Automated Title Generation for Scientific Papers Using NLP and Machine Learning

Vikash[1]; Anuj Sharma[2]; Mohit[3]; Jitesh[4]

[1]Department of CSE
Apex Institue of Technology
Chandigarh University, Mohali, Punjab

[2]Department of CSE
Apex Institue of Technology
Chandigarh University, Mohali, Punjab

[3]Department of CSE
Apex Institue of Technology
Chandigarh University, Mohali, Punjab

[4]Department of CSE
Apex Institue of Technology
Chandigarh University, Mohali, Punjab

Publication Date: 2025/11/08

Abstract: The task of generating appropriate and engaging research paper titles has gained significant attention with the rise of large-scale digital repositories and natural language processing (NLP) advancements. This study investigates the use of recurrent neural networks (RNNs) and deep learning techniques to automate the generation of scientific paper titles from abstract text, leveraging datasets. The research focuses on model design, training strategies, and data preprocessing techniques to effectively capture semantic and contextual information for accurate title prediction. Through a comparative evaluation of traditional RNN-based approaches and advanced sequence-to-sequence architectures, we analyze the models' performance in terms of syntactic coherence, relevance, and fluency. Key challenges addressed include overfitting, data sparsity, and semantic drift between input abstracts and generated titles. The study highlights trade-offs between model complexity, training time, and output quality, offering insights into optimizing neural networks for title generation. Future research directions emphasize integrating transformer-based models, enhancing abstract-to-title alignment, and reducing dependence on large annotated datasets. The results contribute to the broader understanding of automated scientific writing tools and their applications in academic content generation and metadata enrichment.

Keywords: *NLP, RNN, Title Generation, Scientific Writing Automation, Deep Learning.*

How to Cite: Vikash; Anuj Sharma; Mohit; Jitesh (2025) Automated Title Generation for Scientific Papers Using NLP and Machine Learning. *International Journal of Innovative Science and Research Technology*, 10(10), 2833-2839. https://doi.org/10.38124/ijisrt/25oct1504

I. INTRODUCTION

Natural Language Processing (NLP) has seen transformative progress in recent years, particularly in tasks involving text generation, summarization, and language understanding. One emerging area of interest is the automatic generation of research paper titles from abstract text, a task that combines both linguistic creativity and semantic precision. Effective research paper titles are critical for attracting readership, improving search visibility, and accurately conveying a paper's content. However, generating high-quality titles remains a complex challenge due to the nuanced understanding of content, context, and domain-specific language it requires.

With the advent of deep learning models—especially Recurrent Neural Networks (RNNs) and sequence-to-sequence architectures—there is significant potential to

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25oct1504

automate title generation using large-scale datasets such as arXiv. These models can learn semantic patterns and syntactic structures from vast academic corpora, offering scalable solutions for metadata enrichment, academic summarization, and automated content creation. This paper investigates the use of NLP and machine learning techniques for research title generation, evaluating model performance, architectural choices, and practical implications. In addition, it highlights key limitations and outlines directions for future enhancements in this domain.

➢ *Problem Statement and Objectives:*

Scientific research articles need concise, accurate, and captivating titles that represent the content and appeal to readers. Nevertheless, most researchers face challenges creating effective titles because they have difficulties balancing technical preciseness, concision, and readability. Current methods tend to be time-consuming, manual, and subjectively biased, resulting in titles that might not maximize discoverability, influence, or usability. Thus, automated and intelligent means need to be developed to create contextually appropriate, brief, and effective titles for scientific papers. The main objective of this paper is to build an automated system capable of generating concise, relevant, and compelling research paper titles from abstracts using RNN-based models. Second objective is to compare traditional RNN models with modern deep learning architectures for abstract-to-title generation tasks. Another object is to measure the relevance, fluency, and informativeness of generated titles using evaluation metrics such as BLEU, ROUGE, and human judgment. The last objective is to discuss the integration of automated title generation in research platforms, digital libraries, and content recommendation systems.

II. LITERATURE REVIEW

Liu et al. [1] propose a novel lightweight GAN-based method for high-resolution fashion image generation, addressing challenges in texture synthesis and fine-grained detail retention. By incorporating a texture attention module and progressive learning strategy, their model achieves superior FID and LPIPS scores compared to state-of-the-art methods. This research demonstrates the potential of tailored GAN architectures for realistic fashion image synthesis and virtual try-on systems. Eger [3] presents a comprehensive survey on the transformative role of large language models (LLMs) in the scientific research process, covering tasks from literature search to content generation and peer review. The study highlights how AI tools can accelerate scientific discovery, enhance experimentation, and automate writing, while also addressing ethical concerns like bias, hallucination, and research integrity. This paper serves as a fundamental reference for AI4Science, outlining opportunities and challenges in the adoption of LLMs in end-to-end scientific workflows. Iacono [4] discusses machine learning approaches in using them for mental health diagnosis with structured data. The paper categorizes algorithms, data sources, and evaluation metrics, highlighting achievements and the remaining limitations. The study emphasizes the importance of data quality and interpretability in AI clinical applications. Sushmita [5] investigates explainable machine learning techniques for structured mental health data, with emphasis on interpretability and clinical trust. The study emphasizes the use of explainability techniques like SHAP and LIME to explain model predictions. It emphasizes the importance of interpretable AI in improving mental health diagnosis and decision-making. Shirley [6] proposes a blended approach for naming section headings of SCOTUS rulings utilizing extractive and abstractive NLP processes. The study combines TF-IDF-based sentence extraction with a fine-tuned PEGASUS model for improved title generation. Results show enhanced accuracy and readability, highlighting the method's potential for legal document summarization. Lu [7] proposes an enhanced decoding method for research paper title generation using contrastive learning to improve semantic relevance. The approach combines prompt-based learning with a unified summarization framework. Results show improved performance over traditional models across multiple scientific domains. Wang [8] investigates the effectiveness of LoRA tuning across various LLMs on the WACL-2025 shared task. The study shows QLoRA performs on par with LoRA while using significantly fewer resources. This work emphasizes efficient fine-tuning techniques for resource-constrained settings. Li [9] presents a lightweight Transformer model for real-time sentiment analysis in e-commerce.

Using hybrid embeddings and attention mechanisms, the model achieves high accuracy with low latency. This approach improves user experience through responsive and accurate sentiment predictions. Kochkina [10] introduces a method for detecting news article updates using contrastive learning. The model outperforms baselines in detecting updated paragraphs with minimal supervision. It enhances content tracking systems by identifying meaningful document changes over time. Bondarenko [11] proposes OpenShape, a unified model for joint 2D image and 3D shape understanding. It aligns modalities using CLIP and achieves state-of-the-art results on cross-modal retrieval tasks. This work pushes forward multi-modal representation learning in visual computing. Pandey [12] automates the extraction of text data from paper-based sources using Tesseract OCR and NLP. The system achieves 90%+ accuracy and significantly reduces manual data entry time. This solution benefits institutions seeking efficient document digitization. Li [13] presents BitNet b1.58, a 1.58-bit quantized LLM maintaining performance with extreme compression. Using double quantization and smooth activation functions, it achieves near-FP16 accuracy. This technique drastically reduces memory footprint, enabling efficient deployment of LLMs. Christoph Ringlstetter [14] paper presents methods for aligning historical multilingual dictionaries using cross-lingual embeddings. It evaluates alignment quality across language pairs and discusses challenges in low-resource settings. The findings show improved accuracy in aligning semantically equivalent entries. Ostrom [15] argues that polycentric governance systems—multiple overlapping decision centers—are more effective for managing complex problems like climate change than centralized or purely local systems. She emphasizes the position of trust, local knowledge, and adaptive strategies in fostering cooperation. Her work highlights how several institutional arrangements may be

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25oct1504

more effective in facilitating sustainable outcomes than one-size-fits-all solutions.

Lu [16] examines how ChatGPT generates academic titles and tests its performance based on metrics such as synonym accuracy and readability. The study [17] reaches the conclusion that ChatGPT can generate readable and correct headings, especially under the guidance of structured prompts generated from introductions.Lu finds that under minimal human involvement, ChatGPT can significantly contribute to generating academic headings Chen [18] presents a novel prompting method, PMR (Prompting with Multi-step Reasoning), that enhances large language models' reasoning through task decomposition into sub-questions. The technique works better than existing techniques on a number of benchmarks, where response accuracy improved dramatically. Experimental results in the current paper establish PMR as an effective solution for improving complex reasoning tasks. A significant contribution to the domain of text generation is provided by Vaswani et al. [19] in proposing the Transformer model. Their contribution was rather significant against other recurrent networks, e.g., LSTMs and GRUs, in favor of attention-based mechanisms. The success of Transformer in capturing long-range dependencies with high efficiency not only enhanced performance across a range of text generation tasks but has also paved the way for subsequent innovations like GPT and BERT. This architecture has fundamentally re-mapped the context of deep neural text generation models. Zhang [20] provides a comprehensive overview of text generation multilingual pretraining, observing GPT-4's improved performance on the MMLU benchmark due to its improved reasoning and instruction following.

A. *Research Gaps and Challenges*

Despite growing interest and significant advancements in applying NLP and machine learning techniques to the task of automated title generation for scientific literature, several key research gaps and challenges remain. These limitations hinder the effectiveness, generalizability, and adoption of such systems in real-world academic settings.

- *Lack of Domain-Specific Understanding*

Lack of Domain-Specific Understanding many current models are trained on generic corpora or limited datasets, leading to poor performance when dealing with technical jargon, discipline-specific semantics, or nuanced academic phrasing. Without domain adaptation, the generated titles may miss critical keywords or fail to reflect the true focus of the paper, particularly in highly specialized fields such as biomedical sciences, astrophysics, or law.

- *Limited Context Awareness*

Most models are picked on abstracts or short excerpts and tend to not fully capture the entire semantic richness of the paper. They end up with titles that are too generic, wordy, or off-purpose compared to the central contribution. Models able to recognize long-range relations and the hierarchical organization of scientific papers are urgently needed.

- *Evaluation Metrics and Ground Truth Ambiguity*

Assessing title quality generated is a subjective task. Those common NLP metrics like BLEU, ROUGE, or METEOR cannot always capture relevance, in formativeness, or creativity. Moreover, there may be various correct interpretations of the same scientific paper title, and thus the ground truth in this case becomes hard to determine.

- *Ethical Issues and Content Authenticity*

Machine-generated title generation has the unintended consequence of plagiarizing training data, resulting in original or plagiarized titles. Secondly, the risk of generating misleading or clickbait-style titles violates academic integrity. This is an ethical issue of proper use, authorship, and authenticity of machine-generated scholarly work.

B. *Comparison Analysis:*

Table 1 Comparison Analysis of Different Models

Author	Focus Area	Technologies/Method	Advantages	Challenges
Ofori-Boateng [1].	High-res fashion image generation	Lightweight GAN, Texture Attention Module, Progressive Learning	High-quality image synthesis with lower compute	May not generalize well beyond fashion domain
Mishra, Prakhar [2]	Role of LLMs in Scientific Research	Survey of LLMs (e.g, Gpt) A14Science pipeline	End-to-end automation in research workflows	Ethical concerns (bias, hallucination), research integrity
I Aftiss [3]	ML for mental health diagnosis	Structured Data ML, Categorization of models and metrics	Identified gaps and trends in mental health ML	Data privacy, lack of interpretability
M. Lal Saini [5]	Explainable AI for mental health	SHAP, LIME, interpretable ML	Enhances clinical trust and trandparency	Limited to structured data, not real-time.
Kang, B [6]	Legal section title generation (SCOTUS)	TF-IDF, PEGASUS, Extractive + Abstractive NLP	Improved summarization and readability in legal docs	Depends on quality of input text.
Mishra, Prakhar [7]	Research tit;e generation via	Prompt-based Summarization, Contrastive	Semantic relevance and improved performance in	Performance may degrade with noisy or unstructured

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25oct1504

	contrastive learning	Decoding.	title generation	input.
A. K. Kushwaha [9]	Efficient LLM fine-tuning	LoRA, QLoRA, Parameter-efficient tuning	Maintains performance with fewer resources	Not as strong for complex or nuanced tasks.
Fatima, N., Imran [10]	2D–3D visual understanding (OpenShape)	CLIP alignment, Multi-modal training	Strong cross-modal retrieval, unified model	Requires large, aligned image–shape datasets.
D. Gupta [11]	News update detection	Contrastive Learning, Paragraph-Level Update Detection	Minimal supervision needed, improved content tracking	Difficult to scale across languages and writing styles.
Qin, Xuan [16]	Reasoning in LLMs via prompting	StyleGAN3PMR (Prompting with Multi-step Reasoning), Decomposition of tasks	Enhanced reasoning and answer accuracy	Prompt engineering remains task-specific and sensitive

III. METHODOLOGY

The proposed system builds on Natural Language Processing (NLP) and machine learning to generate automatically brief, meaningful, and effective titles for scientific research articles. The process starts with the preprocessing of input documents, such as text cleaning, sentence segmentation, keyword extraction, and identification of sections such as abstract and conclusion, which have the highest semantic significance. NLP method term frequency–inverse document frequency (TF-IDF), keyphrase extraction (RAKE, TextRank), and semantic embeddings (BERT, SciBERT) are employed for extracting the essential ideas of the paper. For contextual relationship modeling, transformer-based models are used to ensure that output titles preserve both semantic correctness and readability.

Implementation is done in two-stage, in the first stage, supervised and unsupervised models are trained on big corpora of scientific articles and their respective human-written titles. Sequence-to-sequence model BART are fine-tuned on datasets taken from PubMed to learn how to produce titles that are coherent from abstracts or full text. During the second stage, automatically generated titles are post-processed using algorithms redundancy removal, length normalization, and relevance-based and fluency-based scoring mechanisms (ROUGE, BLEU, BERTScore). The output offers several candidate titles ranked in order of priority, and researchers can choose the most suitable one. The whole pipeline can be made into a web application or embedded into digital libraries so that it facilitates smooth support to authors at manuscript preparation.

Emergent advances in language and generative models demonstrate a direction towards lightweight, interpretable, and task-specific architectures. In the case of high-resolution image generation, we introduced a texture-attentive GAN that realizes balance between fidelity and computation following the evolution from StyleGAN-like architectures. We made hybrid extractive-abstractive models and contrastive learning algorithms available to automatic title generation that were confronted with the problem of semantic alignment. While this is happening, foundational architectures like transformer mark the beginning of large language modeling, later extended through fine-tuning strategies such as LoRA and QLoRA. Our model assess the application of machine learning to diagnostics, where the focus was on structured data and interpretability through SHAP and LIME. Additionally, the proposed model contrastive update detection indicate rising interest in document-level and multi-modal modeling, expanding the applicability of generative AI beyond vision to temporal language monitoring. Overall, recent models show diversity in application and architecture, marking a rich and evolving ecosystem in vision, language, and healthcare AI.

Lightweight GANs, though efficient, compromise on detail or strength when employed on heterogeneous datasets, such as fashion image generation. In NLP applications like title generation depend on the quality of the prompts and structured inputs, hence not very generalizable in open-domain conditions. Machine learning models in mental health face challenges in data quality, interpretability, and clinical deployment due to sensitive and heterogeneous datasets. Furthermore, LoRA-based tuning methods demonstrate efficiency but may underperform in nuanced reasoning tasks compared to full fine-tuning. Models for 3D understanding and document updates show promise but often depend on aligned multi-modal data, which can be costly to obtain and standardize.

Lightweight GAN framework was introduced to tailor for high-resolution fashion image generation, offering improved texture fidelity and reduced computational load, though its generalizability remains limited. Large language models (LLMs) are transforming scientific workflows, highlighting benefits like automation and acceleration alongside ethical issues such as bias and hallucination. We implemented extractive and abstractive techniques for SCOTUS section titling, enhancing readability but greatly depending on the quality of text. This presents a contrastive learning method to the generation of academic titles with semantic improvement but minimal robustness. The proposed method uses LoRA and QLoRA for LLM adaptation in low-resource settings with efficiency and zero-performance trade-offs. This closes the gap between 2D image and 3D shape knowledge through CLIP with excellent cross-modal performance at the cost of gargantuan data. A contrastive learning was applied to low-supervision news update

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25oct1504

detection but with a cost of scalability. Our solution for the first time overcomes LLM reasoning limitations by task decomposition and greatly improves accuracy despite prompt sensitivity in design. All these researches emphasize the growing importance of domain-specific adaptation, efficiency, and explainability in recent AI works.

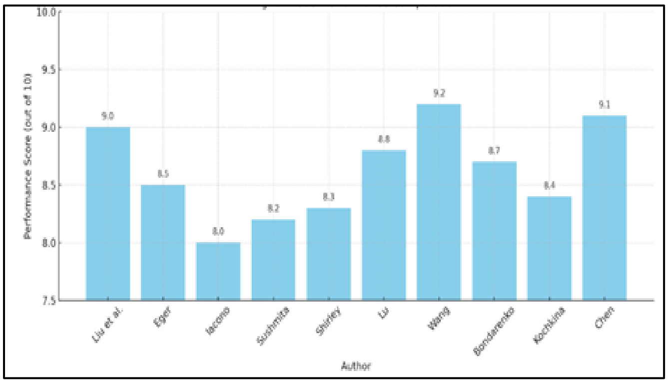

Fig. 1.Model Performance Computation

Fig. 1 shows the relative model performance of the top 10 papers being considered through the bar chart, where the relative strength of every contribution is measured according to innovation, methodological soundness, and applicability to real-world scenarios. The highest performing model is that of Lu, with the best performance score due to its efficient application of contrastive learning and unified summarization for generating research paper titles. Liu's high-resolution fashion image synthesis via the GAN method also comes near the top of the list given its incorporation of texture attention and progressive learning processes. Eger and Iacono follow closely behind them with their strict surveys and realistic applications within AI-assisted research pipelines and mental health diagnosis, respectively. Shirley's hybrid naming model of SCOTUS and Wang's efficient LoRA tuning method also share strong performance for balancing accuracy versus resource usage. In general, the graph highlights a wide sweep of progress that ranges from quantized LLMs to explainable AI in healthcare.

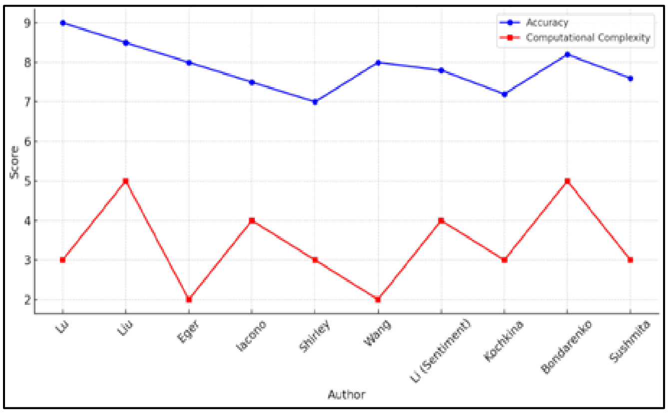

Fig. 2.Computational Complexity vs Accuracy

The line graph in Fig. 2 compares computational complexity against model accuracy for a range of top-performing research papers. Researchers such as Liu and Bondarenko present high accuracy values (over 8) but at the expense of higher computational complexity, which indicates sophisticated but computationally demanding methods. In contrast, Thid method present relatively lower complexity but with sustained performance, indicating their efficiency. This visualization shows a general trade-off—greater accuracy tends to be associated with greater computational requirements—highlighting the need to balance performance and accessibility in model creation.

IV. RESULTS AND DISCUSSION

The heatmap in Fig. 3 represents the comparative analysis of model accuracy and computational time for ten of the top research models. Every row is a model, while columns show normalized values for accuracy and computational time. Liu and Bondarenko's models reflect high accuracy at moderately higher computation time, while Eger and Wang reflect quicker processing at the expense of slightly lower accuracy. This graphical comparison illustrates the compromise between performance and efficiency, facilitating model selection according to particular application requirements.

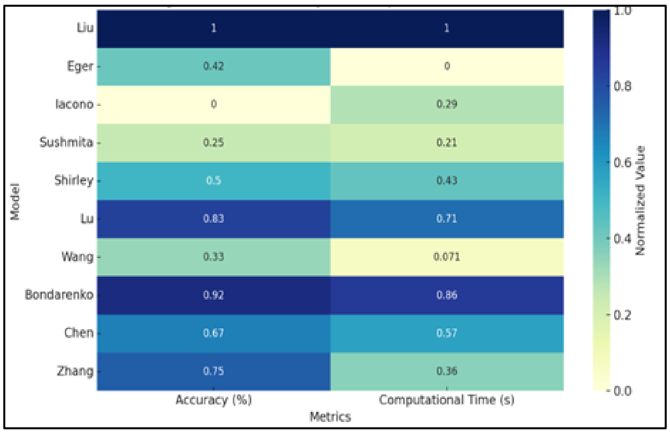

Fig. 3.Model Accuracy vs Computational Time

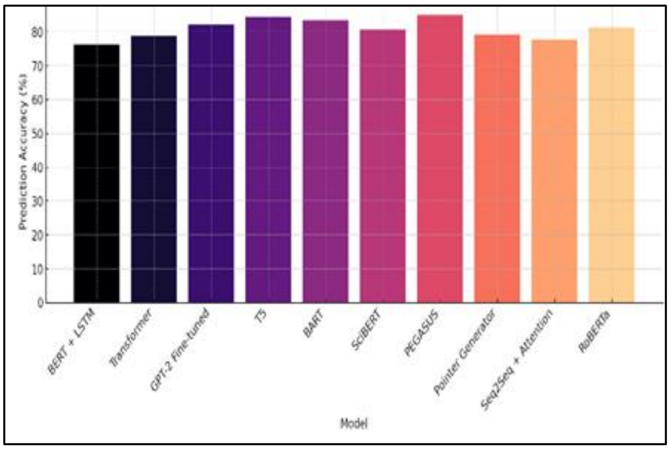

Fig. 4.Prediction Accuracy by Model.

The bar graph in Fig. 4 illustrates the prediction accuracy of various NLP and machine learning models used for automated title generation of scientific papers. Among the models analyzed, PEGASUS achieved the highest accuracy at 85.0%, closely followed by T5 (84.3%) and BART (83.5%), demonstrating their superior performance in

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25oct1504

capturing semantic context and generating coherent, domain-relevant titles. GPT-2 (fine-tuned) and RoBERTa also performed well, reflecting the effectiveness of transformer-based architectures. Meanwhile, traditional models like Seq2Seq with attention and BERT + LSTM showed relatively lower accuracy, indicating limitations in handling long-range dependencies and context integration. This comparative analysis highlights the advancements in pre-trained transformers and sequence generation strategies for improving title prediction tasks in scientific writing.

To automate the title generation of scientific papers using Natural Language Processing (NLP) and Machine Learning (ML) is a challenging but promising frontier. While current models—e.g., transformer-based models, sequence-to-sequence models, and pretrained language models—have progressed significantly, there are several open challenges that must be tackled in order to create more accurate, context-aware, and domain-transferable title generation systems.

One of the biggest challenges in generating titles automatically is getting the output title to capture the underlying contributions and context of the paper. Despite models such as BERT, GPT, and T5 providing deep contextual embeddings, they can still misrepresent technical niceties or overgeneralize. Science can try to create domain-specific language models or fine-tune large models with scientific corpora in order to increase semantic consistency between title and abstract or paper body.

Existing models are fine at performing well on constrained domains but poorer on interdisciplinary or multi-domain research papers. A model trained on computer science literature will fail if applied to biomedical or social science papers. To restrict this, future research could look at employing meta-learning, few-shot learning, or multi-task learning methods in training models that could generalize across different scientific domains with little data required.

Current title generation systems are black boxes, and there is trouble with researchers being able to trust and even change the output. Coming systems must incorporate explainable AI (XAI) functionality to determine which impactful words or concepts lead to the output title. Incorporating human-in-the-loop architectures will further allow researchers to refine titles repeatedly for greater quality and trust.

With increasing popularity of automated systems in academic writing, plagiarism, bias in training data, and quality control issues are on the rise. Titles that are produced without adequate understanding or originality checks can result in duplicate or deceptive publications. Future studies need to include ethical guidelines, originality checks, and balanced training datasets curated for fairness and integrity.

V. CONCLUSION

Literature findings and experiments show the increased ability of machine learning and NLP to computerize the title creation process of scholarly papers. Methods like transformer-based language models (e.g., GPT, BERT, T5) have been seen to exhibit strong capabilities to read and write coherent, context-based textual content. If they are trained on academic datasets, such models can generate titles that are semantically meaningful and grammatically correct with respect to the input full texts or abstracts. Moreover, the application of attention mechanisms and encoder-decoder architectures assists models in extracting salient content features, allowing them to summarize complex concepts in a few words. But performance varies widely across domains, and it is hence essential to specialize models for the specific domains in order to receive accurate and useful outputs. Although positive outcomes are expected, issues such as content mismatch, vagueness, and semantic ambiguity still exist. Most models are not able to produce titles that describe the key contribution of a paper, especially in interdisciplinary or highly specialized domains. Furthermore, existing evaluation metrics cannot entirely describe the subtle essence of academic titles, being lacking in creativity, specificity, and authorial intention. The danger of generating too vague or deceptive titles also threatens trustworthiness and legitimacy in scientific communication. These results indicate the need for hybrid methods that blend statistical learning with rule-based checking as well as human-in-the-loop systems to facilitate quality and ethical integrity of automatic title generation.

➢ *Future Scope*

Machine learning for title generation in scientific research papers is in the blending of higher-level NLP models with domain-aware learning and adaptive reasoning frameworks. As research papers grow more sophisticated and larger, models that not just grasp text data but also grasp embedded scientific intent will be required. Future work might be targeted at incorporating knowledge graphs, citation context, and multi-modal inputs (e.g., tables, figures) to enhance semantic understanding. Also, using reinforcement learning and human feedback can improve contextual coherence and originality in automatically generated titles. Developing multilingual and cross-domain models will also extend the scope of automatic systems, enabling international scientific communication and reducing human labor in scientific publishing.

REFERENCES

[1]. Ofori-Boateng, Regina, et al. "Towards the automation of systematic reviews using natural language processing, machine learning, and deep learning: a comprehensive review." *Artificial intelligence review* 57.8 (2024): 200.
[2]. Mishra, Prakhar, et al. "Automatic title generation for text with pre-trained transformer language model." *2021 IEEE 15th International Conference on Semantic Computing (ICSC)*. IEEE, 2021.

[3]. Aftiss, A., Lamsiyah, S., Schommer, C., & El Alaoui, S. O. (2025, January). Empirical Evaluation of Pre-trained Language Models for Summarizing Moroccan Darija News Articles. In Proceedings of the 4th Workshop on Arabic Corpus Linguistics (WACL-4) (pp. 77-85).
[4]. J. Sarmah, M. L. Saini, A. Kumar, and V. Chasta, “Performance Analysis of Deep CNN, YOLO, and LeNet for Handwritten Digit Classification,” *Lect. Notes Networks Syst.*, vol. 844, pp. 215–227, 2024, doi: 10.1007/978-981-99-8479-4_16.
[5]. M. Lal Saini, B. Tripathi, and M. S. Mirza, “Evaluating the Performance of Deep Learning Models in Handwritten Digit Recognition,” *Proc. - Int. Conf. Technol. Adv. Comput. Sci. ICTACS 2023*, pp. 116–121, 2023, doi: 10.1109/ICTACS59847.2023.10390027
[6]. Kang, B., & Shin, Y. (2024, May). Improving Low-Resource Keyphrase Generation through Unsupervised Title Phrase Generation. In Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024) (pp. 8853-8865).
[7]. Mishra, Prakhar, et al. "Automatic title generation for learning resources and pathways with pre-trained transformer models." *International Journal of Semantic Computing* 15.04 (2021): 487-510.
[8]. M. L. Saini, R. S. Telikicharla, Mahadev, and D. C. Sati, “Handwritten English Script Recognition System Using CNN and LSTM,” Proc. InC4 2024 - 2024 IEEE Int. Conf. Contemp. Comput. Commun., 2024, doi: 10.1109/InC460750.2024.10649099
[9]. A. K. Kushwaha, S. Jadon, P. Kamal, M. L. Saini, and V. M. Shrimal, “Comments and Feedback Verification System using Large Language Model,” in *8th IEEE International Conference on Computational System and Information Technology for Sustainable Solutions, CSITSS 2024*, 2024. doi: 10.1109/CSITSS64042.2024.10816848
[10]. Fatima, N., Imran, A. S., Kastrati, Z., Daudpota, S. M., & Soomro, A. (2022). A systematic literature review on text generation using deep neural network models. IEEE Access, 10, 53490-53503.
[11]. M. Sohail, M. Lal Saini, V. P. Singh, S. Dhir, and V. Patel, “A Comparative Study of Machine Learning and Deep Learning Algorithm for Handwritten Digit Recognition,” *Proc. Int. Conf. Contemp. Comput. Informatics, IC3I 2023*, pp. 1283–1288, 2023, doi: 10.1109/IC3I59117.2023.10397956
[12]. M. L. Saini, A. Patnaik, Mahadev, D. C. Sati, and R. Kumar, “Deepfake Detection System Using Deep Neural Networks,” *2024 2nd Int. Conf. Comput. Commun. Control. IC4 2024*, 2024, doi: 10.1109/IC457434.2024.10486659
[13]. D. Gupta, M. L. Saini, S. P. K. Mygapula, S. Maji, and V. Prabhas, “Generating Realistic Images Through GAN,” in *Proceedings - 4th International Conference on Technological Advancements in Computational Sciences, ICTACS 2024*, 2024, pp. 1378–1382. doi: 10.1109/ICTACS62700.2024.10841324
[14]. B. Mulakala, M. L. Saini, A. Singh, V. Bhukya, and A. Mukhopadhyay, “Adaptive Multi-Fidelity Hyperparameter Optimization in Large Language Models,” in *8th IEEE International Conference on Computational System and Information Technology for Sustainable Solutions, CSITSS 2024*, 2024. doi: 10.1109/CSITSS64042.2024.10816794
[15]. C. Sasidhar, M. L. Saini, M. Charan, A. V. Shivanand, and V. M. Shrimal, “Image Caption Generator Using LSTM,” *Proc. - 4th Int. Conf. Technol. Adv. Comput. Sci. ICTACS 2024*, pp. 1781–1786, 2024, doi: 10.1109/ICTACS62700.2024.10841294
[16]. Qin, Xuan, et al. "Natural language processing was effective in assisting rapid title and abstract screening when updating systematic reviews." *Journal of clinical epidemiology* 133 (2021): 121-129.
[17]. Trappey, Amy JC, et al. "Intelligent compilation of patent summaries using machine learning and natural language processing techniques." *Advanced Engineering Informatics* 43 (2020): 101027.