PROPOSAL OF CUSTOMER JOURNEY OF PRODUCTS

Objective:-

Given Customer Historical Data, the need is to create a seamless pipeline that injects customer track history, implements Python, NLP and Gen-AI to derive insights. The insights can be then saved to a different SQL database, or it can be migrated to a Tableau dashboard for interactive visualization.

PHASE 1: Initial POC development and insights to be delivered: -

1.Data cleaning, wrangling and alignment of data functionality.

2.Customer segregation from data.

3.Each customer would have the following insights extracted from the data:

3.1. Complete information of the product purchased.

3.2. Sites navigated or the entire journey of a product through different sites before a product is purchased.

3.3. If a product is added to bag and the checkout is completed the following attributes of the product would be tracked: -

a) Product Name

b) Category

c)Sub-Category

d) Quantity

e) Colour

f) Size

g) Final Application Used to Purchase the Product

h) Applications Checked before the product is finally purchased

i)Timestamp of the product when purchased.

j) Payment Mode

k) Final price

4. If the product is browsed through an application and not purchased, the entire search history of the product needs to be tracked and the application it is finally purchased would be noted.

5.For the products only searched and not purchased, the information would be saved in a different table.

6.The entire insights extracted would be individual customer focused or oriented.

7.Final output would be 2 tables. The first table would contain the details of every product that is purchased by every customer. So a single customer would have multiple entries. The 2nd table would contain information about every product that is searched but not purchased. Here the 2nd table would have time constraint (ex 1 hour as the min time frame).So there would be only one row per customer.

Tech stack

1.Python

2.NLP

3.OPEN AI GPT 3.5

4.Tableu

TABLE 1:-

TABLE 1

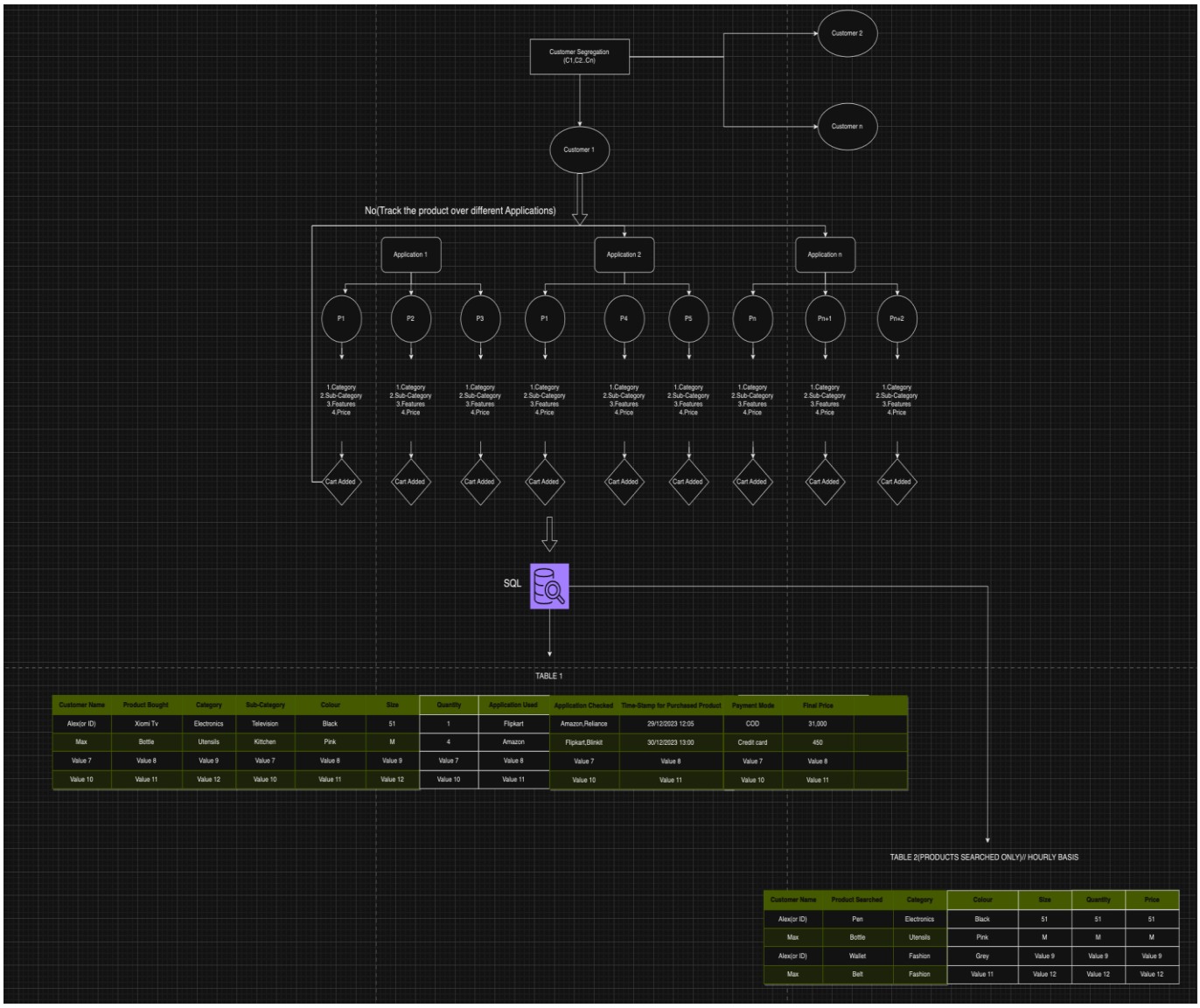
Customer Name	Product Bought	Category	Sub-Category	Colour	Size	Quantity	Application Used	Application Checked	Time-Stamp for Purchased Product	Payment Mode	Final Price
Alex(or ID)	Xiomi Tv	Electronics	Television	Black	51	1	Flipkart	Amazon,Reliance	29/12/2023 12:05	COD	31,000
Max	Bottle	Utensils	Kitchen	Pink	M	4	Amazon	Flipkart,Blinkit	30/12/2023 13:00	Credit card	450

TABLE 2:-

TABLE 2(PRODUCTS SEARCHED ONLY)// HOURLY BASIS

Customer Name	Product Searched	Category	Colour	Size	Quantity	Price
Alex(or ID)	Pen	Electronics	Black	51	51	51
Max	Bottle	Utensils	Pink	M	M	M
Alex(or ID)	Wallet	Fashion	Grey	Value 9	Value 9	Value 9
Max	Belt	Fashion	Value 11	Value 12	Value 12	Value 12

PSEUDO ARCHITECTURE(PHASE 1)



Customer Name	Product Bought	Category	Sub-Category	Colour	Size	Quantity	Application Used	Application Checked	Time-Stamp for Purchased Product	Payment Mode	Final Price
Alex(or ID)	Xiomi Tv	Electronics	Television	Black	51	1	Flipkart	Amazon,Reliance	29/12/2023 12:05	COD	31,000
Max	Bottle	Utensils	Kitchen	Pink	M	4	Amazon	Flipkart,Blinkit	30/12/2023 13:00	Credit card	450
Value 7	Value 8	Value 9	Value 7	Value 8	Value 9	Value 7	Value 8	Value 7	Value 8	Value 7	Value 8
Value 10	Value 11	Value 12	Value 10	Value 11	Value 12	Value 10	Value 11	Value 10	Value 11	Value 10	Value 11

Customer Name	Product Searched	Category	Colour	Size	Quantity	Price
Alex(or ID)	Pen	Electronics	Black	51	51	51
Max	Bottle	Utensils	Pink	M	M	M
Alex(or ID)	Wallet	Fashion	Grey	Value 9	Value 9	Value 9
Max	Belt	Fashion	Value 11	Value 12	Value 12	Value 12

ROADMAP

Week	Process	Actual Timeline
1st January-7th December	Data discovery and Process Formulation	TBD
7-31st January	Implementation of logic (Creation of modules)	TBD
15th January	1 st Progress Discussion	TBD
27th January	Implementation of Feedback and creation of table 1 and table 2.	TBD
31st January	Scaling of the fuctionalities form a single customer to all customers	TBD
1-7th February	Integration testing for all data	TBD

The table shows the tentative dates of completion of the different features along with Data Discovery and the architectural design of the process.

Future Steps:-

1. Once the POC is delivered the integration of the tables should be done with the Power BI dashboard.

1. The process needs to be shifted to Cloud DB.

2.AWS/Azure to be implemented for a seamless injection and insights population.

3.The insights populated needs to be stored. The storage mechanism needs to be decided.

COST OF POC

Task	Cost
OpenAI GPT Token	Subscribed /If institution is happy with the implementation, future cost needs to taken into consideration.
Data extraction and Table formation(POC)	50k*2(2 tables)//
Total Cost	1L
Total Time	1 Month(Approx.)

This is an agreement/initial proposal between MADHURJYA TAMULY and MARKET XCEL(New Delhi).

NAME:MADHURJYA TAMULY APPROVER Sign:-

Place:-PUNE Date:-30/12/2023