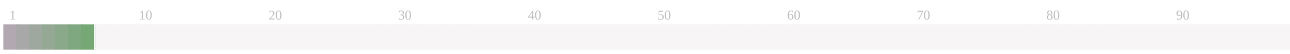
The Report is Generated by DrillBit Plagiarism Detection Software

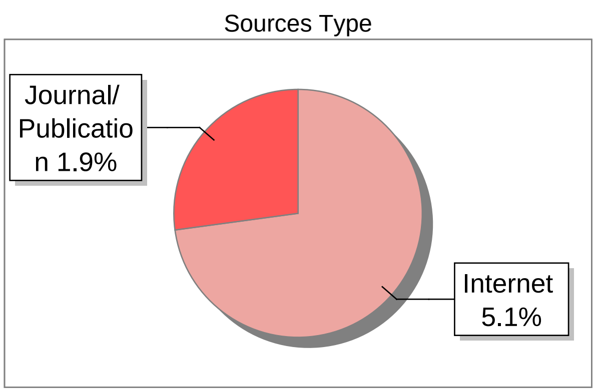
Submission Information

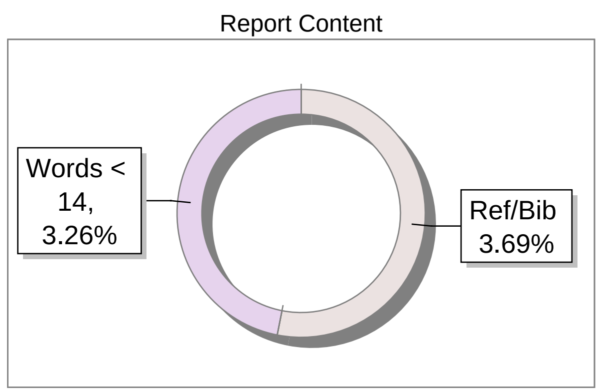
Author Name	GROUP_06_Sreekar
Title	Mini_Project_Group_6_ChatBOT_For_EHR_website
Paper/Submission ID	2536116
Submitted by	nataraj@iiitdwd.ac.in
Submission Date	2024-11-17 21:54:49
Total Pages, Total Words	15, 2057
Document type	Research Paper

Result Information

Similarity **7 %**







Exclude Information

Quotes	Not Excluded
References/Bibliography	Not Excluded
Source: Excluded < 14 Words	Not Excluded
Excluded Source	**0 %**
Excluded Phrases	Not Excluded

Database Selection

Language	English
Student Papers	Yes
Journals & publishers	Yes
Internet or Web	Yes
Institution Repository	Yes

A Unique QR Code use to View/Download/Share Pdf File

DrillBit

DrillBit Similarity Report

7	10	A
SIMILARITY %	MATCHED SOURCES	GRADE

A-Satisfactory (0-10%)
B-Upgrade (11-40%)
C-Poor (41-60%)
D-Unacceptable (61-100%)

LOCATION	MATCHED DOMAIN	%	SOURCE TYPE
1	www.nature.com	2	Internet Data
2	vinija.ai	1	Internet Data
3	Analytical efficiency evaluation of a network mobility management pro, by Nerea Toledo, Mariv- 2012	1	Publication
4	nature.com	1	Internet Data
5	translational-medicine.biomedcentral.com	1	Internet Data
6	Distinguishing between mathematics classrooms in Australia, China, Japan, Korea by Davi-2008	1	Publication
7	Bloom-filter-based request node collaboration caching for named data n by Hou-2018	1	Publication
8	omicstutorials.com	<1	Internet Data
9	ej-math.org	<1	Internet Data
10	epdf.pub	<1	Internet Data

Mini Project Report

on

Chatbot for EHR Website

Submitted by

Challa Sreekar Reddy 21BEC012

Nomula Vamshi Madhav 21BEC028

P.Sai Chandra Lohith 21BEC030

Under the guidance of

Dr Nataraj KS

DEPARTMENT OF ELECTRONICS AND COMMUNICATIONS ENGINEERING

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY DHARWAD

16/11/2024

Certificate

This is to certify that the project, entitled **Chatbot for EHR Website**, is a bonafide record of the Mini Project coursework presented by the students whose names are given below during 7th Semester of 4th year in partial fulfilment of the requirements of the degree of Bachelor of Technology in Electronics and Communication Engineering.

Roll No	Names of Students
21BEC012	Challa Sreekar Reddy
21BEC028	Nomula Vamshi Madhav
21BEC030	P.Sai Chandra Lohith

Dr Nataraj K S
(Project Supervisor)

Contents

List of Figures

1 Introduction

Electronic health record (EHR) chatbots are considered a major development in healthcare technology by using artificial intelligence (AI) and natural language processing (NLP) to enhance patient-provider interactions and facilitate healthcare delivery. Integrating chatbots within EHR are used by healthcare providers to provide more direct and timely modalities for interacting with patients by helping them access their health information, book appointments and get personalized medical advice thus improving both care coordination and patient satisfaction.

Chatbots have gained fame in the healthcare industry for the same reasons they became popular across all tech sectors; They can automate certain aspects of efficiency (remember, automated scheduling), and they offer 24/7 support to an increasingly stretched-to-capacity client base. Studies show that chatbots can support medication adherence, help monitor symptoms and provide patient education and as a result improve health outcomes while lessening the burden on healthcare systems. At the same time, this rapid embrace of chatbots has led to a greater conversation around issues related to data privacy, security and the trustworthiness of information given to patients – highlighting significant demand for enforceable regulatory frameworks.

However, the incorporation of chatbots into EHR systems poses some challenges despite their benefits. Interoperability of various healthcare platforms, medical fact checking, and patient privacy and security are critical issues that must be addressed. Even with the variation in UX, we need further development for AI-powered chatbots to keep their users engaged which means constant testing and iterations on design and usability.

In summary, EHR chatbots represent a paradigm shift in servicing healthcare delivery that could offer healthcare systems exciting new advantages in terms of patient engagement and health or disease management while also presenting novel challenges when it comes to privacy, security needs and ethics of digital (health) practice.

2 Related Work

Large language Models (LLMs) have paved new directions in diverse domains, one of them is healthcare. In this section, we provide an overview of the literature that is relevant to chatbot generation in electronic health records (EHR).

Foundation Models :

The general-purpose LLMs such as ChatGPT and GPT-4 have set the stage for domain-specific LLMs. They are pretrained on huge datasets of text and can generate and understand impressive amounts of human text, making them a base for fine tuning domain-specific models. And open-source models like LLaMA 2,3, Vicuna, Falcon and Mistral make these powerful tools even more broadly available.

Transforming Medical LLMs:

Researchers have created medical LLMs that build on top of the foundation models by pretraining them further using large and diverse medical datasets. As an example, for both FlanPaLM, it applies instruction tuning that better serves medical question answering. This is in addition to the promising results other previous models including Med-PaLM and Med-PaLM 2 have demonstrated when applied to different types of clinical tasks, with a particular emphasis on large corpora of real-world medical data

Open-Source Medical LLMs:

There has been a lot of buzz around open-source models in general, and many of contributions had come from the open-source community to develop medical LLMs. After training on Chinese medical datasets, ClinicalGPT shows good performance in simulations for applications like medical conversation, examination, diagnosis or question answering.BioMistral capitalizes on the Mistral architecture and extra pre-training on PubMed Central, adopting a new strong open-source portrait for an all-purpose Biomedical model. Other examples of well-known open-source models are Aloe, MedAlpaca, ChatDoctor, BioMedGPT-LM-7B

Improving Model Performance:

Even with the improved architecture there is still room to improve both performance and efficiency; merging models, for example, or further quantising remained potential areas well-trodden by past literature. Merging models such as DARE-TIES are one way to combine the knowledge from multiple of them, leveraging complimentary aspects of their weaknesses in order to create a more holistic and stable understanding of the medical domain. Thus, Quantization is effective in reducing the model size so that models can be deployed on resource-constrained devices.

Limitations and Future Directions:

10 Great strides had been made, more limitations must be addressed for the application of LLMs in health care. This potential for factual inaccuracies, logics mistakes and hallucinations will need to be mitigated. Data privacy and security are crucial with sensitive patient information to protect. Similar Work: Future work should seek to address these limitations, create optimized training and trainable-bodied datasets, and deliver sound evaluation techniques so as to ensure responsible and ethical deployment.

3 Methods

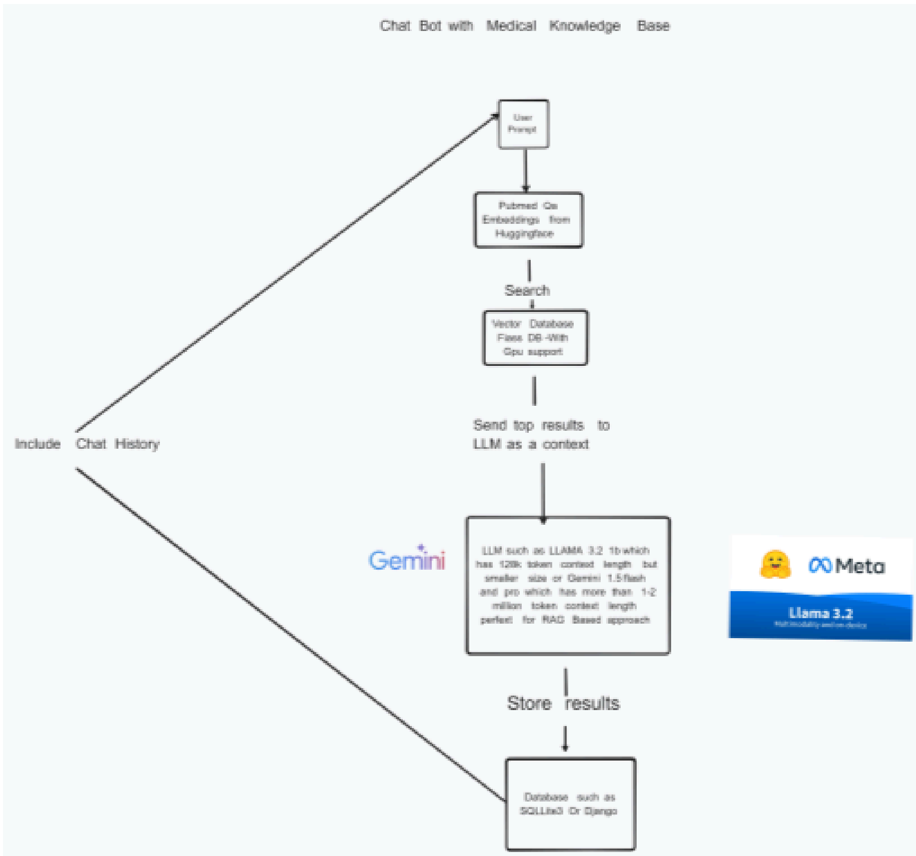


Figure 1. Chat Bot Architecture

User Prompt Processing

It starts with capturing the user prompt (all possible questions and queries for medical information and scheduling request).
The prompt is then passed through the pipeline, which can integrate knowledge bases and processing tools to generate factually correct responses and recommendations.

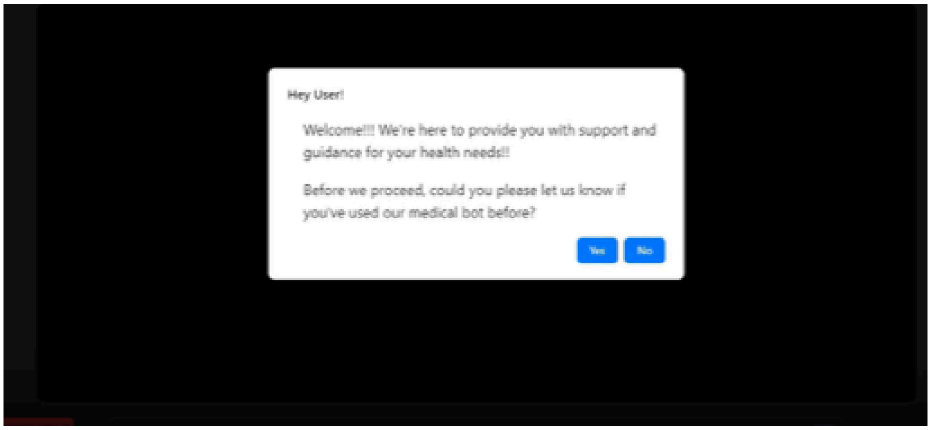


Figure 2. Starting Chatbot

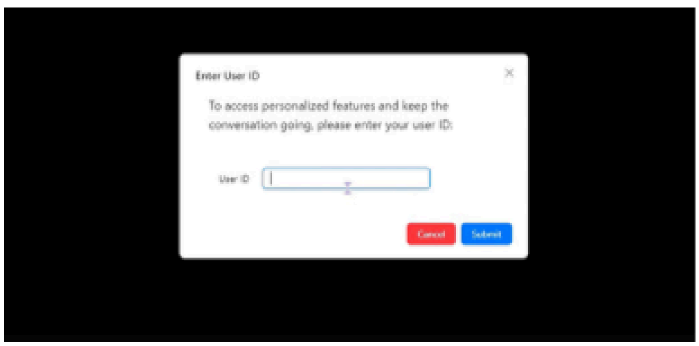


Figure 3. User Login

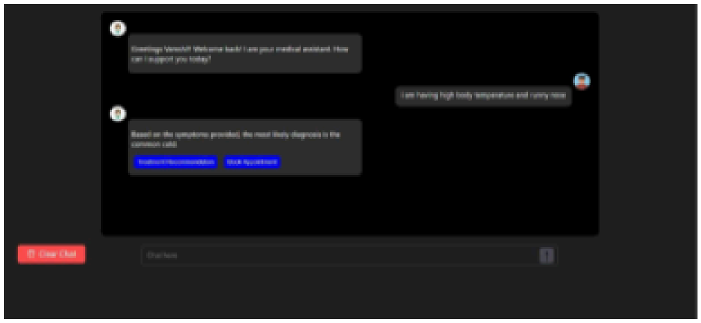


Figure 4. ChatBot Intro

Vector Database with GPU Support:

The embeddings are used to search a vector database, powered by FAISS (Facebook AI Similarity Search) with GPU acceleration. This allows for efficient and rapid retrieval of relevant documents or medical knowledge snippets.

Result Retrieval:

The top-matching results from the vector database are selected to be passed to the Large Language Model (LLM) as context.

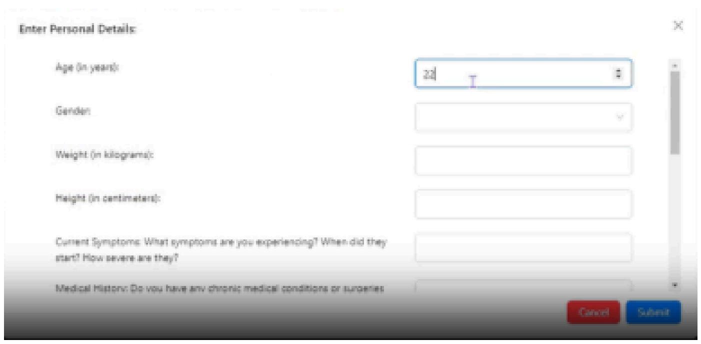


Figure 5. Treatment Recomendations

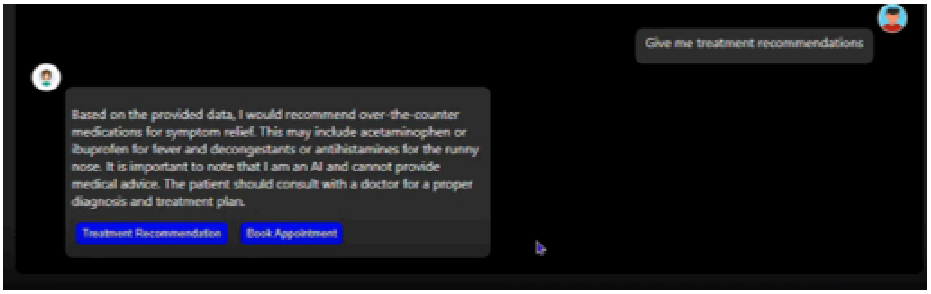


Figure 6. Recommended Treatment

LLMs and their Contextual Integration

These curation results are top-rated and used to provide context based on LLM e.g., 128k token context length for LLaMA 3.2 1B, or if it were until community is decided more later–a potentially better option would be Gemini 1.5 (Flash/Pro), as offers not just 128k but also up a million excess further in total for the same purpose and that seems productive enough here too. These type of 9 models are more suitable for RAG, where we use documents we retrieved to help develop the response.

Using both the controlled synthesis of information from retrieved medical documents and ongoing conversation history, the model can formulate a coherent and contextually relevant response.

Incorporating Chat History

Providing chat history to keep context so that responses are relevant. This enables a conversational experience where the bot can remember past messages and have an ongoing conversation with the user.

Response Storage

The generated responses as well as any relevant data (e.g. MIMIC identifiers) are stored in a structured database format, either SQLite3 or the Django database system. This storage allows one to log and keep track of the interactions – both for future use and for analysis but also creates an organized ledger of providing medical advice or booking.

Scheduling Based on Symptoms

Aside from the medical Q&A, the chatbot also offers a doctor scheduling function where it will recommend appointments according to the user symptoms.
The bot evaluates the symptoms mentioned by the user and recommends a doctor or specialist.

Figure 7. Recommended Doctors

Google Calendar Integration

After an appointment is fixed, the system interfaces with the Google Calendar API and creates a user calendar event. This allows the doctor and patient to keep the appointment details in sync, and sends reminders so that they arrive on time.

4 Conclusion

Advancements in LLM Technology:

Rapid advancement of LLM technology (general and medical-specific) has set the stage for implementation of EHR chatbot ChatGPT and GPT-4 are examples of general-purpose large language models (LLM) that have demonstrated exceptional abilities for natural language understanding and generation. These particular medical LLMs, FlanPaLMMoreover, Med-PaLM and Med-PaLM 2 have proven their capabilities in answering complex medical requirements such as QA, diagnosis and summarization. ClinicalGPT, BioMistral, MedAlpaca, ChatDoctor, BioMedGPT-LM-7B and MediTron-7B are just some of the open-source models that have expanded access to the technology and empowered researchers and developers alike to iterate on these improvements.

Potential Uses for EHR Chatbots:

There were many exciting opportunities to leverage the potential role of large language models in clinical workflows through their application to EHR systems, and thus improving patient outcomes. EHR chatbots can be used for.

Management of data:

Automating clerical duties like entering, retrieving, or summarizing data for administrative parts cutting the burden and increasing efficiency.

Patient education:

Offer patients personalized information regarding their conditions, medications, and treatment options, thereby improving engagement and understanding of care.

Improving Communication:

by enabling a communication platform between healthcare providers and patients to easily book an appointment, request refills as well as send in queries. However, the potential of EHR chatbots is enormous, and several challenges will have to be solved before widespread adoption.

Accuracy and Reliability:

It is of utmost importance to ensure the accuracy and reliability of information provided by these chatbots as medical data is very sensitive in its nature. More work must be done to guard against factual errors, logical flaws, and hallucinations.

Data Privacy and Security:

Privacy preservation, including how sensitive health data are handled securely To meet regulations and be ethical such measures must be implemented stringently

Transparency and Explainability:

Knowing how a chatbot comes up with its responses is crucial for trust and accountability. Creating a verifiable and explainable reasoning for LLM decisions is being explored by researchers.

In summary, the usage potential of LLMs for EHR. As a result, the adoption of less-than-human designovers over EHRsystems embraces beyond this through more elements to integrate them intoclinical workflows through developing sophisticated chatbots that canre-evolutionalize how patientcare is being performed and profoundly transfor mthehealthcare landscape. It is important that research and development continues in this area, even more imperative that the ethical and safety factors surrounding its uses were taken into consideration to ensure its safe use.

References

- A Study of Generative Large Language Model for Medical Research and Healthcare
- MedTrinity-25M: A Large-scale Multimodal Dataset with Multigranular Annotations for Medicine
- ClinicalGPT: Large Language Models Finetuned with Diverse Medical Data and Comprehensive Evaluation
- Aloe: A Family of Fine-tuned Open Healthcare LLMs
- Capabilities of Gemini Models in Medicine
- BioMistral: A Collection of Open-Source Pretrained Large Language Models for Medical Domains
- MED42-V2: A SUITE OF CLINICAL LLMS