High Level Document

Customer Personality Analysis

REVISION NUMBER – 1.2

last date of Revision: 04/03/2023

Authored by: Aditya jain
Aditya Sharma

Document Version Control

Date	Version	Description	Author
02/01/2023	1.0	Abstract Introduction General Description	Aditya Jain/Aditya Sharma
03/01/2023	1.1	Design Flow	Aditya Jain/Aditya Sharma
04/01/2023	1.2	Performance Evaluation Conclusion	Aditya Jain/Aditya Shama

Contents

Document Version Control 2
Abstract 4
1. Introduction 5
1.1 Why this HLD Document? 5
1.2 Scope 5
1.3 Definitions 5
2. General Description 6
2.1 Problem Perspective 6
2.2 Problem Statement 6
2.3 Proposed Solution 6
2.4 Further Improvements 7
2.6 Data Requirements 8
2.7 Tools Used 8
2.8 Constraints 8
2.9 Assumptions 9
3. Design Flow 9
3.1 Model Creation and Evaluation 9
3.2 Deployment Process 9
3.3 Logging 10
3.4 Error Handling 10
4. Performance Evaluation 10
4.1 Reusability 10
4.2 Application Compatibility 10
4.3 Resource Utilization 10
4.4 Deployment 11
5. Conclusion 11

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 Why this HLD Document?

The main purpose of this HLD document is to feature the required details of the project and supply the outline of the Model Creation, Evaluation and Deployment. This additionally provides the careful description on however the complete project has been designed end-to-end.

The HLD will:

- Present 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 architectural design of the project.
- List and describe the non - functional attributes like:
    - o Security
    - o Reliability
    - o Maintainability
    - o Portability
    - o Reusability
    - o Resource Utilization

1.2 Scope

The HLD documentation presents the structure of the system, such as database design, architectural design, application flow and technology architecture. The HLD uses non-technical terms to technical terms that can be understandable to the administrator of the system.

1.3 Definitions

Term	Description
FFP	Customer Personality Analysis
Dataset	Customer Data
Jupyter-Notebook	It is an interactive computational environment, in which you can combine code execution, rich text, mathematics, plots and rich media.
Heroku	Heroku is a platform as a service (PaaS) that enables developers to build, run, and operate applications entirely in the cloud.

2. General Description

2.1 Problem Perspective

Customer Personality Analysis is a detailed analysis of a company's ideal customers. It helps a business to better understand its customers and makes it easier for them to modify products according to the specific needs, behaviors and concerns of different types of customers.

2.2 Problem Statement

Customer personality analysis helps a business to modify its product based on its target customers from different types of customer segments. For example, instead of spending money to market a new product to every customer In the company' database, a company can analyze which customer segment is most likely to buy the product and then market the product only on that particular segment.

The main objective here is -

1. What people say about your product: what gives customers' attitude towards the product.

2.What people do: which reveals what people are doing rather than what they are saying about your product.

2.3 Proposed Solution

To solve the problem, we have created a User interface for taking the input from the user to predict the **Customer behaviors** using our trained ML model after processing the input and at last the output (predicted value) from the model is communicated to the User.

2.4 Further Improvements

We also analysis the data used for training the ML model by considering different labels like Income, Kidhome, Teenhome, Age, Partner, Education Level.

2.5 Technical Requirements

As technical requirements, we don't need any specialized hardware for virtualization of the application. The user should have the device that has the access to the web and the fundamental understanding of providing the input.

2.6 Tools Used

- Python 3.9 is employed because the programming language and frameworks like NumPy, Pandas, Scikit - learn and alternative modules for building the model.
- Jupyter-Notebook is employed as IDE.
- For Data visualizations, seaborn and components of matplotlib are getting used.
- For information assortment prophetess info is getting used.
- Front end development is completed with streamlit by python frame work.
- Flask is employed for each information and backend readying.
- GitHub is employed for version management.
- Heroku is employed for deployment.

2.8 Constraints

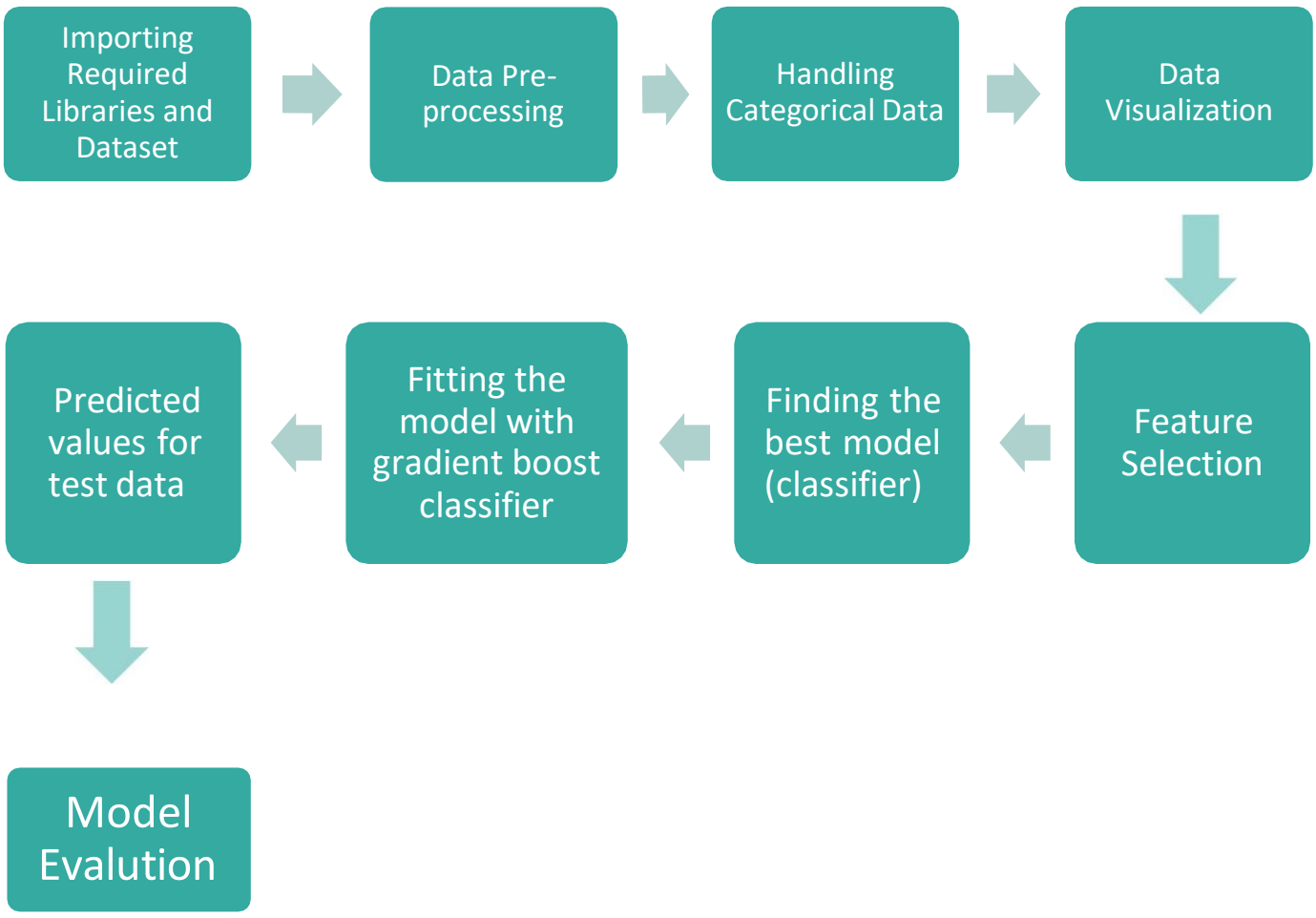
The Customer behaviour answer should be user friendly, as automatic as attainable and also the user should not be needed to understand any of the operating.

2.9 Assumptions

The main objective of the project is to implement the utility cases as for the new dataset that provides the user the ability to predict Customer behaviour . Machine learning model is employed for process the user input for prediction. It additionally assumed that each one aspects of this project have the flexibility to figure along within the approach the designer is expecting.

3. Design Flow

3.1 Modelling Creation and Evaluation



3.2 Deployment Process

3.3 Logging

In logging, at each if an error or an exception is occurred, the event is logged into the system log file with reason and timestamp. These helps the developer to debug the system bugs and rectifying the error.

3.4 Error Handling

Once the error is occurred, the reason is logged into the log file with timestamp to rectify and handle it.

4. Performance Evaluation

4.1 Reusability

The code written and the components used should have the ability to be reused with no problems.

4.2 Application Compatibility

The different parts of the system are communicating or using Python as an interface between them. All the components have its own tasks to perform and it is a job of a Python to ensure proper transfer of data.

4.3 Resource Utilization

When ant task is performed, it'll doubtless use all the process power offered till the process is finished.

4.4 Deployment

The model can be deployed using the any cloud services such as Microsoft Azure, Amazon web services, Heroku, Google cloud, etc.

5. Conclusion

Regardless of the information provided, the results provide actionable ways for retailers to employ a marketing campaign or similar segmentation for their consumers. Despite the usefulness of the analysis as-is, there are numerous routes for improvement and growth. While there was motivation to keep the number of features low, adding a separate feature to account for the recency of the consumer would provide clearer details on whether certain purchase profiles are more common now than in the store's past. On a similar note, finding ways to cluster a customer quicker (such as in one or two visits rather than three) could generate insights into not only the evolutionary aspect of the clustering but potentially also the leakage of customers. Finally, attempting the same analysis with numerous other clustering algorithms such as Gaussian Mixture Models or deep learning would bring about insight into the stability of cluster formation.