A Study on Customer Satisfaction Analysis using Machine Learning

Isaac Varghese[1*] and Rekha Francis C[2]

[1]*Student, Department of Commerce, Christ College (Autonomous), Irinjalakuda, Affiliated to Calicut University*

[2]*Assistant Professor, Christ College (Autonomous), Irinjalakuda, Affiliated to Calicut University*

[*]*Corresponding Author Email: isaacvarghese21@gmail.com*

Abstract

Customer satisfaction plays a vital role in determining business success, customer loyalty, and brand reputation in today's competitive market environment. With the rapid growth of digital platforms, organizations generate large volumes of customer data through reviews, surveys, and service interactions. Traditional methods of analyzing customer satisfaction are often manual, time-consuming, and prone to bias, limiting their effectiveness in providing timely insights. This study focuses on the application of machine learning techniques to analyze customer satisfaction more accurately and efficiently. Machine learning enables the automated processing of both structured and unstructured customer feedback data, allowing organizations to identify sentiment, key satisfaction drivers, and emerging trends. The study involves collecting and preprocessing customer feedback data, extracting relevant features, and applying supervised and unsupervised machine learning algorithms to evaluate customer satisfaction levels. The performance of these models is assessed using accuracy and other evaluation metrics. The findings of this study highlight the effectiveness of machine learning in enhancing customer satisfaction analysis and supporting data-driven decision-making. By leveraging machine learning models, organizations can gain actionable insights to improve customer experience and implement customer-centric strategies.

Keywords: Customer Satisfaction, Machine Learning, Data Analysis, Sentiment Analysis, Customer Feedback, Business Intelligence

1. Introduction

Customer satisfaction is a critical factor that directly impacts business performance, customer loyalty, and brand reputation. In today's highly competitive market, organizations strive to understand and improve customer satisfaction to retain existing customers and attract new ones. With the increasing availability of customer data, machine learning (ML)

offers powerful tools to analyze customer feedback, detect patterns, and predict satisfaction levels with higher accuracy and efficiency than traditional methods. Machine learning techniques enable automated processing of large volumes of structured and unstructured data, such as customer reviews, survey responses, and service interactions. These techniques can classify sentiments, identify satisfaction drivers, and even forecast customer behavior. Leveraging ML not only streamlines analysis but also provides actionable insights for strategic decision-making.

2. **Significance of the study**.

This study is significant for both academic and practical reasons. Academically, it contributes to the growing body of knowledge in the field of machine learning applications in customer relationship management. Practically, it provides a scalable and efficient approach to customer satisfaction analysis that can support decision-making in marketing, product development, and customer service.

Businesses can use the insights gained from ML models to:

* Enhance customer experiences
* Reduce churn
* Improve service quality
* Develop targeted marketing strategies

3. **Literature Review**

Shahid & Sufian (2024) This paper applies machine learning to predict whether airline customers will return, focusing on the impact of emotional connections and user experience. By analyzing sentiment in customer feedback using LIWC and satisfaction ratings, the study builds a predictive model trained on survey data from 17,000 respondents, with a follow-up survey conducted a year later. Among the models tested, XGBoost achieved the highest accuracy at 85%. The results demonstrate the effectiveness of combining sentiment analysis with UX features to understand and enhance customer loyalty in the airline industry.

Kumar & Zymbler (2019) This paper uses machine learning to analyze airline-related tweets on Twitter to enhance customer experience. Tweets are processed using word embedding with GloVe and n-gram techniques to extract features. Classification models,

including SVM and various artificial neural networks (ANNs), were developed to categorize tweets as positive or negative. A convolutional neural network (CNN) model was also implemented and outperformed both SVM and ANN models in classification accuracy. Additionally, association rule mining was applied to uncover meaningful patterns between tweet content and sentiment, offering valuable insights for improving airline customer service.

Siregar & Sinaga (2023) This paper uses a Backpropagation Artificial Neural Network (ANN) to classify customer satisfaction data from Café Alvina. With 70 training and 30 testing samples, four ANN models were evaluated: 5-5-2, 5-10-8-8-2, 5-5-10-2, and 5- 8-10-2. The 5-10-8-8-2 model achieved the best performance, reaching 100% accuracy with very low training and testing MSE values. Other models achieved accuracies between 81% and 98%. The results show that the AN N, especially with deeper architectures, can effectively and accurately classify customer satisfaction data.

Bouzakraoui & Sadiq (2020) This paper introduces a machine learning-based method for detecting customer satisfaction through facial emotion recognition. Geometric features are extracted from facial landmark points by measuring distances that reflect emotional changes from a neutral expression. These features are classified using various algorithms, including SVM, KNN, Random Forest, Adaboost, and Decision Tree. Testing on the JAFFE dataset showed that the SVM classifier achieved the highest accuracy at 98.66%. The proposed approach offers a fast and efficient alternative to traditional customer feedback methods like surveys.

Chen & Li (2020) This paper addresses the challenges of sentiment analysis in customer comments, which often suffer from emotional ambiguity, short length, and vague semantics. To improve accuracy, it introduces agent notes as a complementary user-generated content (UGC) source, offering more detailed and professional insights. The proposed machine learning based method, CAMP, analyzes and matches sentiments and semantics between customer comments and agent notes. It also detects fine-grained aspect mismatches, providing deeper customer insights. CAMP enhances follow-up services and enables faster response times, ultimately boosting customer satisfaction and loyalty.

4. Theoretical and conceptual framework

Customer satisfaction is a critical performance indicator that reflects how well products or services meet or exceed customer expectations, and its analysis has been significantly enhanced through the integration of machine learning (ML). Traditional theories such as Expectancy–Disconfirmation Theory, Equity Theory, and the SERVQUAL model form the conceptual foundation of customer satisfaction analysis. Expectancy–Disconfirmation Theory explains satisfaction as the gap between customer expectations and actual performance, which ML can quantify through sentiment and feedback analysis. Equity Theory emphasizes perceived fairness in the exchange of value between customers and organizations, where ML models can identify fairness perceptions using behavioral and transactional data. The SERVQUAL model evaluates service quality across five dimensions—tangibles, reliability, responsiveness, assurance, and empathy—and machine learning techniques such as text classification and clustering can automate the extraction and categorization of these service quality attributes from customer feedback.

Machine learning techniques further strengthen customer satisfaction analysis through supervised, unsupervised, and deep learning approaches. Supervised learning models, including regression and support vector machines, predict satisfaction scores using labeled historical data, while unsupervised learning techniques such as k-means clustering and association rule mining uncover hidden patterns and satisfaction drivers without predefined labels. Natural Language Processing (NLP) enables real-time sentiment analysis of customer reviews and feedback, offering actionable insights into customer perceptions. Deep learning models, particularly recurrent neural networks (RNNs) and transformer architectures, analyze sequential customer interaction data to predict satisfaction trends and behaviors. By integrating data collection with ML models for classification, regression, and clustering, organizations can generate meaningful insights such as satisfaction scores and sentiment trends, enabling informed decision-making. This synergy between traditional customer satisfaction theories and machine learning methodologies empowers organizations to develop predictive, personalized, and effective customer experience strategies.

5. Objectives of the study

The objectives of the study were:

i. To review customer satisfaction using machine learning techniques

ii. To analyse the scope of machine learning in different sectors

6. Methodology

The study collected primary data from IT professionals in Thrissur, Palakkad, Kochi, and Malappuram.

Sample size: 41 respondents

Tool: Structured questionnaire

Mode: Google Forms

Type of data: Primary data

7. Findings

1. Demographics & Employment:
   - 73.3% of customers are female, and 26.7% are male.
   - 40% are unemployed, 26.7% are employed, and 33.3% are students. - 46.7% have income below ₹10,000, while 10% earn between ₹10,000–₹20,000, 16.7% between ₹20,000–₹30,000, and 26.7% above ₹30,000.
2. Shopping Behavior:
   - 6.7% shop daily, 13.3% weekly, 33.3% monthly, and 46.7% shop rarely.
   - 30% are extremely likely to recommend the store, 60% likely, and 10% not likely.
3. Product Availability:
   - 20% always find what they need, 43.3% most of the time, 30% sometimes, 3.3% rarely, and 3.3% never.
4. Speed of Service:
   - 16.7% very satisfied, 50% satisfied, and 33.3% neutral. No dissatisfaction reported.
5. Promotions and Discounts:
   - 33.3% frequently use them, 43.3% occasionally, while 6.7% always, 10% rarely, and 6.7% never.
   - 70% find offers attractive: 60% agree and 10% strongly agree.
6. Discounts Impact:
   - 66.7% say discounts drew them: 60% agree and 6.7% strongly agree.
   - 26.7% remained neutral, only 6.7% disagreed.

7. Pricing:
    - 63.3% say pricing attracted them: 50% agree and 13.3% strongly agree.
    - 33.3% were neutral and 3.3% disagreed.
8. Product Availability & Store Choice:
    - 66.7% agree availability drew them: 50% agree and 16.7% strongly agree.
    - 26.7% were neutral, and 6.7% disagreed.
9. Convenience:
    - 80% say store location/access is important: 70% agree and 10% strongly agree.
    - 16.7% neutral, 3.3% disagreed.

8. Conclusion

The survey results reflect a growing awareness and acceptance of machine learning across various sectors, with a clear majority of respondents recognizing its potential to transform industries such as finance, education, and manufacturing. Most participants have had positive interactions with ML-based services and believe in its ability to accurately predict consumer preferences. Despite concerns around implementation costs, infrastructure, and ethics, the willingness to upskill and adopt ML tools is evident. The data also highlights the critical role of customer satisfaction and data-driven decision-making in 28 maximizing the impact of machine learning. These insights collectively suggest that with the right strategies and support systems in place, ML integration can significantly enhance organizational productivity and innovation.