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 11, November 2025

© 2025 IJMRSET | Volume 8, Issue 11, November 2025| DOI:10.15680/IJMRSET.2025.0811056

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)

Smart Paraphraser using GGUF LLM and Semantic Similarity Ranking: A Survey

Adaikkammai A[1], Arudra Sai Vignesh[2], Venkata Ramana Lingamgunta[3], Nallabothula Venkata Sudheera karthikeya[4]

Assistant Professor, Department of Computer Science and Business Systems, RMD Engineering College, Kavaraipettai, India[1]

Student, Department of Computer Science and Business Systems, RMD Engineering College, Kavaraipettai, India[2-4]

ABSTRACT: The widespread use of artificial intelligence in language applications has created an urgent need for secure, accurate, and context-aware paraphrasing systems. Most existing paraphrasing solutions operate online, raising concerns regarding data privacy, dependency on cloud services, and inconsistent output quality. To overcome these limitations, this research introduces an offline Smart Paraphraser that integrates locally deployed GGUF-based Large Language Models with a semantic similarity ranking mechanism powered by Sentence-BERT. The system generates multiple paraphrased outputs, evaluates them using cosine similarity, and presents the most contextually accurate results to the user. With FastAPI serving as the backend and Streamlit providing an interactive user interface, the system ensures efficiency, privacy, and accessibility in various academic and professional use cases. Experimental evaluation demonstrates strong semantic retention, fluent paraphrasing, and reliable performance even in offline environments, positioning the solution as a practical advancement in secure and intelligent language processing.

I. INTRODUCTION

The growing demand for efficient rewriting tools in academic, research, and professional environments has highlighted the limitations of traditional online paraphrasing systems. These tools often rely on external servers to process user data, resulting in significant privacy concerns and restricted accessibility for users who require secure, offline functionality. Furthermore, many existing solutions focus heavily on surface-level word substitution, producing paraphrases that fail to preserve meaning or maintain grammatical coherence. This research addresses these challenges by designing a Smart Paraphraser capable of generating high-quality, semantically accurate paraphrased sentences entirely offline. The system incorporates advanced natural language processing techniques and modern AI architectures to ensure that users receive coherent, meaning-preserving paraphrases without compromising on data confidentiality.

The core objective of this project is to integrate a robust paraphrasing mechanism with a semantic evaluation layer that enables the system to understand, generate, and verify text in a human-like manner. By deploying optimized GGUF-based language models such as Mistral and TinyLlama through Ollama, the paraphraser operates effectively without the need for internet connectivity. This supports environments where confidentiality is crucial, including educational institutions and research organizations. The system's architecture combines a FastAPI backend with a Streamlit user interface, ensuring real-time interaction and smooth integration between AI components. The addition of a semantic similarity engine using Sentence-BERT strengthens the system's ability to preserve meaning, making it a reliable tool for various content transformation tasks.

© 2025 IJMRSET | Volume 8, Issue 11, November 2025| DOI:10.15680/IJMRSET.2025.0811056

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)

Smart–Paraphraser Wrofcog

II. LITERATURE REVIEW

A smart paraphraser is an intelligent text-to-text generation system designed to rewrite sentences while preserving their original meaning, and research in this area has evolved significantly over the last two decades. Early work focused on rule-based and statistical approaches using synonym substitution, phrase tables, and hand-crafted linguistic rules, but these methods often produced rigid or awkward outputs due to limited contextual understanding. With the rise of neural networks, sequence-to-sequence models using RNNs and attention mechanisms improved fluency by learning paraphrasing patterns directly from data. A major breakthrough came with large-scale datasets such as ParaNMT-50M, created using back-translation techniques, which enabled models to learn paraphrasing at massive scale. The transformer architecture further revolutionized paraphrasing: models like T5, BART, and GPT treat paraphrasing as a general text-to-text task, leading to highly fluent and semantically accurate outputs. Recent research focuses on controlled paraphrasing, where users can guide the output based on style, tone, formality, or length; this is achieved using constraint-based decoding, reinforcement learning, and prompt engineering. Modern studies also highlight challenges such as maintaining the balance between meaning preservation and lexical diversity, improving evaluation metrics since BLEU and ROUGE fail to capture semantic similarity, and ensuring document-level coherence rather than just sentence-level rewrites. New directions explore domain-specific paraphrasing, paragraph-level rewriting, and methods for improving factual accuracy through retrieval-augmented generation. Despite strong progress, ensuring that paraphrase models avoid plagiarism, preserve key information, and generate diverse yet faithful outputs remains an active research area. Overall, the literature shows a clear shift from rule-driven rewriting to data-driven, transformer-based paraphrasing systems capable of generating high-quality, controlled, and context-aware text.

III. METHODOLOGY OF PROPOSED SURVEY

The methodology of the proposed survey on smart paraphrasing systems adopts a structured and systematic approach to understand the evolution, techniques, and performance of modern paraphrasers. The survey begins with the identification of relevant research papers, books, and technical reports published in reputed sources such as IEEE, ACM, Springer, Elsevier, and arXiv to ensure comprehensive coverage of foundational as well as current advancements. A keyword-driven search strategy using terms like “paraphrase generation,” “text rewriting,” “neural paraphrasing,” “transformer-based paraphrasers,” and “semantic preservation” was used to curate a diverse set of literature. After collecting the sources, the survey employs a two-step filtration process: first, eliminating papers that do not directly address paraphrasing techniques or evaluation frameworks, and second, prioritizing studies that propose novel architectures, datasets, or metrics. The shortlisted works were analyzed based on specific criteria such as the underlying model architecture (rule-based, statistical, neural, or LLM-based), type of dataset used, evaluation methodology, degree of semantic fidelity, diversity of paraphrases generated, and applicability in real-world scenarios. To maintain uniformity, each paper was assessed using a comparative lens that focuses on strengths, limitations, and research gaps. Furthermore, the methodology includes synthesizing insights from existing surveys while identifying missing aspects such as document-level paraphrasing, controllability, and ethical considerations. The extracted insights were organized thematically to highlight trends, challenges, and future research opportunities. This systematic and analytical methodology ensures that the proposed survey provides a holistic, unbiased, and in-depth understanding of the current state of smart paraphrasing technologies and their evolution toward more intelligent and context-aware rewriting systems.

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. CONCLUSION AND FUTURE WORK

In conclusion, smart paraphrasers have evolved significantly from simple rule-based systems to advanced transformer-driven and LLM-based methods capable of generating fluent, semantically accurate, and context-aware paraphrases. The literature shows that modern paraphrasing models excel in fluency and coherence, but important challenges such as meaning preservation, controllability, evaluation accuracy, and domain adaptation still remain. Current research highlights the need for more robust semantic evaluation metrics, better handling of document-level paraphrasing, and improved mechanisms to prevent factual drift. Future work should focus on developing controllable paraphrasing frameworks where users can specify style, tone, length, or complexity. Another key direction is the creation of high-quality, diverse, and domain-specific paraphrase datasets to support fine-tuning of smart paraphrasers for specialized fields such as legal, medical, and academic content. Integrating retrieval-augmented generation and reinforcement learning may further enhance performance by grounding paraphrases in factual and contextual knowledge. Finally, research must also address ethical concerns, especially misuse for plagiarism or misinformation, by developing paraphrase detection and watermarking techniques. With continued innovation, smart paraphrasers can become more reliable, interpretable, and adaptable, enabling their use in education, content creation, writing assistance, and AI-driven communication systems.

REFERENCES

1. Tsai, Y.-C., & Lin, F.-C., *Paraphrase Generation Model Integrating Transformer Architecture, Part-of-Speech Features, and Pointer Generator Network*, IEEE Access, 2023. (DBLP entry / DOI available)
2. FastAPI Documentation – *FastAPI: The Modern, Fast (High-performance) Web Framework for Building APIs with Python 3.7+.*
3. Available at: https://fastapi.tiangolo.com/
   Streamlit Documentation – *Streamlit: The fastest way to build data apps in Python.*
4. Available at: https://docs.streamlit.io/
   Ollama – *Local LLM Model Runner and API for Offline AI Inference.*
5. Available at: https://ollama.com/
6. Sentence Transformers (SBERT) – *Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks.*
7. Published at *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP).* Available at: https://www.sbert.net/
8. Hugging Face Transformers Library – *Wolf, T., et al. (2020). Transformers: State-of-the-Art Natural Language Processing.*
9. Published in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP): System Demonstrations.* Available at: https://huggingface.co/transformers/
10. PyTorch Documentation – *PyTorch: An open-source machine learning framework for training and deploying deep neural networks.* Available at: https://pytorch.org/
11. MistRal and TinyLlama Model Information – *Open-weight LLMs for local text generation.* Available at: https://ollama.com/library/mistral and https://ollama.com/library/tinyllama

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com