to improve model accuracy, speed, and performance, often used in machine learning competitions.	Default parameters used	Accuracy: 0.649464751 F1 Score: 0.596345850
Ada Boost Classifier	AdaBoost classifier combines multiple weak classifiers into a strong one by adjusting weights iteratively, focusing more on misclassified instances to improve overall accuracy.	Default parameters used	Accuracy: 0.670478742 F1 Score: 0.693458629
Gradient Boost Classifier	Gradient Boost classifier builds models sequentially, each correcting errors of the previous ones, by minimizing a loss function, resulting in high predictive accuracy for complex datasets.	Default parameters used	Accuracy: 0.684436619 F1 Score: 0.689547042

Random Forest Classifier	A Random Forest classifier is an ensemble learning method that combines multiple decision trees to improve accuracy, reduce overfitting, and enhance predictive performance.	Default parameters used	Accuracy: 0.67496958 F1 Score: 0.70955689