Ecommerce Shipping Prediction Using Machine Learning

Date	10 July 2024
Team ID	SWTID1720499933
Project Title	Ecommerce Shipping Prediction Using Machine Learning
Maximum Marks	2 Marks

Milestone 1: Project Initialization and Planning Phase

The "Project Initialization and Planning Phase" marks the project's outset, defining goals, scope, and stakeholders. This crucial phase establishes project parameters, identifies key team members,allocates resources, and outlines a realistic timeline. It also involves risk assessment and mitigation planning. Successful initiation sets the foundation for a well-organized and efficiently executed machine learning project, ensuring clarity, alignment, and proactive measures for potential challenges.

Activity 1: Define Problem Statement

Problem Statement: E-commerce needs a real-time delivery predictor using machine learning to consider various parameters for accurate estimates, improving customer experience and efficiency.

Problem Statement Report: Click Here

Activity 2: Project Proposal (Proposed Solution)

The proposed project, aims to leverage machine learning for more accurate product delivery predictions. Using a comprehensive dataset including customer details , warehouse details and product details, the project seeks to develop a predictive model optimizing punctual product delivery predictions. This initiative aligns with the objective to enhance delivery predictor, ultimately improvingcustomer satisfaction and operational efficiency.

Project Proposal Report: Click Here

Activity 3: InitialProject Planning

Initial Project Planning involves outlining key objectives, defining scope, and identifying stakeholders for delivery
predictor. It encompasses setting timelines, allocating resources, and determining the overall project strategy. During this phase, the team establishes a clear understanding of the dataset, formulates goals for analysis, and plans the workflow for data processing. Effective initial planning lays the foundation for a systematic and well-executed project, ensuring successful outcomes.

Project Planning Report:Click Here

Milestone 2: Data Collection and Preprocessing Phase

The Data Collection and Preprocessing Phase involves executinga plan to gather relevantcustomer analytics data from Kaggle, ensuringdata quality throughverification and addressing missing values. Prepro cessing tasks include cleaning, encoding, and organizing the dataset for subsequent exploratory analysisand machine learningmodel development.

Activity 1: Data Collection Plan, Raw Data Sources Identified, Data QualityRe port

The dataset for " Ecommerce Shipping Prediction Using Machine Learning " is sourced from Kaggle. It includes applicant details and financial metrics. Data quality is ensured through thorough verification, addressing missing values,and maintaining adherence to ethical guid elines, establishing a reliablefoundation for predictive modeling.

Data Collection Report:Click Here

Activity 2: Data Quality Report

The dataset for " Ecommerce Shipping Prediction Using Machine Learning " is sourced from Kaggle. It includes customer review details and product metrics. Data quality is ensured through thorough verification, addressing missing values,and maintaining adherence to ethical guid elines, establishing a reliablefoundation for predictive modeling.

Data Quality Report:Click Here

Activity 3: Data Exploration and Preprocessing

Data Exploration involves analyzing the product dataset to understand patterns, distributions, and outliers. Preprocessing includes handling missingvalues, scaling, and e ncoding categorical variables. These crucial steps enhance data quality, ensuring the reliability and effectiveness of subsequent analyses in the delivery predictor project.

Data Exploration and Preprocessing Report:Click Here

Milestone 3: Model Development Phase

The Model Development Phase entails crafting a predictive model for loan approval. It encompasses strategic featureselection, evaluating and selecting models(Random Forest, SVM Classifer,Logistic Regression, KNN, XGB), initiating training with code, and rigorously validating and assessing model performance for informed delivery predictions.

Activity 1: Feature Selection Report

The Feature Selection Report outlines the rationale behind choosing specific features (e.g., Gender, Warehouse Block,Prior Purchases,etc.) for the delivery predictor model. It evaluates relevance, importance, and impact on predictive accuracy, ensuring the inclusion of key factors influencing the model's abilityto predict punctual delivery of product.

Feature Selection Report:Click Here

Activity 2: Model Selection Report

The ModelSelection Report detailsthe rationale behindchoosing Random Forest,SVM Classifier, KNN, and XGB models for timely delivery prediction. It considers each model's strengths
in handlingcomplex relationships, interpretability, adaptability, and overallpredictive performance, ensuringan informed choicealigned with projectobjectives.

Model Selection Report:Click Here

Activity 3: Initial Model Training Code, Model Validation and EvaluationR eport

The Initial Model Training Code employs selected algorithms on the product dataset, setting the foundation for predictive modeling. The subsequent Model Validation and EvaluationReport rigorously assessesmodel performance, employing metrics like accuracyand pr ecision to ensure reliability and effectiveness in predicting timely delivery.

Model Development Phase Template: Click Here

Milestone 4: Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, comparing performance metrics, and justifying the fin al modelselection for enhancedpredictive accuracy and efficiency.

Activity 1: Hyperparameter Tuning Documentation

The Random Forest Classifier model was selected for its superior performance, exhibiting high accuracy. Its ability to handle complex relationships, minimize
overfitting, and optimizepredictive accuracy aligns with projectobjectives, justifying its selection a s the final model.

Activity 2: Performance Metrics Comparison Report

The Performance Metrics Comparison Reportcontrasts the baselineand optimized metricsfor va rious models, specifically highlighting the enhanced performance of the Random Forest Classifier model. This assessment provides a clear understanding of the refined predictive capabilities achieved.

Activity 3: Final Model Selection Justification

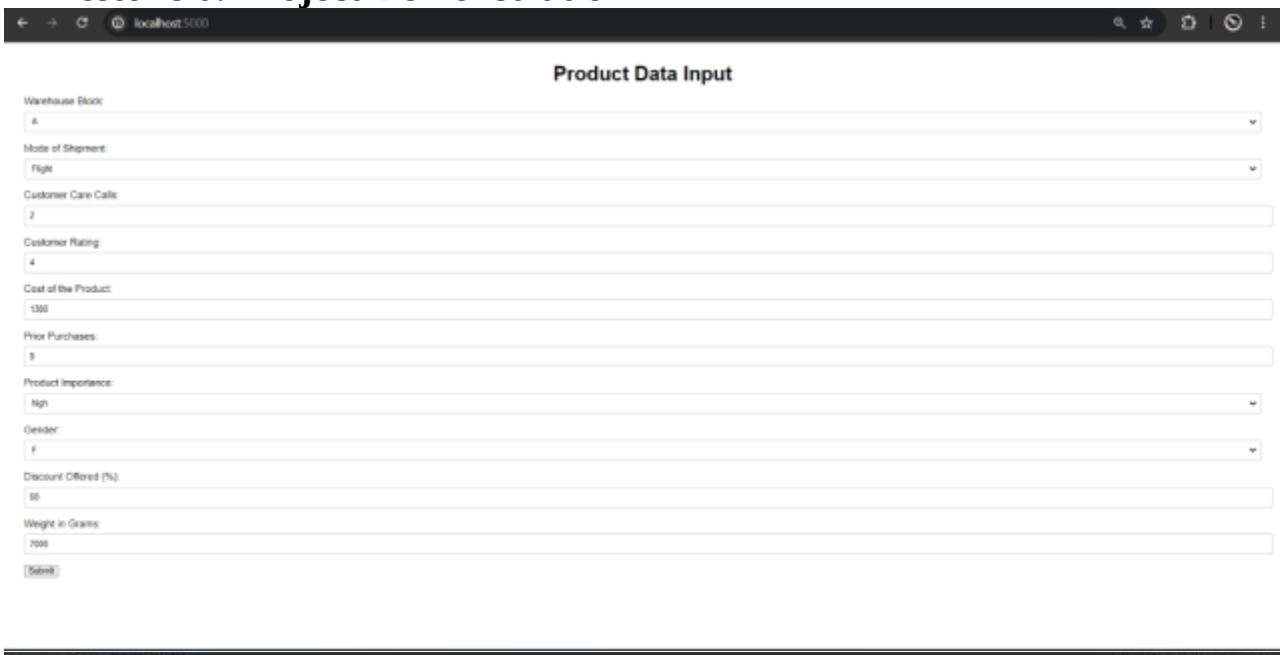
The Final Model Selection Justification articulates the rationale for choosing Random Forest Classifier as the ultimate model. Its exceptional accuracy and ability to handle
complexity alignwith project objectives, ensuring optimal product delivery predictions.

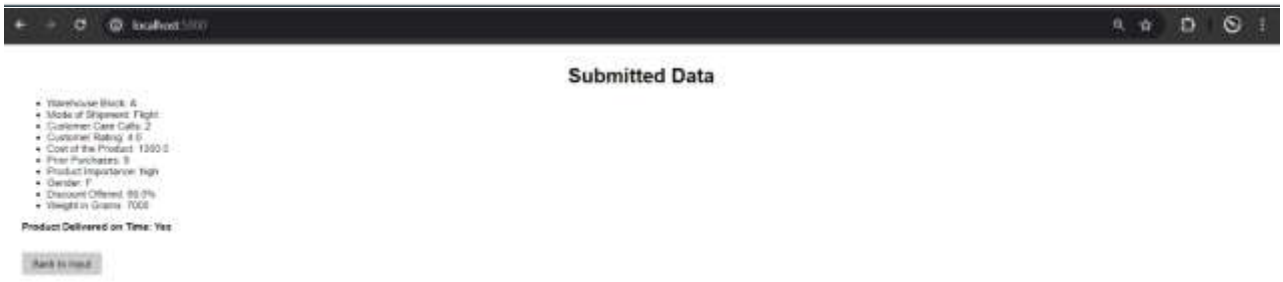
Model Optimization and Tuning Phase Report: Click Here

Milestone 5: ProjectFiles Submission and Documentation

For projectfile access in Github, kindly click the link . **Click Here**

Milestone 6: Project Demonstration





Conclusion:

The model is deployed on local Flask server and is able to predict whether the product is delivered on time or not.