Project Initialization and Planning Phase

Date	15 March 2024
Team ID	SWTID1720116037
Project Title	Ecommerce Shipping Prediction Using Machine Learning
Maximum Marks	3 Marks

Project Proposal (Proposed Solution) report

The project reports aim to e-commerce shipping prediction using machine learning to estimate on-time delivery of products, considering parameters such as package origin, destination, chosen method of shipment, carrier, and potential delays. This can be effectively done using historical information and real-time updates given by carriers. Other variables that go into the model for enhancing its effectiveness include weather conditions and traffic. This functionality not only gives e-commerce the power to give time estimates for delivery but also fuels next-level customer experiences, by making shipping prediction among the most critical tools to meet modern customer satisfaction and drive operational efficiency.

Project Overview	
Objective	The primary objective of the project to develop a machine learning model that will predict with a high degree of accuracy the shipping time for any order an e-commerce orders. This prediction will enhance customer satisfaction, optimize logistics, and reduce operational costs. That will include a check on the history in shipping and some other factors, which may also include origin and destination, the chosen method of shipping by the customer, and delays or problems that may be met along the way of shipment.
Scope	The scope of the work pertains to data collection, preprocessing, model selection, training and validation, and deployment. It would focus on those factors thought to impact shipping times, such as order volume, geographical locations, carrier performance, weather conditions, and time of the year. It simply is confined to only the shipping phase of the e-commerce process .therefore, it will not affect other stages like order processing and inventory management.
Problem Statement	
Description	While developing an e-commerce company, the challenge is how to correctly predict whether the product will arrive at its destination on time. The reason for having dissatisfied customers and consequently a probable loss of revenue is late shipments. For this, we require a

	predictive model involving origin and destination of the package, shipping methods, and probable delays during transit. Our aim should be enhancing delivery accuracy by providing perfect shipment time estimation.
Impact	Accurately predicting on-time delivery for your products in itself poses a giant challenge to your e-commerce business. Delivery time affects customer satisfaction, impacting the entire experience and chances for repeat business. Ensuring that the goods are delivered promptly builds customer loyalty, keeps churning at bay, and hence sets the basis for long-term profitability. Accurate predictions will operationally facilitate better logistics planning, inventory, and resource allocation. Besides, dependable estimates of delivery time could prove to be one of the likely ways of differentiating your brand in a competitive market where reliability is a cardinal need for customers.
Proposed Solution	
Approach	The approach of the e-commerce shipping time prediction system development will involve the following steps: problem formulation, data collection, and integration from order details, customer information, shipping records, and historical data; data preprocessing piping further divided into cleaning, normalization, and feature engineering; division of data for training, validation, and test sets; and exploratory data analysis to know the distributions in data and the relationship between features .Correlation analysis and importance metrics drive feature selection. Subsequently, we train models including Linear Regression, Random Forest, Neural Networks; after this, some other hyper parameter tuning techniques and cross-validation are utilized. We will use metrics such as RMSE and MAE to evaluate model performance that will guide us on which to choose out of the models. Next in line is deployment to the e-commerce platform, the last mile for provisioning real-time time-to-ship predictions. Provide a customer experience that includes an accurate delivery time estimate and real-time updates.
Key Features	The key features of the proposed solution is real-time integration of weather, traffic, and carrier delay data to drive up the accuracy of the prediction; an API for the seamless delivery time prediction in real-time; and user-friendly customer and business dashboards for shipment tracking and performance metrics.

Resource Requirements

Resource Type	Description	Specification/Allocation

Hardware		
Computing Resources	CPU/GPU specifications, number of cores	AMD Ryzen 5 5500H with Radeon Graphics
Memory	RAM specifications	16 GB
Storage	Disk space for data, models, and logs	1 TB SSD
Software		
Frameworks	Python frameworks	Flask
Libraries	Additional libraries	Scikit-learn, Pandas, Numpy, Collections, mathplotlib, seaborn, missingno, pickle
Development Environment	IDE, version control	Jupyter Notebook, Git, Spyder
Data		
Data	Source, size, format	Kaggle dataset, 10999 observations of 12 variables.