High Level Design (HLD)

Clustering for Customer Segmentation and understanding

PAREKH KATHAN CHIRAGKUMAR

Document Version Control

Date Issued	Version	Description	Author
07/05/2023	1.0	Abstract, Introduction, General Description	Kathan
15/05/2023	1.1	Design Detail, API, Deployment	Kathan
05/06/2023	1.2	Final Revision	Kathan

Contents

Abstract

This study proposes a machine learning-based approach for segmenting mall customers. The objective is to group customers into distinct segments based on demographic information, purchasing patterns, and visitation behavior. A comprehensive dataset comprising customer attributes and transactional data is collected. Various machine learning algorithms, including clustering and classification techniques, are employed to analyze the dataset and identify meaningful customer segments. The results reveal the presence of distinct segments with unique traits, such as spending patterns and visitation frequency. The findings provide valuable insights for mall owners and retailers to tailor marketing campaigns and enhance the overall customer experience. Additionally, the study explores the potential of incorporating additional data sources, such as social media and mobile app usage, to improve segmentation accuracy. The application of machine learning in mall customer segmentation enables data-driven decision-making and personalized experiences in the retail industry.

1 Introduction

1.1 Why this High-Level Design Document?

The purpose of this High-Level Design (HLD) Document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions before coding and can be used as a reference manual for how the modules interact at a high level.

The HLD will:

- Present all of the design aspects and define them in detail
- Describe the user interface being implemented
- Describe the hardware and software interfaces
- Describe the performance requirements
- Include design features and the architecture of the project
- List and describe the non-functional attributes like:

  -Security
  -Reliability
  -Maintainability
  -Portability
  -Reusability
  -Application compatibility
  -Resource utilization
  -Serviceability

1.2 Scope

The HLD documentation presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses non-technical to mildly-technical terms which should be understandable to the administrators of the system.

2 General Description

2.1 Product Perspective & Problem Statement

Not all customers are same. To know which group is your customer and their preferences is a big part for success in your business. Unsupervised machine learning can help marketers to know their audience globally and engage them with their products accordingly.
Here we can classify millions of people's interests through their social media activity and also through other surveys online & offline and cluster them in specific group of their interest.

2.2 Tools used

Python Programming, Jupyter Notebook, MS Excel ,Numpy, Pandas, Seaborn

3 Deployment