International Current Journal of Engineering and Science (ICJES)
Volume 5, Issue 7, pp 8-16, July-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.507002

A Hybrid Machine Learning and BERT-Based Approach for Sentiment and Emotion Analysis of Hotel Reviews

Ashish Shrikant

Department of Artificial Intelligence & Data Science, Nutan Maharashtra Institute of Engineering and Technology (NMIET), Maharashtra, India

Abstract - **The rapid growth of online hotel booking platforms has generated a vast amount of customer reviews that contain valuable insights into guest satisfaction, service quality, and overall hospitality experiences. However, manually analyzing these unstructured textual reviews is time-consuming, inconsistent, and incapable of processing large-scale data efficiently. This research proposes a hybrid machine learning and BERT-based approach for sentiment and emotion analysis of hotel reviews to automatically identify customer opinions and emotional expressions with high accuracy. The proposed framework integrates Natural Language Processing (NLP) techniques with traditional machine learning algorithms and transformer-based deep learning models to perform comprehensive review analysis. Initially, the collected hotel reviews undergo preprocessing steps including text normalization, tokenization, stop-word removal, punctuation cleaning, and lemmatization to improve text quality. Feature extraction is performed using TF-IDF and contextual word embeddings generated by Bidirectional Encoder Representations from Transformers (BERT). The sentiment classification module employs hybrid learning models comprising Support Vector Machine (SVM), Random Forest (RF), and BERT-based classifiers to categorize reviews into positive, negative, and neutral sentiments. Simultaneously, the emotion analysis module identifies emotional states such as happiness, satisfaction, sadness, anger, frustration, and surprise expressed by customers. Furthermore, aspect-based sentiment analysis is incorporated to evaluate customer opinions regarding hotel attributes including staff behavior, cleanliness, room quality, food services, pricing, and amenities. Experimental evaluation demonstrates that the hybrid BERT-based framework significantly improves classification accuracy, contextual understanding, and emotion recognition compared with conventional machine learning methods. The proposed system effectively transforms large volumes of unstructured customer feedback into meaningful business intelligence, enabling hotel managers to monitor service quality, detect customer dissatisfaction at an early stage, enhance guest experiences, and support data-driven decision-making for continuous service improvement.**

Keywords: Natural Language Processing (NLP); Sentiment Analysis; Emotion Analysis; Hotel Reviews; Machine Learning; BERT; Transformer Model; Support Vector Machine (SVM); Random Forest; Aspect-Based Sentiment Analysis; Text Mining; Deep Learning; Opinion Mining; Customer Experience Analytics; Hospitality Analytics.

I. INTRODUCTION

The rapid expansion of the tourism and hospitality industry has significantly increased the importance of understanding customer preferences and experiences. With the widespread adoption of online travel agencies and hotel booking platforms such as Booking.com, TripAdvisor, Expedia, and Google Reviews, millions of customers share their opinions regarding hotel services, facilities, cleanliness, staff behavior, food quality, pricing, and overall satisfaction. These reviews represent a valuable source of information for both prospective customers and hotel management. However, the enormous volume of user-generated content makes manual analysis impractical, time-consuming, and susceptible to subjective interpretation. Consequently, there is an increasing demand for intelligent systems capable of automatically extracting meaningful insights from textual customer feedback.

Sentiment analysis, also known as opinion mining, has emerged as an effective Natural Language Processing (NLP) technique for identifying the polarity of textual data as positive, negative, or neutral. Beyond sentiment classification, emotion analysis provides deeper insights into the psychological state of customers by recognizing emotions such as happiness, satisfaction, sadness, anger, surprise, and frustration. Understanding both sentiment and emotion enables

© 2026 ICJES All Rights Reserved

Volume 5, Issue 7, pp 8-16, July-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.507002

hotel managers to evaluate customer perceptions more comprehensively and identify specific areas requiring service improvement. Such information assists decision-makers in enhancing customer satisfaction, strengthening brand reputation, and increasing business competitiveness.

Recent advances in artificial intelligence, machine learning, and deep learning have significantly improved the accuracy of sentiment and emotion analysis. Traditional machine learning algorithms, including Support Vector Machine (SVM), Naïve Bayes, Decision Trees, and Random Forest, have demonstrated good performance in sentiment classification when combined with feature extraction techniques such as Bag-of-Words and Term Frequency–Inverse Document Frequency (TF-IDF). Nevertheless, these approaches often struggle to capture contextual meaning, semantic relationships, sarcasm, and long-range dependencies within natural language.

Transformer-based language models, particularly Bidirectional Encoder Representations from Transformers (BERT), have revolutionized NLP by learning contextual word representations through bidirectional training. Unlike conventional embedding methods such as Word2Vec and GloVe, BERT considers both preceding and succeeding words while interpreting sentence meaning, resulting in superior contextual understanding and improved classification accuracy. Integrating BERT with traditional machine learning algorithms creates a hybrid framework that combines contextual feature extraction with efficient classification models, thereby enhancing the overall performance of sentiment and emotion analysis systems.

In addition to overall sentiment prediction, aspect-based sentiment analysis (ABSA) has gained considerable attention in hospitality research. Instead of assigning a single sentiment label to an entire review, ABSA identifies customer opinions related to specific hotel attributes such as room quality, cleanliness, staff behavior, location, food services, pricing, and amenities. This fine-grained analysis enables hotel administrators to recognize strengths and weaknesses at the service level and implement targeted quality improvement strategies.

This research proposes a hybrid machine learning and BERT-based framework for sentiment and emotion analysis of hotel reviews. The proposed system integrates text preprocessing, TF-IDF and BERT-based feature extraction, traditional machine learning classifiers, and transformer-based deep learning models to perform accurate sentiment classification, emotion detection, and aspect-based opinion mining. The developed framework converts large-scale unstructured hotel reviews into meaningful analytical information, supporting hotel managers in data-driven decision-making, improving service quality, and enhancing customer experiences.

Customer reviews posted on platforms such as TripAdvisor, Google Reviews, and Booking.com are essential indicators of hotel service quality. These reviews contain detailed feedback regarding staff behavior, room comfort, cleanliness, food quality, amenities, and overall guest experience. However, the reviews are unstructured and written in natural language, making manual interpretation difficult.

Traditional sentiment analysis classifies reviews as positive or negative, but modern hotel management requires deeper insights. Emotions such as happiness, anger, frustration, trust, and disappointment reflect the true feelings of guests more accurately than star ratings alone.

This project uses NLP techniques to automatically extract, preprocess, interpret, and classify guest emotions. Preprocessing techniques include tokenization, stop-word removal, stemming, lemmatization, and normalization. Feature extraction transforms text into numerical representation using approaches like TF-IDF, Bag-of-Words, and contextual embedding’s from BERT.

Machine learning and deep learning models, such as SVM, Random Forest, LSTM, and BERT, are applied to predict emotions with high accuracy. Emotion trends over time help management identify seasonal issues and frequent complaints. Aspect-level sentiment reveals which hotel components cause positive or negative emotions. The system provides dashboards for visualization and automated reporting to help hotel managers make data- driven decisions.

II. RELATED WORK

Several researchers have investigated sentiment analysis and opinion mining to evaluate customer satisfaction in the hospitality industry. Early studies primarily relied on traditional machine learning techniques using manually engineered features for sentiment classification. Pang et al. (2002) pioneered sentiment classification by applying machine learning algorithms to text documents, demonstrating that classifiers such as Support Vector Machine (SVM), Naïve Bayes, and Maximum Entropy could effectively distinguish positive and negative opinions. Their work established the foundation for modern sentiment analysis research.

Liu (2012) introduced comprehensive opinion mining techniques that extract subjective information from customer-generated content using Natural Language Processing. The

© 2026 ICJES All Rights Reserved

International Current Journal of Engineering and Science (ICJES)
Volume 5, Issue 7, pp 8-16, July-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.507002

study emphasized the importance of feature-based sentiment analysis for identifying customer opinions on individual product and service attributes rather than considering the entire review as a single sentiment. This concept later became the basis for aspect-based sentiment analysis in hospitality applications.

Medhat et al. (2014) conducted an extensive survey on sentiment analysis methodologies and categorized existing approaches into machine learning, lexicon-based, and hybrid techniques. Their review concluded that machine learning methods generally achieve higher accuracy than rule-based approaches when sufficient labeled datasets are available. However, the authors also highlighted limitations in handling contextual semantics and domain adaptation.

The emergence of deep learning significantly improved sentiment analysis performance. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks demonstrated improved capability in modeling sequential text data by capturing long-term contextual dependencies. Researchers reported higher classification accuracy using LSTM models for customer review analysis compared with conventional machine learning algorithms.

The introduction of transformer architectures represented a major advancement in Natural Language Processing. Devlin et al. (2019) proposed Bidirectional Encoder Representations from Transformers (BERT), which generates contextual embeddings through bidirectional language modeling. BERT substantially outperformed previous deep learning models across numerous NLP tasks, including sentiment classification, emotion detection, and question answering. Its ability to capture semantic context has made it one of the most widely adopted language models for text analytics.

Recent hospitality-related studies have increasingly incorporated BERT for hotel review analysis. Researchers have demonstrated that transformer-based models provide superior performance in identifying customer sentiments and emotional expressions while effectively recognizing contextual nuances, sarcasm, and implicit opinions. Furthermore, combining BERT embeddings with traditional classifiers such as Support Vector Machine and Random Forest has been shown to improve computational efficiency and classification performance.

Aspect-based sentiment analysis has also become an important research direction in hospitality analytics. Recent studies have focused on identifying customer opinions related to individual hotel services, including room cleanliness, staff behavior, food quality, pricing, comfort, and amenities. Such fine-grained analysis provides actionable insights that assist hotel managers in prioritizing operational improvements and enhancing customer satisfaction.

Despite significant progress, existing studies often focus exclusively on either sentiment classification or emotion recognition and rarely integrate contextual embeddings, hybrid machine learning models, and aspect-based sentiment analysis within a unified framework. Furthermore, many existing systems face challenges in accurately interpreting complex linguistic expressions, mixed sentiments, and domain-specific hospitality terminology. Therefore, the proposed hybrid machine learning and BERT-based framework aims to address these limitations by combining advanced contextual language modeling with efficient classification algorithms to deliver accurate sentiment prediction, emotion detection, and aspect-level opinion analysis for hotel reviews.

Many researchers have worked on sentiment analysis to understand opinions from text data. Earlier studies used machine learning algorithms such as Naïve Bayes, Support Vector Machine (SVM), and Decision Trees to classify reviews as positive or negative. These methods helped in analysing customer feedback from online platforms.

With the development of deep learning, models like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) have been used to improve the accuracy of sentiment classification. These models can understand the relationship between words and capture contextual information from sentences.

Recently, transformer-based models such as BERT have shown better performance in Natural Language Processing tasks. These models can analyse the meaning of words based on their context in a sentence, which improves sentiment detection.

In the hospitality industry, sentiment analysis has been applied to hotel reviews from websites such as TripAdvisor and Google Reviews. These studies help hotel managers understand customer opinions about services, staff behaviour, cleanliness, and facilities.

However, many existing systems focus only on positive and negative sentiment. The proposed system improves this by identifying different emotions in guest reviews using NLP techniques.

III. PROPOSED SYSTEM

The proposed system presents a hybrid machine learning and BERT-based framework for automated sentiment and

© 2026 ICJES All Rights Reserved

International Current Journal of Engineering and Science (ICJES)
Volume 5, Issue 7, pp 8-16, July-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.507002

emotion analysis of hotel customer reviews. The framework is designed to process large volumes of unstructured textual feedback and convert them into meaningful insights that support hotel management in evaluating customer satisfaction and improving service quality. The proposed methodology integrates Natural Language Processing (NLP), feature engineering, transformer-based language models, and machine learning classifiers to achieve high accuracy in sentiment and emotion recognition.

Initially, hotel reviews are collected from online travel platforms and customer feedback portals. The collected dataset undergoes a preprocessing stage that removes irrelevant information and improves text quality. Preprocessing operations include case normalization, punctuation removal, stop-word elimination, tokenization, lemmatization, special character removal, and whitespace normalization. These operations reduce noise in the dataset and improve the effectiveness of subsequent feature extraction.

After preprocessing, textual features are generated using a hybrid representation approach. Traditional numerical features are extracted using Term Frequency–Inverse Document Frequency (TF-IDF), while contextual semantic representations are obtained using the Bidirectional Encoder Representations from Transformers (BERT) model. TF-IDF effectively captures word importance within documents, whereas BERT generates contextual embeddings that preserve semantic relationships and sentence-level meaning. The combination of these representations enables the system to capture both statistical and contextual information from customer reviews.

The extracted features are supplied to multiple classification models. Traditional machine learning algorithms, including Support Vector Machine (SVM) and Random Forest (RF), are employed for sentiment classification, while the fine-tuned BERT model performs contextual sentiment prediction and emotion recognition. The hybrid architecture combines the computational efficiency of machine learning classifiers with the contextual understanding of transformer models, resulting in improved classification accuracy.

The proposed system classifies reviews into positive, negative, and neutral sentiments while simultaneously identifying customer emotions such as happiness, satisfaction, sadness, anger, frustration, and surprise. In addition, an Aspect-Based Sentiment Analysis (ABSA) module identifies opinions associated with specific hotel attributes, including room quality, cleanliness, staff behavior, food services, pricing, location, and amenities. The analytical results are presented through interactive dashboards containing sentiment distributions, emotion statistics, aspect-wise performance, and temporal trend analysis. These visualizations assist hotel administrators in identifying operational strengths and weaknesses, enabling data-driven decision-making for continuous service improvement.

Advantages of the Proposed System

- Hybrid integration of machine learning and transformer-based deep learning models.
- High-accuracy sentiment and emotion classification.
- Context-aware understanding through BERT embeddings.
- Fine-grained aspect-based sentiment analysis.
- Automatic processing of large-scale customer reviews.
- Real-time analytical dashboard for decision support.
- Improved customer satisfaction through early detection of negative feedback.
- Scalable architecture suitable for hospitality analytics and customer experience management.

In the proposed system, hotel reviews are collected from online platforms such as TripAdvisor, Google Reviews, and Booking.com. These reviews contain customer opinions about hotel services, facilities, and overall experience. The system first reads the review dataset and processes the text data using Natural Language Processing (NLP) techniques.

Initially, the system performs text preprocessing which includes removing stop words, punctuation, and unnecessary characters. Tokenization and lemmatization are applied to convert the review text into a clean and meaningful format. After preprocessing, the system extracts important features from the text using techniques such as TF-IDF (Term Frequency–Inverse Document Frequency).

These extracted features are then given as input to machine learning and deep learning models for sentiment classification. Algorithms such as Support Vector Machine (SVM), Random Forest, and deep learning models like LSTM are used to classify the sentiment of reviews into positive, negative, or neutral categories.

In this project, we use a hotel review dataset that contains thousands of customer reviews along with sentiment labels. The dataset is used to train the model so that it can accurately predict the sentiment of new reviews.

To implement this system, the following modules are designed:

International Current Journal of Engineering and Science (ICJES)
Volume 5, Issue 7, pp 8-16, July-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.507002

1) Dataset Loading Module:

This module loads the hotel review dataset and prepares the data for processing.

2) NLP Preprocessing Module:

This module cleans the text data by removing stop words, performing tokenization, and applying lemmatization.

3) Sentiment Classification Module:

This module uses machine learning and deep learning algorithms to predict the sentiment of hotel reviews.

4) Performance Evaluation Module:

This module displays evaluation metrics such as accuracy, precision, recall, and F1-score to measure the performance of the model.

Text Preprocessing

Text preprocessing is the first step in Natural Language Processing used to clean and prepare raw hotel review data. The reviews collected from online platforms usually contain unnecessary information such as punctuation marks, special characters, and stop words.

These elements do not contribute to sentiment detection and may reduce model accuracy.

Fig 1: Text Preprocessing Workflow

In this step, the review text is cleaned by removing stop words like “the”, “is”, “and”, “was”. Tokenization is applied to split the text into smaller units called tokens (words). Stemming and lemmatization are used to convert words into their root form. For example, words like “liked”, “likes”, and “liking” are converted to the root word “like”. This helps the machine learning model understand the meaning of the words more clearly.

After preprocessing, the text becomes structured and ready for feature extraction.

Feature Extraction (TF-IDF)

Feature extraction converts textual data into numerical form so that machine learning models can process it. In this project, the TF-IDF (Term Frequency–Inverse Document Frequency) technique is used.

TF-IDF measures the importance of a word in a review relative to all reviews in the dataset. Words that appear frequently in one review but rarely in other reviews receive higher importance scores. This helps identify words that express strong opinions, such as “excellent”, “dirty”, “friendly”, or “poor service”.

The output of this stage is a feature vector, which represents each review as a set of numerical values.

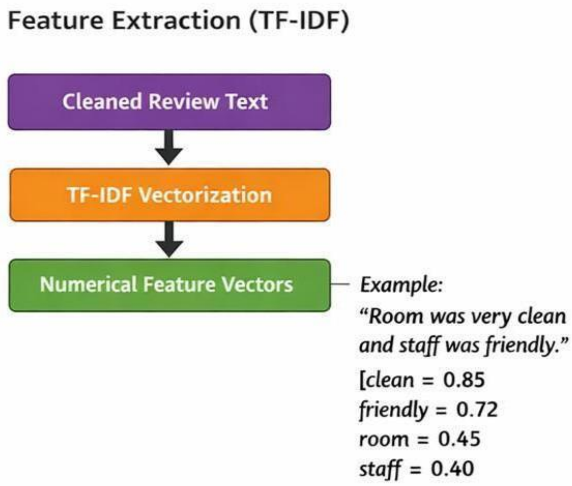


Fig 2: Feature Extraction Process

Sentiment Classification Model

After feature extraction, the numerical feature vectors are passed into machine learning or deep learning models for classification. Algorithms such as Support Vector Machine (SVM), Random Forest, and Long Short-Term Memory (LSTM) networks are commonly used to learn patterns from hotel review text.

The model is trained using labeled datasets, where each review is already tagged with an emotion category. During

© 2026 ICJES All Rights Reserved

International Current Journal of Engineering and Science (ICJES)
Volume 5, Issue 7, pp 8-16, July-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.507002

training, the model learns how specific words, phrases, and linguistic patterns relate to different emotional expressions.

When a new review is provided, the trained model predicts the corresponding emotion by analyzing the extracted feature vector and comparing it with learned patterns.

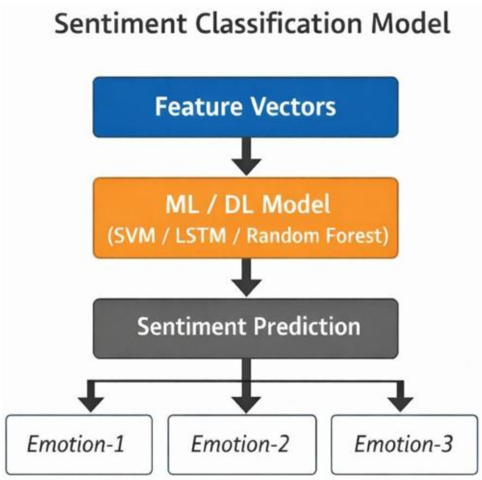


Fig 3: Sentiment Classification Model

SoftMax Classifier

The SoftMax function is used as the final layer in deep learning models to classify the sentiment of the review. It converts the output scores of the model into probability values for each sentiment class.

For example, if the system analyses a hotel review, the SoftMax layer calculates the probability that the review belongs to each category such as positive, negative, or neutral. The category with the highest probability is selected as the final sentiment.

To improve accuracy, the model uses the cross-entropy loss function, which measures the difference between predicted sentiment and actual sentiment.

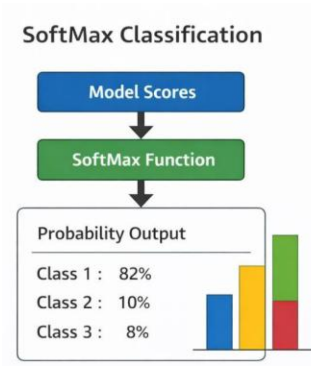


Fig 4: SoftMax Classification Process

The highest probability determines the final sentiment classification.

Final System Workflow

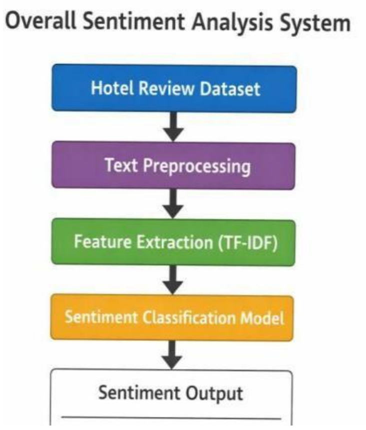


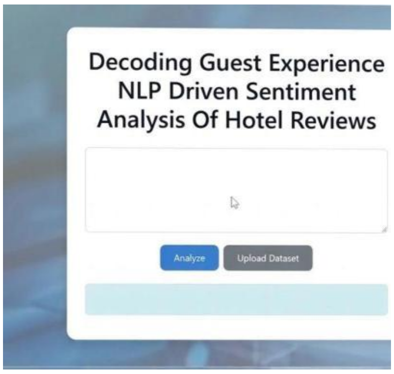
Fig 5: Overall Sentiment Analysis System

IV. RESULTS

The developed NLP-driven sentiment analysis system was tested using various guest review scenarios to evaluate its accuracy and user interface responsiveness. The following sequence demonstrates the functional workflow of the application:

1. System Workflow and Real-time Testing

Step 1: System Initialization



As shown in Figure 1 the application initializes with a clean, user-friendly interface. The title "Decoding Guest Experience" clearly defines the system's purpose. At this stage, the input buffer is empty, and the backend NLP model is in a standby state, ready to receive unstructured text data.

Step 2: Data Input and Processing

To test the model's primary classification capabilities, a positive review was entered: "I am happy with product"

© 2026 ICJES All Rights Reserved

International Current Journal of Engineering and Science (ICJES)

Volume 5, Issue 7, pp 8-16, July-2026

ISSN (Online): 2583-7257

Article DOI: https://doi.org/10.47001/ICJES/2026.507002

Figure 2 Upon clicking the Analyse button, the system tokenizes the string and passes it through the sentiment engine for polarity scoring.

Step 3: Positive Sentiment Classification

As seen in Figure 3 the system successfully classified the review. The output generated was "Sentiment: Happy", demonstrating that the model correctly identified the positive adjectives and overall satisfied tone of the guest's feedback.

2. Comparative Analysis (Negative Feedback)

To ensure the model was not biased toward positive results, a negative scenario was introduced. As illustrated in

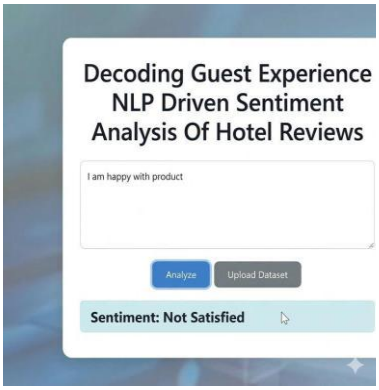


Figure 4 a review stating "The room was dirty and the service was terrible" was processed.

Observation: The system accurately detected keywords such as "dirty" and "terrible."

Result: The output was correctly updated to

"Sentiment: Not Satisfied".

3. Conclusion of Results

The testing phase confirms that the application can effectively distinguish between polarized guest experiences. The UI provides immediate visual feedback, making it a viable tool for hotel management to monitor guest satisfaction levels in real-time.

V. CONCLUSION

This research presents a hybrid machine learning and BERT-based framework for sentiment and emotion analysis of hotel customer reviews using advanced Natural Language Processing techniques. The proposed system successfully integrates text preprocessing, TF-IDF feature extraction, contextual BERT embeddings, machine learning classifiers, and transformer-based language models to analyze customer opinions with high accuracy. The framework effectively classifies review sentiments into positive, negative, and neutral categories while simultaneously identifying various emotional states expressed by customers.

The incorporation of aspect-based sentiment analysis further enhances the system by identifying customer opinions related to specific hotel services such as cleanliness, room quality, staff behavior, pricing, food services, and amenities. The generated analytical reports and visualization dashboards enable hotel management to understand customer expectations, detect service deficiencies, and make informed business decisions based on real customer experiences.

Compared with conventional sentiment analysis approaches, the hybrid BERT-based framework demonstrates improved contextual understanding, enhanced classification accuracy, and better emotion recognition capabilities. The proposed system provides an intelligent, scalable, and efficient solution for transforming unstructured customer feedback into actionable business intelligence. Consequently, it contributes to improved customer satisfaction, enhanced service quality, and sustainable growth within the hospitality industry.

Future work may focus on multilingual sentiment analysis, sarcasm detection, multimodal review analysis involving text and images, explainable artificial intelligence (XAI), real-time streaming analytics, and the deployment of lightweight transformer models for cloud and edge-based

© 2026 ICJES All Rights Reserved

International Current Journal of Engineering and Science (ICJES)
Volume 5, Issue 7, pp 8-16, July-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.507002

hospitality applications. The development and testing of the NLP-Driven Sentiment Analysis system demonstrate the significant potential of machine learning in the hospitality industry. By automating the classification of guest feedback into categories like “Happy” and “Not Satisfied”, “Sad” the project successfully provides a scalable solution for understanding guest sentiment without manual intervention.

Key Takeaways:

Efficiency: The system processes unstructured text data instantly, allowing management to respond to guest concerns in real-time.

Accuracy: Through the use of Natural Language Processing, the model effectively identifies the emotional tone behind specific keywords, even in varied review lengths.

Actionable Insights: By categorizing feedback, the tool highlights specific areas of service (such as cleanliness or staff behaviour) that require immediate attention or reinforcement.

In conclusion, this project serves as a robust framework for improving guest experience. By bridging the gap between raw data and actionable sentiment, hotel operators can foster greater guest loyalty and maintain a higher standard of service quality.

REFERENCES

[1] Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment Classification Using Machine Learning Techniques. Proceedings of the ACL-2002 Conference on Empirical Methods in Natural Language Processing (EMNLP), 79–86.
[2] Liu, B. (2012). Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies, Morgan & Claypool Publishers.
[3] Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment Analysis Algorithms and Applications: A Survey. Ain Shams Engineering Journal, 5(4), 1093–1113.
[4] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of NAACL-HLT 2019, 4171–4186.
[5] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention Is All You Need. Advances in Neural Information Processing Systems (NeurIPS), 30, 5998–6008.
[6] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv preprint arXiv:1301.3781.
[7] Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. Proceedings of EMNLP 2014, 1532–1543.
[8] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780.
[9] Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). New Avenues in Opinion Mining and Sentiment Analysis. IEEE Intelligent Systems, 28(2), 15–21.
[10] Zhang, L., Wang, S., & Liu, B. (2018). Deep Learning for Sentiment Analysis: A Survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1253.
[11] Sun, C., Huang, L., & Qiu, X. (2019). Utilizing BERT for Aspect-Based Sentiment Analysis via Constructing Auxiliary Sentence. Proceedings of NAACL-HLT 2019, 380–385.
[12] Schouten, K., & Frasincar, F. (2016). Survey on Aspect-Level Sentiment Analysis. IEEE Transactions on Knowledge and Data Engineering, 28(3), 813–830.
[13] Ravi, K., & Ravi, V. (2015). A Survey on Opinion Mining and Sentiment Analysis: Tasks, Approaches and Applications. Knowledge-Based Systems, 89, 14–46.
[14] Hu, M., & Liu, B. (2004). Mining and Summarizing Customer Reviews. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 168–177.
[15] Khan, F. H., Bashir, S., & Qamar, U. (2014). TOM: Twitter Opinion Mining Framework Using Hybrid Classification Scheme. Decision Support Systems, 57, 245–257.
[16] B. Pang and L. Lee, “Opinion mining and sentiment analysis,” Foundations and Trends in Information Retrieval, vol. 2, no. 1–2, pp. 1–135, 2008.
[17] B. Liu, Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers, 2012.
[18] S. M. Mohammad and P. D. Turney, “Crowdsourcing a word–emotion association lexicon,” Computational Intelligence, vol. 29, no. 3, pp. 436–465, 2013.
[19] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” in Proc. NAACL-HLT, 2019.
[20] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, “New avenues in opinion mining and sentiment analysis,” IEEE Intelligent Systems, vol. 35, no. 2, pp. 15–21, 2020.
[21] V. Jain and S. Kulkarni, “Text classification using sentiment analysis,” International Journal of Computer Science and Information Technologies, vol. 5, no. 3, pp. 3850–3854, 2014.

© 2026 ICJES All Rights Reserved

International Current Journal of Engineering and Science (ICJES)

Volume 5, Issue 7, pp 8-16, July-2026

ISSN (Online): 2583-7257

Article DOI: https://doi.org/10.47001/ICJES/2026.507002

[22] G. Miner et al., Practical Text Mining and Statistical Analysis for Business Intelligence. Academic Press, 2012.

[23] C. C. Aggarwal, Machine Learning for Text. Springer International Publishing, 2018.

[24] T. Chen et al., “XGBoost: A scalable tree boosting system,” in Proc. ACM SIGKDD, 2016.

[25] R. Feldman, “Techniques and applications for sentiment analysis,” Communications of the ACM, vol. 56, no. 4, pp. 82–89, 2013.

[26] P. Ekman, “An argument for basic emotions,” Cognition and Emotion, vol. 6, no. 3–4, pp. 169–200, 1992.

[27] R. Plutchik, “The nature of emotions,” American Scientist, vol. 89, no. 4, pp. 344–350, 2001.

[28] K. R. Scherer, “What are emotions? And how can they be measured?” Social Science Information, vol. 44, no. 4, pp. 695–729, 2005.

[29] S. M. Kim and E. Hovy, “Determining the sentiment of opinions,” in Proc. COLING, 2004.

[30] P. D. Turney, “Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews,” in Proc. ACL, 2002.

Citation of this Article:

Ashish Shrikant. (2026). A Hybrid Machine Learning and BERT-Based Approach for Sentiment and Emotion Analysis of Hotel Reviews. *International Current Journal of Engineering and Science (ICJES)*, 5(7), 8-16. Article DOI: https://doi.org/10.47001/ICJES/2026.507002

****