

## Model Development Phase Template

Date	10 July 2024
Team ID	SWTID1720084775
Project Title	ECommerce Shipping Prediction Using Machine Learning
Maximum Marks	6 Marks

### Model Selection Report

The following Model Selection Report outlines various models, 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)
LogisticRegression	It is a linear model used for binary classification. It models the probability of a binary outcome based on one or more predictor variables.	The LogisticRegression model is initialized with default hyperparameters.	Accuracy: 0.6467605633802817 F1 Score: 0.64314171883893
KNeighborsClassifier	It is a non-parametric algorithm that classifies a sample based on the majority class among its	n_neighbors=7: The number of neighbors to use.	Accuracy: 0.6473239436619719 F1 Score: 0.6251497005988024

	k-nearest neighbors in the feature space.		
SVC	SVM with a linear kernel finds the hyperplane that best separates the classes in the feature space. It can be used for both classification and regression tasks.	kernel="linear": Specifies the kernel type to be used in the algorithm.	Accuracy: 0.6676056338028169 F1 Score: 0.6428571428571428
GaussianNB	It is a probabilistic classifier based on Bayes' theorem, assuming independence between predictors. It works well with continuous data that follows a normal distribution.	The GaussianNB model is initialized with default hyperparameters.	Accuracy: 0.6698591549295775 F1 Score: 0.6536643026004728
RandomForest Classifier	It is an ensemble learning method that constructs multiple decision trees and merges them to get a more accurate and stable prediction. It uses bagging and feature randomness for better performance.	n_estimators=7: The number of trees in the forest. criterion='entropy': The function to measure the quality of a split. random_state=0: Controls the randomness of the estimator.	Accuracy: 0.643943661971831 F1 Score: 0.6425339366515836
XGBClassifier	It is an efficient and scalable implementation of gradient boosting framework. It builds		Accuracy: 0.6794366197183098 F1 Score: 0.6463642013673089

	<p>decision trees sequentially, where each tree tries to correct the errors of the previous one.</p>	<p>The XGBClassifier model is initialized with default hyperparameters.</p>	
ANN	<p>An Artificial Neural Network (ANN) for binary classification is designed to distinguish between two classes. It consists of interconnected layers of neurons that process input features to learn decision boundaries. The output layer typically uses a sigmoid activation function to produce a probability score between 0 and 1, which is then thresholded to classify inputs into one of the two classes.</p>	<p>Input Layer:</p> <p>units = 12,</p> <p>kernel_initializer = 'random_uniform',</p> <p>activation = 'relu'</p> <p>2 Hidden Layers:</p> <p>units = 12,</p> <p>kernel_initializer = 'random_uniform',</p> <p>activation = 'relu'</p> <p>Output Layer:</p> <p>units = 1 ,</p> <p>kernel_initializer = 'random_uniform',</p> <p>activation = 'sigmoid'</p> <p>Compile:</p> <p>optimizer = 'adam',</p> <p>loss = 'binary_crossentropy',</p> <p>metrics = ['accuracy']</p>	<p>Accuracy:</p> <p>0.7132394366197183</p> <p>F1 Score:</p> <p>0.6418015482054891</p>