Ecommerce Shipping Prediction Using Machine Learning

Milestone 1: Project Initialization and Planning Phase

The Project Initialization and Planning Phase establishes the goals, scope, and stakeholders for our concrete strength prediction project. This phase defines project parameters, identifies team members, allocates resources, and outlines a timeline. It also includes risk assessment and mitigation planning. Effective planning ensures a well-organized project, clear objectives, and proactive strategies for accurate concrete strength predictions.

Activity 1: Define Problem Statement

E-commerce customers rely on promised delivery dates for their online purchases to plan their schedules and meet their needs. However, frequent delays and inaccurate delivery times disrupt these plans, causing frustration. These issues stem from the current shipping prediction system's inability to account for real-time factors such as weather, traffic, and carrier delays. This leads to a loss of confidence in the e-commerce platform's reliability and negatively impacts the overall shopping experience, making customers hesitant to place future orders.

Problem Statement Report: click here

Activity 2: Project Proposal (Proposed Solution)

Ecommerce shipping prediction is the process of estimating whether the product reached on time. which is based on various factors such as the origin and destination of the package, the shipping method selected by the customer, the carrier used for shipping, and any potential delays or issues that may arise during the shipping process. Machine learning models can be used to make accurate predictions about shipping times based on historical data and real-time updates from carriers.

Project Proposal Report: click here

Activity 3: Initial Project Planning

Project planning for the E-commerce Shipping Prediction project involves several key steps: Data Collection and Preprocessing, Exploratory Data Analysis (EDA), Feature Engineering, Model Development and Training, Model Evaluation, and Model Deployment. Essential resources include access to historical shipping data, computing resources for data processing and model training, machine learning libraries (such as

Python, Scikit-learn, TensorFlow, or PyTorch), and data visualization tools (like Matplotlib and Seaborn). Collecting these resources ensures smooth project execution.

Project Planning Report: click here

Milestone 2: Data Collection and Preprocessing Phase

This stage involves gathering and preparing the raw materials for your project. You'll identify the type of data needed, collect it from various sources, and ensure its quality, similar to cleaning and organizing ingredients before cooking. Data cleaning involves removing errors, inconsistencies, or missing values, preparing your data for the next steps.

Activity 1: Data Collection Plan, Raw Data Sources Identified, Data Quality Report

Data will be collected from various e-commerce platforms, including order details, customer demographics, and shipping information. Additional data sources will include historical delivery times and third-party logistics information to ensure comprehensive analysis. The raw data sources for this project include datasets obtained from Kaggle, a popular platform for data science competitions and repositories. The provided sample data represents a subset of the collected information, encompassing variables related to customers.

Data Collection Report: Click here

Activity 2: Data Quality Report

Data quality is crucial for building an accurate predictive model. This activity involves creating a data quality report that assesses data completeness, accuracy, consistency, timeliness, and relevance. The report will identify data quality issues, their severity, and the resolution plan to address them.

Data Quality Report: click here

Activity 3: Data Exploration and Preprocessing

Data exploration involves understanding the data's characteristics and identifying potential patterns through univariate, bivariate, and multivariate analyses to uncover insights and relationships. Additionally, outliers and anomalies are identified and addressed. Data preprocessing prepares the data for modeling, which involves several steps: importing the collected datasets, identifying and addressing missing values, normalizing or scaling the data as needed, creating new features or modifying existing ones to enhance model performance, and storing the cleaned and preprocessed data for future use. Each of these steps ensures that the data is in the best possible shape for the modeling phase, setting the foundation for accurate and reliable predictions.

Data Exploration and Preprocessing Report: click here

Milestone 3: Model Development Phase

This phase is a critical stage in the E-commerce Shipping Prediction project, where we utilize the cleaned and preprocessed data to build and refine machine learning models. The primary goal is to create an effective predictive model that accurately forecasts shipping outcomes based on various features.

Activity 1: Feature Selection Report

We identified and selected the most relevant features from the dataset. Techniques such as correlation analysis, Recursive Feature Elimination (RFE), and feature importance from tree-based models were used to determine the key predictors. The Feature Selection Report documents these techniques and the rationale behind the chosen features.

Feature Selection Report: click here

Activity 2: Model Selection Report

We evaluated multiple machine learning models, including Logistic Regression, XGBoost, Ridge Classifier, KNN, Random Forest, and SVM Classifier. Cross-validation and hyperparameter tuning were employed to compare their performance and select the best model. The Model Selection Report details the comparison process, performance metrics, and the final model selection.

Model Selection Report: clcik here

Activity 3: Initial Model Training Code, Model Validation and Evaluation Report

The selected model was trained on the entire training dataset, validated, and evaluated. We assessed its performance using metrics like accuracy, precision, recall, F1 score, and AUC-ROC, ensuring it meets the project's predictive requirements. The Model Validation and Evaluation Report presents the training code, validation process, and performance results, demonstrating the model's efficacy.

Model Development Phase Template: click here

Milestone 4: Model Optimization and Tuning Phase

Model optimization and tuning are critical steps in developing an efficient and accurate predictive model. This phase involves using new data and feedback to continuously improve the model, optimizing it for low-latency predictions, and achieving better performance and reliability.

Activity 1: Hyperparameter Tuning Documentation

The hyperparameter tuning documentation includes detailed steps of various tuning methods like grid search, random search, and Bayesian optimization. It also specifies the hyperparameters to be tuned, the code and configurations used, and the deployment of the

model. This documentation ensures a systematic approach to finding the best model parameters.

Activity 2: Performance Metrics Comparison Report

This report compares the performance of different models and methods using various metrics such as accuracy, precision, and recall. The comparison is presented both in tabular form and through visualizations like bar graphs, ROC curves, and box plots. The report analyzes the strengths and weaknesses of each model, providing insights for further testing and concluding with the best-performing models.

Activity 3: Final Model Selection Justification

The final model selection is justified by comparing models such as logistic regression, random forest, gradient boosting, and neural networks. The comparison considers trade-offs like accuracy versus interpretability and training time versus performance. The selected model, along with the rationale for its selection and its implications, is documented, ensuring a well-informed decision.

By following these activities, the project aims to refine the predictive model to achieve optimal performance and reliability.

Model Optimization and Tuning Phase Report: click here

Milestone 5: Project Files Submission and Documentation

For project file submission in GitHub, kindly click the link and refer to the flow. Click Here

Milestone 6: Project Demonstration

In the upcoming module called Project Demonstration, individuals will be required to record a video by sharing their screens. They will need to explain their project and demonstrate its execution during the presentation.