International Current Journal of Engineering and Science (ICJES)
Volume 5, Issue 5, pp 32-39, May-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.505005

Natural Language Processing for Guest Experience and Sentiment Analysis in Hospitality

[1]C Lakshmanna, [2]N Renuka Chowdary

[1,2]Department of Computer Science Engineering (Data Science), GATES Institute of Technology, Gooty, Andhra Pradesh, India

Abstract - **The hospitality industry increasingly relies on customer feedback to evaluate service quality and improve guest satisfaction. Hotel reviews contain valuable unstructured textual data that reflects customer opinions, emotions, expectations, and overall experiences. However, manual analysis of large volumes of reviews is time-consuming, inconsistent, and susceptible to human bias. To address these challenges, this study presents a real-time sentiment and emotion analysis system that leverages Natural Language Processing (NLP), machine learning, and deep learning techniques to automatically interpret guest feedback. The proposed system performs comprehensive text preprocessing, including tokenization, stop-word removal, lemmatization, punctuation filtering, and text normalization, to improve data quality. Feature extraction is achieved through TF-IDF and advanced word embedding techniques such as Word2Vec, GloVe, and BERT. A hybrid sentiment-emotion classification framework integrating machine learning algorithms, including Support Vector Machine (SVM) and Random Forest, with deep learning models such as Long Short-Term Memory (LSTM) networks and BERT Transformers is employed to achieve high classification accuracy. In addition, the system performs aspect-based sentiment analysis to evaluate specific service components, including staff behavior, cleanliness, and amenities, while also providing emotion distribution visualization and trend analysis over time. By transforming unstructured customer reviews into meaningful emotional and sentiment insights, the proposed framework enables hotel management to identify service strengths and weaknesses, detect negative feedback at an early stage, enhance customer satisfaction, and support data-driven decision-making. The system demonstrates the potential of AI-driven text analytics to improve service quality and operational efficiency in the hospitality sector.**

Keywords: Natural Language Processing (NLP), Sentiment Analysis, Emotion Detection, Hotel Reviews, Machine Learning, Deep Learning, BERT, LSTM, Aspect-Based Sentiment Analysis, Customer Feedback Analytics.

I. INTRODUCTION

The rapid growth of online booking platforms and travel websites has transformed the hospitality industry by enabling customers to share their experiences through digital reviews. These reviews provide valuable insights into various aspects of hotel services, including staff behavior, room quality, cleanliness, amenities, pricing, and overall customer satisfaction. As the volume of user-generated content continues to increase, manually analyzing thousands of reviews has become impractical, time-consuming, and susceptible to human bias. Consequently, automated sentiment and emotion analysis has emerged as an effective solution for extracting meaningful information from large collections of textual data.

Natural Language Processing (NLP) has become one of the most influential technologies for understanding and interpreting human language. By applying NLP techniques, unstructured text can be transformed into structured information that supports decision-making processes. Sentiment analysis aims to identify the polarity of a review, such as positive, negative, or neutral, while emotion analysis goes a step further by detecting emotional states such as happiness, satisfaction, disappointment, anger, or sadness. Combining sentiment and emotion detection provides a more comprehensive understanding of customer opinions and expectations.

Recent advancements in machine learning and deep learning have significantly improved the performance of text classification systems. Traditional machine learning algorithms such as Support Vector Machine (SVM) and Random Forest have demonstrated effective results in sentiment classification tasks. At the same time, deep learning architectures including Long Short-Term Memory (LSTM) networks and Bidirectional Encoder Representations from Transformers (BERT) have achieved remarkable accuracy by capturing contextual relationships within text data. These technologies enable intelligent systems to understand complex linguistic patterns and improve classification performance.

International Current Journal of Engineering and Science (ICJES)
Volume 5, Issue 5, pp 32-39, May-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.505005

The proposed system integrates NLP, machine learning, and deep learning techniques to develop a real-time sentiment and emotion analysis framework for hotel reviews. The system performs data preprocessing, feature extraction, sentiment classification, emotion recognition, and aspect-based sentiment analysis to generate actionable insights. Furthermore, it provides visualization of emotional trends and service-related aspects, allowing hotel management to identify areas requiring improvement and enhance customer satisfaction. By converting unstructured textual feedback into meaningful analytical information, the proposed framework supports data-driven decision-making and contributes to improved service quality in the hospitality industry. 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

The analysis of customer opinions using Natural Language Processing and machine learning has attracted considerable attention in recent years. Researchers have developed various sentiment analysis models to automatically classify textual data and understand customer preferences across different domains, including e-commerce, healthcare, tourism, and hospitality.

Early sentiment analysis systems primarily relied on lexicon-based approaches, where predefined dictionaries of positive and negative words were used to determine the sentiment of a text. Although these methods were computationally simple, they often struggled with context-dependent meanings, sarcasm, and complex sentence structures. To overcome these limitations, machine learning algorithms such as Naïve Bayes, Support Vector Machine (SVM), and Random Forest were introduced, offering improved classification accuracy through supervised learning techniques.

With the advancement of deep learning, researchers began employing neural network architectures for sentiment classification. Long Short-Term Memory (LSTM) networks demonstrated strong capabilities in capturing sequential dependencies and contextual information within textual data. More recently, transformer-based models such as BERT have achieved state-of-the-art performance by understanding bidirectional contextual relationships between words, making them highly effective for sentiment and emotion detection tasks.

Several studies have also focused on aspect-based sentiment analysis, which evaluates customer opinions regarding specific service attributes such as room quality, staff behavior, food services, cleanliness, and hotel facilities. This approach provides more detailed insights than overall sentiment classification by identifying the strengths and weaknesses of individual service components. Aspect-based analysis has become an important tool for hotel management to improve operational performance and customer experience.

Despite significant progress, many existing systems focus only on sentiment polarity and do not provide detailed emotion recognition or real-time analytical visualization. In addition, several approaches rely on a single classification model, which may limit prediction accuracy. The proposed research addresses these limitations by integrating advanced preprocessing methods, multiple feature extraction techniques, hybrid machine learning and deep learning models, aspect-based sentiment analysis, and emotion visualization within a unified framework. This comprehensive approach aims to provide more accurate and meaningful insights from hotel reviews while supporting strategic decision-making in the hospitality sector.

Volume 5, Issue 5, pp 32-39, May-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.505005

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

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:

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.

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.

International Current Journal of Engineering and Science (ICJES)
Volume 5, Issue 5, pp 32-39, May-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.505005

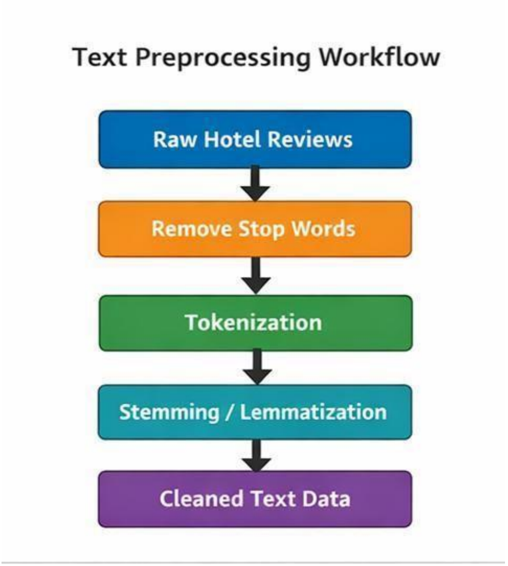

Fig 1: Text Preprocessing Workflow

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.

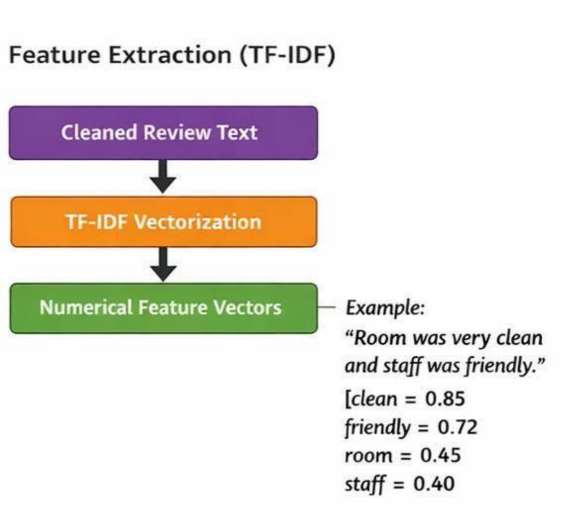

Fig 2: Feature Extraction Process

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 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.

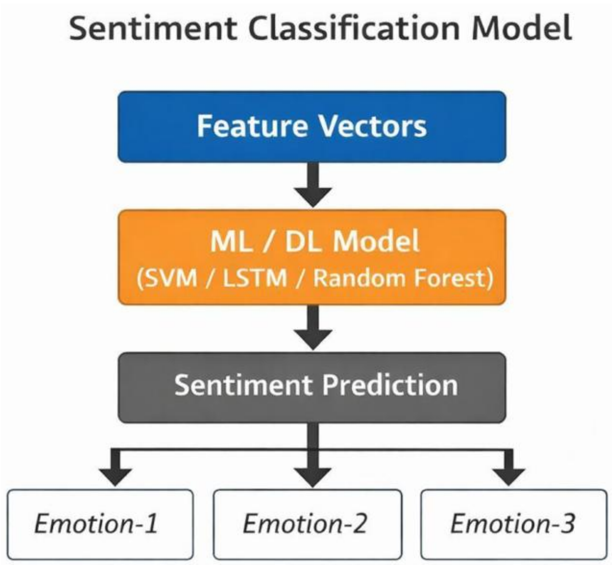

Fig 3: Sentiment Classification Model

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.

International Current Journal of Engineering and Science (ICJES)
Volume 5, Issue 5, pp 32-39, May-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.505005

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

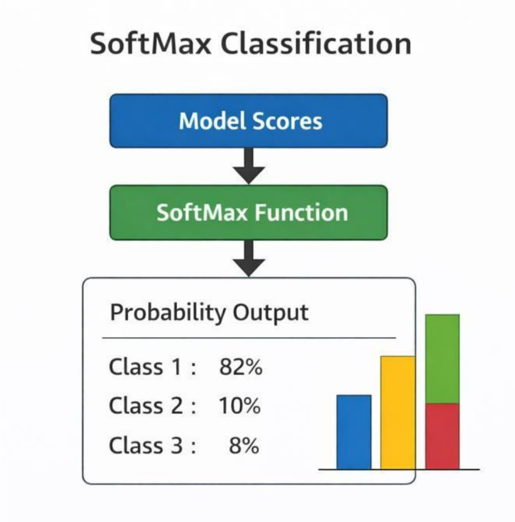

Fig 4: SoftMax Classification Process

The highest probability determines the final sentiment classification.

Final System Workflow

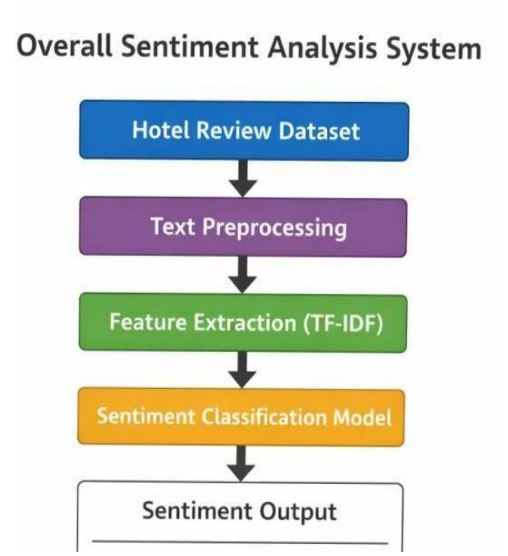

Fig 5: Overall Sentiment Analysis System

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"

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

International Current Journal of Engineering and Science (ICJES)
Volume 5, Issue 5, pp 32-39, May-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.505005

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

The proposed real-time sentiment and emotion analysis system for hotel reviews demonstrates the significant potential of Artificial Intelligence (AI), Natural Language Processing (NLP), machine learning, and deep learning technologies in transforming unstructured customer feedback into meaningful business intelligence. Hotel reviews contain valuable information about guest experiences, expectations, satisfaction levels, and service quality, but the large volume of such data makes manual analysis inefficient and unreliable. The developed framework addresses this challenge by automatically processing textual data through preprocessing techniques such as tokenization, stop-word removal, lemmatization, punctuation cleaning, and text normalization, thereby improving the quality of the input data for analysis.

The system employs advanced feature extraction methods, including TF-IDF, Word2Vec, GloVe, and BERT embeddings, to capture both semantic and contextual information from customer reviews. Furthermore, the integration of traditional machine learning algorithms such as Support Vector Machine (SVM) and Random Forest with deep learning models including Long Short-Term Memory (LSTM) networks and BERT Transformers enhances the accuracy and robustness of sentiment and emotion classification. In addition to identifying overall sentiment, the framework performs aspect-based sentiment analysis to evaluate specific hotel service attributes, such as staff behavior, cleanliness, amenities, and overall guest experience. The visualization of emotional distributions and trend analysis provides management with actionable insights that support strategic planning and service improvement.

The proposed approach enables hotel administrators to detect negative feedback at an early stage, understand customer expectations more effectively, and make informed decisions based on real-time analytical results. By converting large volumes of unstructured text into structured emotional intelligence, the system contributes to improved customer satisfaction, enhanced service quality, and increased operational efficiency. Moreover, the framework reduces the dependency on manual review processes, minimizes human bias, and provides a scalable solution capable of handling continuously growing online review datasets.

The successful implementation of this research highlights the practical application of AI-driven text analytics in the hospitality industry and demonstrates how advanced computational techniques can support customer-centric business strategies. The proposed system can serve as a valuable decision-support tool for hotel management, helping organizations strengthen their reputation and maintain a competitive advantage in the digital era. Future enhancements may include multilingual sentiment analysis, integration with social media and travel platforms, real-time streaming data processing, voice and speech emotion recognition, and the incorporation of explainable artificial intelligence (XAI) techniques to improve model transparency and interpretability. These advancements will further expand the capabilities of intelligent hotel review analysis systems and contribute to the development of smarter and more personalized hospitality services.

International Current Journal of Engineering and Science (ICJES)
Volume 5, Issue 5, pp 32-39, May-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.505005

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] 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.

[2] B. Liu, Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers, 2012.

[3] S. M. Mohammad and P. D. Turney, "Crowdsourcing a word–emotion association lexicon," Computational Intelligence, vol. 29, no. 3, pp. 436–465, 2013.

[4] 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.

[5] 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.

[6] 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.

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

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

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

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

[11] P. Ekman, "An argument for basic emotions," Cognition and Emotion, vol. 6, no. 3–4, pp. 169–200,

1992.

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

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

[14] S. M. Kim and E. Hovy, "Determining the sentiment of opinions," in Proc. COLING, 2004.

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

International Current Journal of Engineering and Science (ICJES)
Volume 5, Issue 5, pp 32-39, May-2026
ISSN (Online): 2583-7257
Article DOI: https://doi.org/10.47001/ICJES/2026.505005

Citation of this Article:

C Lakshmanna, & N Renuka Chowdary. (2026). Natural Language Processing for Guest Experience and Sentiment Analysis in Hospitality. *International Current Journal of Engineering and Science (ICJES)*, 5(5), 32-39. Article DOI: https://doi.org/10.47001/ICJES/2026.505005

****

