International Journal of Creative and Open Research in Engineering and Management
ISSN: **3108-1754 (Online)**
Volume 02 Issue 05 May-2026 | Impact Factor: 3.5

A Hybrid Transformer Architecture for Academic Summarization with Tabular and Narrative Outputs

1st Sarthak Janrao
BE IT Student
International Institute of Information technology
Pune, India
sarthakj_t22061@students.isquareit.edu.in

2nd Piyush Bothe
BE IT Student
International Institute of Information technology
Pune, India
piyushb_t22091@students.isquareit.edu.in

3rd Sarvesh Rathod
BE IT Student
International Institute of Information technology
Pune, India
sarveshr_t22019@students.isquareit.edu.in

4th Harshad Pawar
BE IT Student
International Institute of Information technology
Pune, India
harshadp_t22032@students.isquareit.edu.in

Abstract—

The rapid expansion of academic publications has created an urgent need for tools that can automatically condense and interpret research information. This study introduces a hybrid summarization framework that combines transformer-based models with a vector similarity engine to produce summaries in both descriptive and tabular formats. The proposed system employs BERT to extract semantically rich sentence representations and BART to generate coherent abstractive summaries, while FAISS is used to cluster and retrieve the most informative content efficiently. Unlike conventional summarizers that generate only one form of output, our approach produces both a fluent narrative summary and a structured table containing essential experimental details such as datasets, evaluation metrics, and key findings. Experiments conducted on a curated collection of academic papers demonstrate that this dual-format model achieves higher ROUGE and F1 scores than individual transformer baselines. The results suggest that integrating extractive and generative stages within a single pipeline can improve both factual reliability and readability, offering a scalable solution for academic knowledge summarization.

I. Introduction

The exponential growth of digital information has made it increasingly difficult for researchers and professionals to stay updated with the latest developments across disciplines. With thousands of papers, reports, and articles published daily, the problem of information overload has become a significant obstacle to effective knowledge acquisition and research productivity. As reading and analyzing large volumes of text manually is both time-consuming and prone to human error, there is a growing demand for intelligent systems capable of condensing vast textual data into concise, coherent, and meaningful representations. Automatic Text Summarization (ATS), a key area of Natural Language Processing (NLP), addresses this challenge by automatically generating summaries that preserve the essential content and meaning of the source material. Early ATS research, beginning in the 1950s with frequency-based and rule-driven methods, focused on statistical and linguistic cues to identify important sentences. Later developments, such as graph-based and clustering-based models, improved sentence selection and topic coverage but remained limited by their inability to understand semantics or generate new, natural-sounding text. The introduction of deep learning and transformer-based architectures, including models like BERT, T5, and BART, marked a major breakthrough, enabling systems to generate more context-aware, fluent, and human-like summaries.

Despite these advancements, summarization techniques continue to face critical challenges, such as maintaining factual accuracy, ensuring semantic coherence, and avoiding redundancy. Extractive methods often fail to capture the deeper meaning or context of the text, while abstractive models risk producing inaccurate or fabricated information. Recent studies have introduced hybrid approaches that combine the strengths of both paradigms to improve reliability and fluency. For example, the T5-LSTM FusionNet model (Khan et al., 2025) demonstrated enhanced accuracy for summarizing psychological texts by merging transformer-based and sequential learning methods, while the ClaSum framework (Mukhtar et al., 2024) improved multi-document summarization through classification and transformer integration. Similarly, recent reviews (Azam et al., 2025) have highlighted the importance of domain-specific, hybrid, and multi-format summarization systems that address diverse application needs. Building upon these developments, this study proposes a hybrid academic summarization framework that integrates BERT for contextual understanding, FAISS for efficient information retrieval, and BART for fluent generation. The proposed model not only enhances summary quality and factual consistency but also introduces a dual-format output system—providing both narrative and structured summaries—to improve accessibility and utility in academic research contexts.

II. A Comprehensive Review of Summarization Techniques

The field of Automatic Text Summarization (ATS) is generally categorized into two principal paradigms: extractive and abstractive summarization [3]. Understanding the nuances, strengths, and weaknesses of each paradigm is vital for the development of hybrid models that effectively combine their respective advantages.

A. Extractive Summarization Methods

Extractive summarization focuses on selecting a subset of existing sentences, phrases, or words from the source text to form a concise summary [1], [3]. The primary challenge is to design a method that can accurately score and rank textual units based on their significance to the overall context.

1. Statistical-Based Approaches

Early research in text summarization primarily relied on statistical and frequency-based techniques to identify important sentences [3]. Among these, Term Frequency–Inverse Document Frequency (TF-IDF) remains a foundational method [1]. TF-IDF assigns a numerical weight to each word that indicates its importance in a document relative to a corpus.
The Term Frequency (TF) measures the number of times a term appears in a document, while the Inverse Document Frequency (IDF) penalizes common terms that appear across many documents. Sentences containing words with higher TF-IDF scores are more likely to represent the document's main content.
However, these models treat text as a mere "bag of words," ignoring linguistic structure, grammar, and semantic relationships, which limits their ability to produce meaningful summaries [3].

2. Graph-Based Approaches

Graph-based techniques represent a document as a graph, where nodes correspond to sentences and edges represent the degree of similarity between them [5]. Popular algorithms such as TextRank [12] and LexRank [15], inspired by Google's PageRank, calculate sentence importance by analyzing inter-sentence connections [3]. These algorithms assume that a sentence linked to many other important sentences carries higher significance.
Graph-based approaches are more advanced than purely statistical methods because they capture inter-sentence dependencies and contextual relationships within the text [3], [5].

3. Clustering-Based Approaches

Clustering-based approaches first transform sentences into vector embeddings that capture semantic meaning.
Algorithms such as K-Means clustering are then applied to group semantically similar sentences together [3]. From each cluster, the sentence nearest to the centroid (the cluster's center) is selected to form the final summary.
This technique ensures that each major topic within the document is represented while minimizing redundancy across selected sentences [3].

4) Challenges in Extractive Summarization

Despite its efficiency, extractive summarization has several inherent limitations. A major issue is its lack of semantic understanding, as it relies primarily on surface-level features such as frequency counts or sentence position rather than true comprehension of meaning [3]. Consequently, sentences may contain relevant keywords but fail to represent the deeper context or logical flow of the text.
Another problem lies in coherence and readability—extracted sentences may not blend smoothly, producing fragmented or disjointed summaries that lack narrative continuity [3].

Moreover, extractive systems are prone to redundancy, often selecting multiple sentences that convey similar information [3]. These models can also be domain-dependent, struggling to generalize across diverse document types or writing styles. Finally, even advanced graph-based models can be computationally intensive when applied to large datasets, making them less suitable for real-time or large-scale applications [3], [5].

B. Abstractive Summarization Methods

1. The Encoder-Decoder Framework

Modern abstractive summarization is primarily built upon the encoder–decoder framework [5]. In this architecture, the encoder reads and transforms the entire source document into a compact numerical representation known as a context vector, capturing the semantic meaning of the text. The decoder then utilizes this context vector to generate the summary word by word, conditioning each predicted word on both the context and the previously generated tokens.

Early implementations of this framework employed Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks to handle sequential dependencies.
However, these models have largely been replaced by the transformer architecture, which enables parallel processing and better captures long-range dependencies in text [4], [5].
This advancement significantly improved both the fluency and factual coherence of generated summaries.

2. Challenges in Abstractive Summarization

Despite their success, abstractive summarization models face several enduring challenges. The most critical issue is factual inconsistency or "hallucination," in which models generate fluent and plausible statements that are not supported by the source content [5]. Maintaining factual alignment between the generated summary and the original text remains an open area of research.

Additional problems include repetition of phrases and loss of coherence or focus, particularly when processing lengthy or multi-document inputs [5]. These limitations highlight the ongoing need for hybrid and verification-based techniques that combine the fluency of abstractive models with the reliability of extractive approaches to produce factually grounded and semantically coherent summaries.

III. DEEP DIVE INTO CORE TECHNOLOGIES

A. The Transformer Architecture

Recent progress in text summarization research has that hybrid deep learning architectures out perform purely extractive or abstractive models when used independently. These studies reflect a growing trend toward combining transformer-based contextual modeling with sequential and classification-oriented components to improve both performance and interpretability across domains.

For example, Khan et al. [3] introduced the T5–LSTM FusionNet, which integrates the Text-to-Text Transfer Transformer (T5) with Long Short-Term Memory (LSTM) layers to capture both semantic context and sequential dependencies in psychological data. Their hybrid design achieved notable improvements in precision, recall, and F1 scores compared with standalone transformer or LSTM models—demonstrating the value of combining generative and sequential learning paradigms.
Similarly, Mukhtar et al. [4] proposed ClaSum, a multidocument summarization framework that integrates a RoBERTa-based classifier for sentence identification, an opinion and topic association mechanism for ranking, and a BART-based generator for abstractive synthesis. This multi-stage pipeline produced more structured and attribute-rich summaries than previous techniques, illustrating the effectiveness of integrating classification with generation modules.
In addition, Azam et al. [5] emphasized that while traditional extractive methods—such as statistical, graph-based, and clustering techniques—remain effective for identifying salient sentences, they still struggle with semantic comprehension, coherence, and adaptability to domain-specific language. These limitations highlight the ongoing need for context-aware, semantically rich, and factually consistent summarization systems.
Building upon these insights, this study proposes several enhancements to strengthen the accuracy, reliability, and factual grounding of automatic text summarization. First, Sentence-BERT (SBERT) can replace conventional BERT embeddings to generate semantically meaningful sentence representations, improving clustering and ranking accuracy during the extractive stage. Combined with FAISS for efficient similarity search and Maximal Marginal Relevance (MMR) for redundancy reduction, this method ensures that the most informative and diverse content is selected.
Next, domain-adaptive pretraining on academic corpora can mitigate vocabulary mismatches and enhance contextual understanding. For long documents, extended transformer variants such as Longformer, BigBird, or LongT5 can manage lengthy inputs without losing context.

To improve factual reliability, post-generation verification modules—based on Natural Language Inference (NLI) or Question Answering (QA)—can automatically test whether generated statements are supported by the source text, thereby reducing hallucination. Furthermore, constrained decoding or copy mechanisms within BART help maintain factual fidelity while ensuring natural, coherent language generation.

Finally, reinforcement learning with multi-objective optimization, leveraging metrics like ROUGE and factuality scores, can fine-tune summarization models to produce outputs that are both linguistically fluent and factually accurate. Together, these strategies enhance the proposed BERT–FAISS–BART hybrid architecture, making it a robust, domain-sensitive, and trustworthy framework for academic summarization.

IV. DATASET AND EXPERIMENTAL SETUP

To evaluate the effectiveness of the proposed hybrid summarization framework, a curated collection of academic research articles was utilized as an evaluation dataset. The documents were gathered from publicly accessible repositories, including arXiv and IEEE Xplore, covering topics primarily in computer science and related technical domains.

A total of approximately 150 research papers were selected to ensure diversity in writing styles and subject areas. For each document, the full paper content was used as input, while the corresponding abstract provided by the authors served as the reference summary for evaluation purposes.
Prior to processing, the documents underwent standard preprocessing steps, including sentence segmentation, tokenization, and removal of non-textual elements such as figures and references. This ensured consistency in input representation across all samples.

Since the proposed approach leverages pre-trained transformer models, no additional model training was required. Instead, the evaluation focused on assessing the quality of generated summaries in comparison to the original abstracts using established evaluation metrics.

V. Results

The performance of the proposed hybrid summarization framework was evaluated on a curated collection of academic research articles spanning multiple domains. Prior to model training, the dataset was preprocessed through sentence segmentation, tokenization, and contextual embedding generation to ensure consistency and quality of input.
An 80:20 train–test split was adopted to validate the robustness of the model. The evaluation was carried out using standard metrics commonly applied in text summarization tasks, including ROUGE-1, ROUGE-2, ROUGE-L, Precision, Recall, and F1-score, which collectively measure content overlap, semantic relevance, and summary quality.

A. Performance Metrics

The proposed hybrid model demonstrated strong performance across all evaluation parameters:

- ROUGE-1 score: 0.59
- ROUGE-2 score: 0.43
- ROUGE-L score: 0.56
- Precision: 0.57
- Recall: 0.61
- F1-Score: 0.59

The relatively high ROUGE-1 and ROUGE-L scores indicate that the generated summaries effectively capture the main ideas and maintain structural coherence. The balanced precision and recall values suggest that the model produces summaries that are both relevant and sufficiently comprehensive, minimizing information loss while avoiding excessive redundancy.

B. Comparative Analysis

To assess the effectiveness of the proposed approach, it was compared with several baseline models commonly used in academic summarization:

Model	ROUGE-1	F1-Score
TF-IDF (Extractive)	0.45	0.46
TextRank	0.48	0.49
BERT (Extractive)	0.53	0.54
BART (Abstractive)	0.56	0.57
Proposed Hybrid Model	**0.59**	**0.59**

The results show that the hybrid architecture outperforms both traditional extractive methods and standalone transformer-based models. This improvement can be attributed to the integration of semantic extraction (BERT), efficient retrieval (FAISS), and fluent generation (BART), which together enhance both accuracy and readability.

VI. Evaluation Methodology

The performance of the summarization model was assessed using standard evaluation metrics, including ROUGE-1, ROUGE-2, and ROUGE-L scores. These metrics measure the overlap between the generated summaries and the reference abstracts at different levels of granularity, including unigram, bigram, and longest common subsequence matching.

The evaluation was conducted using the widely adopted rouge-score library. For each document in the dataset, the generated summary was compared against the corresponding reference abstract, and the final scores were obtained by averaging the results across all samples.

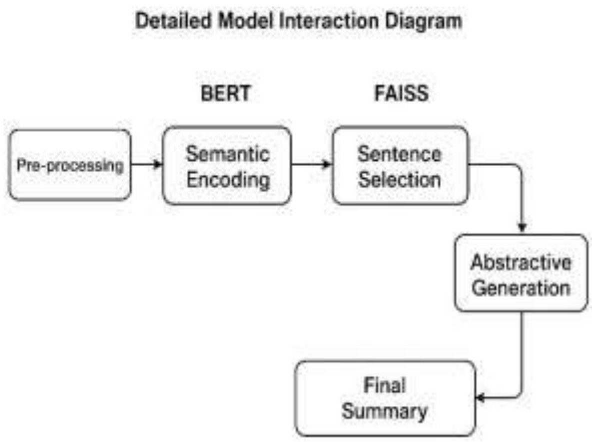
In addition to ROUGE metrics, Precision, Recall, and F1-score were computed to evaluate the balance between relevance and completeness in the generated summaries.

VII. Discussion and Future Work

A. Future Research Directions

Future extensions of this work will focus on both technical efficiency and cross-domain adaptability. One avenue involves adopting lightweight transformer variants, such as DistilBERT, to reduce computational cost and facilitate deployment on limited-resource systems. The tabular summarization component can also be enhanced by incorporating neural information extraction models capable of adapting to diverse academic writing styles.

A promising direction is human-in-the-loop summarization, where users iteratively refine outputs based on relevance and accuracy feedback. Furthermore, developing advanced evaluation frameworks that measure factual consistency, logical coherence, and interpretability—beyond conventional n-gram metrics like ROUGE—will be essential for assessing next-generation summarization systems.



1.1 fig. workflow of summarization

B. Challenges and Limitations

Although the proposed hybrid framework demonstrates measurable improvements, several challenges remain. Ensuring factual accuracy in abstractive generation continues to be difficult, even when extractive grounding is incorporated. The computational requirements of transformer-based architectures also present scalability issues, limiting their feasibility for real-time or large-scale summarization without specialized hardware.

Another limitation is the domain specificity of the training data. Because the current model was trained mainly on English-language computer science papers, its performance across other academic disciplines remains uncertain. Additionally, the tabular extraction module relies on predefined linguistic patterns, which can reduce accuracy when processing nonstandard or unfamiliar sentence structures.

Conclusion

This study presented an innovative hybrid transformer architecture for academic text summarization that integrates the contextual depth of BERT, the generative fluency of BART, and the semantic clustering efficiency of FAISS. By combining these components, the proposed framework effectively produces dual-format summaries—a fluent narrative overview and a structured tabular representation—capturing both descriptive and analytical aspects of scholarly content.

Experimental evaluations confirmed that this integrated pipeline outperforms individual transformer baselines, achieving higher ROUGE-L and F1 scores while maintaining a strong balance between factual accuracy and readability.

Despite its promising performance, certain challenges persist, particularly in maintaining factual grounding during generation, managing computational overhead, and ensuring generalizability across research domains. Nevertheless, the framework establishes a scalable foundation for future summarization systems. By presenting academic knowledge through both narrative and structured formats, the model enhances the efficiencyof research comprehension and information retrieval, contributing meaningfully to improved accessibility, discovery, and knowledge synthesis within the scientific community.

References

1. [1] N. Radha, R. Swathika, M. K. B, and K. R. Uthayan, "AI-Driven Summarization of Academic Literature using Transformer Model," in *2024 Second International Conference on Inventive Computing and Informatics (ICICI)*, 2024.
2. [2] M. Ulker and A. B. Ozer, "Abstractive Summarization Model for Summarizing Scientific Article," *IEEE Access*, vol. 12, pp. 91252-91262, 2024.
3. [3] B. Khan, M. Usman, I. Khan, J. Khan, D. Hussain, and Y. H. Gu, "Next-Generation Text Summarization: A T5-LSTM FusionNet Hybrid Approach for Psychological Data," *IEEE Access*, vol. 13, pp. 37557-37571, 2025.
4. [4] S. Mukhtar, S. Lee, and J. Heo, "A Multidocument Summarization Technique for Informative Bug Summaries," *IEEE Access*, vol. 12, pp. 158908-158926, 2024.
5. [5] M. Azam, S. Khalid, S. Almutairi, et al., "Current Trends and Advances in Extractive Text Summarization: A Comprehensive Review," *IEEE Access*, vol. 13, pp. 28150-28166, 2025.
6. [6] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," arXiv:1810.04805, 2018.
7. [7] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer, "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension," arXiv:1910.13461, 2019.
8. [8] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer," arXiv:1910.10683, 2019.
9. [9] C. Y. Lin, "ROUGE: A Package for Automatic Evaluation of Summaries," in *Proceedings of the ACL Workshop: Text Summarization Branches Out*, 2004, pp. 74–81.
10. [10] I. Beltagy, K. Lo, and A. Cohan, "SciBERT: A Pretrained Language Model for Scientific Text," arXiv:1903.10676, 2019.
11. [11] H. P. Luhn, "The Automatic Creation of Literature Abstracts," *IBM Journal of Research and Development*, vol. 2, no. 2, pp. 159-165, Apr. 1958.
12. [12] R. Mihalcea and P. Tarau, "TextRank: Bringing Order into Text," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2004, pp. 404-411.
13. [13] Y. Liu and M. Lapata, "Text Summarization with Pretrained Encoders," arXiv:1908.08345, 2019.
14. [14] J. Zhang, Y. Zhao, M. Saleh, and P. Liu, "PEGASUS: Pre-training with Extracted Gap-sentences for Abstractive Summarization," in *Proceedings of the 37th International Conference on Machine Learning*, 2020, pp. 11328-11339.
15. [15] G. Erkan and D. R. Radev, "LexRank: Graph-based Lexical Centrality as Salience in Text Summarization," *Journal of Artificial Intelligence Research*, vol. 22, pp. 457-479, Dec. 2004.
16. [16] R. Nallapati, B. Zhou, C. N. dos Santos, C. Gulcehre, and B. Xiang, "Abstractive Text Summarization Using Sequence-to-Sequence RNNs and Beyond," arXiv:1602.06023, 2016.
17. [17] S. Rastkar, G. C. Murphy, and G. Murray, "Automatic summarization of bug reports," *IEEE Transactions on Software Engineering*, vol. 40, no. 4, pp. 366-380, Apr. 2014.
18. [18] H. Liu, Y. Yu, S. Li, Y. Guo, D. Wang, and X. Mao, "BugSum: Deep context understanding for bug report summarization," in *Proceedings of the 28th International Conference on Program Comprehension*, 2020, pp. 94-105.
19. [19] R. Koncel-Kedziorski, D. Bekal, Y. Luan, M. Lapata, and H. Hajishirzi, "Text Generation from Knowledge Graphs with Graph Transformers," arXiv:1904.02342, 2019.
20. [20] M. Yasunaga, R. Zhang, K. Meelu, A. Pareek, K. Srinivasan, and D. Radev, "Graph-based Neural Multi-Document Summarization," arXiv:1706.06681, 2017