**A Capstone Project Report**

**On**

**Customer Behaviour Analysis and Prediction**

***Submitted by***

***M.RAM JEYANTH (192211167)***

***S.HARIHARAN (192211110)***

*Under the guidance of*

# Ms.K. Akshaya

***In partial fulfilment for the completion of the course***

**CSA1668- DATA WAREHOUSING AND DATA MINING FOR DATA ANALYSIS**



**SIMATS ENGINEERING**

**THANDALAM**

**JULY 2024**

# DECLARATION

We, **M. Ram Jeyanth, and S.Hariharan**, students of **Bachelor of Engineering in Computer Science & Engineering**, Department of Computer Science and Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, with this we declare that the work presented in this Capstone Project Work entitled **Customer Behaviour Analysis and Prediction** is the outcome of our Bonafede work and is correct to the best of our knowledge. This work has been undertaken while taking care of engineering ethics.

**M.Ram Jeyanth  
S.Hariharan**

Date: 25.07.2024

Place: Chennai

# Table of Contents

|  |  |
| --- | --- |
| **S.NO** | **TOPICS** |
| 1 | **Abstract** |
| 2 | **Introduction** |
| 3 | **Project Description**    A project implementing an apriori algorithm, Customer behaviour analysis, and data visualization for data mining. |
| 4 | **Problem Description**    Develop a system to identify frequently purchased item sets and generate actionable insights for optimizing Customer satisfaction strategies using the Apriori algorithm and data visualization techniques. |
| 5 | **Tool Description**  User interface  Features |
| 6 | **Operations**  Load and Prepare Data  Apply Apriori Algorithm  Visualize Patterns  Generate Insights Export Results. |
| 7 | **Approach / Module Description / Functionalities**  Import and preprocess transaction data, apply the Apriori algorithm to identify frequent item sets and generate association rules. Visualize patterns, derive actionable insights, and export results for further analysis and reporting. |
| 8 | **Implementation**  Coding |
| 9 | **Output**  Output with Screenshots |
| 10 | **Conclusion** |
| 11 | **References** |

**ABSTRACT**

This project report explores Customer Behaviour Analysis and Prediction, focusing on leveraging data analytics to understand and anticipate customer actions. The study utilizes a comprehensive dataset comprising customer purchase history, demographics, and interaction data to identify patterns and trends. Advanced machine learning techniques, including clustering and predictive modelling, are employed to segment customers and forecast their future behaviours. The findings aim to provide actionable insights for businesses to enhance customer satisfaction, improve retention rates, and optimize marketing strategies. By predicting customer needs and preferences, companies can tailor their offerings more effectively, thereby driving growth and competitive advantage. The project underscores the critical role of data-driven decision-making in modern business environments.

**INTRODUCTION**

Customer behaviour analysis is the systematic study of how individuals make decisions to spend their available resources (time, money, effort) on consumption-related items. It involves understanding what customers buy, why they buy, when they buy, and how they buy. By examining the psychological, social, and economic factors influencing their decisions, businesses can tailor their products, services, and marketing strategies to better meet the needs and preferences of their target audience. This analysis is crucial for creating personalized experiences, fostering customer loyalty, and enhancing overall customer satisfaction. Predictive analytics leverages statistical techniques, machine learning algorithms, and data mining to forecast future customer behaviour based on historical data. By analysing patterns and trends, businesses can predict outcomes such as purchase likelihood, churn rates, and customer lifetime value. Predictive models provide actionable insights that enable companies to anticipate customer needs, optimize marketing campaigns, and allocate resources more efficiently. In an increasingly competitive market, the ability to predict customer behaviour offers a significant strategic advantage, allowing businesses to stay ahead of trends and make informed decisions.

The applications of customer behaviour prediction are vast and varied. Retailers use predictive analytics to recommend products, optimize inventory, and personalize marketing efforts. Financial institutions apply these techniques to assess credit risk, detect fraud, and enhance customer relationship management. In the healthcare sector, predicting patient behaviour can improve treatment plans and patient engagement. Additionally, the travel and hospitality industry leverages customer behaviour predictions to enhance the customer experience through personalized offers and services. Across industries, understanding and predicting customer behaviour helps businesses create more value for both customers and stakeholders.

Despite the numerous benefits, customer behaviour analysis and prediction come with challenges. Data privacy and security concerns are paramount, as businesses must handle sensitive customer information responsibly. Ensuring data quality and accuracy is also crucial, as poor data can lead to flawed predictions. Moreover, developing and maintaining sophisticated predictive models requires substantial expertise and resources. Looking forward, advancements in artificial intelligence and machine learning are expected to enhance predictive capabilities, making them more accessible and effective. As technology evolves, businesses must stay abreast of these developments to continuously improve their customer behaviour analysis and prediction efforts.

**PROJECT DESCRIPTION:**

The project "Customer Behaviour Analysis and Prediction" aims to leverage advanced data analytics and machine learning techniques to understand and forecast customer actions and preferences. By analyzing vast amounts of transactional and demographic data, the research seeks to identify patterns and trends that drive consumer behaviour. The project will develop predictive models to anticipate future customer needs and behaviours, enabling businesses to tailor their marketing strategies and improve customer satisfaction. This study not only enhances the understanding of customer dynamics but also provides actionable insights for strategic decision-making.

**PROBLEM DESCRIPTION:**

Customer behaviour analysis and prediction play a crucial role in understanding and anticipating consumer actions, thereby enabling businesses to tailor their strategies effectively. This research focuses on leveraging advanced data analytics and machine learning techniques to analyse vast amounts of customer data, identifying patterns, and predicting future behaviours. By examining factors such as purchasing habits, demographic information, and engagement metrics, the study aims to provide insights that enhance customer satisfaction, improve retention rates, and drive personalized marketing efforts. The goal is to develop predictive models that offer actionable intelligence for strategic decision-making in a competitive marketplace.

**TOOL DESCRIPTION:**

The Customer Behaviour Analysis and Prediction tool employs sophisticated data analytics and machine learning techniques to scrutinize and interpret consumer behaviour patterns. By utilizing vast amounts of historical purchase data, demographic profiles, and interaction logs, the tool provides in-depth insights into customer preferences and actions. This analysis is crucial for businesses aiming to understand their customer base more profoundly, enabling them to identify key trends and factors that influence buying decisions.

Incorporating predictive analytics, the tool can forecast future customer behaviours and trends with high accuracy. It analyses various data points to predict what products or services a customer is likely to purchase, when they might make a purchase, and through which channels. These predictions enable businesses to tailor their marketing efforts, ensuring that the right message reaches the right customer at the optimal time. Additionally, the tool can segment customers based on predicted behaviours, allowing for more targeted and effective marketing strategies.

Moreover, the tool enhances customer engagement and satisfaction by facilitating personalized experiences. By anticipating customer needs and preferences, businesses can offer customized recommendations, optimize inventory levels, and create targeted promotional campaigns. This proactive approach not only increases sales but also fosters customer loyalty and retention, providing a significant competitive advantage in the market. Furthermore, by reducing the uncertainty in demand forecasting, the tool helps in better resource allocation and minimizes overstock or stockout situations, thereby improving operational efficiency and profitability.

**OPERATIONS:**

* **Load and Prepare Data:** This feature enables users to import and preprocess transaction data from various sources such as CSV files or databases. It ensures the data is clean and formatted correctly for analysis using the Apriori algorithm, allowing for effective and accurate market basket analysis.
* **Apply Apriori Algorithm:** This functionality processes the cleaned transaction data to identify frequent item sets and generate association rules based on user-defined support and confidence thresholds. The algorithm helps in discovering meaningful patterns and relationships between items in the data.
* **Visualize Patterns:** This tool generates visual representations of the identified item sets and association rules using bar charts, scatter plots, and network graphs. These visualizations facilitate the interpretation of complex data patterns and provide actionable insights for optimizing retail strategies.
* **Generate Insights:** Leveraging the results from the Apriori algorithm, this feature generates detailed reports and insights into purchasing behaviours. It highlights significant item relationships and trends, providing recommendations for improving product placement and marketing strategies.
* **Export Results:** This feature allows users to export the analysis results, visualizations, and reports in various formats such as CSV, Excel, and PDF. It also includes options to reset the tool and access information about the software, ensuring users can easily integrate findings into business strategies and reporting systems.

**FUNCTIONALITIES:**

Customer behavior analysis and prediction involves leveraging data-driven insights to understand and anticipate customer actions. This process begins with the collection and integration of data from various sources, including purchase history, online interactions, social media activity, and customer feedback. By consolidating this information, businesses gain a comprehensive view of customer behaviour, enabling more informed decision-making. Advanced data integration techniques ensure that the data is accurate, up-to-date, and readily accessible for analysis, which is crucial for deriving meaningful insights.

Once data is collected, customer segmentation is performed to identify distinct groups within the customer base. This segmentation allows businesses to recognize patterns and tailor their marketing strategies to specific customer needs and preferences. By understanding these segments, companies can create more personalized and effective marketing campaigns, enhancing customer engagement and satisfaction. For instance, high-value customers might receive exclusive offers and personalized communications, while new customers might be targeted with introductory promotions and onboarding assistance. Effective segmentation also helps in resource allocation, ensuring that marketing efforts are directed where they can achieve the highest impact.

Predictive modelling and sentiment analysis are critical tools in this process. Machine learning algorithms are used to analyse historical data and forecast future customer behaviours and trends, allowing businesses to anticipate needs and make proactive decisions. Predictive models can identify potential churn risks, forecast product demand, and optimize inventory management. Sentiment analysis, on the other hand, gauges customer opinions and emotions by analysing textual data from reviews, social media posts, and customer service interactions. This provides valuable insights into their experiences and attitudes, helping businesses address issues promptly and improve their products and services. These functionalities collectively enable businesses to improve customer retention, loyalty, and overall satisfaction, fostering long-term relationships and driving sustainable growth.

**IMPLEMENTATION: (PSEUDO CODE)**

**Initialize Data**:

import pandas as pd

import plotly.express as px

import plotly.graph\_objects as go

data = pd.read\_csv("ecommerce\_customer\_data.csv")

print(data.head())

**Summary Statistics for (numeric and non-numeric columns):**

# Summary statistics for numeric columns

numeric\_summary = data.describe()

print(numeric\_summary)

# Summary for non-numeric columns

categorical\_summary = data.describe(include='object')

print(categorical\_summary)

# Histogram for 'Age'

fig = px.histogram(data, x='Age', title='Distribution of Age')

fig.show()

**Plot Graph:**

# Bar chart for 'Gender'

gender\_counts = data['Gender'].value\_counts().reset\_index()

gender\_counts.columns = ['Gender', 'Count']

fig = px.bar(gender\_counts, x='Gender',

y='Count',

title='Gender Distribution')

fig.show()

# 'Product\_Browsing\_Time' vs 'Total\_Pages\_Viewed'

fig = px.scatter(data, x='Product\_Browsing\_Time', y='Total\_Pages\_Viewed',

title='Product Browsing Time vs. Total Pages Viewed',

trendline='ols')

fig.show()

**Grouped Analysis:**

Try:

Filter rules based on support, confidence, and lift values Depending on file\_type ('csv', 'excel', 'json'):

Save filtered rules to the corresponding file format

Show success dialog with the file name Catch Exception:

Print error message

Show error dialog with the message

**Show Rule Details:**

Try:

Get selected item from Treeview

Retrieve rule details (antecedents, consequents, support, confidence, lift)

Show rule details in a message dialog Catch Exception:

Print error message

Show error dialog with the message **Plot Histograms:**

# Create a bar chart to visualize the customer segments

fig = px.bar(segment\_counts, x='Segment', y='Count',

title='Customer Segmentation by CLV')

fig.update\_xaxes(title='Segment')

fig.update\_yaxes(title='Number of Customers')

fig.show()

# Funnel analysis

funnel\_data = data[['Product\_Browsing\_Time', 'Items\_Added\_to\_Cart', 'Total\_Purchases']]

funnel\_data = funnel\_data.groupby(['Product\_Browsing\_Time', 'Items\_Added\_to\_Cart']).sum().reset\_index()

fig = px.funnel(funnel\_data, x='Product\_Browsing\_Time', y='Items\_Added\_to\_Cart', title='Conversion Funnel')

fig.show()

# Calculate churn rate

data['Churned'] = data['Total\_Purchases'] == 0

churn\_rate = data['Churned'].mean()

print(churn\_rate)

Print the error message

Save filtered rules to the corresponding file format

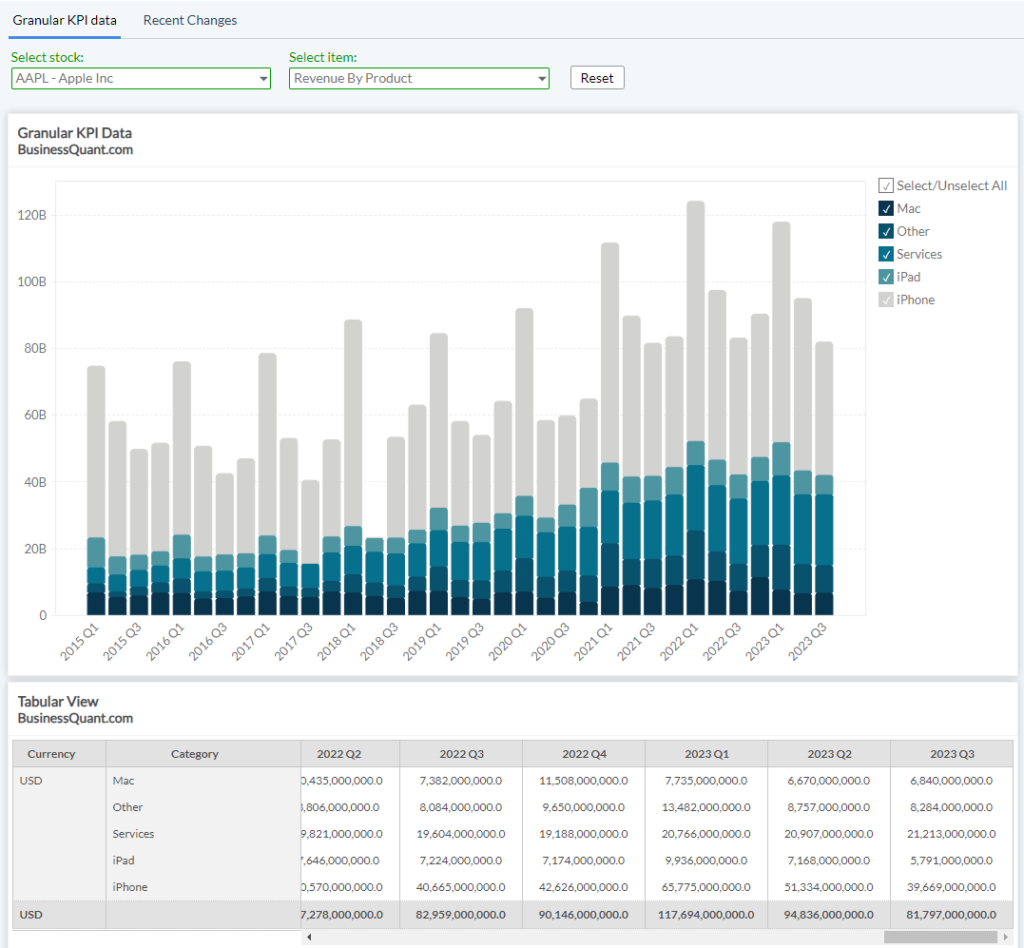
Show success dialog with the file name Catch Exception:

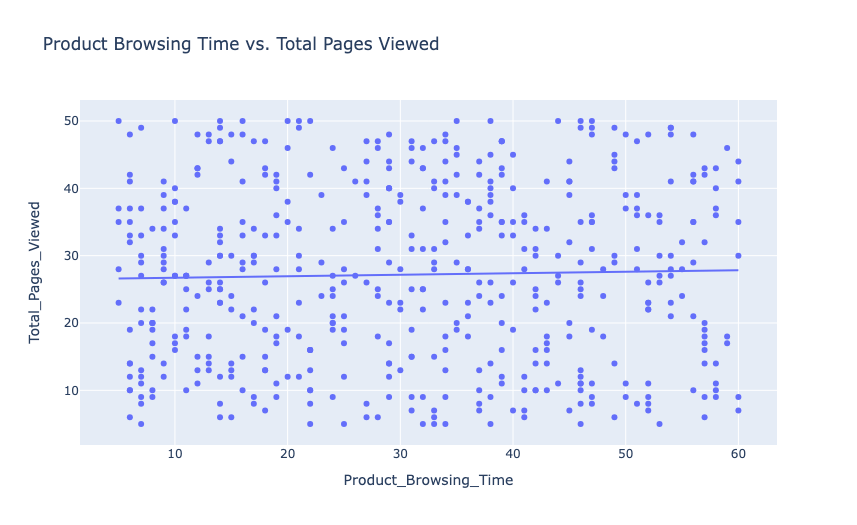
Print error message

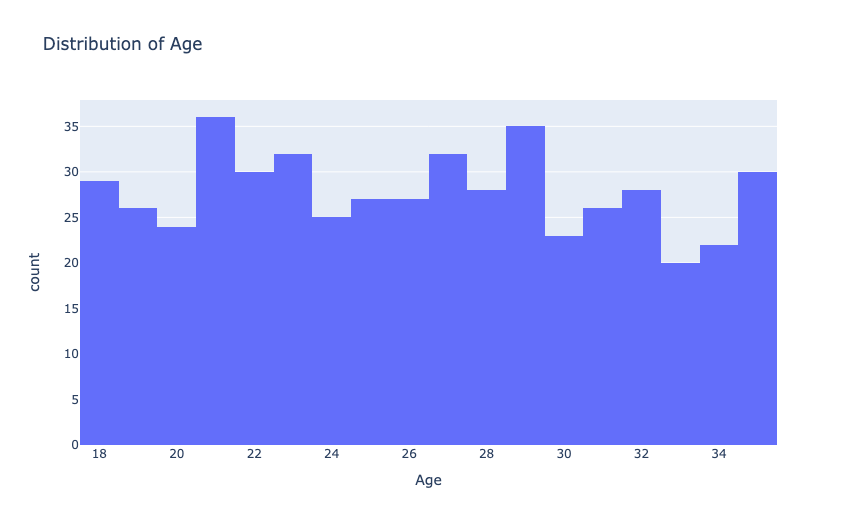
Show error dialog with the message

show error dialog with the message plot

**OUTPUT :**







**CONCLUSION:**

In conclusion, customer behaviour analysis and prediction are essential for understanding and anticipating customer actions. By collecting and integrating data from various sources such as purchase history, online interactions, social media activity, and customer feedback, businesses can create a comprehensive view of customer behaviour. This consolidated data enables more informed decision-making and provides a solid foundation for advanced analytical techniques.

Customer segmentation plays a crucial role in identifying distinct groups within the customer base, allowing businesses to tailor marketing strategies to specific needs and preferences. By recognizing patterns and creating personalized marketing campaigns, companies can enhance customer engagement and satisfaction. Effective segmentation also ensures efficient resource allocation, maximizing the impact of marketing efforts.

Furthermore, predictive modelling and sentiment analysis are indispensable tools for forecasting future behaviours and understanding customer sentiments. Machine learning algorithms help anticipate needs, identify potential churn risks, and optimize inventory management. Sentiment analysis provides insights into customer opinions and emotions, enabling businesses to address issues promptly and improve their offerings. Collectively, these functionalities enhance customer retention, loyalty, and overall satisfaction, fostering long-term relationships and driving sustainable growth. Predictive modelling and sentiment analysis are pivotal in refining customer behaviour strategies. Predictive modelling uses machine learning algorithms to analyse historical data, forecast future behaviours, and anticipate customer needs. This capability allows businesses to proactively address potential challenges, optimize inventory management, and align their offerings with anticipated demand.

**REFERENCES:**

1. Surendro, K. (2019, March). Predictive analytics for predicting customer behavior. In 2019 International Conference of Artificial Intelligence and Information Technology (ICAIIT) (pp. 230-233). IEEE.
2. Kalaivani, D., & Sumathi, P. (2019). Factor based prediction model for customer behavior analysis. International Journal of System Assurance Engineering and Management, 10(4), 519-524.
3. Khade, A. A. (2016). Performing customer behavior analysis using big data analytics. Procedia computer science, 79, 986-992.
4. Li, J., Pan, S., & Huang, L. (2019). A machine learning based method for customer behavior prediction. Tehnički vjesnik, 26(6), 1670-1676.
5. Khodabandehlou, S., & Zivari Rahman, M. (2017). Comparison of supervised machine learning techniques for customer churn prediction based on analysis of customer behavior. Journal of Systems and Information Technology, 19(1/2), 65-93.
6. Raju, S. S., & Dhandayudam, P. (2018, April). Prediction of customer behaviour analysis using classification algorithms. In AIP conference proceedings (Vol. 1952, No. 1). AIP Publishing.
7. Siva Subramanian, R., & Prabha, D. (2021). Customer behavior analysis using Naive Bayes with bagging homogeneous feature selection approach. Journal of Ambient Intelligence and Humanized Computing, 12, 5105-5116.
8. Maheswari, K., & Priya, P. P. A. (2017, March). Predicting customer behavior in online shopping using SVM classifier. In 2017 IEEE international conference on intelligent techniques in control, optimization and signal processing (INCOS) (pp. 1-5). IEEE.
9. Quynh, T. D., & Dung, H. T. (2021). Prediction of customer behavior using machine learning: A case study. In Proceedings of the 2nd International Conference on Human-centered Artificial Intelligence (Computing4Human 2021). CEUR Workshop Proceedings, Da Nang, Vietnam (Oct 2021).
10. Dhandayudam, P., & Krishnamurthi, I. (2013). Customer behavior analysis using rough set approach. Journal of theoretical and applied electronic commerce research, 8(2), 21-33.