International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 4, 2025

Medical report summarization using NLP

[1] Kagitha Monasri, [2]Amara Sri Sai Sudha Amrutha Varshini, [3]Jaggarapu Venkat, [4]Mrs. D. Madhuri

[1,2,3]U. G Student, Dept COMPUTER SCIENCE AND ENGINEERING, St.Ann's College Of Engineering and Technology, Nayunipalli (V), Vetapalem (M), Chirala, Bapatla Dist, Andhra Pradesh – 523187, India

[4]Assistant professor, COMPUTER SCIENCE AND ENGINEERING, St.Ann's College Of Engineering and Technology, Nayunipalli (V), Vetapalem (M), Chirala, Bapatla Dist, Andhra Pradesh – 523187, India

ABSTRACT

This project presents an AI-powered solution for automating medical report summarization using Natural Language Processing (NLP), designed to reduce the documentation workload for healthcare professionals. advanced transformer models and domain- specific medical knowledge, the system extracts key clinical information from complex medical texts and generates concise, accurate summaries to true and exact information. It supports multi- modal data integration like text and lab results and enables real-time report generation to accelerate clinical decision making. The entire process is secured, ensuring patient data privacy while enhancing efficiency and accuracy in healthcare documentation.

***Key Words**: Diagnosis, Medications, Procedures, Symptoms, Tests, Recommendations.*

INTRODUCTION

Currently, medical report summarizing with Natural Language Processing (NLP) is a revolutionary method in healthcare that tackles the difficulty of deriving significant insights from extensive unstructured clinical material. Critical patient data is frequently hidden behind long narratives in medical records like radiology reports, discharge summaries, and doctors' notes. Automated systems can convert long reports into concise, well-structured summaries that highlight important data like diagnosis, treatments, and medications, thanks to NLP approaches, particularly extractive and abstractive summarizing. These systems can precisely comprehend complicated medical terminology and context by utilizing domain-specific language models like natural language processing (NLP), guaranteeing that the summaries maintain clinical relevance and accuracy. The integration of NLP-based summarization into healthcare workflows offers significant
benefits, including improved clinical decision-making, reduced documentation burden, and enhanced accessibility of patient data. Physicians can quickly review summarized patient histories, leading to faster and more informed diagnoses and treatment plans. Additionally, NLP tools support electronic health record optimization by organizing and structuring free text data, making it easier to retrieve and analyze. As healthcare systems increasingly adopt AI-driven solutions, medical report summarization stands out as a key innovation that not only streamlines operations but also contributes to better patient outcomes through timely and accurate information delivery.

LITERATURE SURVEY

[1] Recent years I have explored some related papers to medical report summarization. In that I got some limitations. To check papers, I studied recent papers. In the first paper is Patient's Medical History Summarizer using NLP by
D. Dharrao. But in this paper, there is no extracting the content like procedures, tests separately. In the second paper Adapted Large Language Models for Medical in Clinical Text Summarization by D Van Veen 2023.[2] But in this paper there is no highlighted text in the summarization. In the Third paper is pre-trained language models with domain knowledge for biomedical extractive summarization By Q. Xie et 2023.but not summarize the different medical report what they trained only can summarized.

RELATED WORK

AI-driven medical report summarization offers a transformative solution to the overwhelming volume of clinical documentation in modern healthcare. By leveraging Natural Language Processing (NLP), this technology can automatically extract and condense critical information such as diagnoses, medications, procedures, symptoms, tests, and recommendations from complex medical texts. [3] This not only saves time for healthcare professionals but also enhances accuracy, consistency, and accessibility of patient data. With support for multiple file formats (PDF, Word, TXT), real-time processing, and export capabilities, AI-powered summarization tools streamline workflows, reduce cognitive load, and empower faster, more informed decision-making in clinical settings.

EXISTING SYSTEM

Currently, I choose the paper Adapted Large Language Models for Medical in Clinical Text

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 13, Issue 4, 2025

Summarization by D Van Veen 2023 [4] where there is no highlighted data in the summarized data. from these the existing system should be extracting
information which are procedure, symptoms, medication, procedures, tests, recommendations etc., and also give highlighted content in summarization data. Retrieving and analyzing relevant patient data quickly which are at the history. Important information may be hidden in lengthy, unstructured text.[5] so we may have to gives the highlighted data and makes the structured format as easily understanding way and also patients can directly get the information easily understanding way to the patients doesn't directly interact with the medical staff.

In these proposed a Medical Report Summarization using NLP system uses artificial intelligence techniques to automatically provide succinct summaries of long medical records. [6] Python (Flask) is used for the backend of the system, while HTML, CSS, and JavaScript are used for the front end. It makes use of Natural Language Processing techniques including rule-based key information extraction and TF-IDF-based extractive summarization to find and emphasize important information such as test results, prescriptions, and diagnoses. Via an intuitive web interface, users can input [7] medical records in text or PDF format, and the system will process the data to provide a comprehensible summary. Users can also see or download their summary, retrieve their prior reports.

PROPOSED SYSTEM

SYSTEM ARCHITECTURE

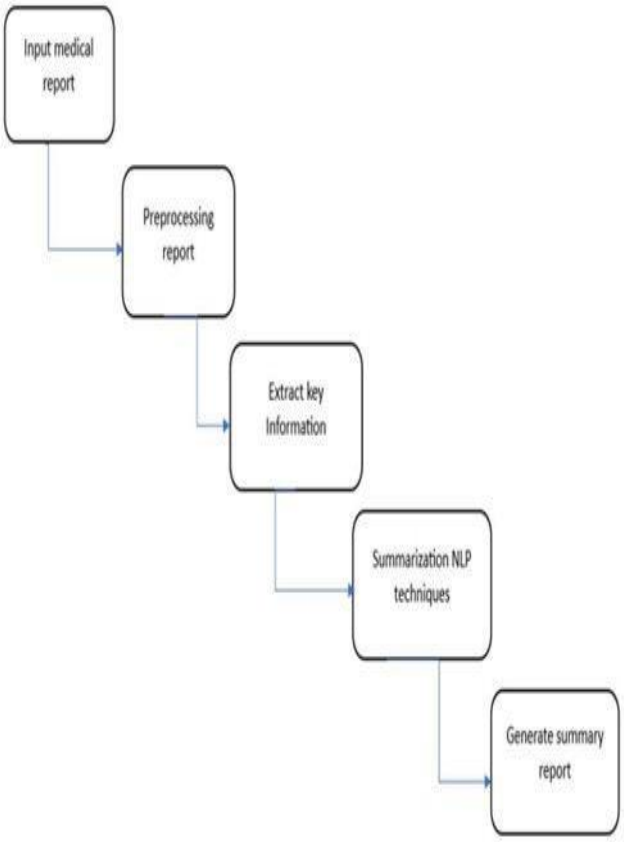

Fig:1 Architecture of medical report summarization using NLP

METHODOLOGY DESCRIPTION

Preprocessing: It should be clean the data and makes to summarize the lengthy context of medical report summarization Extract key Information: It extract the key information which is diagnosis, symptoms, diagnosis, recommendation.
Summarization NLP Techniques: By using NLP Techniques, the TF-ID based on extract information. makes lengthy document report into the summary way.
Displays summary Report: It displays the medical report summarize data and we can download in the format of pdf document. We can see the history of the documents.

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 4, 2025

RESULTS AND DISCUSSION

Fig :2 Home page of the medical report website

It shows the webpage of medical report summarization which shows the basic webpage. it gives the fully information of the website at home page. In this click on try now or get started it goes to another page.

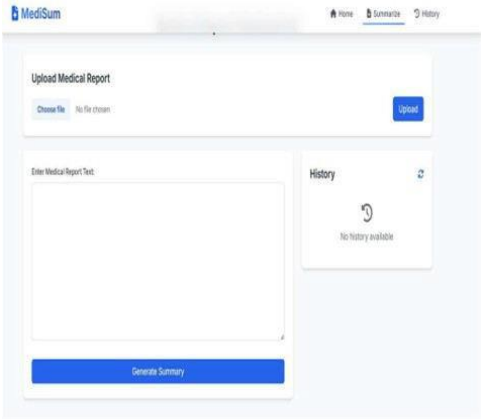

Fig :3 Uploading a medical report

In this page, click on Choose File and upload your medical report from your device. Once selected, the system will automatically process the document and generate a summary.

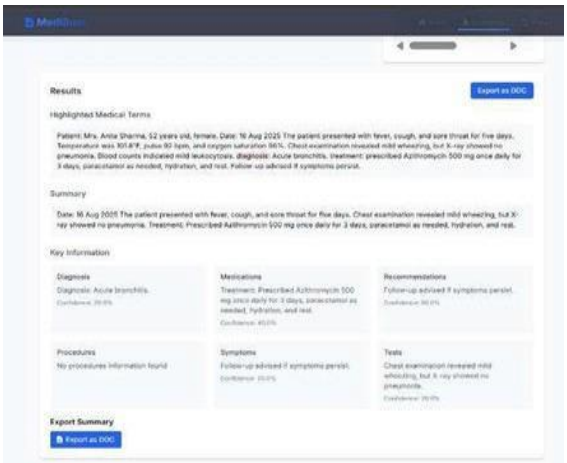

Fig 4: Displays the summary of the medical report

It displays the lengthy report as a summary of the medical report, which may have highlighted the content and extracted key information.

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 13, Issue 4, 2025

CONCLUSION

Reflect AI-driven medical report summarization is a powerful solution to one of healthcare's most persistent challenges: Automating and intelligently condensing clinical notes empowers physicians to focus more on patient care while maintaining high standards of accuracy. It plays a vital role in healthcare, but current practices often lead to time loss, errors, and communication gaps. Unstructured and inconsistent records make it difficult to access important patient information quickly.

FUTURE SCOPE

The future enhancement is Integra on of medical imaging analysis, multi- language medicalsummarization Voice-to-text summarizeon,Multi Document Summarization at a time.

REFERENCE

[1] Harini, D. P. (2013). Scalable and Secure Sharing of Client Medical Records in Cloud. IJCST.

[2] Roshni-1 – NLP-Based Medical Record SummarizaƟon and InformaƟon ExtracƟon • John Snow Labs – Medical Text SummarizaƟon with Spark NLP • AnalyƟcs Vidhya – ExtracƟng Medical InformaƟon from Clinical Text with NLP

[3] Sheshasaayee and Logeshwari – Integrated RFM and LTV SegmentaƟon with Neural Networks

[4] Journal of Physics: ANCOSET 2020 – Data Mining in Healthcare

[5] AI in Healthcare: Simplifying Medical Reports – Rahul Prasath S et al. (IJIRT, 2025)

[6] Medalyze: Lightweight Medical Report SummarizaƟon – Van-Tinh Nguyen et al. (arXiv, 2025) • Biomedical Lay SummarizaƟon – Team MIRAGE (BioNLP Workshop, ACL 2025)

[7] Clinical Text SummarizaƟon with LLM-Based EvaluaƟon – Stanford CS224N Project (2024)

[8] LC Karunya, P. H. (2019). Emergency Alert Security System for Humans. Int. J. Commun. Comput. Technol, 1-5.