
Leveraging Seaborn and XGBoost for Customer Analysis and Early Dissatisfaction Detection

Prof. Sayali Belhe, Siddhesh Kulthe, Chirayu Patil, Krishna Patil, Rushikesh
Pathare, Dnyaneshwar Patil

Department of AI & DS, Vishwakarma Institute of Technology, Pune,
Sayali.belhe@vit.edu, siddheshkulthe43@gmail.com, chirayupatil333@gmail.com,
krishna.patil24@vit.edu, rushikesh.pathare24@vit.edu, dnyaneshwar.patil24@vit.edu

Abstract.

Customer dissatisfaction can significantly affect business survival and purchase decisions. Traditional quantitative surveys measure satisfaction but often fail to capture customers' emotional responses and deeper attitudes. This study analyzes factors influencing purchase decisions using Seaborn for data analysis and XGBoost for prediction, aiming to identify customer dissatisfaction at an early stage and support timely corrective actions.

Keywords. Machine learning, Extreme Gradient Boosting, Customer Satisfaction.

1. INTRODUCTION

A customer is an individual or organization receiving goods or services. Identifying target customers supports effective marketing and customer-focused products (5, 8). XGBoost is an ensemble algorithm that combines decision trees to provide accurate predictions while handling missing data and reducing overfitting through regularization. The prediction is:

$$\hat{y}_i = \varphi(x_i) = \sum_{k=1}^K f_k(x_i)$$

Here, $\hat{y}_i$ is the predicted output, K is the number of trees, and $f_k(x_i)$ is the prediction of the k -th tree. XGBoost builds trees sequentially to correct previous errors and uses L1/L2 regularization and pruning to reduce overfitting. It also supports feature-importance analysis and parallel computing. Seaborn is a Python visualization library that works with Pandas to create clear statistical plots for distributions, correlations, and other relationships.

2. RELATED WORKS

2.1. Literature Review

Machine learning and visualization techniques, particularly Seaborn and XGBoost, are increasingly used for customer analysis and early dissatisfaction detection.

2.2. Seaborn In Customer Analysis

Seaborn helps visualize customer behaviour, preferences, distributions, correlations, and demographic patterns. Its integration with Pandas supports clear and effective data exploration [14].

2.3. *XGBoost for Early Dissatisfaction Detection*

XGBoost provides accurate prediction and identifies subtle behavioural patterns associated with customer dissatisfaction. Regularization and tree pruning improve generalization and reduce overfitting [15].

2.4. *Integration Of Seaborn and XGBoost*

Seaborn helps identify important customer features, while XGBoost uses these features for dissatisfaction prediction. Their integration provides both visual insights and predictive analysis of customer behaviour.

3. OBJECTIVES & SCOPE

3.1. *Objective:*

To identify factors such as price, features, promotions, and brand image that influence purchase decisions and early customer dissatisfaction.

3.2. *Scope:*

The study focuses on predicting early dissatisfaction and understanding how key customer and marketing factors influence purchase decisions for a new business without historical customer data.

4. PROCEDURE

The research examines customer preferences and perceptions using data collected through a simulated business environment. Participants interact with the products and respond, enabling analysis of purchasing behaviour.

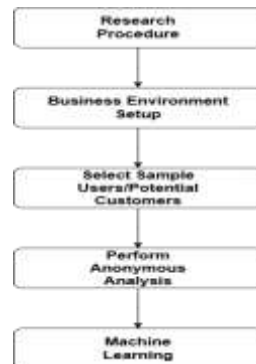


Figure 4.1: Research procedure

4.1. *Business Environment Creation:*

A simulated start-up business was created without existing customer relationships to study the effects of key variables on customer perceptions and purchasing decisions [6].

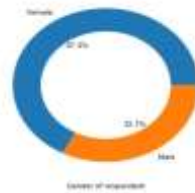


Figure 4.1.1: Gender of Respondents

The sample consisted of 67.3% female and 32.7% male respondents, indicating higher female participation.

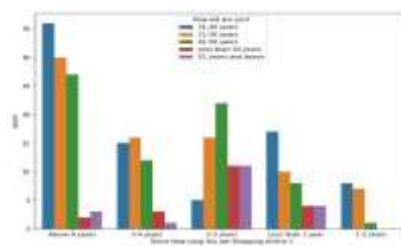


Figure 4.1.2: Shopping and Online Purchase Experience by Age

The 21–30 age group represented the largest proportion of respondents with shopping and online purchasing experience, reflecting their familiarity with digital technologies.



Figure 4.1.3: City of Purchase

The figure presents the geographical distribution of respondents who purchased the product.

5. EXPERIMENTAL RESULTS AND ANALYSIS

The experimental results provide insights into customer purchasing behaviour and the factors influencing purchase decisions. The analysis examines purchase frequency, cart abandonment, shopping duration, and important product features. These relationships help identify patterns of customer engagement and disengagement within the simulated business environment.

6. REFERENCES

- [1] R. L. Oliver, 'A cognitive model of the antecedents and consequences of satisfaction decisions,' J. Marketing Res., vol. 17, no. 4, pp. 460–469, Nov. 1980.

- [2] E. Garbarino and M. S. Johnson, 'The different roles of satisfaction, trust, and commitment in customer relationships,' *J. Marketing*, vol. 63, no. 2, pp. 70–87, Apr. 1999.
- [3] D. Gupta, M. Gupta, R. K. Dhanaraj, B. Balusamy, and P. M. Gupta, Eds., *Computational Intelligence for Autonomous Finance*. Hoboken, NJ: John Wiley & Sons, 2024.
- [4] D. Kang and Y. Park, 'Review-based measurement of customer satisfaction in mobile service: Sentiment analysis and VIKOR approach,' *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1041–1050.
- [5] Bob E. Hayes, 'Measuring Customer Satisfaction and Loyalty,' 2008.
- [6] Trevor Hastie, Robert Tibshirani, and Jerome Friedman, 'The Elements of Statistical Learning,' 2009.
- [7] Stephanie D. H. Evergreen, 'Effective Data Visualization: The Right Chart for the Right Data,' 2016.
- [8] Bob E. Hayes, 'Customer Satisfaction: The Customer Experience Through the Customer's Eyes,' 1998.
- [9] D. P. Verma, D. D. Gupta, and D. M. Gupta, "A comprehensive study on the implications of ESG performance disclosure in the promotion of sustainable development by firms," 2024.
- [10] C. Ranaweera and J. Prabhu, 'The influence of satisfaction, trust and switching barriers on customer retention in a continuous purchasing setting,' *Int. J. ...*

Biographies



Prof. Sayali Belhe received the bachelor's degree in Computer Science and Engineering (CSE) and the master's degree in Information Technology (IT). She is currently pursuing a PhD and working as a faculty member in the field of Computer Science and Engineering.



Siddhesh Kulthe is an Artificial Intelligence and Data Science (AI&DS) Engineer currently working in the IT industry. He is an industry expert with expertise in the field of Information Technology.



Chirayu Patil currently pursuing a bachelor's degree in Artificial Intelligence and Data Science from Vishwakarma Institute of Technology, Pune. He is currently in his third year of study.



Krishna Patil is currently pursuing a bachelor's degree in Artificial Intelligence and Data Science from Vishwakarma Institute of Technology, Pune. He is currently in his third year of study.



Rushikesh Pathare is currently pursuing a bachelor's degree in Artificial Intelligence and Data Science from Vishwakarma Institute of Technology, Pune. He is currently in his third year of study.



Dnyaneshwar Patil is currently pursuing a bachelor's degree in Artificial Intelligence and Data Science from Vishwakarma Institute of Technology, Pune. He is currently in his third year of study.