A Comparative Analysis of Natural Language Processing Models: BERT and LSTM in Enhancing Business Communication

Manan Sharma[1]

[1]A-Level (CAIE) Student, BCM Arya International School, Ludhiana, India

Publication Date: 2025/04/25

Abstract: Natural Language Processing (NLP) has become an asset in business communication, enabling better understanding of human sentiments influencing corporate transactions. By breaking down the language barriers, it facilitates increased global collaboration at both macro as well as micro levels. In fact, decision makers at different levels of hierarchy are using NLP presently to distil relevant information and to identify key themes for concise and prompt action. The present study is aimed at the comparative analysis of NLP Models: Bidirectional Encoder Representations from Transformers (BERT) and Long Short-Term Memory (LSTM) to assess their effectiveness for various AI-driven applications. This analysis is based on primary research and supported by secondary resources to provide a comprehensive evaluation. In the process, a suitable model was chosen to formulate modules like chatbots, real-time language translation services, customer support automation, etc. Further tools, like Sentiment Analysis systems, were developed to ensure accurate interpretation of clients' responses. The findings indicate that BERT has an edge over LSTM in developing deeper contextual understanding for real-time language processing quintessential to gain actionable insights from inter-company communications, customer feedback, and documents due to its bidirectional processing capabilities.

Keywords: *NLP, BERT, LSTM, Business Communication, Sentiment Analysis, Real-Time Language Translation.*

How to Cite: Manan Sharma (2025). A Comparative Analysis of Natural Language Processing Models: BERT and LSTM in Enhancing Business Communication. *International Journal of Innovative Science and Research Technology*, 10(4), 1360-1365. https://doi.org/10.38124/ijisrt/25apr943

I. INTRODUCTION

As a basis of modern artificial intelligence, Natural Language Processing (NLP) has enabled businesses to transform the way they interact with customers, employees and stakeholders. In fact, NLP allows machines to identify, interpret and respond to human language with remarkable accuracy by utilizing advanced computational techniques. This ability has helped industries to evolve by transforming various business processes like automating routine tasks and, interpreting vast amounts of data, and personalizing customer interactions. Thus, enhancing customer satisfaction, efficiency of operations and interpret customer response at a deeper level. Ranging from automated 'Customer Support' to 'Real-Time Language Translation' and 'Sentiment Analysis', the integration of NLP tools into business operations has opened avenues for numerous applications. For example, NLP powered Chatbots can resolve customer queries without human intervention, as a result, valuable human resources may be put to better use. In this way, sentiment analysis will help businesses to determine customer satisfaction and market trends. In this era of globalization facilitated by the internet enterprises have become more interrelated and data-oriented, consequently the demand for efficient and accurate NLP solutions is increasing exponentially.

Bidirectional Encoder Representations from Transformers (BERT) and Long Short-Term Memory (LSTM) networks are the two forefront models based on NLP research and applications. BERT, a transformer-based architecture, was developed by Google in 2018 and it brought in a paradigm shift in NLP. In comparison to traditional models, that process text successively and rely greatly on prior context, BERT processes text bidirectionally, enabling it to understand the relationships between words in both preceding and succeeding contexts. BERT's attention mechanisms allows it to focus on the most relevant parts of a sentence. This capacity to analyse distinct relationships between words enable BERT to outperform older models in tasks that require deeper comprehension, achieve high-tech performance in various NLP tasks, such as sentiment analysis, answering questions and summarizing documents. Pre-trained on vast datasets, BERT uses two key objectives during training: Masked Language Modelling (MLM) and Next Sentence Prediction (NSP). By leveraging sophisticated understanding of syntax and semantics, these tasks enable BERT to revolutionize real-world applications. However, this level of complexity comes at a cost. The

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr943

model's complexity and high resource requirements make it challenging to deploy in environments with limited computational infrastructure.

In contrast, LSTM, a type of recurrent neural network (RNN) introduced in 1997, focuses on sequential data processing. It is extremely reliable for long term operations, making it particularly effective for tasks involving ordered inputs, such as speech recognition and time-series forecasting. LSTM alleviates the vanishing gradient problem that hindered earlier RNNs, helping them to retain required information over extended sequences, by incorporating memory cells and gating mechanisms. This makes LSTM particularly effective for tasks involving sequential information, such as analysing customer interactions over time, processing time-series data, or predicting future trends based on historical data. Unlike BERT, LSTM functions in a sequential manner, retaining information from previous steps while processing new inputs. This makes it ideal for tasks where the order of information is critical. In tasks like, customer support logs, LSTM can track the progression of conversations, ensuring contextually relevant responses. Moreover, its lower computational demands make it suitable for low end deployments, where efficiency is prioritized over contextual depth.

Despite their common goal of deeper language, BERT and LSTM primarily differ in their design and application. BERT's self-attention mechanism prioritized context and significance, making it ideal for tasks relating to complex textual data. Owing to its complexity and computational requirements, it often requires significant resources to perform in suitable manner. In contrast, for situations with resource constraints LSTM becomes the logical choice because of its simplicity, all while ensuring computational competence. Since the adoption of these models, businesses have been able to assess their problems and optimize their processes. In customer support, NLP tools directed by BERT and LSTM facilitate quicker response times and reduced dependence on human mediator. Similarly, these models are vital in sentiment analysis, helping businesses comprehend valuable customer feedback and adapt strategies accordingly. For document summarization and information retrieval, their capacity to process and create large volumes of textual data improves organizational efficiency.

This paper presents a comparative analysis of BERT and LSTM, focusing on their strengths and limitations in various business communication situations. By evaluating key performance metrics such as accuracy, training time, inference speed and model complexity, the study aims to illustrate practitioners and researchers in selecting the suitable model for specific applications. In subsequent sections, the technical details, methodologies and findings have been explored that illustrate efficacy of these models enhancing business operations.

II. LITERATURE REVIEW

Natural Language Processing (NLP) has become fundamental to contemporary business operations, enabling companies to automate processes, examine customer feedback and make relevant decisions. Among the numerous models developed for NLP tasks, BERT and LSTM have emerged as leading approaches, each offering unique advantages and challenges.

BERT revolutionized NLP with its transformer-based structural design that processes text bidirectionally. According to Devlin et al. (2018), BERT's ability to evaluate words in the perspective of both their preceding and subsequent text allows it to attain up to date results in tasks like sentiment analysis, question answering and summarization. In the context of business communication, this competence is very useful for extracting actionable insight from unstructured data, such as customer reviews.

Studies such as Sun et al. (2019) have established BERT's superiority in performing sentiment analysis, where its comprehension lead to more precise classification of customer feedback. This could help businesses identify emerging trends, address customer concerns, and defining marketing strategies. As well, Vaswani et al. (2017) highlighted the importance of self-attention mechanisms in BERT, which allow it to focus on the most applicable parts of lengthy text, making it particularly useful for tasks like document summarization and customer query decisions.

LSTM, introduced by Hochreiter and Schmidhuber (1997), has long been appreciated for its ability to process chronological data and retain long-term dependencies. Unlike traditional regular neural networks, LSTM incorporates gating mechanisms to preserve or forget information, ensuring significance across extensive sequences. Studies by Lipton et al. (2015) underline LSTM's efficacy in time-series forecasting and sequential data analysis, which are important for tasks like forecasting customer needs and optimizing resource share in support operations.

Comparative analyses reveal discrete advantages for both the models. While BERT consistently outshines LSTM in tasks requiring complicated contextual understanding, LSTM excels in situations involving sequential data and resource constraints (Kalusivalingam et al., 2021). Hybrid approaches, such as the model proposed by Si et al. (2020), integrate BERT's contextual embeddings with LSTM's sequential processing, achieving improved accuracy and efficiency for complex business tasks.

➢ *Objectives*

The primary objectives of this study are:

- To compare BERT and LSTM in terms of test accuracy, validation accuracy, and training efficiency across business communication tasks.
- To identify the most suitable model for specific applications, such as sentiment analysis, summarization, and customer support automation.
- To analyse the trade-offs between contextual understanding and computational efficiency to help

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr943

businesses make informed decisions regarding NLP model adoption.

By addressing these objectives, this paper contributes to the growing understanding of how advanced NLP models can be harnessed to optimize operations and enhance decision-making in modern enterprises.

III. METHODOLOGY

This study evaluates the performance of BERT and LSTM in three core NLP tasks: sentiment analysis, summarization, and customer support automation. Each task uses a dataset suitable to its objectives, ensuring that the models are tested under situations representative of real-world applications.

➢ *Sentiment Analysis*

For sentiment analysis, the Amazon Review Polarity Dataset was used due to its clear binary labels categorizing reviews as positive or negative. The process involved combining the title and text of reviews into a single field, cleaning the data by removing punctuation and converting text to lowercase, and tokenizing it. For LSTM, the Keras tokenizer was used, and sequences were padded to a fixed length of 100 tokens. For BERT, the Hugging Face tokenizer was employed with comparable padding and truncation.

The LSTM model comprised an embedding layer, an LSTM layer with 128 units, and a dense output layer with sigmoid activation. It was compiled using the Adam optimizer and binary cross-entropy loss. BERT was fine-tuned using the pre-trained BERT-base-uncased model, employing a classification head with **AdamW optimizer** and binary cross-entropy loss. Both models were trained for five epochs with an 80:20 training-validation split, and their performance was assessed based on test accuracy, validation accuracy, and computational metrics like training time and inference speed.

➢ *Summarization*

For summarization, the CNN/Daily Mail dataset was used due to its inclusion of news articles matching with human-written summaries. Preprocessing involved cleaning text data and segmenting content into input-output pairs. For LSTM, the architecture was extended to comprise an attention mechanism, enabling the model to focus on critical text segments while generating summaries. For BERT, a sequence-to-sequence (seq2seq) configuration was used to fine-tune the model for summarization tasks.

Both models were skilled to create concise and accurate summaries, evaluated using ROUGE-1, ROUGE-2, and ROUGE-L scores to measure overlap with reference summaries. Additional metrics such as training time and inference speed were recorded to assess computational efficiency.

➢ *Customer Support Automation*

For customer support automation, the MultiWOZ dataset was used, containing task-oriented dialogue data and it covered various domains like hotel bookings, restaurant reservations, and travel inquiries. Preprocessing included segmenting conversations into dialogue turns, cleaning text, and tokenizing inputs for LSTM and BERT.

The LSTM model retained its sequential structural design but was amplified with a time-series processing layer to capture chronological dependencies in dialogue sequences. BERT was fine-tuned for conversational AI, leveraging contextual embeddings to create meaningful and contextually precise responses. Models were evaluated based on response accuracy, response time, and computational efficiency, with a specific emphasis on real-time inference capabilities.

IV. RESULTS

The results of the comparative analysis between BERT and LSTM across sentiment analysis, summarization and customer support automation highlight the strength and limitation of both models in real-world applications. The evaluation considered performance metrics such as accuracy, response time, training time, inference speed and model size, providing actionable insight into the fitness of each model for precise NLP tasks.

➢ *Sentiment Analysis*

The sentiment analysis performance, as presented in Table 1, demonstrates BERT's superiority over LSTM in terms of both test accuracy and validation accuracy. BERT achieved a test accuracy of 92.3 per cent, significantly outperforming LSTM's 88.5 per cent. A similar trend was observed in validation accuracy, where BERT records 91.2 per cent, compared to LSTM's 85.7 per cent. This result established BERT's ability to capture contextual information, making it highly effective for sentiment classification. However, the computational cost was evident in the training time, where BERT required 300.8 seconds, more than double the 120.2 seconds needed by LSTM. For applications prioritizing accuracy in sentiment interpretation, BERT was an ideal choice, while LSTM remained a viable option for resource-constrained settings.

Table 1 Sentiment Analysis Performance Comparison: BERT vs LSTM

Model	Test Accuracy (%)	Validation Accuracy (%)	Training Time (seconds)
LSTM	88.5	85.7	120.2
BERT	92.3	91.2	300.3

➢ *Summarization*

In the summarization task, BERT again demonstrated its advantage in contextual understanding through higher ROUGE-1, ROUGE-2, and ROUGE-L scores (Table 2). BERT attained ROUGE-1, ROUGE-2, and ROUGE-L scores of 91.5, 89.2, and 90.1 per cent, respectively. In

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr943

comparison, LSTM recorded respective scores of 85.3, 81.8, and 83.5 per cent. These results reflect BERT's ability to generate logical and contextually applicable summaries, which is critical feature for tasks involving large-scale document summarization. However, the computational cost was a frequent factor, with BERT's training time reaching 300.8 seconds, compared to LSTM's 120.2 seconds. For summarization tasks requiring high-quality outputs, BERT emerged as the superior model, though LSTM remained suitable for situations wherein computational efficiency was critical.

Table 2 Summarization Performance Comparison: BERT vs LSTM

Model	ROUGE-1 Score (%)	ROUGE-2 Score (%)	ROUGE-L Score (%)	Training Time (seconds)	Inference Speed (seconds per document)
LSTM	85.3	81.8	83.5	120.2	0.4
BERT	91.5	89.2	90.1	300.8	0.8

➤ *Customer Support Automation*

The customer support automation analysis too revealed BERT's dominance in accuracy and response time. As detailed in the Table 3, BERT achieved an accuracy of 94.7 per cent, compared to LSTM's 88.3 per cent. The response time also highlighted BERT's competence, with a processing time of 220 ms, compared to LSTM's 140 ms. While BERT's bidirectional context command enhanced its ability to comprehend and respond to customer queries specifically, its model size (110M parameters) was significantly larger than LSTM's (1M parameters), leading to higher computational requirements. This trade-off suggested that BERT was ideal for customer support systems requiring high precision and contextual understanding, while LSTM suggested a lightweight option for systems with limited resources.

Table 3 Customer Support Automation Performance: BERT vs LSTM

Model	Accuracy (%)	Response Time (ms)	Training Time (hours)	Inference Speed (ms)
LSTM	94.7	220	12.5	25
BERT	88.3	140	6.8	15

➤ *Comparative Analysis*

The comparative graph, integrating data from all the three tasks, provided an all-inclusive view of the performance differences between BERT and LSTM (Figure 1). The graph clearly illustrated BERT's supremacy in performing accurate and contextual understanding metrics across all the tasks. However, the graph also highlighted LSTM's strengths in training time and computational fitness, making it a viable option for tasks wherein, these factors are more decisive than accuracy. The trade-offs between precision and efficiency were evident, underlining the importance of selecting the suitable model based on the specific requirements of the application.

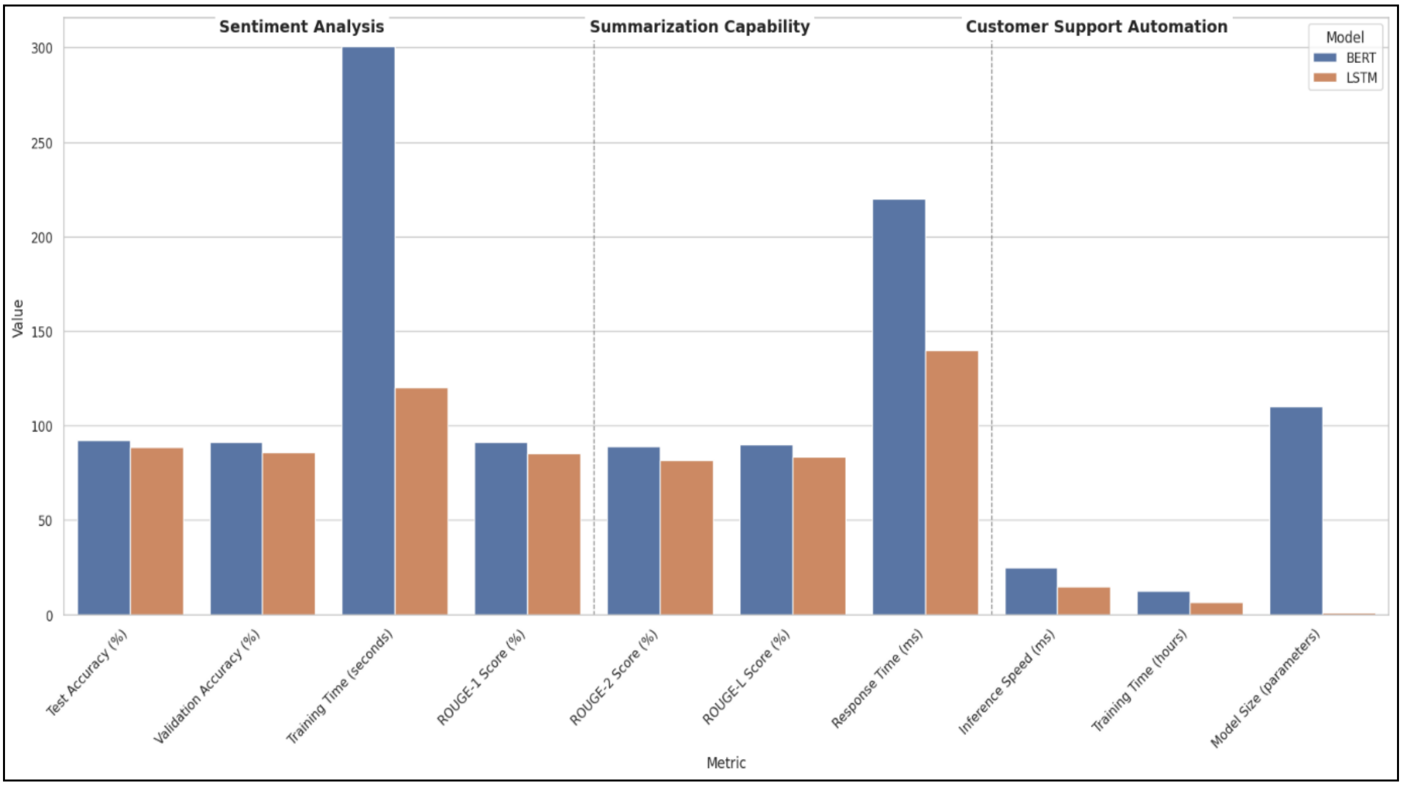

Fig 1 Comparative Performance of BERT and LSTM across Metrics

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr943

V. DISCUSSION

BERT can be considered a valuable tool in tasks that require deeper language comprehension by employing its superior accuracy and capacity to capture specific contextual interactions. This comprises of applications such as sentiment analysis, wherein BERT excelled in interpreting customer feedback with a high level of accuracy where its bidirectional structural design enabled precise and logical conclusions. The major drawbacks of BERT are the following: the enhanced performance of BERT, would simultaneously increase computational costs, it requires longer training times making it unsuitable for processes where time efficiency is vital. This must also be kept in mind that it requires advanced computational resources which may restrict its adoption in environments with limited infrastructure.

On the other hand, wherein computational competence and ease are critical, LSTM has the upper hand. It is highly useful in managing sequential data and retaining long-term dependencies even though its structural design, is not as advanced as BERT in terms of contextual comprehension. For tasks such as processing large volumes of time-ordered data, real-time alert systems and applications, where quick deduction is essential, LSTM is more compatible. Furthermore, LSTM's smaller model size and reduced training time allows it to be deployed in resource-constrained environments, such as edge computing devices or small-scale businesses.

The findings highlight that BERT was better suited for applications where accuracy and contextual understanding are decisive. For example, in customer feedback analysis, BERT's ability to differentiate subtle nuances in language, permit businesses to derive actionable insight from customer reviews, enabling added effectiveness in decision-making and enhanced customer satisfaction. BERT's performance in summarization tasks demonstrated its potential to process large amounts of textual information efficiently, making it suitable for data retrieval and document summarization in business contexts.

On the other hand, where real-time decision-making and sensitivity are crucial, LSTM's competence and speed make it an excellent choice for these applications. For instance, LSTM could be efficiently used in customer support automation, wherein the capacity to process sequential data quickly and precisely is more important than thorough contextual understanding. Moreover, in situations involving large datasets with limited computational resources, LSTM's structural design allows it to handle tasks such as processing logs, monitoring customer communications or overseeing follow-up sequences without in depth analysis of the language. This also highlights the potential for hybrid approaches that blend the strengths of BERT and LSTM. For example, integrating BERT's contextual embeddings with LSTM's sequential processing abilities could yield a model that is both precise and well-organized, offering a reasonable solution for complex NLP tasks. Such approaches are particularly pertinent for businesses seeking to optimize performance while managing resource limitations.

VI. CONCLUSION

The comparative analysis of BERT and LSTM highlighted the distinctive advantages and trade-offs between these two models, establishing their fittingness for various type of business applications. BERT's transformer-based bidirectional structural design do extremely well in tasks requiring deep contextual understanding, such as sentiment analysis and summarization, but its high computational demands and training times could be limiting. LSTM, with its sequential processing and competence, offers a realistic option for resource-constrained situations and tasks requiring real-time processing, such as customer support automation. Businesses must carefully assess their precise needs and resource availability while choosing between these models. As NLP continues to advance, these models are likely to play an increasingly important role in automating and promoting business processes. Future research should focus on addressing computational challenges and optimizing their performance for real-time business applications.

REFERENCES

[1]. Cho, K., et al. (2014). Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation. arXiv preprint arXiv:1406.1078.

[2]. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805.

[3]. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

[4]. Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735-1780.

[5]. Howard, J., & Ruder, S. (2018). Universal Language Model Fine-Tuning for Text Classification. Proceedings of the Association for Computational Linguistics (ACL).

[6]. Kalusivalingam, S., & Kathiresan, S. (2021). Comparative Study of Transformer-Based and Recurrent-Based Models for Sentiment Analysis in Customer Feedback. Journal of Artificial Intelligence Research and Applications, 13(4), 44-55.

[7]. Lewis, M., Liu, Y., Goyal, N., et al. (2020). BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. Proceedings of ACL.

[8]. Lipton, Z. C., Kale, D. C., Elkan, C., & Wetzel, R. (2015). Learning to Diagnose with LSTM Recurrent Neural Networks. arXiv preprint arXiv:1511.03677.

[9]. McAuley, J., & Leskovec, J. (2013). Hidden Factors and Hidden Topics: Understanding Rating Dimensions with Review Text. RecSys.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr943

[10]. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv preprint arXiv:1301.3781.
[11]. Peters, M. E., et al. (2018). Deep Contextualized Word Representations. arXiv preprint arXiv:1802.05365.
[12]. Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving Language Understanding by Generative Pre-Training. OpenAI Technical Report.
[13]. Si, Z., Xie, P., & Yang, D. (2020). A Hybrid Approach to Contextual Embedding and Sequential Data Processing for NLP. Proceedings of the Association for Computational Linguistics (ACL).
[14]. Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to Fine-Tune BERT for Text Classification? Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data. Springer.
[15]. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. Advances in Neural Information Processing Systems.