

## Project Initialization and Planning Phase

Date	12 March 2024
Team ID	SWTID1720089323
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.

<b>Project Overview</b>	
Objective	Predict whether an ecommerce shipment will be delivered on time using machine learning algorithms.
Scope	This project aims to enhance the accuracy of delivery time predictions for ecommerce platforms by leveraging machine learning algorithms. The goal is to improve shipping logistics, reduce delivery delays, and lower costs. Ultimately, this will lead to a better customer experience through more reliable delivery estimates.
<b>Problem Statement</b>	
Description	Current ecommerce shipping processes are plagued by inefficiencies, resulting in delayed deliveries and increased costs. Accurate prediction of delivery times remains a challenge due to various influencing factors. This project addresses these issues by using machine learning to provide more reliable delivery time estimates.
Impact	Implementing this solution will enhance operational efficiency and reduce shipping delays and costs. Accurate predictions will improve customer satisfaction by providing dependable delivery estimates. Ultimately, this will lead to a stronger, more competitive ecommerce platform.

## Proposed Solution

Approach	Leveraging machine learning techniques to predict delivery times accurately by analyzing a variety of factors. This involves preprocessing the data, handling categorical values, managing outliers, and addressing class imbalance using the SMOTE technique. The project includes extensive exploratory analysis and hyperparameter tuning for several models to identify the most effective one.
Key Features	<ul style="list-style-type: none"> <li>• Implementation of machine learning models, including Logistic Regression, XGBoost, Ridge Classifier, KNN, Random Forest, and SVM.</li> <li>• Detailed exploratory data analysis and outlier management.</li> <li>• Use of the SMOTE technique to handle class imbalance.</li> <li>• Hyperparameter tuning for optimal model performance.</li> <li>• Comparison of various models to select the best performing one for accurate delivery time predictions.</li> </ul>

## Resource Requirements

Resource Type	Description	Specification/Allocation
<b>Hardware</b>		
Computing Resources	CPU/GPU specifications, number of cores	12th Gen Intel(R) Core(TM) i5-1235U
Memory	RAM specifications	16.0 GB
Storage	Disk space for data, models, and logs	1 TB SSD
<b>Software</b>		
Frameworks	Python frameworks	Flask
Libraries	Additional libraries	scikit-learn, pandas, numpy, matplotlib, seaborn, imblearn
Development Environment	IDE, version control	Anaconda(Jupyter Notebook, Spyder), Git

Data		
Data	Source, size, format	Kaggle dataset, 10,999, csv

