

Architecture Design (AD)

Customer Personality analysis

REVISION NUMBER – 1.2

Last date of Revision: 12/12/2022

Authored by: Aditya Jain

Aditya Sharma

Document Version Control

Date	Version	Description	Author
12/12/2022	1.0	Abstract Introduction Architecture	Aditya Jain/ Aditya Sharma
12/12/2022	1.1	Architectural Design	Aditya Jain/ Aditya Sharma

Contents

Document Version Control	2
Abstract.....	4
1. Introduction	5
1.1 Why this Architecture Design Document ?	5
2. Architecture	5
3. Architecture Design	5
3.1 Data Collection	6
3.2 Data Description	6
3.3 Importing Data into Database.....	7
3.4 Exporting Data from Database.....	7
3.5 Data Preprocessing	7
3.6 Modelling Process	7
3.7 UI Integration	8
3.8 Data from User	8
3.9 Data Validation	8
3.10 Rendering the Results.....	9
3.11 Deployment	9

Abstract

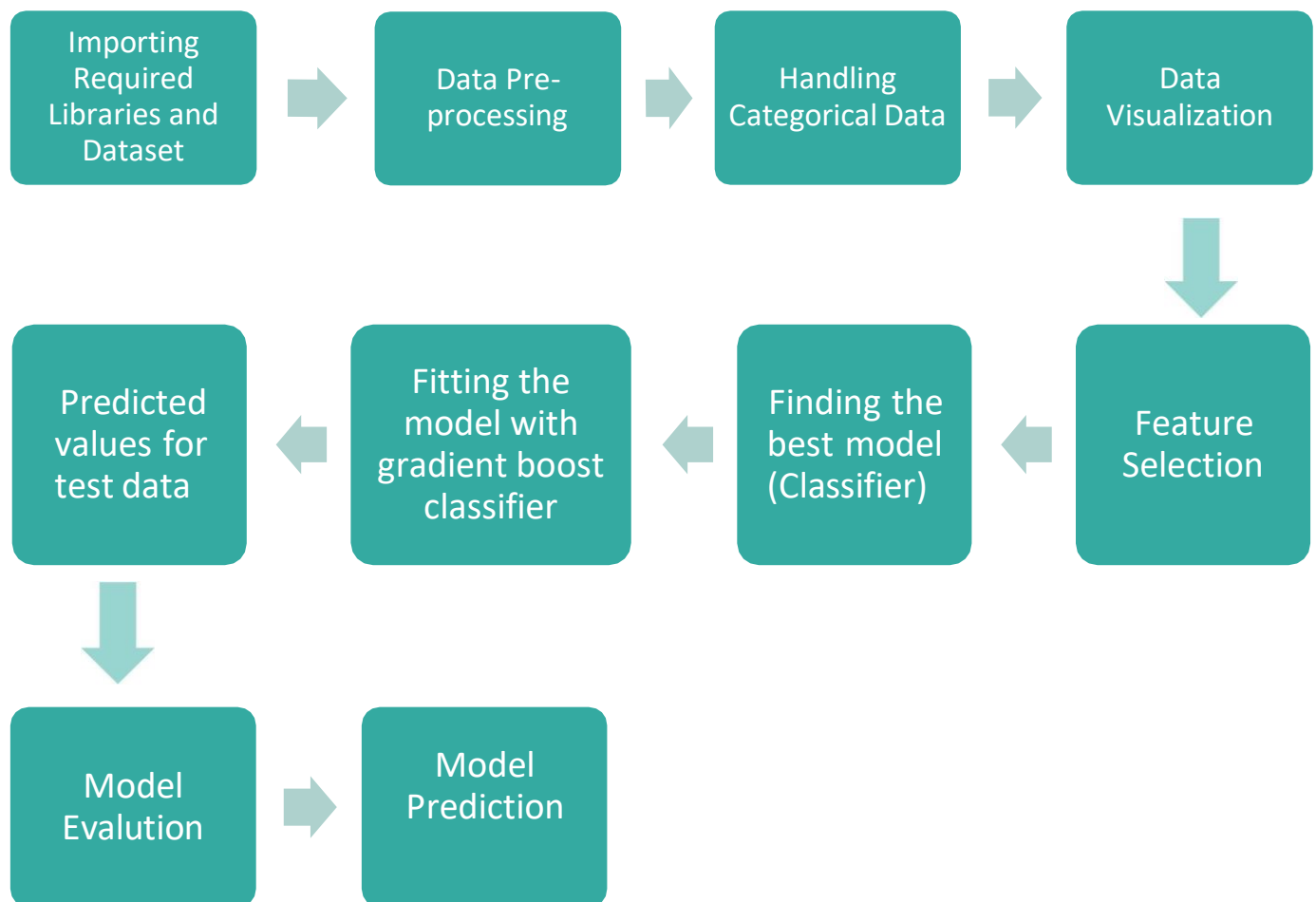
As the legal cannabis industry emerges from its nascent stages, there is increasing motivation for retailers to look for data or strategies that can help them segment or describe their customers in a succinct, but informative manner. While many cannabis operators view the state-mandated traceability as a necessary burden, it provides a goldmine for internal customer analysis. Traditionally, segmentation analysis focuses on demographic or RFM (recency- frequency-monetary) segmentation. Yet, neither of these methods has the capacity to provide insight into a customer's purchasing behavior. With the help of 4Front Ventures, a battle-tested multinational cannabis operator, this report focuses on segmenting customers using cannabis-specific data (such as flower and concentrate consumption) and machine learning methods (K-Means and Agglomerative Hierarchical Clustering) to generate newfound ways to explore a dispensary's consumer base. The findings are that there are roughly five or six clusters of customers with each cluster having unique purchasing traits that define them. Although the results are meaningful, this report could benefit with exploring more clustering algorithms, comparing results across dispensaries within the same state, or investigating segmentations in other state markets.

1. Introduction

1.1 What this Architecture Design Document ?

The main objective of the Architecture design documentation is to provide the internal logic understanding of the flight fare prediction code. The Architecture design documentation is designed in such a way that the programmer can directly code after reading each module description in the documentation.

2. Architecture



3. Architecture Design

3.1 Data Collection

The data for this project is collected from the Kaggle Dataset, the URL for the dataset-:https://raw.githubusercontent.com/amankharwal/Website-data/master/marketing_campaign.csv

Data Description

- People
 - ID: Customer's unique identifier
 - Year_Birth: Customer's birth year
 - Education: Customer's education level
 - Marital_Status: Customer's marital status
 - Income: Customer's yearly household income
 - Kidhome: Number of children in customer's household
 - Teenhome: Number of teenagers in customer's household
 - Dt_Customer: Date of customer's enrollment with the company
 - Recency: Number of days since customer's last purchase
 - Complain: 1 if customer complained in the last 2 years, 0 otherwise
- Products
 - MntWines: Amount spent on wine in last 2 years
 - MntFruits: Amount spent on fruits in last 2 years
 - MntMeatProducts: Amount spent on meat in last 2 years
 - MntFishProducts: Amount spent on fish in last 2 years
 - MntSweetProducts: Amount spent on sweets in last 2 years
 - MntGoldProds: Amount spent on gold in last 2 years
- Promotion
 - NumDealsPurchases: Number of purchases made with a discount

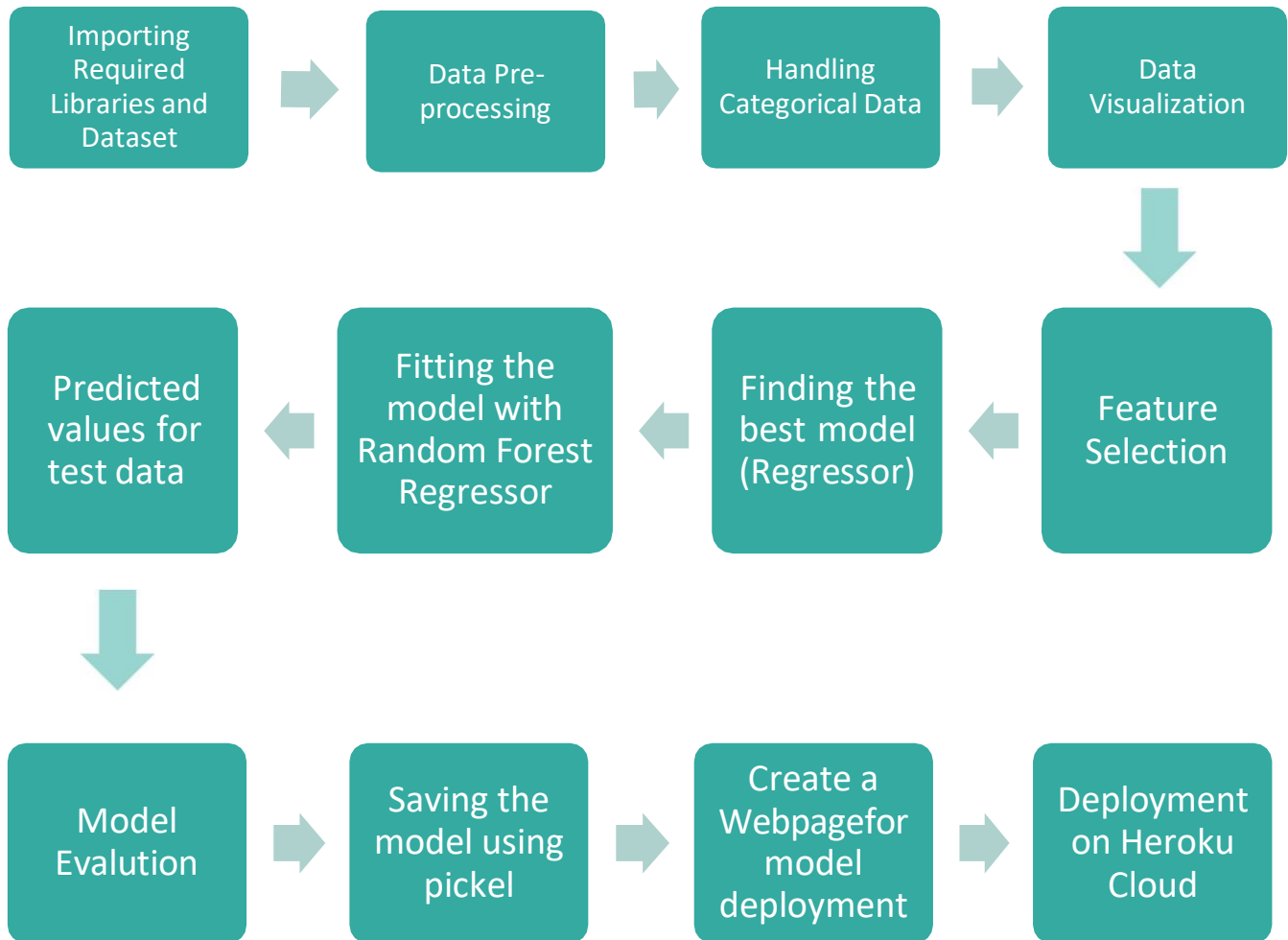
-
- AcceptedCmp1: 1 if customer accepted the offer in the 1st campaign, 0 otherwise
 - AcceptedCmp2: 1 if customer accepted the offer in the 2nd campaign, 0 otherwise
 - AcceptedCmp3: 1 if customer accepted the offer in the 3rd campaign, 0 otherwise
 - AcceptedCmp4: 1 if customer accepted the offer in the 4th campaign, 0 otherwise
 - AcceptedCmp5: 1 if customer accepted the offer in the 5th campaign, 0 otherwise
 - **Response:** 1 if customer accepted the offer in the last campaign, 0 otherwise
-
- Place
 - NumWebPurchases: Number of purchases made through the company's web site
 - NumCatalogPurchases: Number of purchases made using a catalogue
 - NumStorePurchases: Number of purchases made directly in stores
 - NumWebVisitsMonth: Number of visits to company's web site in the last month

3.2 Data Preprocessing

- Checked for info of the Dataset, to verify the correct datatype of the Columns.
- Checked for Null values, because the null values can affect the accuracy of the model.
- Converted all the illegal values into legal values.
- Performed Labeled encoding and One hot Encoding on the desired columns.
- Checking the distribution of the columns to interpret its importance.
- Now, the info is prepared to train a Machine Learning Model.

3.3 Modelling Creation

After preprocessing the data, we visualize our data to gain insights and then these insights are randomly spread and split into two parts, train and test data. After splitting the data, we use Gradient boost classifier to model our data to predict the Restaurant Rating



3.4 UI Integration

We used streamlit for web page and UI for the prediction the customer behavior's.

3.5 Data from User

The data from the user is retrieved from the created streamlit web page.

3.6 Data Validation

The data provided by the user is then being processed by streamlit web file and validated. The validated data is then sent to the prepared model for the prediction.

3.7 Rendering the Results

The data sent for the prediction is then rendered to the web page.

3.8 Deployment

The tested model is then deployed to Heroku. So, users can access the project from any internet devices.