Title : Customer Analytics for Shipment Delivery

Abstract:

This code analyse a dataset containing information about shipment deliveries, including various features such as the mode of shipment, product importance, customer ratings, and whether the shipment was delivered on time or not. The analysis involves data preprocessing, exploratory data analysis using visualizations, and applying several machine learning models to predict if a shipment will reach on time or not. The code aims to identify the factors influencing timely delivery and evaluate the performance of different models in predicting on-time delivery.

Problem Statement:

The goal of this project is to analyse the customer analytics dataset and build a predictive model to determine if a shipment will be delivered on time or not. The analysis involves understanding the relationships between various features, such as the mode of shipment, product importance, customer ratings, and the on-time delivery status. By identifying the key factors influencing timely delivery, the company can make informed decisions to improve their shipping processes and customer satisfaction.

Scope:

The scope of this project includes:

1. Exploratory data analysis to gain insights into the dataset and visualize relationships between features.

2. Data preprocessing, including handling missing values, encoding categorical variables, and scaling numerical features.

3. Applying various machine learning models, such as logistic regression, support vector machines, random forests, and neural networks, to predict if a shipment will reach on time or not.

4. Evaluating and comparing the performance of different models using appropriate metrics, such as accuracy, precision, recall, and F1-score.

5. Identifying the best-performing model and fine-tuning its hyperparameters using techniques like cross-validation and grid search.

6. Analyzing the results and drawing conclusions about the factors influencing on-time delivery and the effectiveness of the predictive models.

Learning Outcomes:

By working through this code, you will learn the following:

1. Data exploration and visualization techniques using libraries like Pandas, Matplotlib, and Seaborn.

2. Data preprocessing steps, including handling missing values, encoding categorical variables, and feature scaling.

3. Implementing various machine learning algorithms for classification problems using scikit-learn.

4. Evaluating model performance using appropriate metrics and techniques like cross-validation and grid search.

5. Interpreting and comparing the results of different models to select the best-performing one.

6. Gaining insights into the factors influencing on-time delivery and their impact on customer satisfaction.

Note: The code provided covers a wide range of topics, including data analysis, visualization, and machine learning. It is recommended to have a basic understanding of Python, data analysis libraries (Pandas, NumPy, Matplotlib, Seaborn), and machine learning concepts before diving into this code.