Machine Learning: Data-driven Customer Segmentation

CSE 475 and 476

Data Mining and Lab

Project proposal

Sheikh Rafia Noor Maisha(19202103037)

Sihab Mahmud(19202103032)

Sourov Karmokar(19202103020)

Md. Monzurul Alam Sajal(19202103007)

Sourov Hazra(19202103036)

Submitted To

Md. Anwar Hussen Wadud

Assistant Professor,Dept. of CSE

Bangladesh University of Business and Technology

September, 2023

Project Title

Machine Learning: Data-driven Customer Segmentation

Project Purpose

The purpose of this project is to use machine learning to develop a data-driven customer segmentation model for company. The model will be used to identify and group customers based on their shared characteristics, such as demographics, purchase behavior, and interests. This information can then be used to improve the company's marketing and sales efforts, as well as to develop more personalized customer experiences.

1.Improved marketing and sales results: By targeting marketing campaigns and sales pitches to specific customer segments, companies can improve their ROI.
2.Increased customer retention: By understanding the needs and wants of different customer segments, companies can develop products and services that are more likely to appeal to them.
3.Personalized customer experiences: By providing customers with personalized recommendations and offers, companies can improve customer satisfaction and loyalty.

Problem Statement

We want to use machine learning to segment our customers into different groups based on their purchase history, demographics, and other factors. This will allow us to better understand our customers and tailor our marketing efforts to their specific needs.

1.Improved understanding of customers
2.Increased customer satisfaction
3.Increased sales
4.Reduced marketing costs

Objectives

The objective of this project is to use machine learning to segment the customer base of a company into groups of customers with similar characteristics.

1.Collect and prepare the data from Kaggle. This includes cleaning the data, removing outliers, and transforming the data into a format that can be used by the machine learning algorithm.
2.Select the machine learning algorithm so in our model we will used unsupervised learning K-Means clustering for better accuracy.
3.Tune the hyperparameters of the algorithm.