Model Development Phase

Date	08 July 2024
Team ID	SWTID1720084775
Project Title	Ecommerce Shipping Prediction using Machine Learning
Maximum Marks	5 Marks

Feature Selection Report

Below is a table where each feature present in the dataset is briefly explained in terms of its significance in our project. Additionally, the process behind the selection of these features for training the model is described.

Feature	Description	Selected (Yes/No)	Reasoning
ID	A unique identifier for each order	No	it is a unique identifier and does not provide any information relevant to predicting shipping outcomes.
Warehouse_block	Indicates the specific block within a warehouse where the product is stored. It helps in tracking and managing inventory within the warehouse.	Yes	Important for understanding how the location within the warehouse might affect the time it takes to prepare and ship a product

Mode_of_shipment	The method by which the product is shipped to the customer, such as via flight, ship, or road	**Yes**	Critical as different shipment methods (air, road, ship) have varying transit times that directly influence delivery speed.
Customer_care_calls	The number of calls made by the customer to the customer care center regarding the shipment	**Yes**	Useful as frequent customer care calls might indicate issues with the shipping process that could delay deliveries.
Customer_rating	The rating given by the customer for the product or service between 1 to 5.	**No**	Customer ratings are not related to the shipping time prediction and are more about product quality or service satisfaction
Cost_of_the_product	The price of the product purchased by the customer(in Rs)	**No**	The product cost is not a factor in determining shipping efficiency or times.
Prior_purchases	The number of previous purchases made by the customer before the current purchase	**No**	Previous purchase history does not affect the current shipping process or times.
Product_importance	The importance level of the product, which could be categorized as low, medium, or high based on various criteria such as demand, value, or necessity.	**Yes**	Relevant as high-importance products might be prioritized and handled differently, potentially affecting shipping times

Gender	The gender of the customer, typically categorized as male or female.	**No**	Dropped as gender is not expected to impact shipping times also it is considered irrelevant to the logistic processes.
Discount_offered	The discount provided on the product, usually represented as a percentage.	**No**	The focus is on predicting shipping times and the discount does not significantly influence shipping duration
Weight	The weight of the product in grams	**Yes**	Significant because heavier items may require special handling and longer shipping times compared to lighter items.
Reached.on.time_Y.N	Indicates whether the product reached the customer on time. 'Y' stands for yes, and 'N' stands for no	**Yes**	This is the final prediction for the Ecommerce Shipping prediction