CAPSTONE PROJECT SUBMISSION

Online Retail Customer Segmentation

The purpose of this analysis is to identify customer purchase behavior, form customer segmentation, and identify customer behavior this research uses customer purchase data.

Customer Segmentation is considered an effective method for managing customers while developing diverse marketing strategies, it is the process of dividing customers into homogeneous and distinct groups. Segmentation could be done according to customer characteristics, which are tracked online helped by certain algorithms. Companies need focusing the target customer than gain maximum profit with a win-win situation for the company-customer. Customer segmentation is one of the solutions to optimize the result of a win-win situation.

The increase in online sales indicates that the way consumers purchase and use financial services has changed. There are unique characteristics of online shopping, such: as each customer's shopping process and activities can be tracked instantaneously and accurately, each customer's order is associated with a delivery address and a billing address, and each customer has an online store account with essential contact and payment information. These enabled online retailers to treat customers personally with an understanding of each customer and to build upon customer-centric business intelligence [7]. Regarding the customer-centric business model, Online retailers are usually concerned with the following common business concerns:

- Who are the most/least valuable customers to the business? What are their distinct characteristics of them?
- What are customers' purchase behavior patterns? Which products/items have customers purchased together often? In which sequence the products have been purchased?

Customer segmentation is the exercise of dividing a company's customers into companies that replicate similarities amongst customers in every group. The aim of segmenting customers is to determine how to relate to customers in every phase to be able to maximize the price of every consumer to the business. The contents of the dataset had features together with invoice no., stock code, description, quantity, unit price, customerID, and country. The problem statement become to construct an unsupervised machine-learning algorithm to carry out consumer segmentation. We started with data wrangling in which we attempted to address null values, and duplicates and carried out feature adjustments. Next, we did a few exploratory data analyses and attempted to attract observations from the capabilities we had withinside the dataset. Next, we formulated a few quantitative elements together with recency, frequency, and monetary referred to as the RFM model for each of the customers. We carried out the KMeans clustering set of rules on those features. We additionally carried out silhouette and elbow approach analysis to decide the optimal no. of clusters which was 3. We noticed clients having excessive recency and occasional frequency and monetary values had been a part of one cluster and clients having low recency and excessive frequency, and monetary values had been a part of any other cluster. We noticed better values of frequency, monetary and low values of recency in figuring out one class and occasional values of frequency, monetary and excessive values of recency in figuring out any other magnificence. However, there may be greater changes to this analysis. One may also select to cluster into greater

numbers relying on company goals and preferences. The labeled feature after clustering may be fed into classification-supervised machine learning algorithms that might expect the lessons for a brand new set of observations. The clustering also can be carried out on a brand new set of features together with the form of products every consumer prefers to shop for often, locating out consumer lifetime price (clv), segmenting on the idea of the term they visit, and plenty greater. As system mastering has grown to be greater of an ART, there may be not anything together proper or wrong. We simplest attempt to get the first-class outcomes that could shape our very last goals. There is, and constantly will be, a want to improve, going forward

Contribution Roles:

i). **Sunanda Debnath**
Data cleaning
Performed Data Preprocessing
General Analysis Perform EDA(Exploratory Data Analysis)
Standardization of features RFM model

ii). **Ajay Tiwari**
Performed Data Preprocessing
General Analysis Perform EDA(Exploratory Data Analysis)
Build Model Such As
K-means clustering

Please paste the GitHub Repo link.
GitHub Link: https://github.com/sunanda-debnath/Online_Retail_Customer_Segmentation

Drive Link:
https://drive.google.com/drive/folders/1A0L6ehAwwj9gW58HeZkMX65n1ZZjttfp?usp=sharing