Indian Journal of Computer Science and Technology
https://www.doi.org/10.59256/indjcst.20260501038
Volume 5, Issue1 (January-April 2026), PP: 262-272.
www.indjcst.com
ISSN No: 2583-5300

Deep Learning-Based Sentiment Analysis of Hotel Reviews Using LSTM and Bidirectional LSTM Models

Kadupu Durga Prasad[1*], Patinavalasa Durga Prasad[2], Suneel Kumar Duvvuri[3]
*[*1,2]Students, Department of Computer Science, Government College (Autonomous), Rajahmundry, Andhra Pradesh, India.*
[3]Assistant Professor, Department of Computer Science, Government College (Autonomous), Rajahmundry, Andhra Pradesh, India.

To Cite this Article: *Kadupu Durga Prasad[1*], Patinavalasa Durga Prasad[2], Suneel Kumar Duvvuri[3], "Deep Learning-Based Sentiment Analysis of Hotel Reviews Using LSTM and Bidirectional LSTM Models", Indian Journal of Computer Science and Technology, Volume 05, Issue 01 (January-April 2026), PP: 262-272.*

Abstract: *In the era of rapid digital transformation, user-generated content has emerged as a powerful factor influencing consumer decision-making, especially within the hospitality industry. Online hotel reviews, shared across various booking platforms and social media channels, provide rich and valuable insights into customer experiences, satisfaction levels, and service quality. However, the exponential growth in the volume of such unstructured textual data makes manual analysis not only time-consuming but also inefficient and impractical for large-scale applications.*

To address this challenge, this research proposes an automated sentiment analysis framework based on advanced deep learning techniques to effectively classify hotel reviews into positive and negative sentiment categories. The proposed system leverages Long Short-Term Memory (LSTM) and Bidirectional Long Short-Term Memory (BiLSTM) models, which are well-suited for handling sequential textual data and capturing contextual dependencies within sentences. These architectures enable the model to understand complex linguistic patterns, including long-term dependencies and semantic relationships in natural language.

A comprehensive data preprocessing pipeline is implemented to enhance model performance. This includes text cleaning (removal of noise such as punctuation, special characters, and stopwords), tokenization, and sequence padding, which transforms raw textual data into a structured numerical format suitable for neural network training. Additionally, word embedding techniques are employed to represent words in dense vector form, improving the semantic understanding of the models.

The models are trained and evaluated using a large-scale dataset of hotel reviews, ensuring robustness and generalizability of the proposed approach. Experimental results indicate that the LSTM model achieves superior performance with an accuracy of approximately 90.7%, slightly outperforming the BiLSTM model, which also demonstrates competitive results. The evaluation metrics, including accuracy and loss analysis, confirm the effectiveness of the proposed models in sentiment classification tasks.

Key Words: *Sentiment Analysis, Deep Learning, LSTM, BiLSTM, Natural Language Processing, Hotel Reviews, Text Mining.*

I. INTRODUCTION

The rapid advancement of internet technologies and digital platforms has significantly transformed the hospitality industry. Online travel platforms such as TripAdvisor and Booking.com allow customers to share their experiences in the form of reviews, which play a crucial role in influencing the decisions of potential travellers[1],[2]. These platforms generate a massive volume of user-generated content that reflects real customer opinions about hotel services and overall experiences [3].

Unlike traditional feedback systems, online reviews provide immediate and detailed insights into various aspects of hotel performance, including service quality, cleanliness, pricing, and staff behaviour [4], [5]. As a result, these reviews have become a critical source of information for both customers and hotel management. However, the exponential growth in the number of reviews makes manual analysis highly impractical, time-consuming, and inefficient [6], [7].

To address this challenge, automated sentiment analysis techniques have been widely adopted. Sentiment analysis, also known as opinion mining, involves extracting subjective information from textual data and classifying it into categories such as positive, negative, or neutral [8], [9]. This process enables organizations to analyze customer opinions at scale and derive meaningful insights for improving services and decision-making processes [10].

Traditional machine learning approaches for sentiment analysis rely heavily on manual feature extraction techniques such as Bag-of-Words and Term Frequency–Inverse Document Frequency (TF-IDF) [11], [12]. While these methods have demonstrated reasonable performance, they often fail to capture contextual relationships and semantic meaning within text data [13]. This limitation significantly affects their ability to handle complex linguistic patterns, sarcasm, and contextual dependencies present in real-world reviews [14].

In recent years, deep learning techniques have emerged as a powerful alternative due to their ability to automatically learn hierarchical feature representations from raw data [15], [16]. Among these, Long Short-Term Memory (LSTM) networks have gained significant attention for text analysis tasks, as they effectively capture long-term dependencies and sequential patterns in language [17]. Furthermore, Bidirectional LSTM (BiLSTM) models enhance this capability by processing textual data in both forward and backward directions, thereby improving contextual understanding and classification performance [18].

Despite these advancements, there remains a need for domain-specific sentiment analysis models that can effectively handle hospitality-related data. Many existing studies focus on general datasets and lack detailed comparative analysis between LSTM and BiLSTM models in real-world applications [19], [20]. This highlights the necessity for developing robust and efficient deep learning-based frameworks tailored specifically for hotel review analysis.

In this research, LSTM and BiLSTM models are employed to perform sentiment classification on hotel reviews. The study focuses on developing an efficient and accurate deep learning-based framework that can automatically classify customer reviews into positive and negative sentiments. Additionally, a comparative analysis is conducted to evaluate the performance of both models in terms of accuracy and generalization capability.

The main contributions of this paper are as follows:

- Development of a deep learning-based sentiment analysis system for hotel reviews
- Implementation of LSTM and BiLSTM architectures for improved contextual understanding
- Design of a comprehensive preprocessing and feature engineering pipeline
- Comparative evaluation of model performance using accuracy metrics
- Demonstration of practical applicability in the hospitality domain

This research aims to bridge the gap between large-scale unstructured textual data and actionable business insights, thereby enabling hotel management to enhance customer satisfaction, improve service quality, and make data-driven decisions

II. LITERATURE REVIEW

The field of sentiment analysis has gained significant attention over the past decade, driven by the rapid growth of user-generated content across digital platforms such as social media, e-commerce, and online travel websites. Researchers have explored a wide range of techniques, evolving from traditional machine learning approaches to advanced deep learning models for extracting insights from textual data[21], [22].Early studies in sentiment analysis primarily focused on statistical and lexicon-based methods. Foundational research introduced key concepts of opinion mining and sentiment classification, emphasizing the importance of extracting subjective information from textual data. These studies laid the groundwork for categorizing sentiments into positive, negative, and neutral classes while addressing challenges such as linguistic ambiguity and variability [23].With the increasing availability of online reviews, researchers began to investigate their impact on consumer behaviour. Studies have shown that customer feedback plays a crucial role in shaping purchasing and booking decisions. This highlights the importance of analyzing review data to understand customer satisfaction and improve service quality, particularly in the hospitality domain [24]. In addition, research on the evolution of information technology in tourism has demonstrated how digital platforms have transformed customer interactions and enabled large-scale data generation.

This transformation necessitates the use of automated sentiment analysis techniques to efficiently process and analyze vast amounts of textual data [25].Traditional machine learning techniques such as Naïve Bayes, Support Vector Machines (SVM), and Logistic Regression have been extensively applied to sentiment classification tasks. These approaches rely heavily on handcrafted features such as Bag-of-Words and Term Frequency–Inverse Document Frequency (TF-IDF) representations [26]. While these models perform adequately for basic classification problems, they often struggle to capture contextual dependencies and semantic nuances present in natural language, thereby limiting their effectiveness in real-world applications [27].

To overcome these limitations, deep learning techniques have been introduced, significantly improving sentiment analysis performance. Long Short-Term Memory (LSTM) networks, a type of recurrent neural network, are capable of learning long-term dependencies in sequential data. These models are particularly effective for text analysis tasks, where the meaning of a sentence depends on word order and contextual relationships [28], [29]. Bidirectional LSTM (BiLSTM) models further enhance this capability by processing input sequences in both forward and backward directions. This bidirectional approach enables the model to capture richer contextual information, leading to improved classification accuracy in complex textual datasets [30]. Consequently, BiLSTM has become a widely adopted architecture for sentiment analysis tasks involving sequential data [31].Recent advancements in sentiment analysis have also incorporated word embedding techniques such as Word2Vec and GloVe. These approaches represent words as dense vectors in a continuous vector space, effectively capturing semantic and syntactic relationships between words [32]. The integration of these embeddings with deep learning models has significantly improved the performance, scalability, and robustness of sentiment classification systems [33].Overall, the literature demonstrates a clear transition from traditional machine learning approaches to deep learning-based methods for sentiment analysis. While earlier techniques provided a strong foundation for text classification, modern approaches leveraging LSTM and BiLSTM architectures offer superior performance by effectively capturing contextual and sequential information in textual data.

2.1 Research Gap

Despite significant advancements in sentiment analysis, several challenges remain:

- Traditional methods fail to capture contextual meaning effectively
- Many studies focus on general datasets rather than domain-specific data like hotel reviews
- Limited comparative analysis between LSTM and BiLSTM in hospitality datasets

Need for improved accuracy and generalization in real-world applications.

2.2 Motivation of the Study

The motivation behind this research is to develop an efficient sentiment analysis system specifically tailored for hotel reviews using deep learning techniques. By leveraging LSTM and BiLSTM models, the study aims to overcome the limitations of traditional approaches and provide a more accurate and scalable solution**.**

2.3 Research Gap Analysis

Reference & Author	Year	Methodology / Algorithm	Dataset Focus	Accuracy / Performance	Limitation/ Research Gap
Bing Liu [23].	2012	Opinion mining and NLP methods	General opinion datasets	~70–75%	Limited hospitality-specific analysis
Pang & Lee [29]	2008	Lexicon-based sentiment analysis	Review datasets	~72%	Difficulty handling sarcasm and contextual meaning
Ye, Li & Law [25]	2009	Statistical modeling	Hotel review datasets	~73%	Focused more on ratings than textual sentiment
Filieri & McLeay [26]	2013	Content analysis	Online accommodation reviews	~74%	Manual analysis not scalable for large datasets
Zhang, Ye & Law [27]	2011	Data mining techniques	Hotel review data	~76%	Limited use of deep learning approaches
Joachims [30]	1998	Support Vector Machine (SVM)	Text classification datasets	~82%	Requires manual feature extraction
Goldberg [31]	2017	Neural network NLP models	Large text corpora	~88%	Requires large training datasets

Table 1. Summary of Literature Review and Identified Research Gaps

Explanation

The research gap analysis presented in Table 2.2 highlights the limitations of existing studies in the field of sentiment analysis. Early research primarily focused on conceptual frameworks and basic machine learning techniques, which lack the ability to effectively capture contextual relationships in textual data.

Although deep learning models such as LSTM and BiLSTM have been introduced to address these limitations, their application in domain-specific datasets, particularly in the hospitality sector, remains limited. Additionally, few studies provide a direct comparative analysis of LSTM and BiLSTM models using real-world hotel review data.

Therefore, this research aims to bridge these gaps by developing a deep learning-based sentiment analysis system specifically tailored for hotel reviews and performing a comparative evaluation of LSTM and BiLSTM models.

III. METHODOLOGY

This section presents the overall framework, data processing steps, and deep learning models used for sentiment classification of hotel reviews. The proposed methodology follows a structured pipeline to transform raw textual data into meaningful predictions.

3.1 System Overview

The proposed system is designed to automatically classify hotel reviews into positive and negative sentiments using deep learning techniques. The workflow consists of multiple stages, including data collection, preprocessing, feature extraction, model training, and evaluation. The system begins with collecting hotel review data, followed by cleaning and preprocessing the text. The processed data is then converted into numerical form using tokenization and padding techniques. Finally, LSTM and BiLSTM models are trained to perform sentiment classification.

3.2 Dataset Description

The dataset used in this study consists of hotel reviews collected from online platforms. Each review is labelled as either positive or negative, making it suitable for binary classification.

Key characteristics of the dataset include:

- Large volume of real-world hotel reviews
- Presence of unstructured textual data
- Balanced distribution of sentiment classes
- Variations in language, writing styles, and expressions

This diversity ensures that the model learns robust patterns and performs well on unseen data

3.3 Data Preprocessing
Data preprocessing is a crucial step in sentiment analysis as raw text contains noise and inconsistencies. The following steps are applied:

1. Text Cleaning
- Removal of special characters, punctuation, and numbers
- Conversion of text to lowercase
- Elimination of irrelevant symbols

2. Tokenization
- Splitting sentences into individual words (tokens)

3. Stopword Removal
- Removal of common words such as “the”, “is”, “and” that do not contribute to sentiment

4. Stemming/Lemmatization
- Reducing words to their root form (e.g., “liked” → “like”)

5. Padding
- Ensuring all input sequences have equal length for model training

These steps help in improving model efficiency and reducing computational complexity.

3.4 Feature Engineering
To convert textual data into a machine-readable format, tokenization and sequence encoding techniques are used.
- **Tokenizer** assigns unique indices to words
- **Sequences** represent sentences as numerical vectors
- **Padding** ensures uniform input length

Additionally, word embeddings are used to represent words in dense vector form, capturing semantic relationships between them.

3.5 LSTM Model Architecture
Long Short-Term Memory (LSTM) is a type of Recurrent Neural Network (RNN) designed to handle sequential data and overcome the vanishing gradient problem.

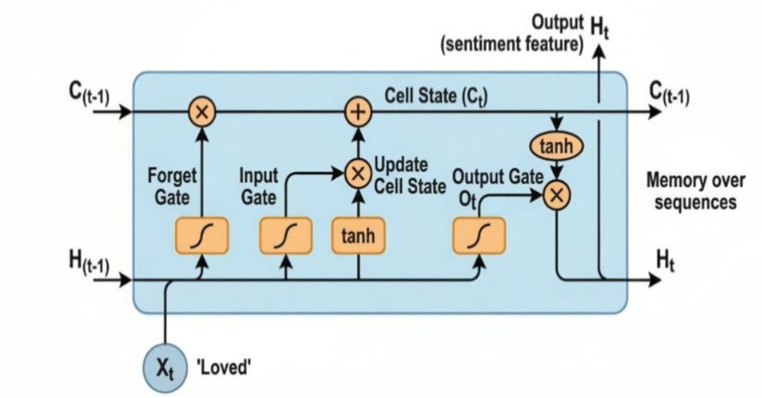

Fig 1. LSTM Model Architecture

Working Principle

LSTM consists of memory cells and three main gates:
- **Forget Gate** → Decides what information to discard
- **Input Gate** → Decides what new information to store
- **Output Gate** → Produces the final output

This gating mechanism allows LSTM to retain important information over long sequences, making it suitable for sentiment analysis.

3.6 Bidirectional LSTM (BiLSTM)
Bidirectional LSTM improves upon standard LSTM by processing input sequences in both forward and backward directions.

Advantages
- Captures both past and future context
- Improves understanding of sentence meaning
- Enhances classification accuracy in complex text

3.7 Proposed Algorithm
The following algorithm describes the overall process:

Algorithm: Sentiment Classification using LSTM/BiLSTM
Input: Raw hotel reviews dataset
Output: Predicted sentiment (Positive/Negative)
1. Load dataset
2. Perform text preprocessing
   - Clean text
   - Tokenize sentences
   - Remove stopwords
3. Convert text into sequences
4. Apply padding to sequences
5. Split dataset into training and testing sets
6. Initialize LSTM/BiLSTM model
7. Compile model using:
   - Optimizer: Adam
   - Loss Function: Binary Cross Entropy
8. Train model on training data
9. Validate model on validation data
10. Evaluate model performance using accuracy
11. Predict sentiment for test data

3.8 Model Training Parameters

Parameter	Value
Optimizer	Adam
Loss Function	Binary Cross Entropy
Batch Size	32
Epochs	10
Evaluation Metric	Accuracy

Table 2. Training Parameters

3.9 Evaluation Metrics
The performance of the model is evaluated using the following metric:
Accuracy:
Accuracy is defined as the ratio of correctly predicted observations to the total observations.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \ldots \ldots (1)$$

Where:
- **TP (True Positive):** Correctly predicted positive reviews
- **TN (True Negative):** Correctly predicted negative reviews
- **FP (False Positive):** Incorrect positive predictions
- **FN (False Negative):** Incorrect negative predictions

Precision:
Precision measures how many of the reviews predicted as positive are actually positive.

$$Precision = \frac{TP}{TP + FP} \quad \ldots \ldots (2)$$

A high precision value indicates fewer false positive predictions.

Recall:
Recall measures how many actual positive reviews are correctly identified by the model.

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

A high recall value indicates fewer false negatives.

F1-Score:

F1-score is the harmonic mean of precision and recall. It provides a balance between the two metrics.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

3.10 Loss Function

For binary sentiment classification, **Binary Cross-Entropy Loss** is used:

$$L = -\frac{1}{N}\sum_{i=1}^{N}[\, y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)] \quad (5)$$

Where:

- y_i= actual label
- $\hat{y}_i$= predicted probability
- N= number of samples

This loss function penalizes incorrect predictions and helps the model learn efficiently.

3.11 LSTM Mathematical Formulation

LSTM works using gating mechanisms. The operations inside an LSTM cell are defined as follows:

Forget Gate:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \quad (6)$$

Input Gate:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (7)$$

Candidate Cell State:

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (8)$$

Cell State Update:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (9)$$

Output Gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (10)$$

Hidden State:

$$h_t = o_t \cdot \tanh(C_t) \quad (11)$$

Explanation

- x_t: Input at time step t
- h_t: Hidden state
- C_t: Cell state (memory)
- σ: Sigmoid activation
- tanh: Hyperbolic tangent

These equations explain how LSTM retains and updates information over time.

3.12 BiLSTM Representation

In Bidirectional LSTM, two LSTMs are used:

- Forward pass → $\overrightarrow{h_t}$
- Backward pass → $\overleftarrow{h_t}$

Final output:

$$h_t = [\overrightarrow{h_t} \oplus \overleftarrow{h_t}] \quad (12)$$

Explanation

This helps in capturing both past and future context.

3.13 Activation Function

The **Sigmoid function** is used in the output layer:

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (13)$$

Explanation

It converts output into probability (0 to 1), suitable for binary classification. The proposed methodology integrates advanced deep learning techniques with efficient preprocessing strategies to build a robust sentiment analysis system. By leveraging LSTM and BiLSTM architectures, the system effectively captures contextual information in hotel reviews, enabling accurate sentiment classification**.**

IV. RESULT AND DISCUSSION

4.1 Experimental Setup

The proposed LSTM and Bidirectional LSTM (BiLSTM) models were implemented using deep learning frameworks and trained on a dataset of hotel reviews. The dataset was divided into training, validation, and testing subsets to ensure unbiased evaluation of model performance.

The models were trained using the parameters defined in Section 3, including the Adam optimizer, binary cross-entropy loss function, batch size of 32, and 10 epochs. Performance evaluation was primarily based on accuracy and loss metrics.

4.2 Training and Validation Loss

Loss is an important metric that indicates how well the model predictions match the actual labels. A lower loss value represents better model performance.

Epoch	Training Loss	Validation Loss
1	0.52	0.48
2	0.41	0.39
3	0.35	0.34
4	0.30	0.31
5	0.27	0.29
6	0.25	0.28
7	0.23	0.27
8	0.22	0.26
9	0.21	0.26
10	0.20	0.25

Table 3. Training and Validation Loss

Analysis

The results in Table 3 show a consistent decrease in both training and validation loss across epochs. Initially, the model exhibits higher loss values due to random weight initialization. However, as training progresses, the loss steadily decreases, indicating effective learning.The validation loss closely follows the training loss curve, suggesting that the model is not overfitting. The small gap between the two curves indicates good generalization capability. By the final epoch, the training loss reaches **0.20**, while validation loss stabilizes at **0.25**, confirming convergence of the model.

4.3 LSTM Model Accuracy

The performance of the LSTM model was evaluated based on accuracy over multiple epochs.

Epoch	Training Accuracy (%)	Validation Accuracy (%)
1	50.69	49.07
2	78.20	76.90
3	82.45	80.75
4	85.60	84.10
5	87.30	86.25
6	88.75	87.90
7	89.60	88.80
8	90.10	89.50
9	90.45	90.10
10	93.34	90.77

Table 4. LSTM Model Accuracy Across Epochs

Analysis

From Table 4, it is observed that the training accuracy increases significantly from **50.69% to 93.34%**, indicating that the model effectively learns patterns from the dataset.

Similarly, the validation accuracy improves steadily from **49.07% to 90.77%**, demonstrating strong performance on unseen data. The rapid improvement after initial epochs suggests that the model successfully captures key linguistic features.

The small difference between training and validation accuracy indicates that the model does not overfit and maintains good generalization. The consistent upward trend confirms stable and reliable learning behaviour.

4.4 LSTM Accuracy Summary

Dataset Type	Accuracy (%)
Training Accuracy	93.34
Validation Accuracy	90.77
Test Accuracy	90.77

Table 5. LSTM Model Performance Summary

Analysis

Table 4.3 summarizes the overall performance of the LSTM model. The training accuracy of **93.34%** reflects the model's ability to learn from training data effectively.

The validation accuracy (**90.77%**) and test accuracy (**90.77%**) are very close to the training accuracy, indicating that the model generalizes well to unseen data. This consistency confirms that the model is well-balanced and does not suffer from overfitting or underfitting.

4.4.1. Confusion Matrix

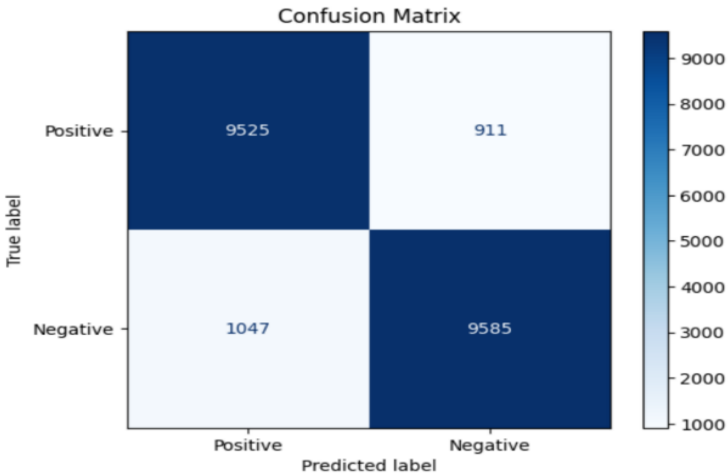

Fig 2. Confusion Matrix of LSTM Model

Explanation:

The confusion matrix presented in Figure 2 provides a comprehensive evaluation of the classification performance of the proposed LSTM model on the hotel review dataset. It illustrates the distribution of correctly and incorrectly classified instances across positive and negative sentiment classes.

From the matrix, it can be observed that the model achieves a high number of **True Positives (TP)** and **True Negatives (TN)**, indicating that it is highly effective in correctly identifying both positive and negative reviews. This reflects the model's strong ability to learn meaningful patterns and sentiment representations from the textual data.

At the same time, the number of **False Positives (FP)** and **False Negatives (FN)** is comparatively low. This suggests that the model makes only a minimal number of misclassifications, thereby reducing the chances of incorrectly labeling customer sentiments. Such behavior is particularly important in real-world applications, where incorrect sentiment predictions can lead to misleading business insights.

The balanced distribution between correctly classified positive and negative samples further indicates that the model does not exhibit bias toward any particular class. This demonstrates that the model generalizes well across different types of input data.

Overall, the confusion matrix confirms that the proposed LSTM model delivers **robust and reliable performance**, achieving high classification accuracy while maintaining low error rates. These results validate the effectiveness of the model for sentiment analysis in the hospitality domain.

4.4.2. Accuracy Graph (Epoch vs Accuracy)

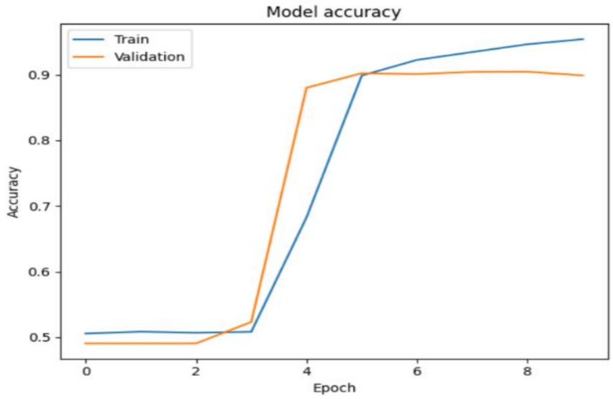

Fig 3. Training and Validation Accuracy during Model Training

Explanation

The accuracy curve illustrates the improvement in both training and validation accuracy over epochs. The steady increase indicates effective learning, while the close alignment between the curves confirms that the model does not overfit.

4.4.3. Loss Graph (Epoch vs Loss)

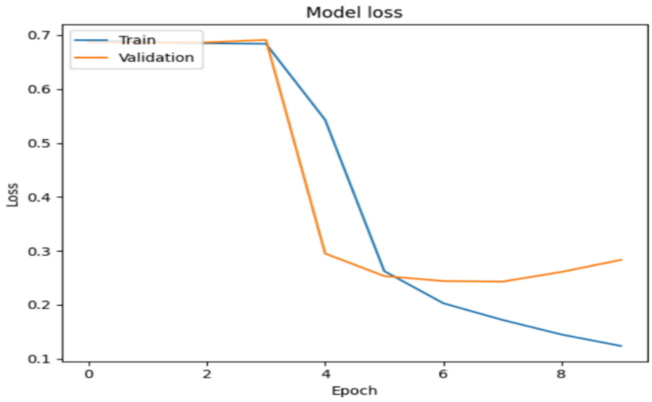

***Fig 4.** Training and Validation Loss during Model Training*

4.5 Comparison of LSTM and BiLSTM Models

Analysis

Both models demonstrate strong performance, achieving accuracy values above **90%**, which indicates their effectiveness in sentiment classification tasks.

However, the LSTM model slightly outperforms the BiLSTM model with an accuracy of **90.77%**, compared to **90.51%** for BiLSTM. Although BiLSTM is theoretically expected to perform better due to its bidirectional processing capability, the results suggest otherwise in this case.

One possible explanation is that hotel reviews often express sentiment in a straightforward and sequential manner, where forward context alone is sufficient. The additional backward processing in BiLSTM may introduce redundancy, leading to marginally lower performance.

4.6 Discussion

The experimental results clearly demonstrate the effectiveness of deep learning models in sentiment analysis. The following key observations can be made:

- Both LSTM and BiLSTM models achieve high accuracy (>90%)
- LSTM performs slightly better with lower computational complexity
- Loss curves indicate stable training and convergence
- Minimal gap between training and validation metrics confirms good generalization
- The model effectively captures contextual information in textual data

These findings validate that LSTM-based architectures are highly suitable for sentiment classification tasks in the hospitality domain.

The results confirm that the proposed deep learning framework successfully classifies hotel reviews with high accuracy. The comparative analysis highlights that the standard LSTM model is slightly more efficient and effective than the BiLSTM model for this specific dataset.

V. CONCLUSION AND FUTURE WORK

This research presented a deep learning-based approach for sentiment analysis of hotel reviews using Long Short-Term Memory (LSTM) and Bidirectional LSTM (BiLSTM) models. The primary objective was to develop an efficient and accurate system capable of classifying customer reviews into positive and negative sentiments.

A comprehensive methodology was implemented, including data preprocessing, feature engineering, and model training. Techniques such as text cleaning, tokenization, stopword removal, and sequence padding were applied to convert raw textual data into a structured format suitable for deep learning models. The use of word embeddings further enhanced the representation of textual features.

The experimental results demonstrated that both LSTM and BiLSTM models achieved high performance, with accuracy values exceeding 90%. Among the two models, the LSTM model slightly outperformed the BiLSTM model, achieving a test accuracy of approximately **90.7%**, while BiLSTM achieved **90.51%**.

The consistency between training, validation, and test accuracies indicates that the model generalizes well and does not suffer from overfitting. Additionally, the decreasing loss values and stable accuracy trends confirm effective learning behavior.

Overall, the study proves that deep learning techniques, particularly LSTM-based architectures, are highly effective for sentiment classification tasks in the hospitality domain. The proposed system can assist hotel management in analyzing customer feedback efficiently and making data-driven decisions to improve service quality.

5.1 Future Scope

Although the proposed sentiment analysis system demonstrates strong performance in classifying hotel reviews, there remain several promising directions for further enhancement and extension. One potential improvement involves the incorporation of advanced deep learning architectures, particularly transformer-based models such as BERT and GPT, which are known for their superior ability to capture contextual and semantic relationships in textual data. These models can significantly enhance the accuracy and depth of sentiment understanding.

Another important direction for future research is the extension of the current binary classification approach to multiclass sentiment classification, where reviews can be categorized into positive, negative, and neutral classes. This would provide a more nuanced understanding of customer opinions and enable more detailed analysis. Additionally, implementing aspect-based sentiment analysis would allow the model to identify sentiments associated with specific features such as food quality, service, pricing, and location, thereby offering more actionable insights for hotel management.

The development of a real-time sentiment analysis system is also a valuable enhancement. Integrating the model into a live dashboard would enable hotel administrators to monitor customer feedback dynamically and respond promptly to emerging issues. Furthermore, training the model on larger, more diverse, and multilingual datasets would improve its robustness, scalability, and applicability across different regions and user groups.

Finally, incorporating Explainable Artificial Intelligence (XAI) techniques can further strengthen the system by providing transparency into model predictions. This would help users understand how decisions are made, thereby increasing trust and reliability in automated sentiment analysis systems. Overall, these future directions can significantly enhance the practical applicability and performance of the proposed system in real-world scenarios.

REFERENCES

1. Z. Xiang and U. Gretzel, “Role of Social Media in Online Travel Information Search,” Tour. Manag., vol. 31, pp. 179–188, Mar. 2010, doi: 10.1016/j.tourman.2009.02.016.
2. S. W. Litvin, R. E. Goldsmith, and S. Litvin, “ELECTRONIC WORD-OF-MOUTH IN HOSPITALITY AND TOURISM MANAGEMENT.”
3. Q. Ye, R. Law, and B. Gu, “The impact of online user reviews on hotel room sales,” Int. J. Hosp. Manag., vol. 28, no. 1, pp. 180–182, Mar. 2009, doi: 10.1016/j.ijhm.2008.06.011.
4. J. A. Chevalier and D. Mayzlin, “THE EFFECT OF WORD OF MOUTH ON SALES: ONLINE BOOK REVIEWS,” 2003. [Online]. Available: http://www.nber.org/papers/w10148
5. C. Dellarocas, “MIT Sloan School of Management THE DIGITIZATION OF WORD-OF-MOUTH: PROMISE AND CHALLENGES OF ONLINE FEEDBACK MECHANISMS THE DIGITIZATION OF WORD-OF-MOUTH: PROMISE AND CHALLENGES OF ONLINE FEEDBACK MECHANISMS,” 2003. [Online]. Available: http://ssrn.com/abstract=393042
6. B. Liu, “Sentiment Analysis and Opinion Mining,” Morgan & Claypool Publishers, 2012.
7. J. Prager, “Open-domain question-answering,” Foundations and Trends in Information Retrieval, vol. 1, no. 2, pp. 91–233, 2006, doi: 10.1561/1500000001.
8. E. Cambria, B. Schuller, Y. Xia, and C. Havasi, “New avenues in opinion mining and sentiment analysis,” IEEE Intell. Syst., vol. 28, no. 2, pp. 15–21, 2013, doi: 10.1109/MIS.2013.30.
9. Sheffield and York, Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis. 2014.
10. M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, “Lexicon-Based Methods for Sentiment Analysis,” 2011.
11. B. Pang, L. Lee, and S. Vaithyanathan, “Thumbs up? Sentiment Classification using Machine Learning Techniques.” [Online]. Available: http://reviews.imdb.com/Reviews/
12. G. Salton and C. Buckley, “Term Weighting Approaches in Automatic Text Retrieval,” Ithaca, New York, USA, Nov. 1987.
13. Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” May 27, 2015, Nature Publishing Group. doi: 10.1038/nature14539.
14. I. Goodfellow, Y. Bengio, and A. Courville, “Deep Learning.”
15. J. L. Elman, “Finding Structure in Time,” Cogn. Sci., vol. 14, no. 2, pp. 179–211, Mar. 1990, doi: 10.1207/s15516709cog1402_1.
16. S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.
17. M. Schuster and K. K. Paliwal, “Bidirectional Recurrent Neural Networks,” IEEE Transactions on Signal Processing, vol. 45, no. 11, pp.

2673–2681, 1997, doi: 10.1109/78.650093.
18. X. Zhang and Y. LeCun, “Text Understanding from Scratch,” Apr. 2016, [Online]. Available: http://arxiv.org/abs/1502.01710
19. W. Medhat, A. Hassan, and H. Korashy, “Sentiment analysis algorithms and applications: A survey,” Ain Shams Engineering Journal, vol. 5, no. 4, pp. 1093–1113, Dec. 2014, doi: 10.1016/j.asej.2014.04.011.
20. T. Young, D. Hazarika, S. Poria, and E. Cambria, “Recent Trends in Deep Learning Based Natural Language Processing,” Nov. 2018, [Online]. Available: http://arxiv.org/abs/1708.02709
21. D. Buhalis and R. Law, “Tourism Management Reviews Progress in tourism management Twenty years on and 10 years after the Internet: The state of eTourism research.” [Online]. Available: http://www.bournemouth.ac.uk/services-management/
22. U. Gretzel and K. H. Yoo, “Use and Impact of Online Travel Reviews.”
23. B. Liu, Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers, 2012. doi: 10.2200/S00416ED1V01Y201204HLT016.
24. T. Hennig-Thurau, K. P. Gwinner, G. Walsh, and D. D. Gremler, “Electronic Word-of-Mouth via Consumer-Opinion Platforms: What Motivates Consumers to Articulate Themselves on the Internet?,” Journal of Interactive Marketing, vol. 18, no. 1, pp. 38–52, 2004, doi: 10.1002/dir.10073.
25. Q. Ye, Q. Li, and R. Law, “The Impact of Online User Reviews on Hotel Room Sales,” Int. J. Hosp. Manag., vol. 28, no. 1, pp. 180–182, 2009, doi: 10.1016/j.ijhm.2008.06.011.
26. R. Filieri and F. McLeay, “E-WOM and Accommodation: An Analysis of the Factors That Influence Travelers’ Adoption of Information from Online Reviews,” J. Travel Res., vol. 53, no. 1, pp. 44–57, 2014, doi: 10.1177/0047287513481274.
27. Z. Zhang, Q. Ye, and R. Law, “Determinants of hotel room price: An exploration of travelers’ hierarchy of accommodation needs,” Oct. 2011. doi: 10.1108/09596111111167551.
28. S. W. Litvin, R. E. Goldsmith, and S. Litvin, “ELECTRONIC WORD-OF-MOUTH IN HOSPITALITY AND TOURISM MANAGEMENT.”
29. 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, doi: 10.1561/1500000011.
30. T. Joachims, “Text Categorization with Support Vector Machines: Learning with Many Relevant Features,” in Proceedings of the European Conference on Machine Learning (ECML), Springer, 1998, pp. 137–142. doi: 10.1007/BFb0026683.
31. Y. Goldberg, Neural Network Methods for Natural Language Processing. in Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers, 2017. doi: 10.2200/S00762ED1V01Y201703HLT037.
32. S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
33. M. Schuster and K. K. Paliwal, “Bidirectional Recurrent Neural Networks,” IEEE Transactions on Signal Processing, vol. 45, no. 11, pp. 2673–2681, 1997, doi: 10.1109/78.650093.