

Model Development Phase Template

Date	July 2024
Team ID	739663
Project Title	Estimating the stock keeping units 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
Linear Regression Classifier	The linear regression classifier is often selected for estimating the stocks 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 Regression	The Decision Tree Regression 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 SKUs logistics.
Random Forest Regression	The Random Forest Regression is ideal because it combines the strength of multiple decision trees, offering high accuracy, robust performance against over-fitting , and the ability to handle large and complex datasets, ensuring reliable predictions in varied SKUs 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 SKUs 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 stock data, making it suitable for real-time prediction and adaptability to changing stock estimating conditions.
XGB Classifier	The XGB due to its superior performance in handling large datasets, capability to capture complex relationships in data, robustness against over-fitting, and ability to optimize predictive accuracy through boosting techniques, ensuring reliable and efficient stock logistics predictions.
Ridge Classifier	The Ridge Classifier due to its ability to handle multi-collinearity in feature variables, regularization to prevent over-fitting, and suitability for datasets where predictors are correlated, ensuring stable and reliable predictions in stock logistics scenarios.