Low Level Design (LLD)

Customer Segmentation

Revision Number - 1.2

Last Date of Revision - 29/1/2022

AKASH

Document Control

Date	Version	Description	Author
21/1/2022	1.0	Introduction, Problem Statement	Akash Shetty
22/1/2022	1.1	Dataset Information, Architecture Description	Akash Shetty
29/1/2022	1.2	Final Revision	Akash Shetty

Contents

1. Introduction

1.1 What is Low Level Design Document?

The goal of the Low-level design document (LLDD) is to give the internal logic design of the actual program code for the Customer Segmentation Analysis . LLDD describes the class diagrams with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document.

2. What is Scope?

Low-level design (LLD) is a component-level design process that follows a step- by-step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

3. Project Introduction

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 activity and also through other surveys online & offline and cluster them in specific group of their interest .

2. Problem Statement

The goal of this project is to analyse to predict the group of cluster our customers belong, based on a combination of features that describes their spending and shopping habits. To achieve the goal, we used a data set that is formed by taking into consideration some of the information of 9000 individuals. The problem is based on the given information about each individual we have to calculate that whether that individual one of four different groups

3. Dataset Information

Dataset

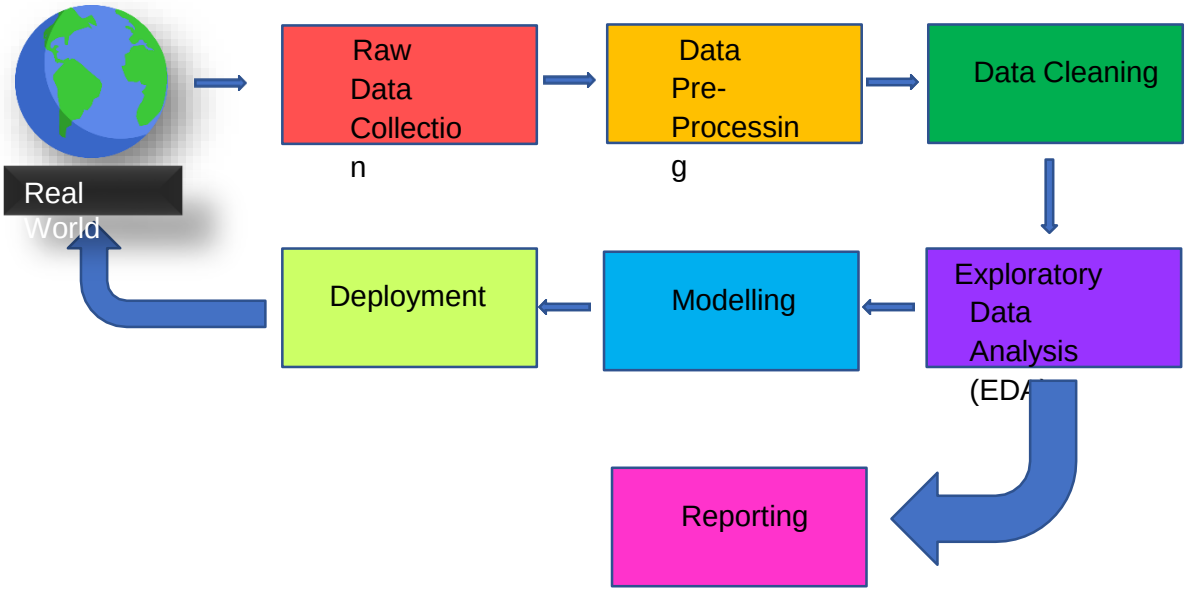
The sample Dataset summarizes the usage behavior of about 9000 active credit card holders during the last 6 months. The file is at a customer level with 18 behavioral variables

Variables of Dataset

- Balance
- Balance Frequency
- Purchases
- One-off Purchases
- Installment Purchases
- Cash Advance
- Purchases Frequency
- One-off Purchases Frequency
- Purchases Installments Frequency

- Cash Advance Frequency
- Cash Advance TRX
- Purchases TRX
- Credit Limit
- Payments
- Minimum Payments
- PRC Full payment
- Tenure
- Cluster

4. Architecture



1. Architecture Description

1. **Raw Data Collection**
   The Dataset was taken from internet.

2. **Data Pre-Processing**
   Before building any model, it is crucial to perform data pre-processing to feed the correct data to the model to learn and predict. Model performance depends on the quality of data feeded to the model to train.
   This Process includes-
   a) Handling Null/Missing Values
   b) Handling Skewed Data
   c) Outliers Detection and Removal

3. **Data Cleaning**
   Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.
   a) Remove duplicate or irrelevant observations
   b) Filter unwanted outliers
   c) Renaming required attributes

4. **Exploratory Data Analysis (EDA)**
   Exploratory Data Analysis refers to the critical process of performing initial investigations on data to discover patterns, spot anomalies, test hypothesis and to check assumptions with the help of summary statistics and graphical representations.

5. **Reporting**
   Reporting is a most important and underrated skill of a data analytics field. Because being a Data Analyst you should be good in easy and self-explanatory report because your model will be used by many stakeholders who are not from technical background.
   a) High Level Design Document (HLD)
   b) Low Level Design Document (LLD)
   c) Architecture
   d) Wireframe
   e) Detailed Project Report
   f) Power Point Presentation

6. **Modelling**
   Data Modelling is the process of analysing the data objects and their relationship to the other objects. It is used to analyse the data requirements that are required for the business processes. The data models are created for the data to be stored in a database. The Data Model's main focus is on what data is needed and how we have to organize data rather than what operations we have to perform.

7. **Deployment**
   We created a Streamlit app