Collaborative AI Healthcare Solution Development (MedicAI)

Gökçenaz Akyol, Evren Çağılcı, Utku Bora

Patients struggle to find the right specialists, and doctors face overwhelming medical literature. An AI system on a user-friendly website aids diagnosis, medication choice, and literature access with Literature Review, Drug Recommendation, and Disease Detection models.

Problem Definition

In the modern healthcare environment, several persistent challenges continue to affect both patients and healthcare professionals. **Patients often struggle to determine which specialist to consult** based on their symptoms, leading to delays in receiving appropriate care and potentially resulting in ineffective or incorrect treatment. This confusion can significantly hinder timely and effective healthcare delivery.

Healthcare professionals face the daunting task of staying updated with the extensive and continually expanding body of medical literature. The sheer volume of new research and advancements can be overwhelming, making it difficult for practitioners to keep abreast of the latest developments. Additionally, **prescribing the correct medication for various diseases** adds another layer of complexity, which can compromise patient care and increase the risk of medication errors.

Introduction

To address these issues, a comprehensive system has been developed in a **user-friendly website format**. This system offers separate logins for doctors, patients, and administrators. Depending on their login, users can access different models specifically designed to assist them. These models include a Literature Review Model, a Drug Recommendation Model, and a Disease Detection Model.

The Literature Review Model assists healthcare professionals in efficiently accessing relevant medical literature, staying current by continuously integrating the latest information. **The Drug Recommendation Model** aids doctors in selecting the most appropriate medications for various diseases, reducing uncertainty and enhancing the quality of patient care. **The Disease Detection Model** helps both patients and doctors by accurately identifying diseases based on symptoms, allowing timely and precise treatment.

Literature Review

Various AI models have been developed to address specific healthcare challenges, such as ChatGPT by OpenAI and ClinicalBERT, each with their unique capabilities and limitations.

ChatGPT is a general-purpose language model capable of engaging in a wide range of conversations but lacks the specialized focus required for medical applications.

ClinicalBERT, on the other hand, has been fine-tuned on clinical notes from electronic health records to improve its understanding of medical language and its ability to predict hospital readmissions. However, ClinicalBERT's primary focus is on understanding clinical notes and predicting readmissions, rather than providing comprehensive medical information or assisting with drug recommendations and disease detection from symptoms.

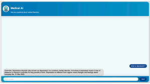
Methodology

Literature Review Model: A new dataset was created from 15 medical books on various subjects. The books were divided into chunks and question-answer pairs were generated, resulting in approximately 13,000 pairs. This dataset was used to train a conversational AI model based on DialoGPT from Hugging Face. A web scraping function was developed to extract relevant information from web pages, focusing on Wikipedia to ensure reliable and structured information.

Drug Recommendation Model: A dataset containing 161,528 entries about various drugs and their ratings for different conditions was used. This dataset included columns for medical conditions, drug names, and ratings. The drug names were encoded into numerical labels, and the model was trained using the RLHFlow/ArmoRM-Llama3-8B-v0.1 model from Hugging Face.

Disease Detection Model: The "Symptom-Disease Dataset" from Hugging Face was used, containing training and test datasets mapping symptoms to diseases. The data was tokenized using the BERT tokenizer, and a BERT-based model for sequence classification was trained.

Website Development: A user-friendly website was developed using Flask, HTML, and CSS, allowing patients, doctors, and medical students to access the models. The website featured a chat interface where users could input questions and receive answers from the trained models.



Results

The Literature Review Model demonstrated a moderate degree of semantic similarity with an average cosine similarity score of 0.6154 and an average Euclidean distance of 2.5342 between the provided answers and the ground truth. The model's predictive capability was effective, indicated by an average perplexity score of 0.9984. Relevancy assessment using the OpenAI API confirmed the correctness of responses.

The Disease Detection Model showed significant improvements across training epochs. In the first epoch, the model achieved a loss of 4.3195 and an accuracy of 37.40%, with a validation accuracy of 80.20%. By the third epoch, the loss decreased to 1.2851, and accuracy increased to 83.33%, with a validation accuracy of 84.10%. These results indicate effective learning and robust generalization to new data, making the model reliable for disease diagnosis based on symptom inputs.