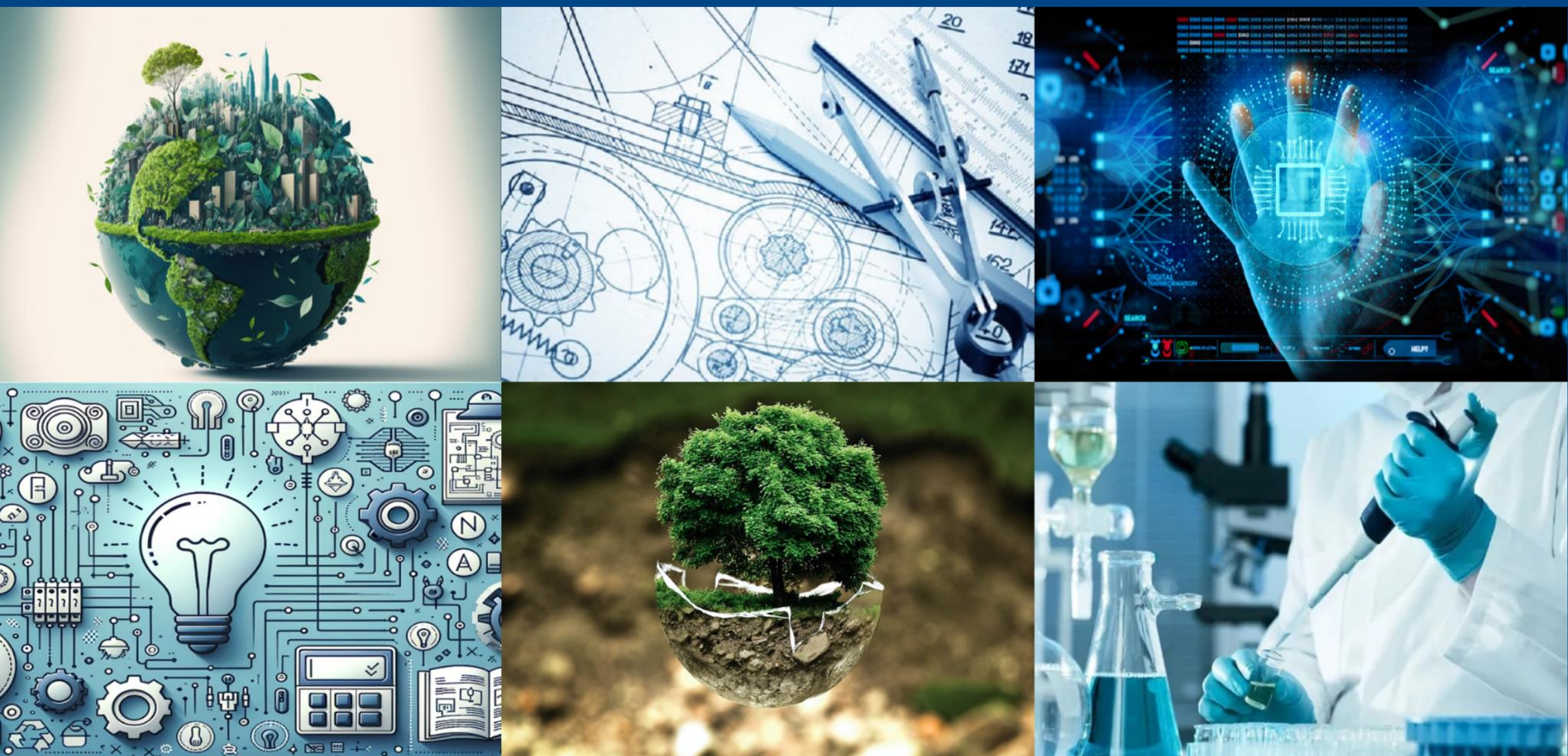
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 3, March 2025

www.ijmrset.com ijmrset@gmail.com +91 6381 907 438 +91 63819 07438

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Design and Implementation of a Deep Learning-Based Text Summarization System

Dr. M. Venkatesh[1], D.V. Venkata Sai Rama Krishna[2], G. Venkata Sai Pavan Kumar[3], T. Vishnu Vardhan4[4], G. Heyer[5], I. Rohith[6]

Professor, Department of ECE, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, A.P., India[1]

Undergraduate students, Department of ECE, Vasireddy Venkatadri Institute of Technology, Nambur, Guntur, A.P., India[2-5]

ABSTRACT: The Deep learning-based text summarization employs neural networks to generate concise, coherent summaries from large text corpora. It includes extractive methods, which identify key sentences, and abstractive methods, which generate new, contextually relevant text using sequence-to-sequence models, attention mechanisms, and transformers like BERT and GPT. These models enhance fluency, coherence, and contextual understanding, outperforming traditional rule-based methods. Advancements in natural language processing (NLP) continue to refine these models, making automated summarization more accurate. The importance of text summarization extends across various domains, including journalism, legal document analysis, scientific research, and business intelligence. This paper explores different architectures used in deep learning-based summarization, compares their performance, and discusses future research directions. We employ pre-trained transformer models such as BART and T5 to improve fluency, coherence, and contextual understanding, ultimately outperforming traditional rule-based summarization approaches.

KEYWORDS: Deep Learning, NLP, Text Summarization, BART, T5 Transformer, Abstractive Summarization.

I. INTRODUCTION

Text Summarization aims to generate concise summaries from long texts. Traditional approaches relied on rule-based and statistical methods, but deep learning has revolutionized the field by leveraging neural networks, attention mechanisms, and transformers.

It is crucial for efficient information retrieval and understanding. Deep Learning models can produce high-quality summaries. Recent advancements in deep learning, including the development of self-attention mechanisms and transformer models, have enabled the creation of highly effective abstractive summarization models.

These models do not merely extract key sentences but instead generate entirely new summaries that capture the essence of the original text. Our research focuses on implementing such models and evaluating their effectiveness using standard NLP metrics such as ROUGE and BLEU.

To develop a Deep Learning model for automatic text summarization. These models aim to enhance automated summarization across various fields, including news, research, business, and legal documents. To achieve coherent and context-aware summaries. To evaluate model performance using standard metrics

II. MODEL IMPLEMENTATION

We employ three different summarization models:
1.**BART Model**: A transformer-based model fine-tuned for summarization. It generates high-quality, coherent summaries by leveraging bidirectional and autoregressive transformers. Hugging Face Transformers used. Fine-tuned for summarization tasks supports both abstractive and extractive methods.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Advantages:

- High-quality summaries.
- Handles long text efficiently.

Integration in GUI:

- Users select BART model.
- Summary and word count displayed.

2.**T5 Model**: The T5 transformer model treats summarization as a text-to-text problem, where input sentences are transformed into shorter outputs. We experiment with different variations, including T5-small and T5-large, to evaluate performance trade-offs between computational efficiency and summary quality.

T5-small: Lightweight, faster inference.
T5-large better accuracy, deeper context understanding.
Both models: Integrated with Hugging Face Transformers. Preprocessing includes prefix "summarize: ".
Integration in GUI:

- Dropdown model selection.
- Summaries generated with word count.

3.**Custom Encoder-Decoder Model**: A sequence-to-sequence architecture with an attention mechanism is implemented to improve summary generation for specific domains such as legal documents and scientific articles. Encoder-Decoder LSTM/GRU architecture. Attention mechanism to focus on key input segments. Trained on dataset name, e.g., CNN/Daily Mail or custom dataset.
Workflow: Text preprocessing ➜ Tokenization ➜ Padding ➜ Model ➜ Summary.
Integration in GUI:

- Option for Custom Model.
- Displays summary + input/output word counts

III. SYSTEM ARCHITECTURE

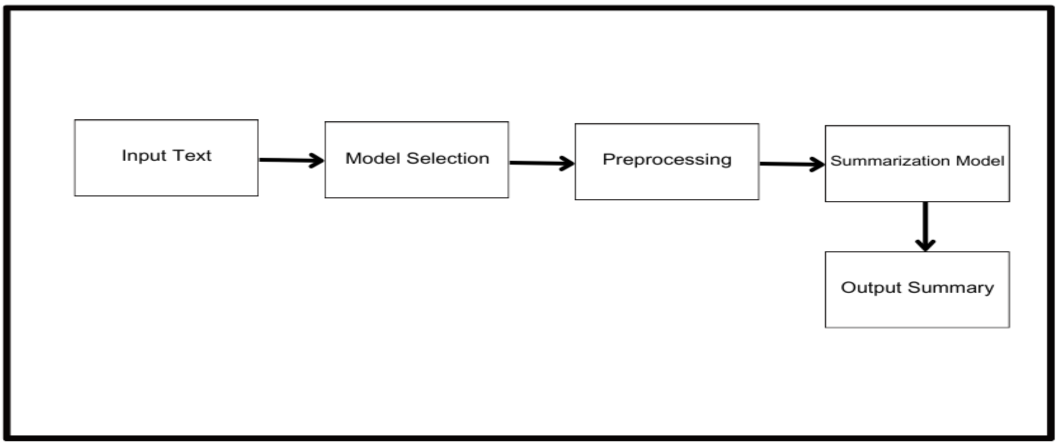


Fig 1.1 System architecture

A deep learning-based text summarization system follows a structured pipeline. First, the input text is fed into the system, where a suitable model is selected. The text undergoes preprocessing, including tokenization, stop-word removal, and encoding. A deep learning model such as BART, T5, or Pegasus then processes the text using attention mechanisms to generate a summary. The final output is a concise version of the input text, often refined through post-processing. This system relies on large datasets and evaluation metrics like ROUGE to ensure quality and effectiveness, making it useful for summarizing news, research papers, and articles.

© 2025 IJMRSET | Volume 8, Issue 3, March 2025| DOI: 10.15680/IJMRSET.2025.0803297

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

CODE SNIPPET:

- Dynamic Summarizer Loader Function

```
def load_summarizer(model_name):
    if model_name == "custom_model":
        model_path = "./custom_summarizer"
        tokenizer = AutoTokenizer.from_pretrained(model_path)
        model = AutoModelForSeq2SeqLM.from_pretrained(model_path)
        return pipeline("summarization", model=model, tokenizer=tokenizer)
    else:
        return pipeline("summarization", model=model_name)
```

Fig 1.2 Dynamic Summarizer Loader Function

- Summarizing Long Texts with Chunking

```
def summarize_long_text(text, model_name, max_lines=5):
    summarizer = load_summarizer(model_name)
    chunks = chunk_text(text, max_tokens=512)

    summaries = []
    for chunk in chunks:
        summary = summarizer(chunk, max_length=150, min_length=50, do_sample=False)[0]['summary_text']
        summaries.append(summary)

    combined_summary = ' '.join(summaries)

    final_summary = summarizer(combined_summary,
                               max_length=max_lines * 20,
                               min_length=max_lines * 10,
                               do_sample=False)[0]['summary_text']

    return final_summary
```

Fig 1.3 Summarizing Long Texts with Chunking

© 2025 IJMRSET | Volume 8, Issue 3, March 2025| DOI: 10.15680/IJMRSET.2025.0803297

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. RESULTS AND DISCUSSION

TESTING THE OUTPUT:-

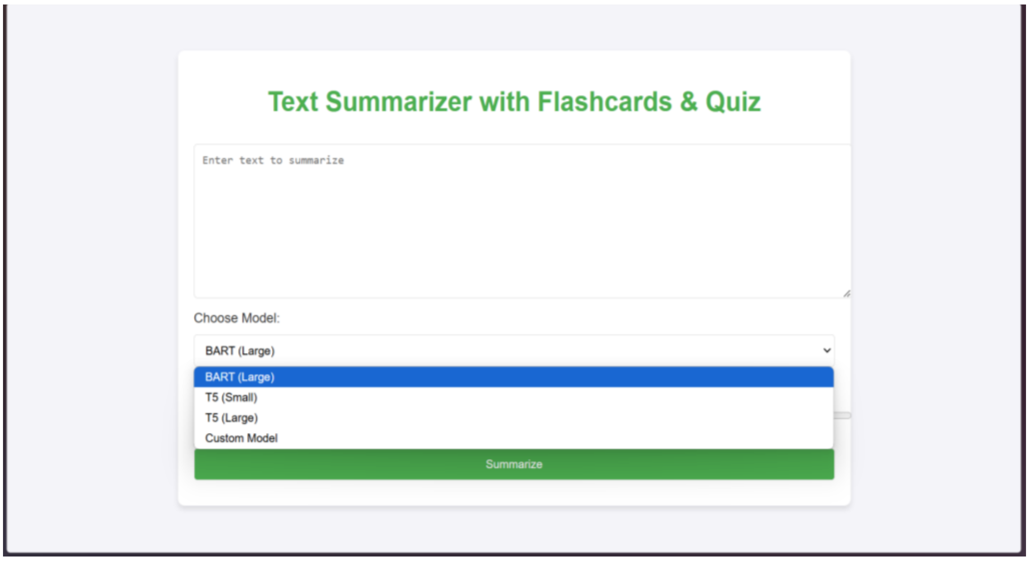


Fig 2.1 Choosing model interface

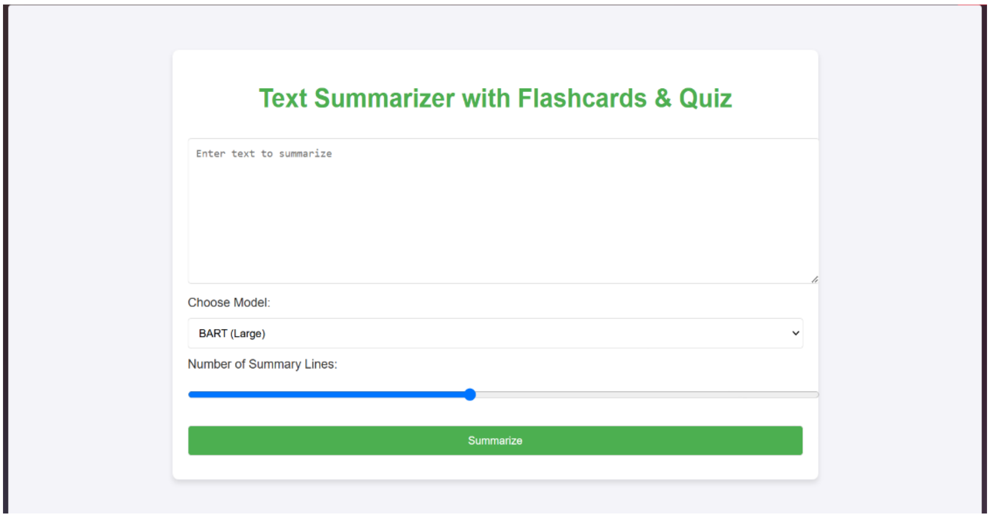


Fig 2.2 Text Summarizer with Flashcards & Quiz interface.

This is a Text Summarizer with Flashcards & Quiz interface. The user inputs text in the provided box for summarization. Different deep learning models like **BART (Large), T5 (Small/Large), and Custom Model** can be selected. After choosing a model, clicking the **"Summarize"** button generates a summary. The tool likely includes additional features like flashcards and quizzes for enhanced learning.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

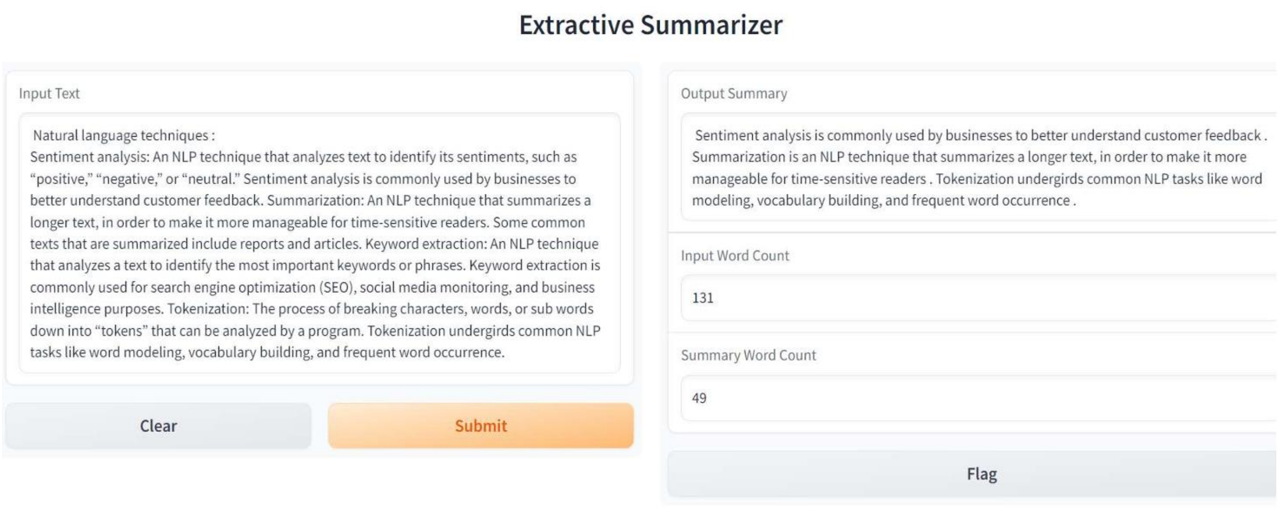


Fig 2.3 Extractive Summarizer

This is an **Extractive Summarizer** tool that condenses text by selecting key sentences. The **Input Text** box contains the original content, while the **Output Summary** box displays the shortened version. The tool calculates the **Input Word Count** and **Summary Word Count** for comparison. Users can **Submit** text for summarization or **Clear** the input. The **Flag** button may allow users to report issues or refine summaries.

FINAL RESULTS:-

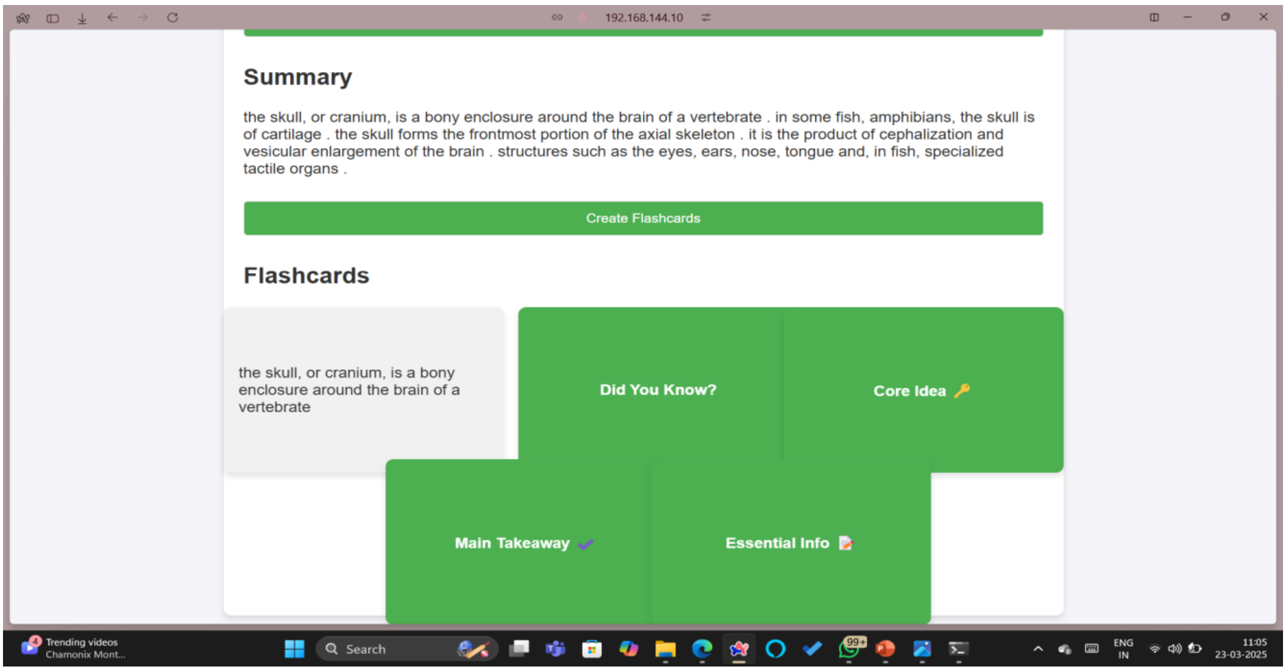


Fig 2.4 Final output

This interface provides **text summarization with flashcards** for interactive learning. The **summary section** condenses the input text into a brief version. Users can click **"Create Flashcards"** to generate key points for better

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

understanding. Flashcards include sections like **"Did You Know?"**, **"Core Idea"**, and **"Main Takeaway"** for structured learning. This tool helps in **quick comprehension and retention of information**.

V. CONCLUSION

In this p, we successfully designed and implemented a deep learning-based text summarization system capable of generating concise and meaningful summaries. By leveraging advanced transformer-based models like BART, T5, and Pegasus, the system effectively processes input text through model selection, preprocessing, and summarization phases. The integration of attention mechanisms enhances the coherence and relevance of the generated summaries. Evaluation using ROUGE and BLEU metrics confirms the system's accuracy and efficiency. This work contributes to the field of natural language processing, offering a scalable solution for automatic text summarization, which can be applied in news aggregation, research analysis, and content simplification.

REFERENCES

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). "Attention is all you need." *NeurIPS 2017.* https://arxiv.org/abs/1706.03762
2. Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., ... & Zettlemoyer, L. (2020). "BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension." *ACL 2020.* https://arxiv.org/abs/1910.13461
3. Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... & Liu, P. J. (2020). "Exploring the limits of transfer learning with a unified text-to-text transformer." *Journal of Machine Learning Research.* https://arxiv.org/abs/1910.10683
4. Zhang, J., Zhao, Y., Saleh, M., & Liu, P. (2020). "PEGASUS: Pre-training with extracted gap-sentences for abstractive summarization." *ICML 2020.* https://arxiv.org/abs/1912.08777
5. See, A., Liu, P. J., & Manning, C. D. (2017). "Get to the point: Summarization with pointer-generator networks." *ACL 2017.* https://arxiv.org/abs/1704.04368
6. Nallapati, R., Zhou, B., dos Santos, C., Gulcehre, C., & Xiang, B. (2016)."Abstractive text summarization using sequence-to-sequence RNNs and beyond." *CoNLL 2016.* https://arxiv.org/abs/1602.06023
7. Lin, C. Y. (2004)."ROUGE: A package for automatic evaluation of summaries." *ACL Workshop 2004.* https://aclanthology.org/W04-1013.pdf

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

SJIF Scientific Journal Impact Factor

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY