Data Collection and Preprocessing Phase

Date	09 July 2024
Team ID	SWTID1720499933
Project Title	Ecommerce Shipping Prediction Using Machine Learning
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Report:

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Data Collection Plan:

Section	Description
Project Overview	The machine learning project aims to predict loan approval based on applicant information. Using a dataset with features such as gender, marital status, income, and credit history, the objective is to build a model that accurately classifies loan status (approved or denied), facilitating efficient and informed decision-making in the lending process.
Data Collection Plan	● Search Kaggle.
Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Kaggle The provided sample data represents a subset of the collected information, encompassing variables such as Warehouse_block, Mode_of_Shipment, Customer_care_calls, Customer_rating, Cost_of_the_Product, Prior_purchases, Product_importance, **Gender,** Discount_offered, Weight_in_gms

for machine learning analysis.

Raw Data Sources Report:

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	The dataset comprises package (Warehouse bloc,weight of block,Priority,cost) and customer details (gender,customer rating),.	https://www.kaggle.com/datasets/prachi13/customer-analytics?resource=download	CSV	124 kB	Public