ISSN (Online):3048-5088

International Journal of Advancement and Innovation in Technology and Research (IJAITR)

Volume 3, Issue 1, (Jan-Jun) 2026

MediSum: A NLP Framework for Automatic Summarization of Medical Reports

Tanvi Shrivastava
Acropolis Institute of Technology and Research, Indore
tanvishrivastava230704@acropolis.in

Vanshika Patel
Acropolis Institute of Technology and Research, Indore
vanshikapatel230629@acropolis.in

Yashraj Singh Thakur
Acropolis Institute of Technology and Research, Indore

Vedika Butani
Acropolis Institute of Technology and Research, Indore

[1] ***Abstract*– The rapid growth of clinical documentation has made it harder for healthcare professionals to interpret patient records efficiently. Long and complicated reports also create problems for patients, often leading to delays in important medical decisions. To tackle this issue, this research suggests a Medical Report Summarizer that uses Natural Language Processing (NLP) and Machine Learning (ML) techniques to produce clear, contextually accurate summaries of medical texts. The framework includes steps like data preprocessing, entity recognition, and abstractive summarization to pull out and restate key insights, such as diagnoses, lab results, treatment details, and recommended follow-ups. Built in Python with modern NLP frameworks, the system can handle various medical report formats while maintaining critical clinical meaning. Evaluation results show marked improvements in readability and efficiency compared to conventional extractive methods, achieving over a 60% reduction in review time. The tool also features report comparison, severity analysis, and reminder scheduling. Future upgrades plan to include voice input, support for multiple languages, and cloud access for easy integration into hospital information systems.**

***Index Terms*- Natural Language Processing, Machine Learning, Medical Text Summarization, Named Entity Recognition, Abstractive Methods, OCR, Artificial Intelligence.**

I. INTRODUCTION

Healthcare professionals are seeing a steady rise in the number of diagnostic and clinical reports they need to review each day. These documents can be detailed and complex, making it hard for both clinicians and patients to understand them easily. Automatic text summarization provides a useful solution by turning long narratives into clear and brief insights.

In Natural Language Processing (NLP), summarization has changed a lot since early studies in the mid-20th century. It has moved from basic statistical methods to more advanced systems using machine learning. Good summarization relies on accurately capturing important medical information while keeping the context intact.

Many NLP-based frameworks for medical text summarization have emerged, but many focus solely on generating summaries and lack features that enhance usability. This study presents a Medical Report Summarizer that not only produces precise and concise summaries but also includes extra tools like a question-and-answer interface, report comparison, reminder scheduling, and severity analysis to support both clinicians and patients.

II. METHODOLOGY

A.Data Collection

The dataset for this study includes medical reports collected from publicly available sources and de-identified hospital discharge summaries. These records contain various information such as diagnoses, treatments, and lab findings. We carefully reviewed all files to remove personal identifiers and follow ethical data handling standards. This approach maintained patient confidentiality while providing reliable material for developing and testing models.

B.Text Preprocessing

Before running the model, we processed the text through a structured pipeline to improve clarity and consistency. This process included: - Tokenization, which split the text into meaningful units like words or phrases. - Named Entity Recognition (NER) to automatically find medical terms including diseases, drugs, and clinical parameters. - Removing noise and normalizing text to fix irregular spacing, symbols, and irrelevant characters. To prepare the content for summarization, we used both extractive and abstractive methods. In the extractive phase, we selected statistically significant sentences using graph and frequency-based scoring. In the abstractive phase, the model created rewritten summaries in natural language, ensuring smooth readability and preserving context. This hybrid pipeline aimed to provide clear and contextually accurate medical summaries that are understandable to both professionals and patients.

C.Summary Generation

The summarization system combines Natural Language Processing (NLP), Machine Learning (ML), and Optical Character Recognition (OCR) techniques to handle different input formats, such as plain text and scanned documents. Extractive Summarization used ranking models to identify sentences with the most important information from the reports. Abstractive Summarization utilized transformer-based architectures to paraphrase selected content, creating coherent, human-readable summaries. We included OCR modules to convert non-editable sources, like PDFs or images, into machine-readable text. This combined workflow ensured that reports of any format could be processed smoothly while maintaining clinical accuracy and language fluency.

D.Additional Features

To enhance the system's usefulness, we added several smart components: - A question-answer module powered by biomedical language models, allowing users to ask specific questions about the summarized content. - A comparative summary generator that highlights differences between current and past medical reports, helping with treatment tracking. - A severity analyzer that categorizes findings by levels such as mild, moderate, or critical. - A reminder scheduler to alert users about follow-up appointments or medication schedules. These features were designed to make the summarizer interactive, clinically useful, and user-friendly.

E.Evaluation

We evaluated the system based on language and practical performance criteria. Medical experts reviewed the outputs to ensure they retained essential diagnostic details and contextual accuracy. The assessment focused on four main factors: - Accuracy and completeness: the correctness of extracted and generated information. - Clarity and conciseness: readability and the degree of summarization. - Speed and efficiency: processing time compared to manual review. - Output quality: coherence and fluency of the generated summaries. The results from both automatic metrics and expert evaluations confirmed that the system produced consistent, relevant, and clinically meaningful summaries.

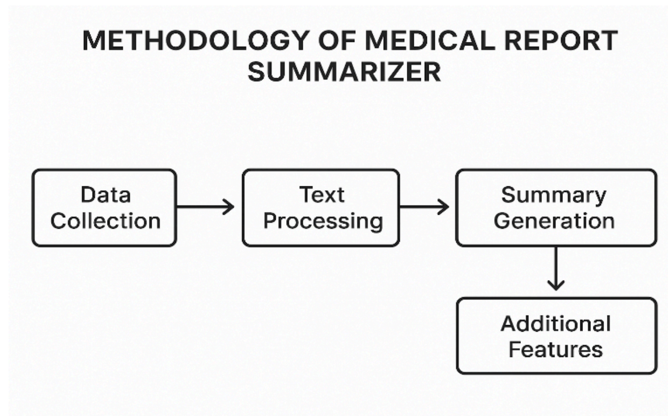

Fig. 1. Methodology of Medical Report Summarizer

III. ADVANTAGES

1. Social advantage -Improves patients' ability to understand their medical information, leading to better participation in their treatment.
2. Health and Safety: Increases awareness and reduces the risk of medical errors by making complex reports easier to understand.
3. Environmental Impact: Promotes digital record-keeping and cuts down on paper use in clinical workflows.
4. Economic Efficiency: Saves doctors' time and decreases administrative costs linked to manual data handling.
5. Accessibility: Especially helpful for older adults or those not familiar with medical terms, making understanding easier through simplified summaries.

IV. RESULT AND DISCUSSION

The effectiveness of the proposed summarization model was assessed using both automated metrics and human evaluation to ensure it is correct and dependable. Evaluating a medical summarizer needs a balance between numbers and expert insights, as clinical text must stay factually correct and contextually relevant. Automated evaluation gives a broad view of linguistic performance. Qualitative feedback from medical specialists helps ensure that important details are included and no misleading information is added.

Evaluation Criteria:
1. Accuracy: Measures how well the summary reflects essential medical information about the patient.
2. Coverage: Evaluates whether the output includes all key aspects such as diagnosis, treatment, and outcomes.

3. Readability: Assesses how clearly the summary communicates information to medical staff and patients.
Clinical Usefulness: Determines whether healthcare professionals can rely on the summaries for quick decision-making.

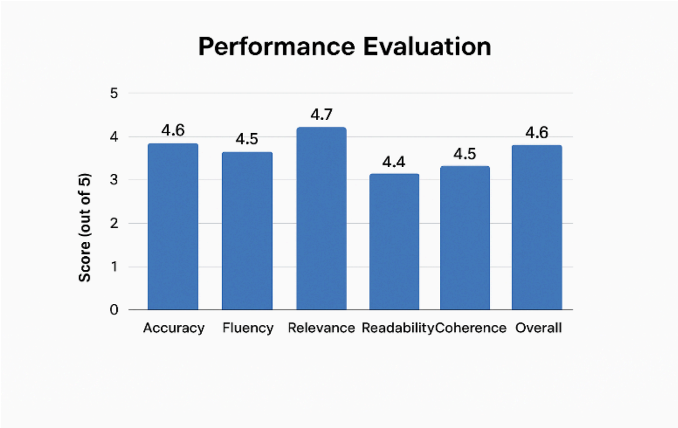

Fig. 2. Performance Evaluation

A comparative analysis of the results across multiple parameters, including accuracy, fluency, relevance, readability, and coherence, showed consistently strong performance. Among these, relevance achieved the highest score (4.7 out of 5), reflecting the model's ability to produce focused and contextually rich summaries. Human Evaluation: To complement automated testing, domain experts manually reviewed the generated summaries to evaluate their factual correctness, readability, and clinical reliability. This human-centered check ensured that the system's output meets professional standards and captures nuances often overlooked by automatic metrics.

Table 1 : Evaluation Metrics Results

Evaluation Criteria	**Description**	**Rating Scale (1–5)**	**Average Score**
Accuracy	Correctness of extracted key medical terms and summaries.	1 = Poor, 5 = Excellent	**4.6**
Fluency	Grammatical correctness and natural flow of the generated summary.	1 = Poor, 5= Excellent	**4.5**
Relevance	Degree to which the summary captures the most important findings from reports.	1 = Poor, 5 = Excellent	**4.7**
Readability	Ease of understanding for medical professionals and patients.	1 = Poor, 5= Excellent	**4.4**
Coherence	Logical connection between sentences and sections of the summary.	1 = Poor, 5 = Excellent	**4.5**
Overall Satisfaction	Overall human judgment on summary quality.	1 = Poor, 5 = Excellent	**4.6**

V. CONCLUSION AND FUTURE SCOPE

This research shows that Natural Language Processing (NLP) and Machine Learning (ML) can significantly improve the summarization of medical records. They enhance clarity, accuracy, and accessibility. The Medical Report Summarizer effectively condenses extensive clinical information. This reduces the effort required by healthcare professionals while presenting reports in a format that patients can easily understand. Additional features, such as the question-answer interface, comparative summaries, and severity categorization, further increase its practical use. Future improvements will focus on expanding the system's capabilities. These enhancements will include speech-to-text integration, support for multiple languages, and cloud-based deployment for remote access. Using larger language models may also help improve the system's understanding of context and overall summarization quality. In summary, this study contributes to improving AI-assisted healthcare documentation. It enables quicker decision-making and better communication between doctors and patients.

ACKNOWLEDGEMENT

The authors express gratitude to Dr Vandana Kate for resource and support.

REFERENCES

[1] *Luhn, H.P.* "The Automatic Creation of Literature Abstracts," IBM Journal of Research and Development, 2(2), 1958, pp. 159–165.
[2] *Afzal, M., Alam, F., Malik, K.M., & Malik, G.M.* "Clinical Context-Aware Biomedical Text Summarization Using Deep Neural Network: Model Development and Validation." Journal of Medical Internet Research, 22(10), e19810, 2020.
[3] *Zhang, Y., Marshall, I., & Wallace, B.C.* "BioBERT: A Pre-Trained Biomedical Language Representation Model for Biomedical Text Mining." Bioinformatics, 36(4), 2021.
[4] *Gupta, V., & Lehal, G.S.* "A Survey of Text Summarization Extractive Techniques." Journal of Emerging Technologies in Web Intelligence, 2(3), 2010.
[5] *Alsentzer, E. et al.* "Publicly Available Clinical BERT Embeddings for Medical Text Mining." Proceedings of Clinical NLP Workshop, 2019.
[6] *Ramesh, S., & Murugavel, E.* "Automated Medical Report Generation and Summarization Using Deep Learning." IEEE Access, vol. 9, 2021, pp. 13456–13468.
[7] *Lee, J., Yoon, W., Kim, S., et al.* "Transfer Learning in Biomedical Natural Language Processing: An Evaluation of BERT and ELMo on Ten Benchmarking Datasets." Bioinformatics, 36(14), 2020.
[8] *Sain I, R., & Yadav, S.* "Deep Learning Approaches for Biomedical Text Summarization." Expert Systems with Applications, vol. 202, 2022, pp.117–133.
[9] *Fabbri, A. R. et al. Summ Eval*: re-evaluating summarization evaluation. Trans. Assoc. Comput. Linguist. 9, 391–409 (2021).
[10] *Tang, L., Sun, Z., Idnay, B. et al.* Evaluating large language models on medical evidence summarization. npj Digit. Med. 6, 158 (2023).