AI Doctor - A Specialized Medical Chatbot

Contents

Executive Summary

The AI Doctor project represents an innovative application of artificial intelligence in the healthcare domain. Designed to provide users with accurate and relevant medical information, the AI Doctor functions as a specialized chatbot capable of addressing a range of medical queries. This report outlines the development, methodology, and impact of the AI Doctor, highlighting its potential to enhance user access to medical information and its alignment with ethical and regulatory standards.

Introduction

In the era of digital transformation, artificial intelligence (AI) has made significant strides in various domains, including healthcare. The AI Doctor project is a sophisticated application designed to provide medical information and guidance through an AI-powered chatbot. This project leverages a combination of advanced machine learning models, natural language processing techniques, and a comprehensive dataset to deliver accurate and relevant medical advice to users. The core of the system is built on a Retrieval-Augmented Generation (RAG) framework, which integrates both retrieval-based and generation-based approaches to ensure high-quality responses.

The bot aims to assist users in getting general medical advice by analyzing their questions, retrieving relevant medical information, and delivering concise yet informative responses. It is not intended to replace medical professionals but to act as a first point of contact for medical inquiries.

Objectives

The primary objectives of the AI Doctor project are to develop an AI-powered chatbot capable of interacting with patients and providing medically relevant information. The system is designed to:

1. **Interpret and Understand Symptoms**: Using natural language processing (NLP) techniques, the AI Doctor can comprehend patient descriptions of symptoms and health concerns.

2. **Provide Preliminary Diagnoses**: By referencing a vast database of patient-doctor dialogues and medical literature, the AI Doctor can offer possible diagnoses.

3. **Recommend Treatment Options**: Based on the diagnosed condition, the AI Doctor suggests treatment protocols and medications.

4. **Ensure Ethical Compliance**: The system is built with strict adherence to medical ethics, ensuring it advises patients to seek professional medical advice.

5. **Enhance Accessibility**: The AI Doctor aims to provide medical information to underserved populations with limited access to healthcare professionals.

Datasets Used

The foundation of the AI Doctor system lies in its access to a comprehensive and diverse set of datasets. The primary sources of data include:

1. **Patient-Doctor Dialogues**: Several datasets from Hugging Face were used, such as "HealthCareMagicWithSummary-100k," [1] "HealthCareMagic-100k-Chat-Format-en," [2] and others. [3] [4] These datasets contain extensive dialogues between patients and doctors, offering valuable insights into medical consultations.

2. **Medical Literature**: Additional datasets such as "Disease_Database" [5] were incorporated to provide information on various diseases, symptoms, and treatments.

3. **SOAP Summaries**: The "medical-dialogue-to-soap-summary" [6] dataset was utilized to structure conversations in the SOAP (Subjective, Objective, Assessment, and Plan) format, which is standard in medical documentation.

Methodology

Data Preprocessing

To create a unified and coherent dataset, the following steps were undertaken:

- **Data Cleaning**: The raw data was inspected for inconsistencies, missing values, and duplicates. Necessary corrections and adjustments were made to ensure data integrity.

- **Data Merging**: The various datasets were concatenated, ensuring a seamless blend of patient-doctor dialogues, disease information, and treatment protocols.

- **Text Chunking**: The dialogues were chunked into manageable pieces of text.

Embedding Generation

The core of the AI Doctor's understanding capabilities is built on sentence embeddings, which are vector representations of text that capture semantic meaning. The following processes were implemented:

- **Sentence Transformers**: The **"neuml/pubmedbert-base-embeddings"** model, a variant of BERT fine-tuned on biomedical literature, was employed to generate embeddings for each chunk of text.

- **Batch Processing**: Given the large volume of data, embeddings were generated in batches to optimize memory usage and processing time.

Vector Indexing

To facilitate quick and accurate retrieval of relevant medical information, the AI Doctor uses a vector database. The steps involved include:

- **Pinecone Integration**: The embeddings were stored in a Pinecone vector database, allowing for efficient similarity searches.
- **Upserting**: Data was upserted in the Pinecone index, ensuring that each text chunk was indexed with a unique identifier and its corresponding embedding.
- **Query Mechanism**: When a patient query is received, the AI Doctor generates an embedding for the query and searches for the most similar chunks in the Pinecone index, retrieving the top results.

Retrieval-Augmented Generation (RAG) Pipeline

The AI Doctor employs a Retrieval-Augmented Generation (RAG) pipeline, which combines retrieval and generation capabilities:

- **Retrieval**: The most relevant text chunks are retrieved from the Pinecone index based on their similarity to the patient query.
- **Cross-Encoder Scoring**: The retrieved chunks are re-ranked using a cross-encoder model, which scores them based on their relevance to the query.
- **Response Generation**: The retrieved relevant chunks of information are combined to form a context and is used by a language model like Llama3-8b to generate a coherent and accurate response. Prompt engineering is employed to guide the model's focus on the most relevant medical details, ensuring that the responses provided are both contextually appropriate and aligned with the user's query.

User Interface and Interaction

The AI Doctor is designed with a user-friendly interface:

- **Streamlit Frontend**: The project uses Streamlit to create a simple and intuitive user interface where patients can enter their queries and view responses.
- **PDF Upload**: Users can upload PDF documents containing medical reports, and the AI Doctor will extract and analyze the text to provide relevant advice.
- **Chat History**: The interface maintains a chat history, allowing users to review previous interactions and continue the conversation seamlessly.

Challenges

Data Quality and Consistency

One of the significant challenges faced during the project was ensuring the quality and consistency of the data. Medical data often comes from various sources, each with its format and structure. Merging these datasets required careful handling to prevent data corruption and ensure consistency.

Ethical Considerations

Providing medical advice through an AI system carries significant ethical responsibilities. The AI Doctor must avoid giving potentially harmful advice and should always recommend seeking professional medical attention. This requires the system to be built with robust ethical guidelines.

Real-Time Performance

Given the large volume of data and the complexity of the queries, ensuring real-time performance was a challenge. The system needs to process user inputs, retrieve relevant information, and generate a response within a few seconds to maintain usability.

Potential Implications

Accessibility to Healthcare Information

The AI Doctor has the potential to significantly improve access to healthcare information, especially in remote and underserved areas. Patients who lack access to professional medical advice can use the AI Doctor to gain insights into their symptoms and possible conditions, bridging the gap between patients and healthcare professionals.

Empowering Patients

By providing patients with detailed information about their symptoms and possible conditions, the AI Doctor empowers them to make informed decisions about their health. This can lead to better health outcomes and a more proactive approach to healthcare.

Reducing the Burden on Healthcare Systems

The AI Doctor can reduce the burden on healthcare systems by handling minor queries and providing general advice. This allows healthcare professionals to focus on more complex cases, improving overall efficiency and patient care quality.

Team Contributions

Fakhra Amjad

- **Collecting Dataset:** Searched for relevant datasets.
- **Researching Encodings:** Suggested using "neuml/pubmedbert-base-embeddings" as encodings for vectors because of the model's accuracy on medical datasets.
- **Chunking and Encoding:** Devised initial attempts to encode the datasets using relevant encoding.
- **RAG Pipeline:** Implemented Mixtral and Llama 3.1 initially to test and run RAG chain.
- **PDF Function:** Implemented PDF reading function and integrated it into the RAG chain.
- **Documentation:** Prepared a detailed report about the project.

Muhammad Qasim Alias Haseeb

- **Collecting Dataset:** Searched for relevant datasets.
- **Chunking and Encoding:** Combined the datasets resulting in the needed chunks and encoded those.
- **Vector Database Pipelines:** Implemented the pipeline to encode the chunks batchwise and store them into the Pinecone Vector Database with their respective metadata (Compressed Strings).
- **Prompt Engineering**: Applied various techniques to make LLM return relevant results, it included adding multiple prompts and restructuring the pipeline.
- **Integration with Frontend:** Integrated the pipeline with Streamlit frontend.
- **Deployment:** Deployed the complete application on Streamlit Cloud.

Ali Raza

- **Collecting Dataset:** Searched for relevant datasets.
- **Researching Frontend Frameworks:** Suggested using Streamlit for frontend after researching other possible solutions.
- **Frontend Development:** Developed a responsive and User-friendly Streamlit UI specifically tailored to handle Medical Context Chats.
- **Troubleshooting and Debugging:** Tested the application on various instances and customized the UI as per demand.
- **Integration:** Helped making integration process swift and easy through continuous involvement.

Future Work

Integration with Electronic Health Records (EHR)

Future iterations of the AI Doctor could be integrated with Electronic Health Records (EHR) systems, allowing it to provide more personalized and accurate advice based on a patient's medical history. This would enhance the AI Doctor's diagnostic capabilities and allow for more tailored treatment recommendations.

Continuous Learning and Improvement

The AI Doctor should be designed to continuously learn and improve based on new data and feedback. Incorporating mechanisms for updating its knowledge base with the latest medical research and patient interactions will ensure that it remains accurate and up to date.

Conclusion

The AI Doctor project represents a significant step forward in the application of artificial intelligence in healthcare. By providing patients with access to medical information and preliminary diagnoses, the system has the potential to improve healthcare accessibility, empower patients, and reduce the burden on healthcare systems. However, the project also presents several challenges, including ensuring data quality and maintaining real-time performance. Future work will focus on enhancing the system's capabilities, expanding its reach, and ensuring it remains a valuable tool for both patients and healthcare professionals.

The AI Doctor is not intended to replace professional medical advice but to serve as a supplementary tool, offering valuable insights and guidance to those in need. As the field of AI in healthcare continues to evolve, projects like AI Doctor will play an increasingly important role in shaping the future of medicine.

References

[1] hakeematyab, "HealthCareMagicWithSummary-100k," [Online]. Available: https://huggingface.co/datasets/hakeematyab/HealthCareMagicWithSummary-100k.

[2] RafaelMPereira, "HealthCareMagic-100k-Chat-Format-en," [Online]. Available: https://huggingface.co/datasets/RafaelMPereira/HealthCareMagic-100k-Chat-Format-en.

[3] lavita, "ChatDoctor-HealthCareMagic-100k," [Online]. Available: https://huggingface.co/datasets/lavita/ChatDoctor-HealthCareMagic-100k.

[4] ruslanmv, "ai-medical-chatbot," [Online]. Available: https://huggingface.co/datasets/ruslanmv/ai-medical-chatbot.

[5] FreedomIntelligence, "Disease_Database," [Online]. Available: https://huggingface.co/datasets/FreedomIntelligence/Disease_Database.

[6] omi-health, "medical-dialogue-to-soap-summary," [Online]. Available: https://huggingface.co/datasets/omi-health/medical-dialogue-to-soap-summary.