VOICE ENABLED PATIENT DOCUMENTATION AND ASSISTANCE SYSTEM

20BCE0339 | Saraswathi Bavadharini S| Prof. Manoov R| SCOPE

Introduction

The objective is to provide a user-friendly and efficient solution that addresses the challenges faced by medical professionals in managing patient information and documentation, thereby contributing to the advancement of healthcare technology and patient care delivery.

Motivation

The project's motivation lies in its ability to optimize workflow processes, enhance information accessibility, and promote inclusivity within healthcare settings. By leveraging advanced technologies to address these challenges, the project aims to empower healthcare professionals to deliver more efficient, effective, and patient-centered care.

Scope of the Project

The project aims to develop a user-friendly voice-enabled application for healthcare professionals, integrating advanced technologies like NLP and speech recognition. It will streamline documentation processes, offer intuitive interaction, and generate accurate patient details and document summaries. Ensuring seamless integration with existing systems and robust privacy measures, the application will undergo rigorous testing and continuous improvement based on user feedback.

Methodology

- Setup and Imports: The code begins with importing necessary libraries such as Streamlit for creating the web interface, OpenAI for AI-powered responses, and others for text-to-speech and speech recognition functionalities.It also initializes the OpenAI API with the appropriate key.
- Function Definitions: Several functions are defined to perform specific tasks.These include functions for transcribing audio to text, generating AI responses based on prompts, converting text to speech, and summarizing text documents.
- Main Functionality: The main() function is where the core functionality of the web application resides.It sets up the Streamlit interface, including a sidebar for selecting actions.Depending on the selected action, it either retrieves patient details by processing audio input or uploads patient documents and generates summaries for them.
- Audio Processing: The code utilizes speech recognition to transcribe spoken words into text.This functionality is primarily used for retrieving patient details by capturing the patient's name or ID through a microphone.
- Document Summarization: For uploaded patient documents, the code employs a text summarization model to condense the content into shorter, more manageable summaries. It supports multiple file formats such as PDF, DOCX, and plain text.
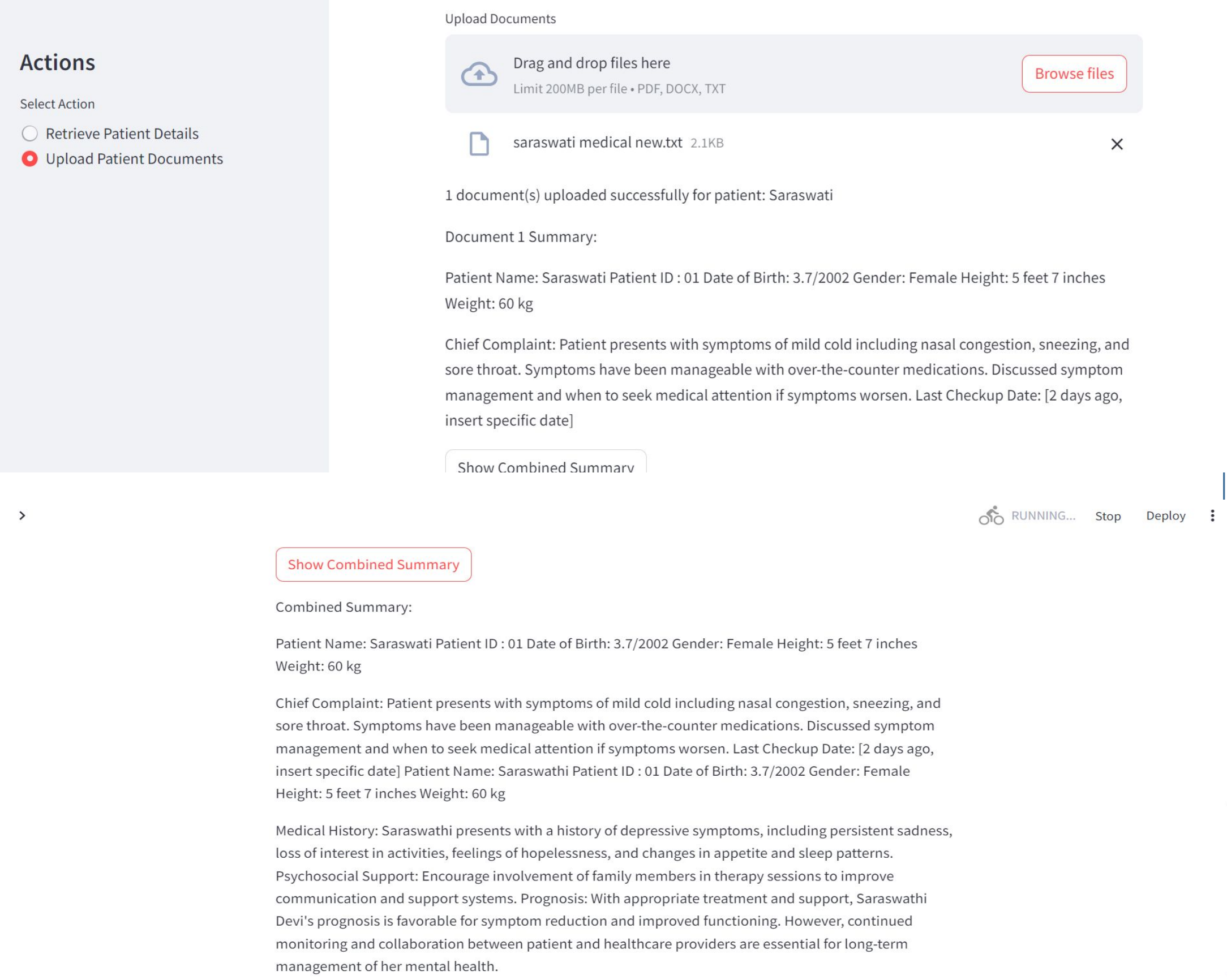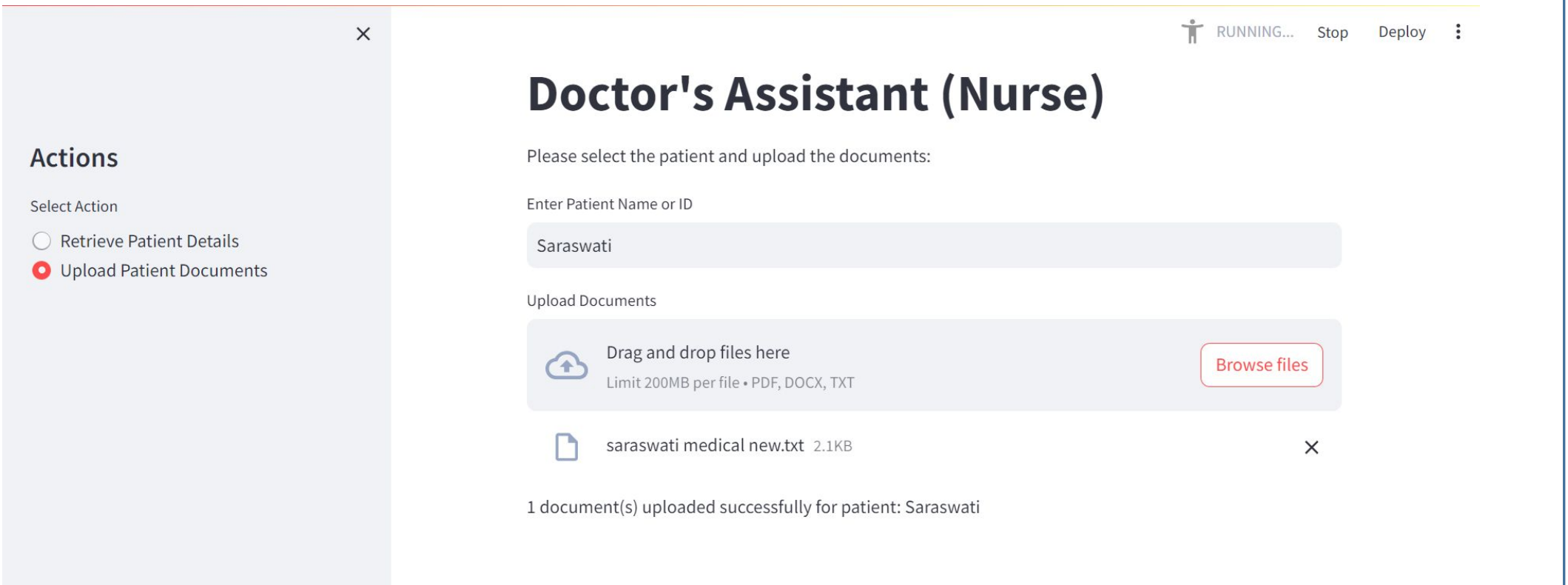
Based on the outlined methodology, the developed application promises several significant achievements. Firstly, the incorporation of speech recognition technology facilitates swift and hands-free retrieval of patient details, thereby enhancing the efficiency of healthcare professionals. This feature enables seamless access to patient information by simply speaking the patient's name or ID. Additionally, the application's capability to upload and summarize patient documents in various formats streamlines the process of reviewing extensive medical records. By providing concise summaries of important information within these documents, the application aids in efficient decision-making and understanding of patient histories. Moreover, the integration of OpenAI's advanced language model empowers the application to generate detailed responses to user queries or prompts, further enhancing the capabilities of healthcare professionals. The web-based nature of the application ensures accessibility from any internet-enabled device, facilitating seamless integration into existing healthcare workflows and enabling professionals to access patient data and document summaries conveniently from any location. Overall, the application contributes to improved patient care by reducing the time and effort required for data retrieval and document analysis, allowing healthcare professionals to focus more on delivering quality care to their patients.
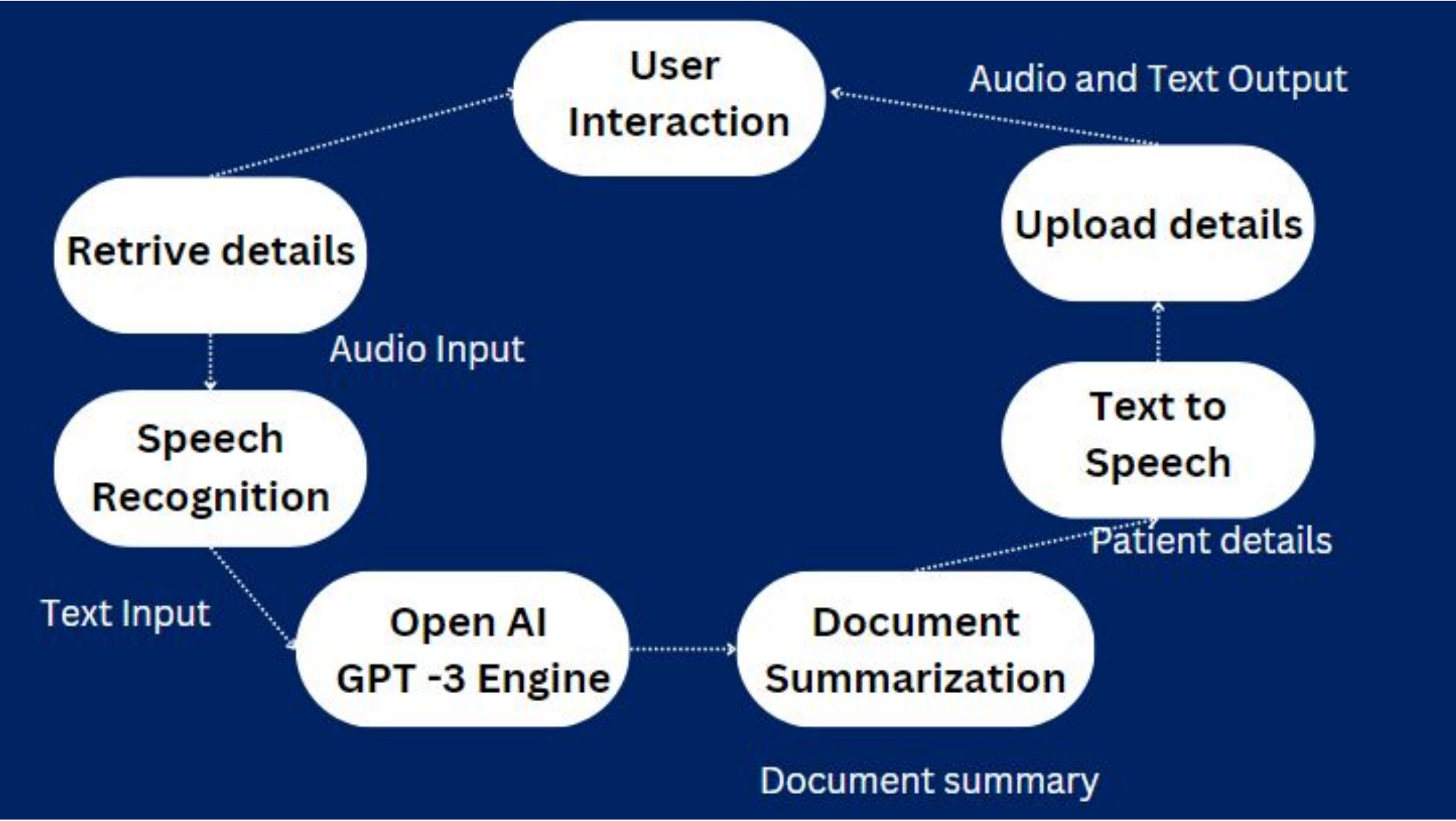
Results

The implemented system successfully enables medical professionals to:

1. Retrieve patient details by voice command.

2. Upload medical documents for summarization.

3. Obtain summarized document content in both text and audio formats.

4. Experience streamlined workflow and enhanced efficiency in accessing patient information.

Conclusion

The Doctor's Assistant (Nurse) project exhibits satisfactory performance across key functionalities.	The speech recognition module effectively transcribes audio inputs, facilitating seamless interaction with the system. Similarly, the text summarization module generates concise summaries, aiding healthcare professionals in decision-making.
The data generally supports the hypothesis that integrating advanced technologies enhances healthcare workflows.	Through efficient speech recognition and accurate text summarization, the project streamlines processes and improves access to patient information. This aligns with the hypothesis that technology integration enhances healthcare delivery.
Future work should focus on refining accuracy and robustness in speech recognition and enhancing natural language understanding.	Fine-tuning speech recognition models for improved accuracy and exploring advanced NLP techniques can enhance system performance. Additionally, continued user feedback and iterative refinement are essential for optimizing user experience and ensuring long-term effectiveness.

References

- Juhi A, Pipil N, Santra S, Mondal S, Behera JK, Mondal H. The Capability of ChatGPT in Predicting and Explaining Common Drug-Drug Interactions. Cureus. 2023 Mar 17;15(3):e36272. doi: 10.7759/cureus.36272. PMID: 37073184; PMCID: PMC10105894.
- Lee P, Bubeck S, Petro J. Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. N Engl J Med. 2023 Mar 30;388(13):1233-1239. doi: 10.1056/NEJMsr2214184. PMID: 36988602.
- Chi-Chun Lee, Emily Mower, Carlos Busso, Sungbok Lee, Shrikanth Narayanan,Emotion recognition using a hierarchical binary decision tree approach,Speech Communication,Volume 53, Issues 9–10,2011,Pages 1162-1171,ISSN 0167-6393.