Project Initialization and Planning Phase

Date	10 July 2024
Team ID	SWTID1720086535
Project Title	Ecommerce Shipping Prediction Using Machine Learning
Maximum Marks	3 Marks

Project Proposal (Proposed Solution) template

This project proposal outlines a solution to address a specific problem. With a clear objective, defined scope, and a concise problem statement, the proposed solution details the approach, key features, and resource requirements, including hardware, software, and personnel.

Project Overview	
Objective	The primary objective of this project is to predict the shipping of products to customers using machine learning techniques, ensuring proper tracking of the products.
Scope	The project aims to develop a system that accurately predicts if products will reach their destination on time, considering factors like origin, destination, shipping method, carrier, and potential delays. Using machine learning models trained on historical data and real-time updates, the system will account for weather, traffic, and other external factors. The objective is to provide reliable delivery estimates, enhancing customer satisfaction and trust in e-commerce businesses by improving the overall customer experience.
Problem Statement	
Description	E-commerce businesses face challenges in providing accurate delivery estimates to their customers due to various unpredictable factors such as traffic, weather, and carrier performance. Inaccurate delivery predictions can lead to customer dissatisfaction and a loss of trust. Therefore, there is a need for a robust system that can leverage historical data and real-time updates to accurately predict shipping times, account for external variables, and provide reliable delivery estimates to improve the overall customer experience.
Impact	Social Impacts: Accurate delivery estimates enhance customer experience by reducing uncertainty and increasing transparency. They

	optimize logistics, lowering unnecessary trips and emissions, thus reducing environmental impact. Additionally, they alleviate stress for delivery workers by optimizing routes and workloads, creating a better work environment, and reducing turnover. Business Impacts: Providing accurate delivery estimates boosts customer confidence, reducing cart abandonment and increasing sales and revenue. It also improves operational efficiency by optimizing routes and reducing transportation, labor, and inventory management costs. Implementing machine learning-based delivery prediction offers a competitive advantage over businesses without accurate and transparent delivery estimates.
Proposed Solution	
Approach	Machine learning models for ecommerce shipping prediction work by training on historical delivery data to identify patterns and predict future delivery times. The process involves data collection, preprocessing to handle inconsistencies, and feature engineering to create relevant variables. The models, such as linear regression or gradient boosting machines, are then trained and validated. These models can incorporate real-time data like traffic and weather to adjust predictions dynamically, improving accuracy and customer satisfaction.
Key Features	The primary goal of the system should be to provide accurate delivery estimates to customers, considering factors such as mode of shipment, cost, warehouse details, and other relevant variables. Customers should receive real-time updates on their delivery status, including any delays or changes to the estimated delivery time, with the system adjusting estimates based on the most current information. The system should employ machine learning models to predict delivery times based on historical data and relevant variables, with these models being continually trained and optimized for improved accuracy. Additionally, the system must be scalable, capable of handling large volumes of orders and calculating delivery estimates quickly and accurately for many orders simultaneously.

Resource Requirements

Resource Type	Description	Specification/Allocation
Hardware		
Computing Resources	CPU/GPU specifications,	11th Gen Intel(R) Core(TM)

	number of cores	i5-11300H @ 3.10GHz
Memory	RAM specifications	8.00 GB
Storage	Disk space for data, models, and logs	1 TB SSD
Software		
Frameworks	Python frameworks	Flask
Libraries	Additional libraries	scikit-learn, pandas, numpy, pickle, seaborn, matplotlib, xgboost
Development Environment	IDE, version control	Jupyter Notebook, Git
Data		
Data	Source, size, format	Kaggle dataset, 10,999, csv