IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, November 2025
Impact Factor: 7.67

Custom Experience Analysis in the Hospitality Service Sector using AI-Based Hotel Review Analytics

Kshitij Dixit
Independent Researcher
kshitijdixit102@gmail.com

Abstract: *The hotel industry has experienced rapid online portal development which has produced enormous volumes of customer review data that businesses can use to assess public perceptions and make their operational choices. Hoteliers can achieve better service delivery and higher customer satisfaction by using online hotel reviews to identify customer needs through advanced review analysis techniques. The proposed study presents an AI-driven hotel review analytics model to analyze customer experience in the hospitality industry. The proposed approach will execute text preprocessing, followed by sentiment labeling, TF-IDF feature extraction, and supervised ML-based sentiment classification. A TripAdvisor hotel reviews dataset consisting of 20,491 records is used to test different models for their performance evaluation. The XGBoost model exhibits robust results that surpass the performance of BERT, SVM, and LSTM models, while the MLP model obtains the highest performance with an F1 score of 95.85 and an accuracy of 93.00%. The comparative outcomes prove the effectiveness and consistency of the offered method regarding sentiment analysis. This study is valuable because it offers an effective AI-driven approach for extracting meaningful insights from customer feedback at scale. This solution helps organizations make data-driven decisions and improve personalized services and customer experience management in the hospitality sector.*

Keywords*:* Hotel Review Analytics, Sentiment Analysis, Customer Experience Analysis, TF-IDF, Machine Learning, Hospitality Service Sector

I. INTRODUCTION

People today prefer to socialize with others through online platforms. Users submit their hotel experience evaluations together with their personal opinions to travel websites. The massive data collection provides major benefits to tourism associations and organizations which aim to improve customer satisfaction while their businesses strive for increased profitability. The textual customer reviews might be seen as subjective or objective. Statements that are subjective include things like personal beliefs, perspectives, and opinions, whereas statements that are objective include things like facts, evidence, and measured observations. The digital technologies and online resources that have developed at a fast pace have transformed the ways people plan their travel and book accommodations [1]. Online travel websites serve as the primary platform through which customers search for hotels, compare hotel services and prices, and complete their reservations for both business trips and leisure travel [2]. Once customers have stayed at a hotel, they often post reviews online which serve as a source of information about the hotel's location service facilities cleanliness and overall customer satisfaction [3]. These are user-generated reviews serving as a valuable and ever-expanded body of customer feedback [4]. For hotel operators, proper analysis of large volumes of textual data is critical for understanding customer expectations, identifying weaknesses in service, and assessing competitiveness in an ever-saturated hospitality industry. Customer reviews will usually have some subjective and objective expressions [5]. Personal emotions, opinions, and attitudes are communicated by subjective content, and factual descriptions and quantifiable observations are

ISSN 2581-9429 IJARSCT

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, November 2025
Impact Factor: 7.67

communicated through objective content [6]. Sentiment analysis methods are common for deriving customers' emotional inclinations and satisfaction from review text, using their subjective content, followed by sentiment polarity [7]. The conventional sentiment analysis techniques, sentiment dictionary-based and machine learning techniques, are characterized by a number of limitations, including high maintenance costs, relying on labelled data, and a lack of understanding of the context [8]. Sentiment analysis using deep learning has shown improved results in text semantics and contextual information [9] [10] with recent developments in artificial intelligence [11]. Consequently, hotels may get operational insights, enhance service quality, and boost customer satisfaction levels via customer experience analysis in the hospitality service business.

The fast development of hotel booking and reviewing websites has led to a huge number of customer reviews that cannot be analyzed in a traditional way. The reviews deliver essential information about customer satisfaction, service quality, and customer expectations, which serve as vital components for tracking progress in the hospitality sector. The unstructured nature of text review creates major obstacles that prevent researchers from performing standard analysis methods. The study demonstrates how AI-powered sentiment analysis algorithms need to process large volumes of review data in order to extract valuable insights. The research uses machine learning hotel review analysis to assist hoteliers in understanding customer perceptions, which helps them enhance service quality and make data-driven decisions for better customer experiences. This study has made several main contributions, which are as follows:

- Based on real-world TripAdvisor hotel reviews data, which has 20,491 customer reviews, analyzes customer sentiment and experience.
- Applies full text preprocessing and TF IDF feature extraction, which utilizes review data in an unstructured form well.
- Trains binary sentiment models on hotel reviews under the machine learning models.
- The confusion matrix, F1 score, recall, accuracy, and precision are among the most popular metrics used to assess the efficacy of a model.
- Determines the Multilayer Perceptron model as the most effective method based on the mass of experimental research.

A. Justification and Novelty

The growing number of online hotel reviews is both an opportunity and a challenge for the hospitality industry, as valuable customer information may be hidden in unstructured text data. The existing methods of traditional analysis do not have the capacity to process and analyze such massive data efficiently, explaining why AI-based sentiment analysis strategies are necessary. This paper proposes combining effective feature representation with TF-IDF and SMOTE to enhance sentiment classification. Moreover, the paper presents a systematic experimental analysis of various ML and DL models, where the Multilayer Perceptron is found to be the most effective model to analyze customer experience. The proposed framework functions as an effective and expandable system that enables data-driven decision processes and tailored service improvement in the hospitality industry.

B. Structure of the Paper

This is the remaining structure of the paper. After a brief literature review in Section II, the proposed methodology is laid out in Section III. Experimental data and analysis are reviewed in Section IV. Finally, Section V presents a summary of the key results and discusses potential future directions.

II. LITERATURE REVIEW

Recent research highlights the application of sentiment analysis in order to know customer opinion and feedback on the internet. The literature highlights the importance of text preprocessing and classification methods for distinguishing between positive and negative sentiments, as well as between unreliable reviews. These works, in general, indicate that sentiment analysis assists in extracting better insights and making informed decisions based on large amounts of textual data content.

Budiharja, Juarsa and Karmagatri, (2025) aimed to analyze consumer reviews on Online Travel Agent (OTA) platforms regarding a five-star hotel, namely Sheraton Bandung Hotel & Towers, using the Naïve Bayes algorithm and sentiment analysis. A total of 1,367 reviews were collected from Traveloka, TripAdvisor, and Agoda. The study results show that the majority of reviews, namely 79.1 %, are positive, while the rest are negative reviews [12]. Zabi Ur Rahaman and Giri, (2025) approach identifies comments of type NR and PR from customer sentiments on services that are offered by hotels. Moreover, this method enables access to comments and does analysis of sentences. The methodologies apply SA to divide big sentences into small sentences and then classify them into NR and PR, and finally analyze the emotion of the big sentence into NR or PR. The implementation is done using the MNB and BNB algorithms. The MNB Algorithm produces results with 81.5% and is found to be more accurate [13].

Jadhav et al., (2024) used the dataset, which included 9479 reviews, both fake and legitimate. Text preprocessing and feature engineering are used, and the text is then translated into numeric vector using the TF-IDF vectorization and Count Vectorization approach. Acc, prec, rec, and F1score are some of the common metrics used to train and assess the methods on the dataset. The outcomes reveal that all three algorithms perform effectively at distinguishing between genuine and fraudulent reviews. However, the SVM model has a maximum accuracy of 82% and is thus one of the most suitable algorithms to identify fraudulent reviews in the hotel industry [14]. Abhyudhay et al. (2024) seek to provide a thorough grasp of sentiment analysis as it pertains to hotel reviews, exploring preprocessing and using cutting-edge sentiment analysis methods. The SVM-based sentiment analysis algorithm performed well in classifying hotel evaluations as either positive or negative. The evaluation metrics indicate an accuracy of 82%. The model's ability to differentiate between positive and negative feelings is further shown by visualisations of the ROC curve and confusion matrix [15].

Pramudya and Alamsyah, (2023) utilize three randomly selected reviews as inputs for both the classification model and the recommendation system. The findings demonstrate that BERT outperformed RoBERTa in review classification, achieving an accuracy score of 0.8963 and a macro F1 score of 0.83. On the other hand, when constructing a review-based recommendation, RoBERTa proved superior to BERT, with the highest cosine similarity score of 0.99917 [16]. Kumar and Uddin Haider (2023) analysed the political Twitter dataset for sentiment to determine whether the Indian political party is receiving more positive or negative feedback. Researchers often make use of TextBlob, a sentiment analyzer that relies on a lexicon, to forecast sentiment. This study employs ensemble learning, which combines fundamental ML and DL techniques such as SVM and LSTM with boosting methods, to predict sentiment, as the lexicon-based methodology fails to provide sufficient accuracy. The results of the accuracy and ROC-AUC measurements suggest that ensemble learning, which integrates SVM and LSTM, outperforms other ML techniques. The ROC-AUC score is 93%, and the accuracy is 87% [17].

III. METHODOLOGY

The proposed methodology uses AI techniques to analyze hotel reviews to assess customer experience in the hospitality service industry. The TripAdvisor reviews are firstly processed with text cleaning, normalization, tokenization, and lemmatization, and then TF IDF-based feature extraction with unigrams and bigrams is performed. The data is categorized under binary sentiment labels, and divided into training and testing subsets and balanced with SMOTE. The most efficient model for sentiment classification is then determined by training and evaluating the supervised learning models in accordance with the standard performance measures, as illustrated in Figure 1.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, November 2025
Impact Factor: 7.67

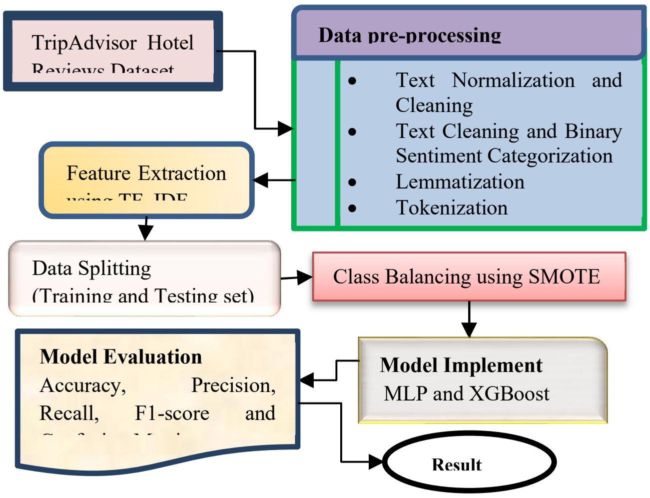

Fig. 1. Flowchart of the Proposed AI-Based Hotel Review Analytics for Custom Experience Analysis

A. Data Collection and Analysis

The TripAdvisor Hotel Reviews data, which is available in Kaggle[1] is a dataset of 20,491 customer records with a shape of 20,491 × 2. It consists of a text attribute, Review, containing customers' opinions and experiences regarding hotel services, and a numerical attribute, Rating, presented as an integer, representing overall customer satisfaction. This dataset is appropriate for sentiment analysis because the review texts are unstructured, while the rating data is structured.

Fig. 2. Word Cloud Visualization of Positive and Negative Hotel Reviews

The word clouds of the most commonly used words in the positive and the negative reviews of the hotels on TripAdvisor are presented in the Figure. 2. Words used in positive comments like room, hotel and stay portray satisfaction, whereas words used in negative comments like service, staff, and bad depict the most important areas of customer dissatisfaction in the hospitality business.

[1] https://www.kaggle.com/datasets/andrewmvd/trip-advisor-hotel-reviews

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, November 2025
Impact Factor: 7.67

B. Data Preprocessing

The pre-processing process in this study focused on cleaning and normalizing hotel review text to improve sentiment analysis performance. It included cleaning and normalizing hotel reviews through lowercasing, removal of HTML tags, URLs, stop-words, and non-alphabetic characters, followed by tokenization and lemmatization using NLTK to prepare the text for sentiment analysis.

1) Text Normalization and Noise Removal

Basic text preprocessing methods include lowercasing, removal of HTML tags, removal of URLs and removal of stop words, which are employed to normalize text data and eliminate noise. Lowercasing normalizes the word representations, whereas the elimination of HTML tags and URLs results in only meaningful and visible text. Stop words when the, is, and and are often words that are common and have a small semantic value to the words and can be eliminated to reduce the dimensionality and enhance the computational efficiency. All of these preprocessing procedures contribute to textual features and increase the effectiveness of the downstream text analysis and machine learning models..

2) Text Cleaning and Binary Sentiment Categorization

Non-alphabetic characters are stripped to reduce noise and enhance text clarity. Hotel review ratings are transformed into binary sentiment scores: ratings of three and above are positive, and ratings of three and below are negative, enabling easy binary classification.

3) Lemmatization

The word lemmatize is its root, which means to strip a word of its ending to a form that has no other meaning. Lemmatization takes a predetermined dictionary to take into account the context of words when reduction is done. In this study, lemmatization is used to normalize the text in hotel reviews, decrease the redundancy of vocabulary and enhance the consistency of the features, which increases the accuracy of the sentiment classification models.

4) Tokenization

In order to analyze text, it should be divided into small units like words or sentences. The concept of tokenization breaks down text into meaningful messages that are easier to execute other processing procedures feature extraction. The present study employs a tokenization process to transform textual hotel review content into unit-sized words that can be systematically cleansed, normalized, and processed into numerical features. This is necessary to have an exact representation of customer views and be able to effectively use TF-IDF feature extraction and supervised learning models to classify sentiment.

C. Feature Extraction using TF–IDF

The dataset's textual reviews are transformed into numerical characteristics usable by ML techniques via the use of TF-IDF. The approach includes the calculation of Term Frequency (TF) of each word that represents the frequency with which a term is present in a particular review as compared to the number of terms in the review. At the same time, Inverse Document Frequency (IDF) attaches less weight to those words of the dataset that are frequent in the dataset, so that common but less informative terms will have little impact on the numerical expression. TF-IDF is an algorithm that uses a combination of the number of times the word appears in the text and the number of documents where it appears to emphasize terms that are informative and particular to the review, and the frequency with which irrelevant words appear to lower. The cleaned reviews are converted into a TF-IDF matrix of 20491x 5000 using unigrams and bigrams with a maximum of 5,000 features which is fed in supervised sentiment classification models.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 1, November 2025 Impact Factor: 7.67

D. Data Splitting

Training and assessment data are identified during this phase. The training data is divided from the overall dataset, with 80% being used for training and 20% for testing. "Positive" and "negative" are the additional classifications assigned to the data.

E. Class Balancing using SMOTE

SMOTE is a widely used technique that creates synthetic samples of underrepresented classes in order to rectify unbalanced data. This method will be used by determining the data points belonging to the minority group and synthetically creating new samples along the line segments between them. In comparison with the traditional oversampling techniques that recreate the existing samples, SMOTE creates new unique samples, and it decreases the probability of the model overfitting. Figure. 3 shows the class distribution before and after SMOTE, revealing a strong imbalance with Class 1 prevailing over Class 0. Following SMOTE, the minority group is overrepresented, achieving a balanced distribution that facilitates more effective, unbiased model training.

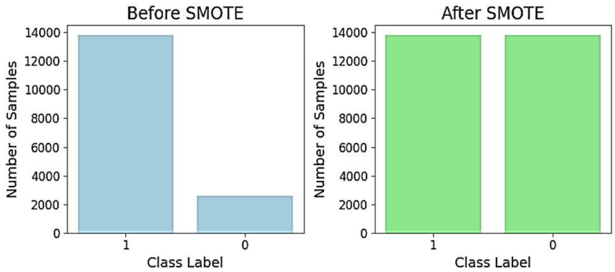

Fig. 3.Class Distribution Before and After SMOTE

F. Proposed Model

The choice of ML models in this paper has been informed by their strengths and their ability to handle various issues in the dataset and the prediction task. Each model has its own strengths, which address different problems in the dataset and the prediction goal. The sentiment classification procedure used multiple algorithms, each with different strengths and properties. Naive Bayes is also selected based on its simplicity and efficiency, which are best applied in the analysis of texts, as it is a probabilistic algorithm.

Multilayer Perceptron Network (MLP)

The MLP was among the first DNNs to be employed for the purpose of predicting short-term traffic flow. MLPs are the most basic variety of DNNs. An input layer, hidden layers, an output layer, and nonlinear activation functions comprise a conventional multilayer feedforward ANN [18]. This approach achieves great computational performance and minimal computational cost by using just fully linked layer matrix operations. The hidden layers of an MLP handle the signal input, while the input layer normalises features and accepts data. According to Equation (1), the nonlinear activation function φ translates the summation function (xw + b) to the output value y, and the output layer uses this information to make decisions or predictions. The input vector (x), weighted vector (w), bias (b), and output value (y) are represented by the components in Equation (1), in that order [19]. The MLP model's structure is seen in Figure 4.

$$y = \varphi(xw = b) \qquad (1)$$

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 1, November 2025

Impact Factor: 7.67

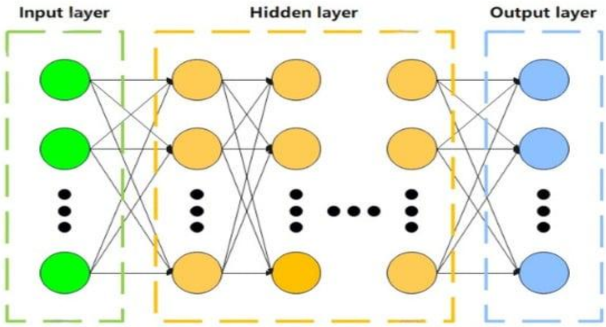

Fig. 4. Structure of the MLP

XGBoost

XGBoost implements the gradient augmentation approach well. The gradient-based alternative may be rigorously developed for precision and optimization, even if no mathematical breakthroughs occur in this specific instance [20]. A linear representation is used, and the newborn tree may be a strategy that utilizes multiple artificial intelligence algorithms to assess whether a susceptible rookie would result in a trustworthy rookie to increase the model's accuracy.

Performance Metrics

The results of the experiment must be compared using a range of indicators. The likelihood that the classifier will provide accurate predictions is known as the accuracy rate. The percentage of a document analysis that is accurate across all datasets for text classification characteristics is known as the recall rate. Several metrics, including F1score, accuracy, recall, and prec, are used to evaluate performance. Accuracy is defined as the fraction of samples for which the classifier produced the correct classification according to Equation (2). Recall is defined as the total number of samples that the classifier correctly recognized as positives, as shown in Equation (3). The proportion of classifier-predicted positive samples that are true positives is called precision, as shown in Equation (4). The F1-score yields a balanced average by combining accuracy and recall, as defined in Equation (5). The aforementioned equations can be used to compute various ML metrics, including TP, FP, TN, and FN.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

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

$$F1-Score = 2 * \frac{(Precision*Recall)}{Precision+Recall} \quad (5)$$

IV. RESULT ANALYSIS AND DISCUSSION

To perform custom experience analysis in the hospitality service industry, all experiments are conducted on a laptop running Windows 11 with 512 GB of storage. Table I demonstrates that the Multilayer Perceptron model obtains the greatest performance, with an accuracy of 93.00 % and an F1score of 95.85, making it dependable for sentiment classification. The efficacy of XGBoost in hotel review sentiment analysis is further supported by its 92.48% acc and 95.60 F1score.

Table 1: Performance Evaluation of MLP and XGBoost Models for Hotel Review

Matrix	MLP	XGBoost
Accuracy	93.00	92.48
Precision	96.36	94.89
Recall	95.34	96.32
F1-Score	95.85	95.60

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, November 2025
Impact Factor: 7.67

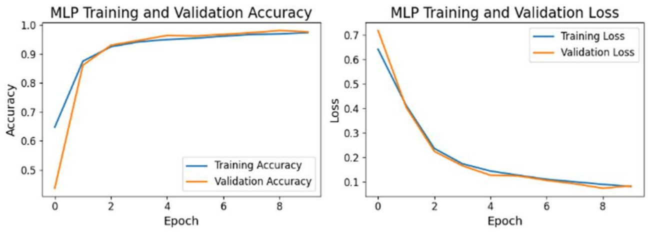

Fig. 5. MLP Training and Validation Performance

Figure. 5 illustrates the training and validation accuracy and loss on 10 epochs of the MLP model and indicates a high rate of improvement in accuracy and a steady decline in loss, with nearly parallel curves on either side of the line showing the model has stable learning and good generalization.

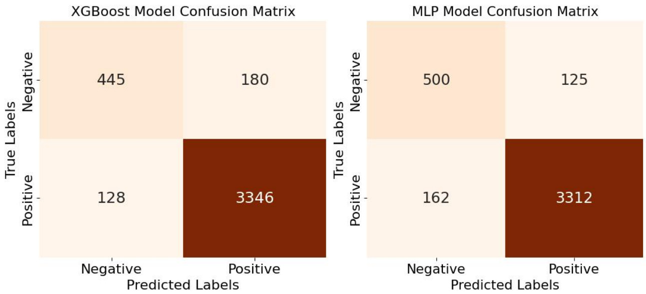

Fig. 6. Confusion Matrices of XGBoost and MLP Models.

Figure. 6 shows the confusion matrices for the XGBoost and MLP models. XGBoost achieves the best performance on correctly classified positive reviews, making 3,346 positive predictions. This discovery shows that the model is quite good at detecting positive customer sentiment in analytics of hotel reviews.

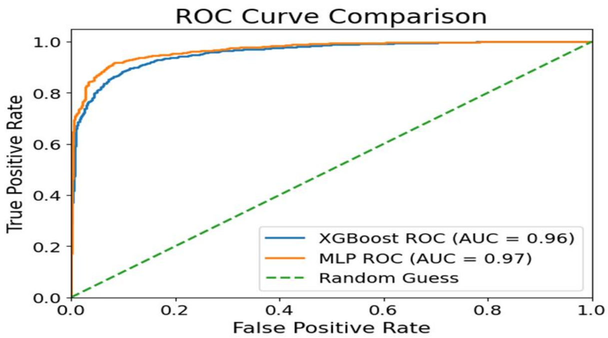

Fig. 7. ROC Curve Comparison of XGBoost and MLP Models.

Figure. 7 is a comparison of the ROCcurves of the XGBoost and MLP models in the hotel review sentiment classification. The MLP model has a higher area under the curve (AUC = 0.97) than XGBoost (AUC = 0.96), indicating that the MLP is slightly more discriminative, but both models are significantly better than random guessing.

A. COMPARISON AND DISCUSSION

Table II presents a comparative performance analysis of several ML and DL models for analyzing customer experiences in the hotel service industry. The evaluation uses four metrics which are acc, prec, rec and F1score. The multilayer perceptron model performs best overall, with an accuracy of 93.00 and an F1score of 95.85; XGBoost comes next, with an F1score of 95.60. The BERT model shows competitive performance across all measures, whilst the support vector machine exhibits moderate performance. The LSTM achieves the lowest scores, indicating it is less effective for sentiment analysis in this context.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, November 2025
Impact Factor: 7.67

Table 2: Performance Comparison of Models for Custom Experience Analysis in the Hospitality Service Sector

Models	Accuracy	Precision	Recall	F1-Score
MLP	93.00	96.36	95.34	95.85
XGBoost	92.48	94.89	96.32	95.60
BERT [21]	89.7	89.3	89.1	89.3
SVM [22]	84.44	80.19	80.19	81.23
LSTM [23]	61	63	61	61

The proposed hotel review analytics system uses supervised ML algorithms to determine customer sentiment in the hospitality industry. The system divides hotel reviews into different sentiment categories. The Multilayer Perceptron (MLP) model emerged as the leading model because it achieved both the highest accuracy and F1 score performance results. The experimental outcomes confirm that the framework successfully uses the MLP model to evaluate consumer input and achieve data-driven decisions that improve service quality and customer satisfaction. Nevertheless, the study is limited to binary sentiment analysis on a single dataset and does not capture the contextual or fine-grained emotional nuances in customer reviews, leaving room for improvement.

V. CONCLUSION AND FUTURE ENHANCEMENTS

Sentiment analysis of hotel reviews is performed by labelling them as positive or negative. In this study, an AI-driven hotel review analytics framework to analyze customer experience data in the hospitality service industry is introduced, involving systematic preprocessing of texts, TF IDF feature extraction, and supervised ML models. The experimental findings on the TripAdvisor dataset revealed that the Multilayer Perceptron model could be trained to achieve a minimum accuracy of 93.00% and a maximum F1 score of 95.85, whereas XGBoost also showed strong classification performance. Although the results of this study are positive, the research is limited to binary sentiment classification and a single dataset, and TF-IDF representation is not able to capture semantic and contextual information across all aspects. Future research can be involved in the multi-class/aspect-based sentiment analysis, the usage of more sophisticated DL models, including CNN LSTM hybrids, RoBERTa, and DistilBERT, and the testing on several hospitality datasets to enhance generalization, scalability, and real-world applicability.

REFERENCES

[1] S. P. Kalava, "The Role of AI in Reinventing Hospitality Safety Measures After COVID-19," *J. Artif. Intell. Cloud Comput.*, vol. 3, no. 1, pp. 1–3, Feb. 2024, doi: 10.47363/JAICC/2024(3)E153.

[2] C. Patel, "Customer Experience Optimization Using Machine Learning : A Systematic Review," *ESP J. Eng. Technol. Adv.*, vol. 3, no. 4, pp. 176–187, 2023, doi: 10.56472/25832646/JETA-V3I8P120.

[3] P. De Pelsmacker, S. van Tilburg, and C. Holthof, "Digital marketing strategies, online reviews and hotel performance," *Int. J. Hosp. Manag.*, vol. 72, pp. 47–55, Jun. 2018, doi: 10.1016/j.ijhm.2018.01.003.

[4] D. Gavilan, M. Avello, and G. Martinez-Navarro, "The influence of online ratings and reviews on hotel booking consideration," *Tour. Manag.*, vol. 66, pp. 53–61, Jun. 2018, doi: 10.1016/j.tourman.2017.10.018.

[5] R. Dattangire, R. Vaidya, D. Biradar, and A. Joon, "Exploring the Tangible Impact of Artificial Intelligence and Machine Learning: Bridging the Gap between Hype and Reality," in *2024 1st International Conference on Advanced Computing and Emerging Technologies (ACET)*, IEEE, 2024, pp. 1–6. doi: 10.1109/ACET61898.2024.10730334.

[6] S. Anis, S. Saad, and M. Aref, "Sentiment Analysis of Hotel Reviews Using Machine Learning Techniques," 2021, pp. 227–234. doi: 10.1007/978-3-030-58669-0_21.

[7] N. Prajapati, "The Role of Machine Learning in Big Data Analytics: Tools, Techniques, and Applications," *ESP J. Eng. Technol. Adv.*, vol. 5, no. 2, 2025, doi: 10.56472/25832646/JETA-V5I2P103.

[8] D. Patel, "AI-Enhanced Natural Language Processing for Improving Web Page Classification Accuracy," vol. 4, no. 1, pp. 133–140, 2024, doi: 10.56472/25832646/JETA-V4I1P119.

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 1, November 2025

Impact Factor: 7.67

[9] S. Thangavel, S. Srinivasan, S. B. V. Naga, and K. Narukulla, "Distributed Machine Learning for Big Data Analytics: Challenges, Architectures, and Optimizations," *Int. J. Artif. Intell. Data Sci. Mach. Learn.*, vol. 4, no. 3, pp. 18–30, Oct. 2023, doi: 10.63282/3050-9262.IJAIDSML-V4I3P103.

[10] D. Chi, T. Huang, Z. Jia, and S. Zhang, "Research on sentiment analysis of hotel review text based on BERT-TCN-BiLSTM-attention model," *Array*, vol. 25, p. 100378, Mar. 2025, doi: 10.1016/j.array.2025.100378.

[11] H. P. Kapadia and K. B. Thakkar, "Generative AI for Real-Time Customer Support Content Creation," *J. Emerg. Technol. Innov. Res.*, vol. 10, no. 12, pp. 36–43, 2023.

[12] J. R. Budiharja, M. A. Juarsa, and M. Karmagatri, "NaïVe Bayes-Based Data Analysis on Online Travel Agency's Guest Review of Sheraton Bandung Hotel & Towers," in *Proceedings - 3rd International Conference on Artificial Intelligence and Machine Learning Applications: Healthcare and Internet of Things, AIMLA 2025*, 2025. doi: 10.1109/AIMLA63829.2025.11041674.

[13] K. Zabi Ur Rahaman and M. Giri, "Enhancing Hotel Services in the Industry 6.0 Era: Optimizing Customer Satisfaction through Sentiment Analysis," in *2025 6th International Conference for Emerging Technology, INCET 2025*, 2025. doi: 10.1109/INCET64471.2025.11140059.

[14] A. Jadhav, M. Walvatkar, J. Bagde, and P. Kalyan, "Sentiment Analysis and Authenticity Assessment of Hotel Reviews," in *Proceedings of the 2nd International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics, ICIITCEE 2024*, 2024. doi: 10.1109/IITCEE59897.2024.10467514.

[15] S. R. Abhyudhay, G. M. Aditya, A. K. Upadya, A. Naik, and P. Ushashree, "Customer Feedback and Sentiment Analysis for Hotel Services," in *International Conference on Distributed Computing and Optimization Techniques, ICDCOT 2024*, 2024. doi: 10.1109/ICDCOT61034.2024.10516070.

[16] Y. G. Pramudya and A. Alamsyah, "Hotel Reviews Classification and Review-based Recommendation Model Construction using BERT and RoBERTa," in *2023 6th International Conference on Information and Communications Technology, ICOIACT 2023*, 2023. doi: 10.1109/ICOIACT59844.2023.10455890.

[17] S. Kumar and M. T. Uddin Haider, "Sentiment Analysis of Political Party Twitter Data using Ensemble Learning Classifier," in *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2023, pp. 1–5. doi: 10.1109/ICCCNT56998.2023.10308149.

[18] R. Liu and S. Y. Shin, "A Review of Traffic Flow Prediction Methods in Intelligent Transportation System Construction," 2025. doi: 10.3390/app15073866.

[19] K. C. Ke and M. S. Huang, "Quality prediction for injection molding by using a multilayer perceptron neural network," *Polymers (Basel).*, 2020, doi: 10.3390/polym12081812.

[20] S. Sakri *et al.*, "An Improved Concatenation of Deep Learning Models for Predicting and Interpreting Ischemic Stroke," *IEEE Access*, vol. 12, pp. 53189–53204, 2024, doi: 10.1109/ACCESS.2024.3386220.

[21] Q. Gao and J. A. Esquivel, "Tripadvisor Hotel Review Text Mining Based on RoBERTa and LDA Models," in *Proceedings of the 2024 7th International Conference on Computer Information Science and Artificial Intelligence*, 2024, pp. 337–341. doi: 10.1145/3703187.3703244.

[22] D. S. Garg and M. B. Ashar, "Sentiment Analysis of Hotel Reviews Using Machine Learning: A Study on Model Performance Under Noisy Conditions," *Int. J. Sci. Res. Eng. Manag.*, vol. 09, no. 09, pp. 1–9, Sep. 2025, doi: 10.55041/IJSREM52538.

[23] Y. Zhang *et al.*, "A Predictive Model Based on TripAdvisor Textual Reviews: Early DestinationRecommendations for Travel Planning," *Sage Open*, vol. 14, no. 2, pp. 1–15, Apr. 2024, doi: 10.1177/21582440241246434.