

Model Development Phase Template

Date	July 2024
Team ID	740295
Project Title	Ecommerce Shipping Prediction using Machine Learning
Maximum Marks	5 Marks

Model Selection Report

In the model selection report for future deep learning and computer vision projects, various architectures, such as CNNs or RNNs, will be evaluated. Factors such as performance, complexity, and computational requirements will be considered to determine the most suitable model for the task at hand.

A model selection report outlines the process of evaluating and choosing the most suitable machine learning model for a specific task, detailing criteria such as performance metrics, computational efficiency, interpretability, and suitability for the dataset's characteristics to justify the final model choice.

Model Selection Report:

Model	Description
Logistic Regression Classifier	The logistic regression classifier is often selected for ecommerce shipping prediction due to its simplicity, interpretability, and effectiveness in handling binary classification problems. It provides probabilistic predictions, making it easy to understand and implement, while performing well with large datasets and requiring less computational power compared to more complex models.
Decision Tree Classifier	The Decision Tree Classifier is chosen due to its ability to handle non-linear relationships, interpretability in decision-making processes, and robustness in handling diverse types of data relevant to shipping logistics.
Random Forest Classifier	The Random Forest Classifier is ideal because it combines the strength of multiple decision trees, offering high accuracy, robust performance against overfitting, and the ability to handle large and complex datasets, ensuring reliable predictions in varied shipping scenarios.
Support Vector Machine Classifier	The SVM due to its effectiveness in handling high-dimensional data, ability to capture complex relationships between variables, and robustness in achieving high accuracy even with smaller datasets, making it suitable for precise shipping logistics predictions.

K-Nearest Neighbors Classifier	The K-NN for its simplicity in implementation, flexibility in handling various types of data, and effectiveness in capturing local patterns in shipping data, making it suitable for real-time prediction and adaptability to changing shipping conditions.
XGBoost Classifier	The XGBoost due to its superior performance in handling large datasets, capability to capture complex relationships in data, robustness against overfitting, and ability to optimize predictive accuracy through boosting techniques, ensuring reliable and efficient shipping logistics predictions.
Ridge Classifier	The Ridge Classifier due to its ability to handle multicollinearity in feature variables, regularization to prevent overfitting, and suitability for datasets where predictors are correlated, ensuring stable and reliable predictions in shipping logistics scenarios.