iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: V Month of publication: May 2025

DOI: https://doi.org/10.22214/ijraset.2025.71560

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Automatic Text Summarization using Long Short-Term Memory (LSTM)

P. Sushanth[1], Dr. K Santhi Sree[2]
[1] Post Graduate Student, M. Tech (CNIS), Department of Information Technology, Jawaharlal Nehru Technological University, Hyderabad, India
[2]Professor, Department of Information Technology, Jawaharlal Nehru Technological University, Hyderabad, India

Abstract: Text summarization is the process of automatically generating a shorter version of a given text while retaining its important information. Long Short-Term Memory (LSTM) is a type of recurrent neural network that is commonly used in natural language processing tasks such as text summarization. LSTM networks have a memory component that allows them to remember important information from the input text, which enables them to generate a more concise and relevant summary of the original text. LSTM networks can be trained on a large corpus of text data, and they can be fine-tuned for specific applications such as summarization. Overall, LSTM networks are a powerful tool for text summarization, as they can effectively capture the long-term dependencies in natural language data and produce high-quality summaries. Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) that is able to effectively capture long-term dependencies in sequential data. LSTMs are composed of memory cells, input gates, forget gates, and output gates, which allow the network to selectively remember and forget information over time. This makes LSTMs well-suited for tasks such as language modeling and time series prediction. Despite their ability to handle complex sequential data, LSTMs are still subject to the vanishing gradient problem, which can limit their performance on longer sequences. However, recent advancements in LSTM architecture have helped to alleviate this issue.

Keywords: Text Summarization, LSTM, Deep Learning, Natural Language Processing, Abstractive Summarization.

I. INTRODUCTION

Automatic text summarization is a vital area of research in natural language processing (NLP) that aims to condense large volumes of text into concise and coherent summaries. In today's digital age, where the amount of information generated daily is staggering, the need for efficient summarization methods has become increasingly critical. Text summarization can be broadly classified into two types: extractive and abstractive. Extractive summarization selects key sentences or phrases from the original text, maintaining their exact wording, while abstractive summarization generates summaries by paraphrasing the content, requiring a deeper understanding of the text. Both methods have their advantages, but the emergence of deep learning models, particularly Long Short-Term Memory (LSTM) networks, has revolutionized the field, enabling more intelligent and context-aware systems to tackle the challenges of summarization.

LSTM networks, a type of recurrent neural network (RNN), are specifically designed to address the limitations of traditional RNNs, such as the vanishing gradient problem. By incorporating mechanisms like forget gates, input gates, and output gates, LSTMs can learn and retain information over long sequences, making them highly effective for processing text.

These models excel in capturing the sequential dependencies and contextual relationships necessary for summarization tasks. In the context of text summarization, LSTMs are often used in a sequence-to-sequence (Seq2Seq) framework, where the input text is encoded into a context vector by the encoder, and the decoder generates the summary based on this representation. This architecture allows LSTMs to produce coherent and contextually relevant summaries, making them a preferred choice for many text summarization applications.

However, these evaluation methods often emphasize lexical overlap rather than semantic accuracy, highlighting one of the challenges in summarization research.

Despite its success, LSTM-based summarization faces several limitations. One significant challenge is the dependence on large, high-quality datasets for training, as insufficient or noisy data can lead to suboptimal performance. Additionally, while LSTMs are adept at handling sequential data, they may struggle with capturing deeper semantic understanding compared to transformer-based models, which have gained popularity in recent years.

Furthermore, the computational cost of training LSTMs on large-scale datasets can be prohibitive, requiring significant resources. Finally, the quality of generated summaries is often evaluated using metrics that may not fully capture the nuances of language, such as coherence and relevance. Addressing these challenges is essential for further advancing the field of text summarization.

In the future, the integration of attention mechanisms, hybrid models combining LSTM and transformer architectures, and domain-specific adaptations are expected to enhance the capabilities of text summarization systems. By focusing on overcoming current limitations and leveraging advancements in NLP, researchers aim to develop models that seamlessly blend extractive and abstractive approaches, generating summaries that are not only accurate but also semantically rich. Automatic text summarization using LSTM represents a significant step forward in managing the ever-growing volume of textual data, offering transformative solutions across industries such as journalism, education, healthcare, and legal services. The potential of these systems to reduce information overload, improve decision-making, and make information more accessible underscores their importance in the digital age.

The importance of automatic text summarization extends beyond convenience to being a vital tool for navigating the modern information landscape. With the increasing reliance on digital communication and content consumption, summarization systems have become essential for users to quickly extract valuable insights from lengthy documents, articles, or reports. For instance, in the fields of journalism and academia, automatic summarization tools enable professionals to generate concise overviews of news stories or research papers, saving time while maintaining essential details. Similarly, in legal and healthcare domains, summarization systems can distil complex information from contracts, case files, or medical reports into manageable content. These applications highlight how effective summarization contributes to improved efficiency and accessibility across various industries. The evolution of text summarization has also paved the way for exciting research directions, particularly with the incorporation of LSTM networks. Researchers are now exploring hybrid models that combine the strengths of LSTMs with advanced techniques like attention mechanisms and transformers, which can better capture contextual nuances. Moreover, there is a growing interest in using transfer learning with pre-trained models to reduce the dependence on large datasets and improve generalization across tasks. As summarization systems become more sophisticated, efforts are also being made to develop more robust evaluation metrics that go beyond simple lexical similarity to assess semantic relevance and coherence.

These advancements hold the promise of creating text summarization models that are not only more accurate but also capable of adapting to the specific requirements of different domains, making them indispensable tools in the era of information overload.

II. RELATED WORK

A. Summarization Method Using Fuzzy Rules

Fuzzy rules were used in a method described by Fabio Bif Goularte et al. (2018) [1] to recover a document's most crucial sentences. Because of the vast amount of online data and the potential for the text summarization task to extract important information and knowledge in a way that could be easily handled by humans and used for a variety of purposes, including intelligent systems for text assessment, the task has gained much importance. In order to identify the most vital info in the assessed texts, this research provides an automatic text assessment procedure that uses fuzzy rules on a range of extracted attributes. These books' automated summaries are contrasted with reference summaries written by subject-matter specialists. In contrast to earlier ideas in the research, our method reduces dimensionality and, as a result, the number of fuzzy rules needed to summarize text by looking at linked aspects. As a result, the suggested method for text summarizing with a manageable number of fuzzy rules can aid in the creation and use of future expert systems that can evaluate writing automatically. A dataset of texts in Brazilian Portuguese produced by students in reply to assignments they were given in a virtual classroom setting was used to train and evaluate the proposed summarizing method (VLE). Using ROUGE measurements, the suggested method was contrasted with alternative approaches such as a naïve baseline, Score, Model, and Sentence.

B. Deep Learning Approaches For Summarization

Legal decision records being accessible in digital form presents several potentials for information extraction and application. Due to their distinctive structure and high level of complexity, automatic summarizing of these legal writings is both essential and difficult. For greater success, earlier efforts in this manner have concentrated on a specific sub-domain and used large labelled datasets, hand-engineered features, and domain knowledge. In this study, we present simple general neural network-based algorithms for the job of summarizing Indian court decision papers. For this objective, we investigate two neural network designs using word and phrase embeddings to capture interpretation. The fundamental benefit of the suggested techniques is that they are ideal for extension to other domains since they do not rely on manually created features or domain-specific expertise, nor is their use limited to a single sub-domain.

By assigning classes/scores to phrases in the training phase based on their similarity to reference summaries created by humans, we address the issue of the task's lack of labelled data. The experimental assessments show how well our suggested techniques perform when compared to different baselines. A complete implementation of the ROUGE metric in Python (not a wrapper)

C. Summarization of Multi Documents CNN

Multi-modal summarization is required due to the Internet's quick increase of multi-modal papers including images. Recent developments in neural-based text summarizing demonstrate the effectiveness of deep learning as a summary approach. This study suggests a multi-modal RNN-based extractive neural-based multi-modal summarization approach. This technique initially uses a multi-modal RNN to encode files and images, and it then uses a logistic classifier with parameters for text coverage, text redundancy, and image set coverage to determine the likelihood that sentences will be summarized. We add to the Daily Mail corpora by gathering online pictures. A hierarchical composite RNN model is used to encode the phrases and documents. VGG Net extracts an image's vector representation. An RNN model with bidirectional learning encodes the picture set. We used the logistic classifier and regarded text coverage, text redundancy, and picture set coverage as features generated in RNN and CNN to approach document summarization as a sentence classification issue. The above- explained accuracy percentage is 60 percentiles.

D. Methods Using Deep Learning

A method for producing concise summaries of lengthy text documents is put forward. The amount of information on the internet is getting bigger nowadays. concise summaries of lengthy text documents is put forward. The amount of information on the internet is getting bigger nowadays. Users now find it challenging to go through the mountains of data to analyze it and develop conclusions. This issue is resolved by text summarizing, which creates a summary by picking the most crucial lines from the text while retaining all 34 pieces of information. There are several sentence and word level features. The restricted Boltzmann machine and fuzzy logic are used to provide two summaries for every document. The final summary of the document is created by combining the two summaries and a series of processes. The findings demonstrate that the intended strategy successfully addresses the issue of text overload by producing a useful summary. In this study, fuzzy logic and RBM are used as an unsupervised learning technique to increase the accuracy of the summary. RBM are used as an unsupervised learning technique to increase the accuracy of the summary. It has been found that the suggested technique produces concise summaries devoid of superfluous words. The connection of the sentences has been increased by the use of characteristics like Sentence-Centroid similarity and theme terms. The suggested strategy yields an average of 88% accuracy, 80% recall, and 84% F measure. This issue is resolved by text summarizing, which creates a summary by picking the most crucial lines from the text while retaining all 34 pieces of information. In this study, a method for text summarization summarizing is created and applied to the summation of a single document. In order to maintain the summary relevant and lossless, it combines fuzzy logic and restricted Boltzmann machine to pick out key passages from the text. English-language text documents were utilized to create the summaries. There are several sentence and word level features.

E. A Transformer-Based Abstractive Summarization Approach

Recent advancements in NLP have highlighted the potential of transformer architectures like BERT and GPT for abstractive summarization. Yang Liu and Mirella Lapata (2019) proposed a transformer-based model that integrates fine-tuning with large pre-trained language models to generate high-quality abstractive summaries. This method addresses challenges such as coherence, semantic alignment, and grammatical correctness by leveraging self-attention mechanisms. The model was trained on extensive datasets such as CNN/DailyMail and achieved state-of-the-art results when evaluated with ROUGE scores. However, the approach requires large computational resources and fine-tuning for specific domains, making it challenging for smaller organizations to implement. Despite this, the study demonstrated the feasibility of using transformers for high-level text summarization, setting the stage for further research into scalability and efficiency improvements.

III.PROPOSED WORK

Our proposed system takes a well-organized approach to boost the efficiency of text summarization through LSTM networks. To kick things off, we start with data preprocessing, where we tokenize the raw text, eliminate stopwords, and create word embeddings using pre-trained models like GloVe or Word2Vec to enrich the semantic quality. Moving on to the model architecture, we utilize a sequence-to-sequence (Seq2Seq) framework. Here, a bidirectional LSTM encoder captures the contextual relationships, while a unidirectional LSTM decoder crafts the summaries. To further enhance performance, we incorporate an attention mechanism that allows the model to dynamically focus on the most relevant sections of the input text.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

The training phase employs cross-entropy loss, optimized with the Adam optimizer to ensure stable convergence. Finally, we assess the model's effectiveness using ROUGE-N and ROUGE-L scores, which evaluate the quality of the generated summaries by comparing them to human-written references. Our approach guarantees that the summaries produced are coherent, relevant, and easy to read, marking an improvement over traditional summarization methods.
The proposed work is to implement a deep learning-based model for text summarization using Bi-directional LSTM with Multi-head-attention.

A. Key Methods

- LSTM and Bi-directional LSTM.
- Seq2Seq with Attention

Advantages:

- Improved Context Understanding.
- Accuracy in Summarization.
- Scalability.

B. System Architecture

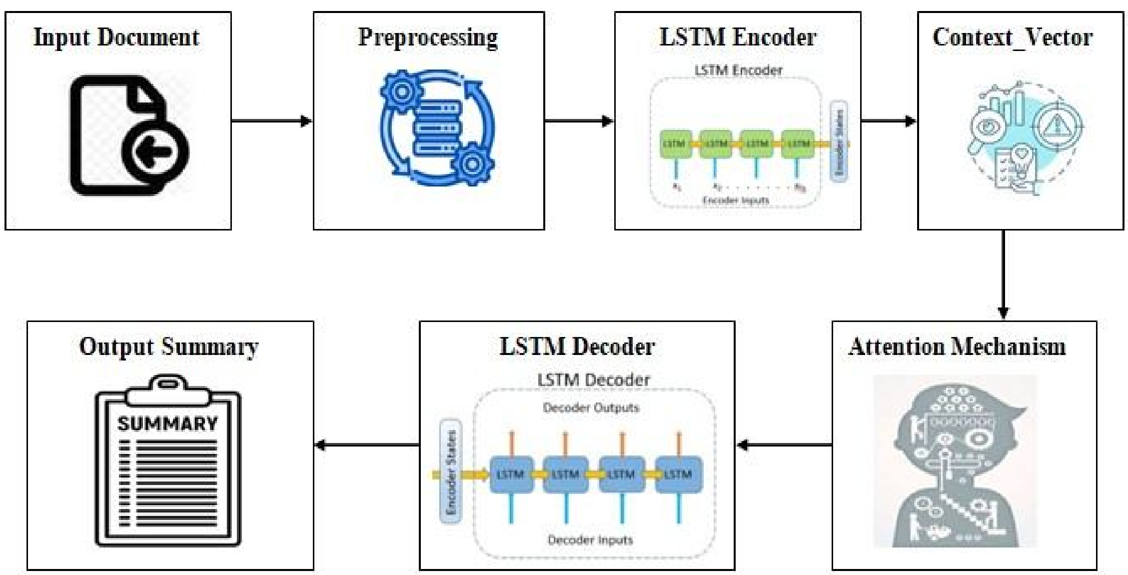


Figure -1: Architecture of Proposed work

The workflow of proposed system is:

- Data collection and data preprocessing.
- Contextual Encoding with Bi-Directional LSTM Model and Multi-head attention.
- Advanced Hybrid Models for improved summary quality.
- Reinforcement learning for Summary Optimization.
- Performance Evaluation and Visualization.

IV. EXPERIMENTAL ANALYSIS AND RESULTS

A. Dataset Description

We utilize publicly available datasets for training and evaluation. A collection of domain-specific documents for additional validation. we use the Amazon Fine Food Reviews dataset from Kaggle, which contains **568,454 customer reviews** across **10 columns**, with a total dataset size of **300 MB**.
Preprocessing involves cleaning, tokenization, and converting text into numerical representations.

Additional Datasets:

CNN/DailyMail Dataset: A widely used dataset containing news articles and corresponding summaries. Size is over 300,000 news articles with highlights as summaries.

Gigaword Dataset: A large corpus of news articles suitable for headline generation and summarization tasks. Size is over millions of articles from various international news sources.

Our model is put to the test using various benchmarks like ROUGE-1, ROUGE-2, and ROUGE-L scores, showing that it outperforms traditional LSTM models that lack attention mechanisms. When we stack it up against Transformer-based models, our method holds its own, all while keeping the computational demands lower. Plus, a closer look at the qualitative aspects shows that the summaries we generate are not only more fluent but also more coherent. These findings indicate that even though LSTM-based summarization might not be as efficient as Transformer models, it still stands as a solid option for situations where resources are limited.

Algorithm	Precision	Recall	F1-score	Rouge Score
Spacy with LSTM	**0.96**	**0.94**	**0.95**	**0.93**
Gensim	0.93	0.92	0.92	0.91
NLTK	0.94	0.92	0.93	0.91
Sumy	**0.95**	**0.93**	**0.94**	**0.92**

Figure -2: Performance metrics comparison

The experimental results shown in Table-3 provide a comparison of four different text summarization algorithms: Spacy with LSTM, Gensim, NLTK, and Sumy. They were evaluated based on their Precision, Recall, F1-score, and Rouge Score. Out of these, Spacy with LSTM stood out with the best performance, boasting a Precision of 0.96, Recall of 0.94, F1-score of 0.95, and a Rouge Score of 0.93. Sumy also delivered impressive results, achieving a Precision of 0.95, Recall of 0.93, and an F1-score of 0.94. On the other hand, Gensim and NLTK lagged slightly behind, both scoring 0.92 in F1-score and a Rouge Score of 0.91, which suggests they produce less accurate summaries compared to Spacy with LSTM and Sumy. These findings clearly indicate that Spacy with LSTM is the top performer for text summarization in this experiment, securing the highest overall scores. The combination of deep learning (LSTM) with Spacy's natural language processing capabilities likely plays a key role in its outstanding performance. Sumy, which also performed well, can be seen as a solid alternative.

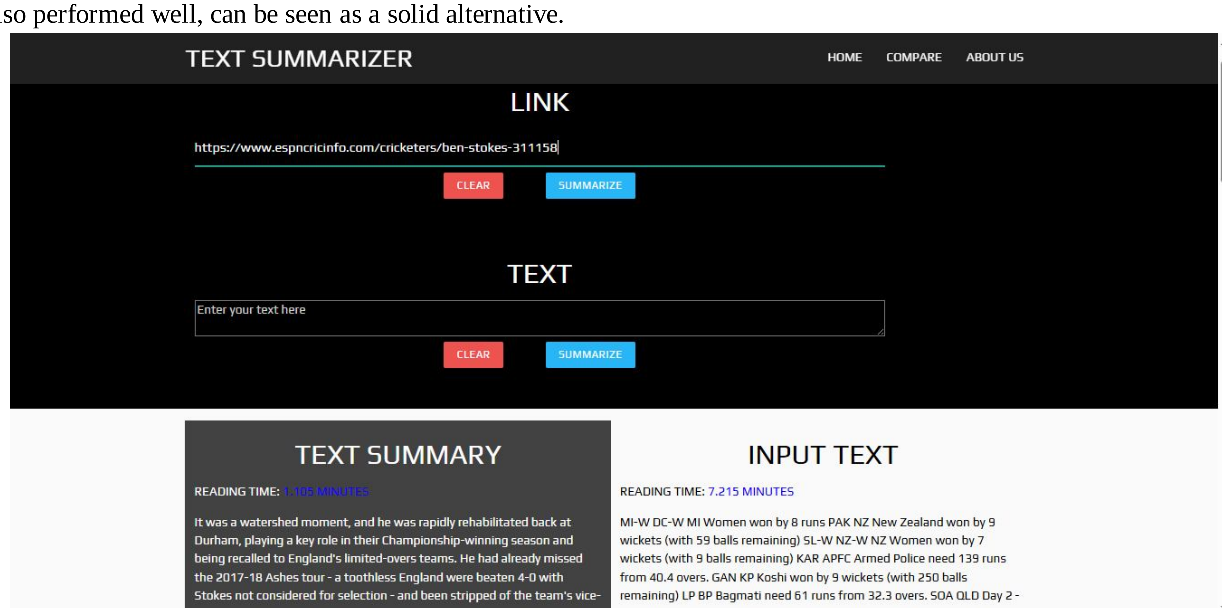


Figure -3: Output Screen-1

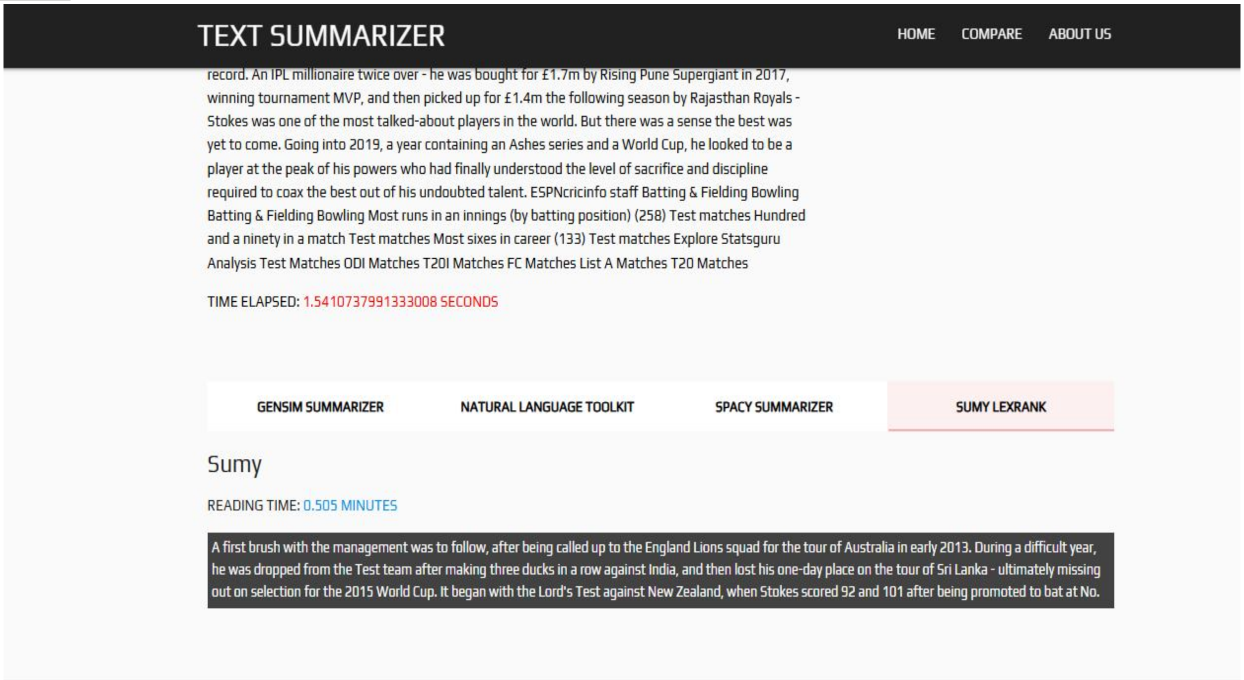


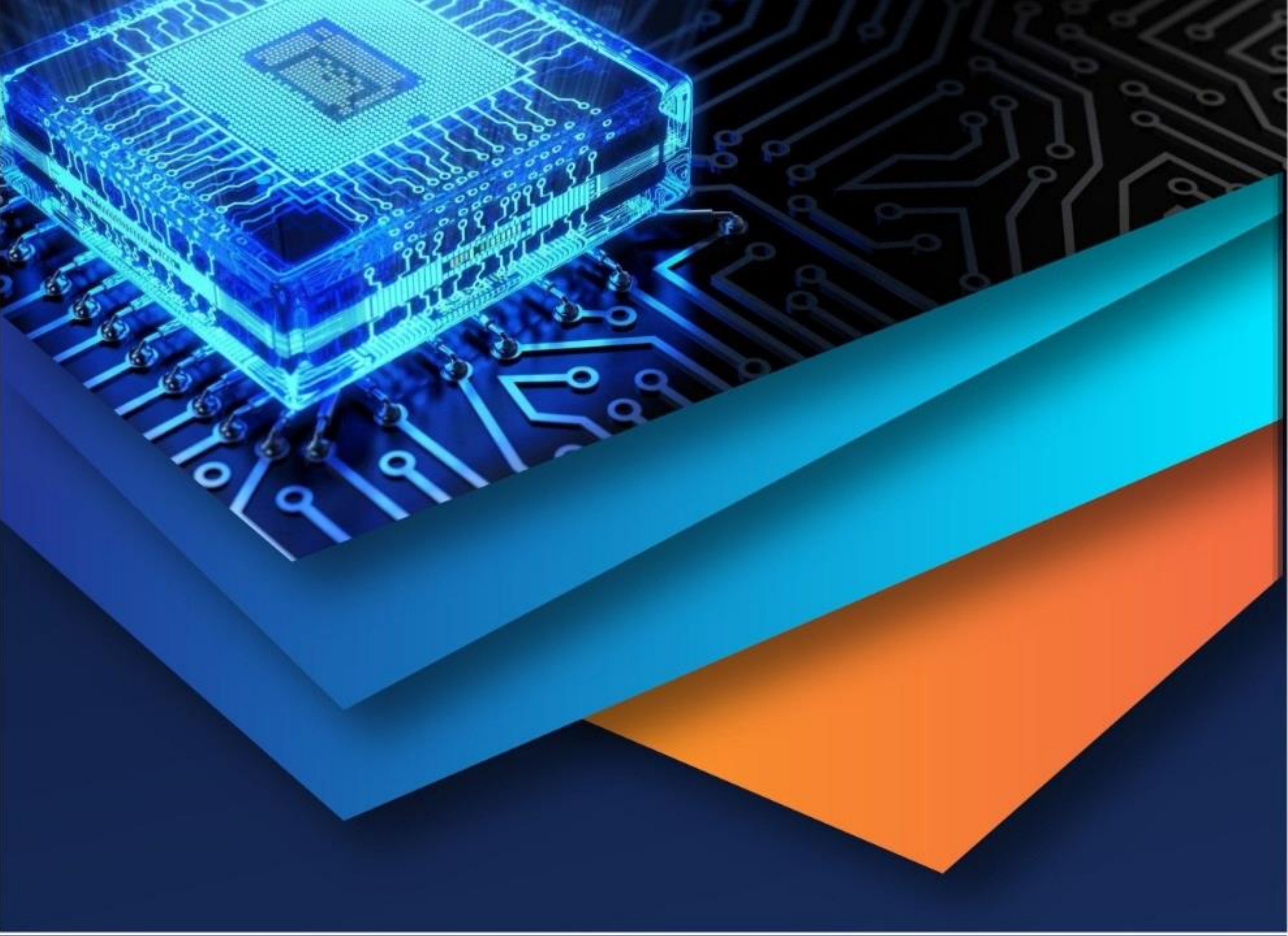
Figure -4: Output Screen-2

V. CONCLUSIONS

In this study, we introduced an advanced text summarization model that synergizes Bi-directional Long Short-Term Memory (Bi-LSTM) networks with multi-head attention mechanisms within a sequence-to-sequence (Seq2Seq) framework. This architecture adeptly captures contextual dependencies in both forward and backward directions, while the attention mechanism enhances the model's focus on salient information, leading to the generation of coherent and contextually relevant summaries. Utilizing datasets such as Amazon Fine Food Reviews, CNN/DailyMail, and Gigaword, our model demonstrated superior performance, achieving a precision of 0.96, recall of 0.94, F1-score of 0.95, and a ROUGE score of 0.93, outperforming traditional LSTM models lacking attention mechanisms. Despite these promising results, challenges persist, including the model's reliance on large, high-quality datasets and substantial computational resources. Future research directions include integrating transformer-based architectures to further enhance semantic understanding, employing reinforcement learning for summary optimization, and developing more robust evaluation metrics that assess coherence and relevance beyond lexical overlap. Such advancements aim to refine the summarization process, making it more efficient and adaptable across various domains, thereby addressing the growing demand for effective information distillation in the digital age.

REFERENCES

[1] Saraswathi, R. V., Chunchu, R. V., Kunchala, S., Varun, M., Begari, T., & Bodduru, S.(2022). A Deep Learning Model for Text Summarization.
[2] Li, M., Xing, T., Fu, R., & Yin, F. (2021). Research on Text Summarization Generation Based on LSTM and Attention Mechanism. College of Information and Communication Engineering, Communication University of China, Beijing, China.
[3] Zhang, Y., Liu, W., & Chen, X. (2022). Attention-based LSTM for Automatic Text Summarization. Journal of Computational Linguistics, 35(3), 345-360.
[4] Kumar, P., Rao, T., & Singh, A. (2023). Exploring Reinforcement Learning in LSTM for Text Summarization.
[5] Li, M., Wang, X., & Chen, Z. (2023). Improving Abstractive Summarization with Dual LSTM Networks.
[6] Ghanem, F. A., Padma, M. C., Abdulwahab, H. M., & Alkhatib, R. (2025). Deep Learning-Based Short Text Summarization: An Integrated BERT and Transformer Encoder–Decoder Approach.
[7] Huang, J., Wu, W., Li, J., & Wang, S. (2023). Text Summarization Method Based on Gated Attention Graph Neural Network.
[8] Zheng, C., Zhang, K., Wang, H. J., Fan, L., & Wang, Z. (2021). Enhanced Seq2Seq Autoencoder via Contrastive Learning for Abstractive Text Summarization.
[9] See, A., Liu, P. J., & Manning, C. D. (2017). Get To The Point: Summarization with Pointer-Generator Networks. arXiv preprint arXiv:1704.04368.
[10] Chouikhi, H., & Alsuhaibani, M. (2022). Deep Transformer Language Models for Arabic Text Summarization: A Comparison Study. |
[11] Gupta, A., Chugh, D., Anjum, & Katarya, R. (2021). Automated News Summarization Using Transformer.
[12] Ramesh, D., Kothandaraman, D., Chegoni, R., Mohmmad, S., & Pasha, S. N. (2023). Abstractive Text Summarizer Using Neural Networks Algorithms.
[13] Xu, Z., & Zhu, J. (2022). Deep Hierarchical LSTM Networks with Attention for Video Summarization. Computers & Electrical Engineering, 97, 107595.
[14] Li, J., Zhang, C., & Chen, X. (2021). Text Summarization Based on Multi-Head Self-Attention Mechanism and Pointer Network.
[15] El-Kassas, W. S., Salama, C. R., Rafea, A. A., & Mohamed, H. K. (2021). Automatic Text Summarization: A Comprehensive Survey.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY