Low Level Design (LLD)

Customer Segmentation

Revision Number: 1.0

Atufa Shireen

Document Version Control

Date Issued	Version	Description	Author
5th Oct 2021	1.0	First Version of Complete LLD	Atufa Shireen

Table of Contents

Abstract

As per pareto principle 80% of the revenue is derived from the 20% of the customers, and hence retaining the available customers is less expensive than onboarding new customers.
The customer analysis and segmentation can help in predicting the future transactions of the customers, their behaviour towards business products, relationship strength with the business, customer retention and churn rate, which can also help in targeted marketing.

1 Introduction

1.1 Why this Low-Level Design Document?

The purpose of this document is to present a detailed description of the Customer Segmentation System. It will explain the purpose and features of the system, the interfaces of the system, what the system will do, the constraints under which it must operate and how the system will react to external stimuli. This document is intended for both the stakeholders and the developers of the system and will be proposed to the higher management for its approval.

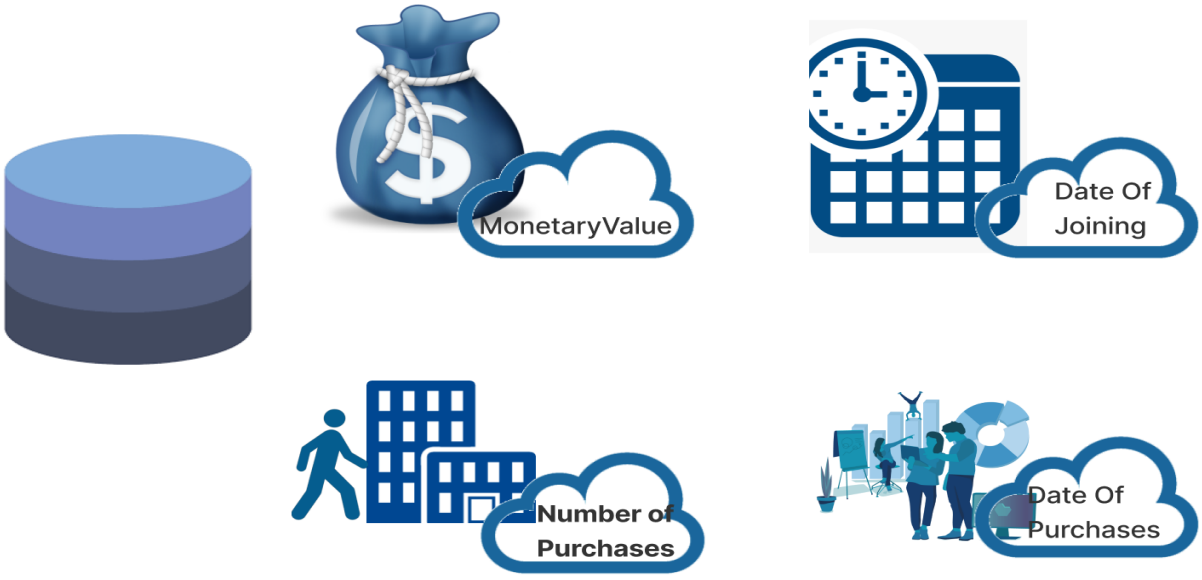
The main objective of the project is to create segments of the customers based on their similarity in behaviour using their previous purchasing data.

Customer segmentation is a vital step in businesses that helps in:

- Understanding customers behaviour towards business products.
- Targeted marketing.
- The future monetary value expected from customers.
- The probability of leaving the business.
- Expected number of transactions in future.

Customer segmentations needs information such as:

- When the customer has joined.
- When or how much the customer has interacted with business.
- Amount spent on transactions.



1.2. Scope

This software system will be a web application that will be designed to create segments of the customers based on their purchase history, and analyse the customer's behaviour. This is also important for targeted marketing. This is a system designed to create segments of customers based on their date of joining the business, number of purchases and amount spent on purchases.

1.3. Constraints

We will be segmenting customers based on only their purchasing behaviour. .

1.4. Risks

Document specific risks that have been identified or that should be considered.

1.5. Out Of Scope

Delineate specific activities, capabilities, and items that are out of scope for the project.

2 Technical specifications

2.1. Input schema

Feature name	Datatype	Size	Null/Required
Date Of Joining(Age)	String or Datetime	--	Required
Number Of Purchases(Frequency).	Integer	--	Required
Total Amount spent.(Monetary)	Integer	--	Required
Number Of Times Customer interacted with products(Recency)	Integer	--	Required

2.2 Creating Segments

- The system displays the option of adding the data.
- The user needs to add data with 4 important features labelled as Age or date of joining, number of purchases or Frequency Value, Total Amount or Monetary Value, Recency or Last Date Of Purchase.
- The system should be able to create and represent segments to the user with visualisations.

2.3 Logging

- We should be able to log every activity done by the user.
- The System identifies at what step logging is required.
- The System should be able to log each and every system flow.
- Developers can choose logging methods. You can choose database logging/ File logging as well.
- System should not be hung even after using so many loggings. Logging just because we can easily debug issues so logging is mandatory to do.

2.4 Database

- System needs to store every request into the database and we need

to store it in such a way that it is easy to retrain the model as well.
- The system stores each and every data given by the user or received on request to the database. Databases can be chosen from MongoDB/ MySQL.

2.5 Deployment

- AWS (Amazon Web Services)

3 Technology stack

Front End	HTML/CSS/JS/React
Backend	Python Django
Database	MongoDB/MySql
Deployment	AWS

4 Proposed Solution

4.1. Segmentation With Clustering Algorithm

The K-Means algorithm is used for classifying customers into segments using their Recency, Frequency and Monetary Values.

- **Elbow Method**
  This method helps in identifying the number of clusters(or segments) required by plotting Within Cluster Sum of Squared Errors (WSS) vs number of clusters.
- **Silhouette scores**
  help in determining the closeness of the clusters.

4.2. Estimation Of Future Transactions/Monetary Value

4.2.1. BTYD Model

The Buy Till You Die model is built on 4 metrics

- Recency
- Frequency
- Monetary
- Age (days or time period since the customer has joined).

There are various BTYD models such as Pareto/GGG (PGGG) Model, Pareto/NBD (PNBD) Model, NBD model, Pareto / NBD,BG / NBD etc., but in this project we'll be using the BG/NBD Model.

4.2.2. BG/NBD Model

This model will help in determining

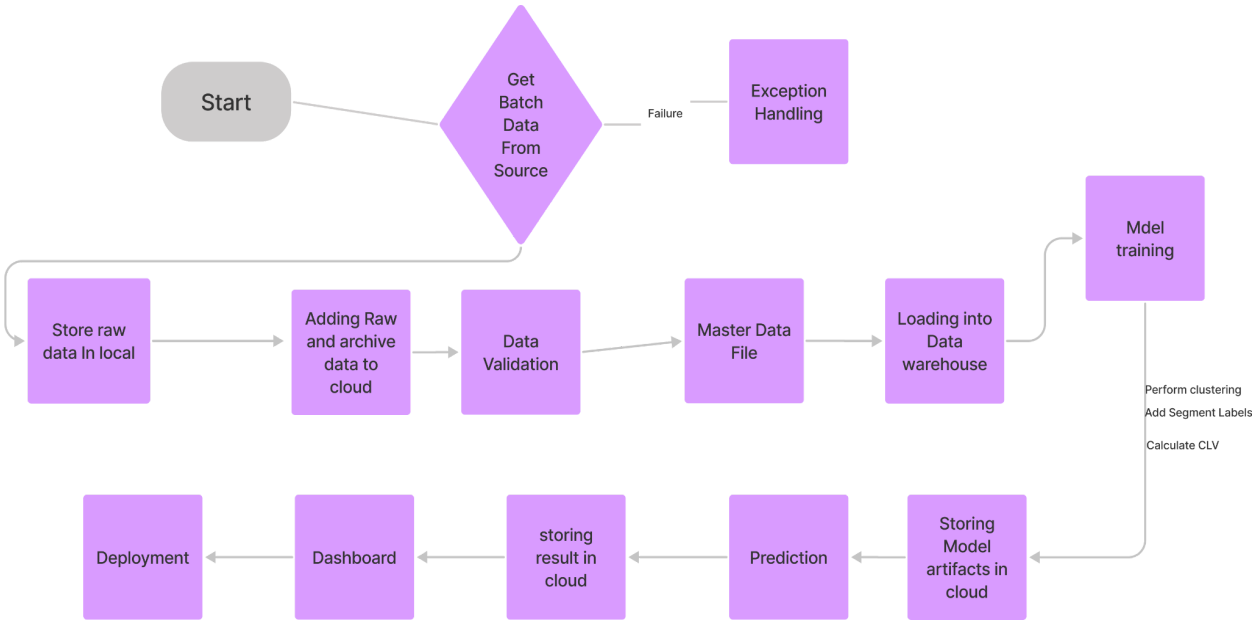
- the probability of observing future transactions.
- the expected number of transactions.
- the probability of a customer becoming inactive.

4.2.3. Gamma-Gamma Model

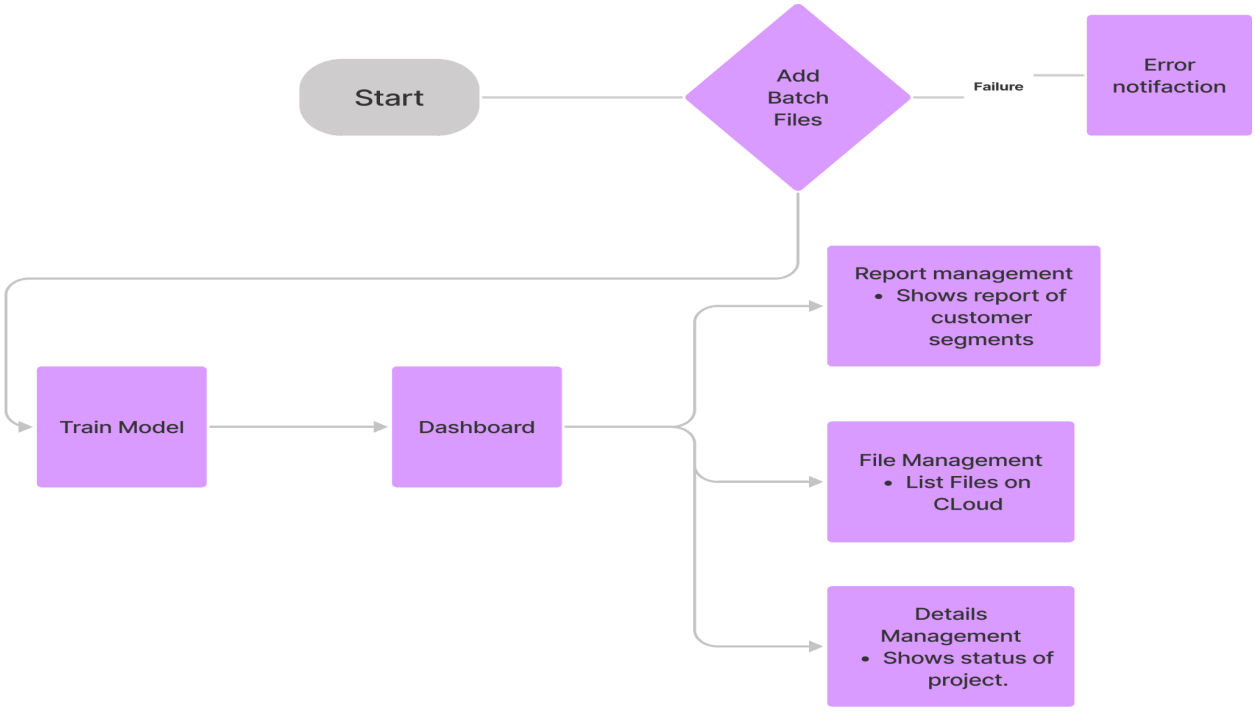
This model will help in calculating

- Average transaction value of each customer
- Long term value of each customer (CLV)

5 Model training/validation workflow



6 User I/O workflow



9. Key performance indicators (KPI)

Key indicators display a summary of the person's health and physique as compared to a normal individual with similar basic traits.

1. Number of times the customer has made purchases.
2. Number of times the customer has visited the business.
3. Date of joining and leaving of customers from the business.
4. Revenue or sale derived from customers.
5. Churn rate of customers.
6. Customer Lifetime Value.