

Machine Learning Techniques for Customer Satisfaction Prediction using Online Reviews

Ms. Shivani Jain

Assistant Professor, Department of Computer Science and Applications

Mandsaur University, Mandsaur

madhav.sharma@meu.edu.in

Abstract: Customer satisfaction is a key indicator of a successful modern business, making accurate customer satisfaction prediction an important task in Business Intelligence. With the rapid growth of online reviews, machine learning techniques can effectively analyze customer feedback and identify sentiment patterns associated with satisfaction. This study proposes an XGBoost-based machine learning (ML) approach for customer satisfaction prediction using the IMDb Movie Reviews Dataset containing 50,000 labeled reviews. The proposed methodology includes text preprocessing, spell correction, stop-word and punctuation removal, tokenization, TF-IDF feature extraction, z-score normalization, and an 80:20 training-testing split. The extracted textual features are subsequently provided to the XGBoost model for sentiment-based prediction. Experimental results demonstrate that the proposed model achieves 99.98% accuracy, 99.99% precision, 99.96% recall, and 99.97% F1-score, indicating highly effective classification performance. The results show that the suggested method can help businesses improve customer experience, happiness, and loyalty by providing trustworthy analysis of internet reviews.

Keywords: Sentiment Analysis, IMDB Review Dataset, Machine Learning Models, Data Preprocessing, Model Performance Evaluation.

I. INTRODUCTION

Consumer-generated evaluations are becoming more widely available on social media, online marketplaces, and review websites, including hotels, thanks to the quick development of digital platforms [1]. These reviews offer insightful information about consumer satisfaction and experiences [2][3]. Customer satisfaction is an important factor in determining the success and sustainability of businesses in today's competitive environment. It reflects the extent to which customers are satisfied with the products, services, and experiences provided by an organization [4][5]. Understanding customer satisfaction enables businesses to improve their products, services, and customer relationships, thereby increasing customer loyalty, sales, and revenue [6]. Traditionally, satisfaction has been evaluated through interviews, focus groups, questionnaires, and surveys. However, these methods are often time-consuming, costly, difficult to scale, and less effective when large volumes of customer feedback must be analyzed [7][8].

The rapid growth of internet services, smartphones, social media, and e-commerce platforms has generated enormous amounts of customer-generated data in the form of ratings, comments, complaints, and online reviews [9]. Online shopping provides customers with convenient opportunities to purchase products and share their experiences through ratings and written feedback[10]. Online reviews, which offer useful information on product quality, delivery performance, service quality, and purchase experiences, have grown in importance as a type of electronic word-of-mouth communication. While negative reviews can highlight shortcomings in a product or service and point out areas that need development, positive ones can inspire new buyers [11][12][13].

Sentiment analysis enables the identification and classification of opinions and emotions expressed in customer-generated text [14]. Customer feedback can be classified into positive, negative, and neutral categories, while advanced



approaches can identify fine-grained or aspect-level sentiments related to specific products and services [15][16]. Analyzing online reviews helps businesses understand customer behavior, identify product strengths and weaknesses, improve service quality, and enhance customer loyalty [17]. However, extracting meaningful information from unstructured reviews remains challenging because of noisy text, negation, high-dimensional features, and imbalanced sentiment classes [18][19].

In this context, Machine Learning (ML) and Deep Learning (DL) techniques provide effective approaches for automatically analyzing large volumes of online reviews and predicting customer satisfaction [20]. ML and DL models can learn meaningful patterns from customer reviews and classify their satisfaction levels efficiently. The suggested method seeks to transform unstructured consumer input into actionable insights by combining Natural Language Processing (NLP) with Machine Learning (ML) and Deep Learning (DL) techniques. This would assist organizations in understanding customer expectations, improving goods and services, and supporting data-driven decision-making [21][22] [23].

A. Motivation and Contribution

This study is driven by the increasing number of online customer evaluations and the importance of properly analyzing them to evaluate consumer satisfaction. Traditional feedback methods are time-consuming and often ineffective at processing large-scale unstructured textual data. Therefore, an automated machine learning approach is needed to extract meaningful sentiment patterns and generate accurate satisfaction predictions. This study aims to develop an effective XGBoost-based framework to support data-driven decisions and improve customer experience and loyalty. This research offers several key contributions as listed below:

- Developed an XGBoost-based framework for accurate customer satisfaction prediction from online movie reviews.
- Applied comprehensive text preprocessing, including spell correction, stop-word removal, punctuation removal, and tokenization.
- Extracted meaningful textual features using TF-IDF to represent customer reviews numerically for machine learning.
- Applied z-score normalization to standardize the extracted features and improve model learning.
- Outperformed KNN, SVM, BERT, and LSTM models as well.
- Provided a reliable approach for customer feedback analysis, supporting improved customer experience, satisfaction, and long-term loyalty.

B. Justification and Novelty

The proposed approach is justified by the growing volume of online reviews, which provides valuable information for understanding customer satisfaction and sentiment. The novelty of this work is that it combines large text preprocessing, TF-IDF feature extraction, z-score normalization, and XGBoost into a single prediction mechanism. The suggested model outperforms state-of-the-art methods in terms of prediction accuracy while also successfully capturing key trends in review data. The results demonstrate its potential as an accurate and reliable solution for automated customer satisfaction analysis.

C. Organization of the Paper

The remainder of this paper is structured into five sections: Section II discusses existing research on customer satisfaction analysis and prediction, while Section III describes the dataset, preprocessing techniques, and proposed model. Section IV reports the experimental outcomes and comparative performance, and Section V summarizes the conclusions and outlines potential directions for future work.



II. LITERATURE REVIEW

Existing research on Customer Satisfaction Analysis and Prediction was systematically reviewed to find previously done approaches, datasets and research gaps. To summarize, the recent research is presented in Table 1, including adopted models, datasets, major outcomes and key limitations.

B. Srinivas and M.Suresh, (2026) build a machine learning system that can be understood by humans to forecast multi-category consumer happiness with the use of the Social Network Ads Dataset. In order to test six different classification algorithms—Logistic Regression, DT, RF, Gradient Boosting, AdaBoost, and SVM with RBF kernel—they build nonlinear interaction characteristics. The best result is achieved by SVM-RBF, which has an accuracy rate of 90.0% and a macro-F1 score of 0.8256 ± 0.012 . According to feature analysis, the three most important predictors are age, expected salary, and the interaction between the two, whereas gender plays a relatively minor role. The framework provides actionable insights for improving social media advertising strategies and data-driven marketing [24].

A. Parupalli (2025) A deep learning-based system that also features sentiment analysis to produce very accurate results in the prediction of customer satisfaction. After integrating advanced text preparation methods grounded on the AraCust dataset, the system uses a Long Short-Term Memory (LSTM) model to detect contextual and sequential patterns in the text. It is determined by its metrics performance and also demonstrates that it outperforms other traditional machines on predictive accuracy and generalization ability. The LSTM algorithm achieved an accuracy of 97.03%, precision of 97.12%, recall of 93.34% and F1-score of 95.19% which is significantly superior to the traditional classifiers including XGBoost, Random Forest and SVM. The strength of the model and its capability to discriminate is also proved by such indicators as analysis of confusion matrix and ROC curve (AUC = 0.96) [25].

S. A. Rajagukguk, (2025) presents a machine learning framework integrating behavioral intention analytics and advanced feature engineering for improved CLV prediction. Using 300 customer records from a digital service provider, they evaluate eight classification and regression models. Feature engineering includes Satisfaction Gap Score, weighted Loyalty Index, and K-means-based customer segmentation. XGBoost achieves the best classification performance with 94.2% accuracy and 0.963 AUC-ROC, while Random Forest Regressor achieves an R^2 of 0.923. Feature analysis shows that behavioral features outperform demographic variables, with Loyalty Index and Satisfaction Gap Score as the strongest predictors [26].

M. Haris, Al-Khowarizmi, and O. Nugroho (2024) use the integration models IndoLEM and IndoBERT. The field of natural language processing (NLP) with an emphasis on Indonesian can benefit from these two approaches. While IndoLEM handled label-based question grouping and data vector representation in the IndoLEM process, IndoBERT understood context and made better predictions. Testing was carried out in two stages, with the results showing an increase in accuracy from 83% in the first test to 90% in the second test [27].

P. Wangkiat and C. Polprasert, (2023) examine ML performance for predicting customer satisfaction scores using the Olist Brazilian e-commerce sales dataset. Satisfaction is categorized into four classes: Low, Average, Good, and Excellent. They develop features based on delivery duration and average product rating and evaluate RF, LR, and K-NN against a baseline model. RF achieves the best performance, with average precision, recall, and macro-F1 scores of 0.34, 0.36, and 0.32, respectively. Product rating features are the most important, with mean and SD achieving importance scores of 0.313 and 0.087, respectively [28].

K. M. B. A. Joolfoo, R. A. Jugurnauth, and M. B. A. Joolfoo, (2022) demonstrate how ML can predict customer satisfaction regarding call-drop quality in telecom services. Using a Kaggle dataset, they develop a Random Forest Classifier based on call rating, call drop, and number of subscribers. Using data from September to November, the model evaluates call quality and customer ratings through recall, precision, accuracy, and F1-score. The developed model achieves 91% accuracy in predicting customer satisfaction with call quality [29].



Table I. Recent Research on Customer Satisfaction Analysis and Prediction using Online Reviews

Author	Proposed Work	Results	Key Findings	Limitations & Future Work
B. Srinivas and M. Suresh (2026)	Developed an interpretable ML framework for multi-class customer satisfaction prediction using demographic and purchasing-power features. Evaluated six ML classifiers.	SVM-RBF: 90.0% accuracy, Macro F1 = 0.8256 ± 0.012 .	Age and estimated salary were the most influential predictors; gender had minimal impact.	Focuses on demographic and purchasing data rather than review text. Future work can incorporate review text and NLP features.
A. Parupalli (2025)	Proposed a deep learning-based customer satisfaction prediction system using sentiment analysis and LSTM on the AraCust dataset.	Accuracy: 97.03%, Precision: 97.12%, Recall: 93.34%, F1-score: 95.19%, AUC = 0.96.	LSTM captured contextual and sequential patterns and outperformed traditional ML classifiers.	Can be extended using larger, diverse review datasets and transformer-based models.
S. A. Rajagukguk (2025)	Developed an ML framework for Customer Lifetime Value prediction using feature engineering, Satisfaction Gap Score, Loyalty Index, and K-means.	XGBoost: 94.2% accuracy, 0.963 AUC-ROC; RF Regressor: $R^2 = 0.923$.	Loyalty Index and Satisfaction Gap Score were stronger predictors than demographic variables.	Small dataset (300 records) and no primary focus on review text. Future work can integrate textual feedback.
M. Haris, Al-Khowarizmi, and O. Nugroho (2024)	Integrated IndoLEM and IndoBERT for Indonesian NLP to classify customer questions and improve contextual understanding.	Accuracy improved from 83% to 90%.	IndoBERT improved contextual understanding and customer-response prediction.	Focuses on customer questions rather than satisfaction from reviews. Future work can apply it to multilingual review sentiment and satisfaction prediction.
P. Wangkiat and C. Polprasert (2023)	Used ML to predict four customer satisfaction classes from the Olist Brazilian e-commerce dataset using engineered features.	RF: Precision = 0.34, Recall = 0.36, Macro F1 = 0.32; best recall = 0.43.	Product-rating mean and SD were the most important features (0.313 and 0.087).	Uses structured attributes and ratings rather than raw review text. Future work can apply NLP and sentiment analysis to reviews.
K. M. B. A. Joolfoo, R. A. Jugurnauth, and M. B. A. Joolfoo (2022)	Developed an RF model to predict telecom customer satisfaction using call rating, call drops, and subscriber information.	RF achieved 91% accuracy.	Call ratings and quality parameters can predict telecom customer satisfaction.	Limited to telecom call-quality data and three months of observations. Future work can use larger datasets and review sentiment.



Research gaps: The existing work has mainly focused on structured customer data, churn prediction and domain-specific satisfaction analysis, with few studies addressing large-scale online review text. Many methods require using complex deep learning models or have relatively poor predictive capability. Limited work has explored the integration of comprehensive text preprocessing, TF-IDF feature extraction, and XGBoost for customer satisfaction prediction. Therefore, this study proposes an efficient XGBoost-based framework to improve the accuracy and reliability of satisfaction prediction from online reviews.

III. RESEARCH METHODOLOGY

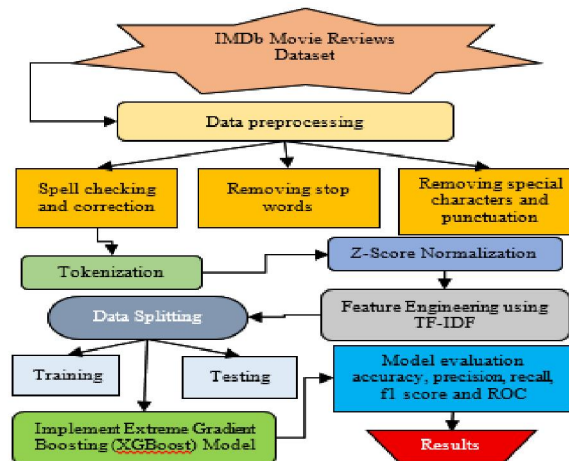


Fig.1. Proposed Framework for Analyzing and Predicting Customer Satisfaction from Online Reviews using ML. The proposed methodology involves collecting and preprocessing the IMDb movie review dataset through text cleaning, spell correction, stop-word and punctuation removal, tokenization, and TF-IDF feature extraction, followed by z-score normalization and an 80:20 train-test split. For sentiment-based customer satisfaction prediction, the retrieved features are then fed into the suggested XGBoost model. Confusion matrices, accuracy, precision, recall, F1-score, and ROC-AUC are the last metrics used to assess the model's proficiency. Figure 1 shows the suggested pipeline for utilizing machine learning to analyze and forecast consumer happiness based on online reviews.

The following section provides a detailed description of each step involved in the proposed methodology:

Data Gathering and Analysis

The IMDb Movie Reviews Dataset, compiled by Andrew Maas, was used for customer satisfaction analysis and prediction from online reviews, containing 50,000 reviews, divided equally between training and testing sets. The reviews include positive and negative sentiments based on IMDb ratings and were prepared for subsequent preprocessing and model development. Bar plots and heatmaps were used to visualize sentiment distribution and examine relationships among extracted features:

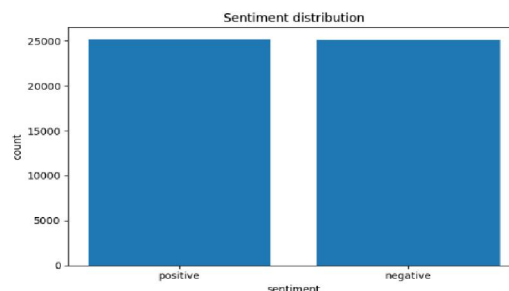


Fig. 2. Distribution of Positive and Negative Sentiments in the IMDB Dataset



Figure 2 presents the distribution of positive and negative sentiments in the IMDb dataset, with both sentiment classes having nearly equal numbers of reviews. The balanced distribution of the dataset suggests that it is appropriate for sentiment classification without bias.



Fig 3. Correlation Matrix Heatmap

Figure 3 presents the relationship between the review and sentiment variables in the IMDb dataset. The color intensity represents the strength and direction of the correlation between the variables. The values are close to zero, indicating a weak linear relationship between the displayed variables. This suggests that the review text itself does not exhibit a direct numerical correlation with the sentiment label without further feature extraction.

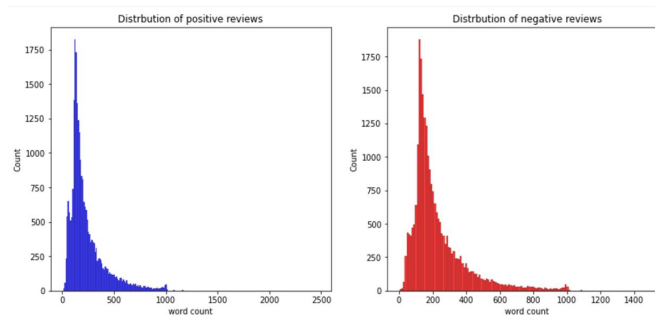


Fig. 4 Histograms of Word Count Distribution in Positive and Negative IMDB Reviews

Figure 4 illustrates the word-count distribution of positive and negative IMDb reviews. Both distributions are right-skewed, with most reviews containing relatively few words and fewer reviews having very high word counts. The similar patterns suggest that the reviews of both sentiment classes share similar review-length characteristics.

B. Data Pre-Processing

The IMDB Movie Reviews Dataset was utilized for data preparation, including data concatenation, cleaning, and feature engineering. The preprocessing process involved spell checking and correction, removal of stop words, elimination of special characters and punctuation, and tokenization. Finally, data labeling and normalization were performed to prepare the dataset for subsequent analysis and model development. The major preprocessing steps are summarized as follows:

- **Spell checking and correction:** To enhance the accuracy of the sentiment analysis model, it is imperative to correct spelling mistakes that might be present in the dataset. Python offers several libraries such as Spellchecker and TextBlob that provide functionalities for spell-checking and automatic corrections.
- **Removing stop words:** Common words like "the", "and", "is", known as stop words, are filtered out to reduce dataset dimensionality and boost model performance. Tools for this include NLTK and spaCy in Python.
- **Removing special characters and punctuation:** The dataset might include special characters and punctuation that do not contribute to sentiment analysis, hence are eliminated using regular expressions or Python's string library.



- **Tokenization:** This involves segmenting text into tokens, typically single words, which is a crucial step for feature engineering, enabling the creation of a bag of words or n-grams.

C. Z-Score Normalization

Data Normalization is a technique used to transform or standardize data to achieve a more uniform distribution. The most widely used methods for data normalization are rescaling or min-max normalization and z-score normalization. The z-score normalization method, which has a mean of 0 and a standard deviation of 1, was used in this investigation. This scaling technique transforms values centered around the average, with a unit standard deviation. The z-score normalization is defined by Equation (1).

$$E' = \frac{E - \bar{M}}{\sigma_M} \quad (1)$$

Where, E' and E are new and old for each data entry, $\bar{M}$ is the mean, and σ_M is the standard deviation.

E. Feature Engineering using TF-IDF

Feature engineering converts unstructured textual data into meaningful numerical features suitable for machine learning models in NLP and sentiment analysis. Term Frequency-Inverse Document Frequency (TF-IDF) measures the importance of words within a document relative to the entire corpus. TF highlights frequently occurring terms, while IDF assigns greater importance to words that are relatively rare across documents. In this study, TF-IDF is used to extract significant textual features and improve the predictive performance of the sentiment analysis model.

F. Data Splitting

The dataset was split into 80% training data for model development and 20% test data for evaluating classification performance. The test data was used to classify previously unseen instances and assess the model's effectiveness.

F. Proposed Extreme Gradient Boosting (XGBoost) Model

In this study, a supervised machine learning approach based on Extreme Gradient Boosting (XGBoost) is proposed for predicting customer satisfaction. XGBoost is an ensemble learning technique that combines multiple decision trees sequentially to improve predictive performance. The model minimizes a predefined loss function while incorporating regularization to control model complexity and reduce overfitting. For regression-based customer satisfaction prediction, XGBoost estimates the target value by combining the outputs of multiple decision trees. The mathematical formulation of the proposed XGBoost regression model is expressed in Equation (2):

$$y = f(x) \quad (2)$$

where y is the predicted property price, x is the vector of input features (such as square footage, number of bedrooms, etc.), and $f(x)$ is the XGBoost model that predicts y based on x . In order to calculate $f(x)$, XGBoost constructs a network of decision trees that have been trained to minimize the loss function representing the MSE. The model combines the predictions from multiple decision trees to arrive at a final prediction. The general form of the XGBoost regression model can be expressed as (3):

$$y = \sum_{k=1}^K f_k(x) \quad (3)$$

where $f_k(x)$ is the prediction of the k th decision tree, and K is the total number of decision trees in the ensemble. Each decision tree generates a prediction based on the learned values assigned to its leaf nodes. The final XGBoost prediction for an input (x) is obtained by combining the outputs of all trees in the ensemble, where each tree contributes to the overall prediction.

The proposed XGBoost model was configured with 300 estimators, maximum depth of 6, learning rate of 0.05, subsample of 0.8, and column sampling of 0.8. The `min_child_weight`, `gamma`, `Regale`, and `regaled` values were set to 1, 0, 0.01, and 1.0, respectively. The model used binary: logistic as the objective function, log loss as the evaluation metric, and 42 as the random state.

Evaluation Metrics

The effectiveness of the proposed model was evaluated using several standard performance metrics. A confusion matrix was generated to summarize the classification outcomes by presenting the numbers of correctly and incorrectly



classified instances for each class. The matrix provides four fundamental values: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). These values were subsequently used to calculate the key evaluation metrics, including accuracy, precision, recall, and F1-score, as presented below (4) to (7):

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (3)$$

$$recision = \frac{TP}{TP + FP} \quad (4)$$

$$ecall = \frac{TP}{TP + FN} \quad (5)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

Accuracy is the percentage of samples the model accurately classifies. Precision is the percentage of true positive predictions among all predictions labeled positive, indicating the accuracy of positive predictions. Recall is the percentage of positive instances that are correctly predicted. The F1 score is a harmonic mean of precision and recall, ranging from 0 to 1 and indicating an overall balanced score. ROC Curve illustrates the model's classification performance across different decision thresholds by plotting the TPR, also known as sensitivity or recall, against the False Positive Rate (FPR), which is calculated as 1 – specificity.

IV. RESULTS AND DISCUSSION

This section describes the experimental setup and performance of the proposed model. The experiments were conducted on an Intel Core i9-13900K CPU, 64 GB DDR5 RAM, and NVIDIA RTX 4090 GPU with 24 GB VRAM, running Windows 11 Pro. Data processing, model creation, visualization, and analysis were all carried out in Python using NumPy, Matplotlib, Seaborn, Scikit-learn, and XGBoost. The IMDb Movie Reviews Dataset was used to train and assess the proposed XGBoost model utilizing metrics like as accuracy, precision, recall, and F1-score. As presented in Table II the model achieved 99.98% accuracy, 99.99% precision, 99.96% recall, and 99.97% F1-score, demonstrating effective and reliable customer satisfaction classification from online reviews.

Table II. Classification results of proposed model on Customer Satisfaction Analysis and Prediction Through Online Reviews

Matrix	Extreme Gradient Boosting (XGBoost)
Accuracy	99.98
Precision	99.99
Recall	99.96
F1-score	99.97

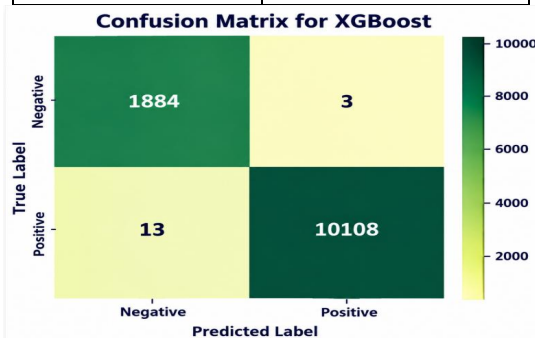


Fig 5. Confusion Matrix for the XGBoost Model



Figure 5 presents the classification results for the negative and positive classes. The model correctly classified 1,884 negative and 10,108 positive samples, with only 3 and 13 misclassifications, respectively. These results demonstrate the high classification accuracy and strong predictive performance of the XGBoost model.

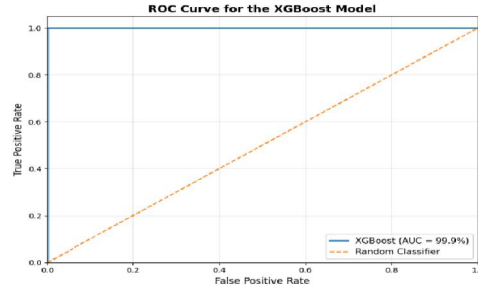


Fig 6. ROC Curve for Xgboost Model

Figure 6 shows the XGBoost model's ROC Curve, which shows how well it can differentiate between the positive and negative categories. The model achieves an excellent AUC of 99.9%, indicating highly accurate classification performance. The curve approaches the top-left corner, indicating a very low false-positive rate and a high true-positive rate.

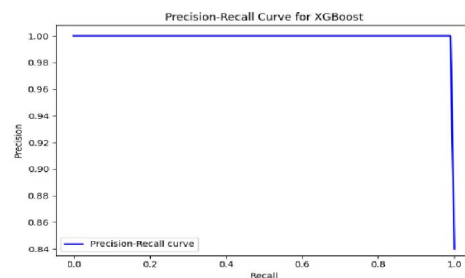


Fig 7. Precision-Recall Curve for XGBoost

Figure 7 illustrate curve demonstrates that the XGBoost model maintains nearly 100% precision across a wide range of recall values. This indicates strong classification performance with very few false-positive predictions. The results confirm the model's high reliability in distinguishing between the target classes.

A. Comparative Analysis

The usefulness of the proposed XGBoost model may be evaluated by comparing its performance with other current deep learning and machine learning models, as shown in Table 3. The KNN model achieved 89% accuracy, 82% precision, and 56% recall, while SVM obtained 90% accuracy, 87% precision, 90% recall, and 88% F1-score. BERT demonstrated improved performance with 90.75% accuracy, 89.87% precision, 91.87% recall, and 90.85% F1-score, whereas LSTM achieved 91% accuracy, 89% precision, and 92% recall. In comparison, the proposed XGBoost model achieved the highest overall performance, with 99.98% accuracy, 99.99% precision, 99.96% recall, and 99.97% F1-score, demonstrating its superior capability for predicting customer satisfaction from online reviews.

Table iii. Comparison of Different Machine learning and deep learning Models for Predicting Customer Satisfaction from Online Reviews

Model	Accuracy	Precision	Recall	F1-score
KNN[30]	89	82	56	-
SVM[31]	90	87	90	88
BERT[32]	90.75	89.87	91.87	90.85
LSTM [33]	0.91	0.89	0.92	-
Proposed XGBoost	99.98	99.99	99.96	99.97



The proposed XGBoost model demonstrates superior predictive performance, achieving 99.98% accuracy in customer satisfaction prediction. Its boosting-based ensemble structure effectively captures complex patterns in online review data while reducing prediction errors and improving generalization. Therefore, XGBoost provides a highly accurate and reliable approach for automated customer satisfaction analysis.

V. CONCLUSION AND FUTURE STUDY

Accurate customer satisfaction prediction is vital to organizations whose goals are to enhance customer experience, reduce churn, and enhance long-term loyalty. The increasing volume of customer interaction data provides valuable opportunities to identify behavioral patterns and predict satisfaction levels effectively. Based on the experimental results, the proposed XGBoost model achieved an exceptional accuracy of 99.98%, outperforming KNN (89%), SVM (90%), BERT (90.75%), and LSTM (91%). These results demonstrate that the proposed model can effectively learn important patterns from online reviews and provide highly accurate sentiment-based customer satisfaction prediction. Therefore, the proposed XGBoost approach provides a robust and efficient solution for analyzing customer feedback and can support organizations in making informed decisions to improve customer experience and retention.

The IMDB dataset and the sentiment-based classification of movie reviews might not accurately capture the varied customer reviews in the real world. Future work can incorporate multilingual, domain-specific datasets and advanced NLP or hybrid models to improve generalization and prediction performance.

REFERENCES

1. H. Singh, "The Metaverse Revolution: Transforming the Future of Retail," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 11, no. 2, pp. 109–115, Mar. 2025, doi: 10.32628/CSEIT251112391.
2. K. Dixit, "Custom experience analysis in the hospitality Service Sector using AI-based Hotel Review Analytics," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 5, no. 1, pp. 720–729, November, 2025, doi: 10.48175/IJARSCT-29699.
3. R. K. Kanneganti, "Impact of Custom Interface Widgets on User Efficiency in Content Navigation Platforms," *Eastasouth J. Inf. Syst. Comput. Sci.*, vol. 1, no. 01, pp. 179–186, Aug. 2023, doi: 10.58812/esiscs.v1i01.1121.
4. Á. Aldunate, S. Maldonado, C. Vairetti, and G. Armelini, "Understanding customer satisfaction via deep learning and natural language processing," *Expert Syst. Appl.*, vol. 209, p. 118309, Dec. 2022, doi: 10.1016/j.eswa.2022.118309.
5. O. Alsemaree, A. S. Alam, S. S. Gill, and S. Uhlig, "Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions," *Heliyon*, vol. 10, no. 9, p. e27863, May 2024, doi: 10.1016/j.heliyon.2024.e27863.
6. A. H M, B. T, S. Tanisha, S. B, and C. C. Shanuja, "Customer Churn Prediction Using Synthetic Minority Oversampling Technique," in *2023 4th International Conference on Communication, Computing and Industry 6.0 (C2I6)*, IEEE, Dec. 2023, pp. 01–05. doi: 10.1109/C2I659362.2023.10430989.
7. K. M. Ahmad et al., "Aspect-specific Parsimonious Segmentation via Attention-based Graph Convolutional Network for Aspect-Based Sentiment Analysis," *Knowledge-Based Syst.*, vol. 300, p. 112169, Sep. 2024, doi: 10.1016/j.knosys.2024.112169.
8. V. Chang, K. Hall, Q. Xu, F. Amao, M. Ganatra, and V. Benson, "Prediction of Customer Churn Behavior in the Telecommunication Industry Using Machine Learning Models," *Algorithms*, vol. 17, no. 6, p. 231, May 2024, doi: 10.3390/a17060231.
9. P. Atandoh, F. Zhang, D. Adu-Gyamfi, P. H. Atandoh, and R. E. Nuhoho, "Integrated deep learning paradigm for document-based sentiment analysis," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 35, no. 7, p. 101578, Jul. 2023, doi: 10.1016/j.jksuci.2023.101578.



10. N. Bhatt, "A Hybrid Artificial Intelligence Model for Product Recommendation in Online Retail System," in 2026 2nd International Conference on Artificial Intelligence Systems (AIS), IEEE, May 2026, pp. 33–39. doi: 10.1109/AIS69919.2026.11621290.
11. S. B. Toksöz and G. Işık, "Evaluating the Effectiveness of Different Machine Learning Approaches for Sentiment Classification," *Iğdır Üniversitesi Fen Bilim. Enstitüsü Derg.*, vol. 13, no. 3, pp. 1496–1510, Sep. 2023, doi: 10.21597/jist.1292050.
12. J. G. Blodgett, K. L. Wakefield, and J. H. Barnes, "The effects of customer service on consumer complaining behavior," *J. Serv. Mark.*, vol. 9, no. 4, pp. 31–42, Oct. 1995, doi: 10.1108/08876049510094487.
13. H. H. Chang, Y.-H. Wang, and W.-Y. Yang, "The impact of e-service quality, customer satisfaction and loyalty on e-marketing: Moderating effect of perceived value," *Total Qual. Manag. Bus. Excell.*, vol. 20, no. 4, pp. 423–443, Apr. 2009, doi: 10.1080/14783360902781923.
14. S. Devarakonda, R. Lingam, and V. Challa, "Confidence-Gated RAG for Adaptive Retrieval in Sequential Agents," in *ICLR 2026 Workshop on Logical Reasoning of Large Language Models*, ICLR, 2026, pp. 01–10, Apr.
15. J. Kim and C. Lim, "Customer complaints monitoring with customer review data analytics: An integrated method of sentiment and statistical process control analyses," *Adv. Eng. Informatics*, vol. 49, p. 101304, Aug. 2021, doi: 10.1016/j.aei.2021.101304.
16. E. Kauffmann, J. Peral, D. Gil, A. Ferrández, R. Sellers, and H. Mora, "A framework for big data analytics in commercial social networks: A case study on sentiment analysis and fake review detection for marketing decision-making," *Ind. Mark. Manag.*, vol. 90, pp. 523–537, Oct. 2020, doi: 10.1016/j.indmarman.2019.08.003.
17. U. Gani Joy, K. E. Hoque, M. Nazim Uddin, L. Chowdhury, and S.-B. Park, "A Big Data-Driven Hybrid Model for Enhancing Streaming Service Customer Retention Through Churn Prediction Integrated With Explainable AI," *IEEE Access*, vol. 12, pp. 69130–69150, 2024, doi: 10.1109/ACCESS.2024.3401247.
18. S.-W. Lee, G. Jiang, H.-Y. Kong, and C. Liu, "A difference of multimedia consumers' rating and review through sentiment analysis," *Multimed. Tools Appl.*, vol. 80, no. 26–27, pp. 34625–34642, Nov. 2021, doi: 10.1007/s11042-020-08820-x.
19. B. Kantheti and C. A. P. De Araujo, "Decoupling Privacy Noise from Optimization in Transformer Forecasting," *Mach. Learn. Knowl. Extr.*, vol. 8, no. 6, p. 156, Jun. 2026, doi: 10.3390/make8060156.
20. Y. Piris and A.-C. Gay, "Customer satisfaction and natural language processing," *J. Bus. Res.*, vol. 124, pp. 264–271, Jan. 2021, doi: 10.1016/j.jbusres.2020.11.065.
21. H. Yakubu and C. K. Kwong, "Forecasting the importance of product attributes using online customer reviews and Google Trends," *Technol. Forecast. Soc. Change*, vol. 171, p. 120983, Oct. 2021, doi: 10.1016/j.techfore.2021.120983.
22. K. N. Lemon and P. C. Verhoef, "Understanding Customer Experience Throughout the Customer Journey," *J. Mark.*, vol. 80, no. 6, pp. 69–96, Nov. 2016, doi: 10.1509/jm.15.0420.
23. S. Al-Otaibi et al., "Customer Satisfaction Measurement using Sentiment Analysis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 2, 2018, doi: 10.14569/IJACSA.2018.090216.
24. B. Srinivas and M. Suresh, "Assessing Advertising Importance in Social Media-Driven Buying and Selling Decisions through Customer Satisfaction Prediction," in *2026 IEEE Guwahati Subsection Conference (GCON)*, IEEE, Jun. 2026, pp. 1–8. doi: 10.1109/GCON69192.2026.11648297.
25. A. Parupalli, "Predicting Customer Satisfaction Through Sentiment Analysis in CRM Using Machine Learning," in *2025 5th International Conference on Artificial Intelligence and Signal Processing (AISP)*, IEEE, Nov. 2025, pp. 1–5. doi: 10.1109/AISP68263.2025.11396157.
26. S. A. Rajagukguk, "Machine Learning-Based Customer Lifetime Value Prediction Framework: An Integrated Approach Using Behavioral Intention Analytics in Digital Service Ecosystems," in *2025 International*



- Conference on Information Technology and Computing (ICITCOM), IEEE, Aug. 2025, pp. 240–245. doi: 10.1109/ICITCOM66635.2025.11265047.
27. M. Haris, Al-Khowarizmi, and O. Nugroho, “Integration of IndoLEM and IndoBERT Models in Real-time Question Prediction in Customer Service Support,” in 2024 12th International Conference on Cyber and IT Service Management (CITSM), IEEE, Oct. 2024, pp. 1–6. doi: 10.1109/CITSM64103.2024.10775632.
 28. P. Wangkiat and C. Polprasert, “Machine Learning Approach to Predict E-commerce Customer Satisfaction Score,” in 2023 8th International Conference on Business and Industrial Research (ICBIR), IEEE, May 2023, pp. 1176–1181. doi: 10.1109/ICBIR57571.2023.10147542.
 29. K. M. B. A. Joolfoo, R. A. Jugurnauth, and M. B. A. Joolfoo, “Application of Machine Learning in Predicting Customer Satisfaction of Telecom Service Providers,” in 2022 4th International Conference on Emerging Trends in Electrical, Electronic and Communications Engineering (ELECOM), IEEE, Nov. 2022, pp. 1–10. doi: 10.1109/ELECOM54934.2022.9965212.
 30. S. Bhoir, D. Dubey, and M. Dongare, “An Overview of Optimization Techniques Utilized in Sheet Metal Blanking Processes,” *Int. J. Adv. Res. Sci. Commun. Technol. Int. Open-Access, Double-Blind, Peer-Reviewed, Ref. Multidisciplinary Online J.*, vol. 3, no. 1, pp. 83–87, 2023, doi: 10.48175/568.
 31. D. D. Nur Cahyo et al., “Sentiment Analysis for IMDb Movie Review Using Support Vector Machine (SVM) Method,” *Inf. J. Ilm. Bid. Teknol. Inf. dan Komun.*, vol. 8, no. 2, pp. 90–95, 2023, doi: 10.25139/inform.v8i2.5700.
 32. R. D. Putra and A. Sunyoto, “The Application of BERT in Sentiment Analysis of IMDB Movie Reviews,” *J. Soc. Res.*, vol. 5, no. 2, pp. 614–622, 2026, doi: 10.55324/josr.v5i2.3011.
 33. S. K. Singh and N. Singla, “Sentiment Analysis on IMDB Review Dataset,” *J. Comput. Mech. Manag.*, vol. 2, no. 6, pp. 18–29, December, 2023, doi: 10.57159/gadl.jcmm.2.6.230108.18.

