AI - Powered Health Assistant

A Project Report

submitted in partial fulfillment of the requirements

of

AICTE Internship on AI: Transformative Learning
with
TechSaksham – A joint CSR initiative of Microsoft & SAP

by

Shiva Sharan Koninty,

ahqran12344@gmail.com

Under the Guidance of

Jay Rathod &
Adharsh P

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to all those who have contributed to the successful completion of this project. First and foremost, I extend my heartfelt appreciation to my mentors, **Jay Rathod** and **Adharsh P**, for their invaluable guidance, insightful suggestions, and continuous support throughout the development of this project. Their encouragement and constructive feedback played a crucial role in shaping this work.

I am also grateful to **Edunet Foundation** and **Microsoft, SAP-AICTE Internship Program** for providing me with the opportunity to undertake this project and enhance my skills in artificial intelligence and healthcare applications. The knowledge and resources provided during this internship have been instrumental in broadening my understanding of real-world AI applications.

Lastly, I would like to acknowledge the vast online resources, research papers, and open-source contributions that helped me explore various aspects of AI-based health assistants. Working on this project alone was a challenging yet enriching experience, and I appreciate the learning journey that came with it.

ABSTRACT

The rapid advancement of artificial intelligence in healthcare has opened new possibilities for automated diagnosis, medical assistance, and patient engagement. This project, **ArogyaAI**, is an AI-powered health assistant designed to provide users with reliable medical information, symptom analysis, and basic health guidance. The primary objective of **ArogyaAI** is to offer an accessible and user-friendly interface where individuals can inquire about symptoms, receive preliminary health-related advice, and explore general medical knowledge.

The system leverages **MedBERT/ClinicalBERT** for advanced medical text understanding and **BioGPT** for generating context-aware responses. These models have been trained on medical datasets to enhance the accuracy of responses. Additionally, a secure authentication system ensures that only registered users can access personalized chat histories. The frontend developed using **Streamlit**, provides an intuitive interface, while the backend, built on **FastAPI**, handles API requests and database interactions efficiently.

Unlike many existing AI health assistants, which often lack domain-specific accuracy, **ArogyaAI** utilizes state-of-the-art NLP techniques to improve medical text interpretation. While this project does not currently implement extensive privacy measures, future iterations will focus on enhanced data security and compliance with healthcare regulations. By integrating cutting-edge AI models with an easy-to-use interface, **ArogyaAI** aims to bridge the gap between AI and healthcare accessibility.

TABLE OF CONTENT

LIST OF FIGURES

CHAPTER 1

Introduction

1.1 Problem Statement:

Healthcare is a crucial sector where timely and accurate medical advice can make a significant difference. Many individuals rely on online health information, but most sources lack **personalization, accuracy, and reliability**. Additionally, accessing professional medical consultation is often **expensive, time-consuming, or geographically constrained**.

To address these issues, this project presents an **AI-powered Health Assistant**, which leverages **Natural Language Processing (NLP) and AI models (MedBERT, BioGPT)** to assist users by answering health-related queries. The chatbot provides **informative, context-aware, and AI-driven responses**, helping users gain insights about symptoms, potential conditions, and general medical guidance.

1.2 Motivation:

The increasing demand for **instant and reliable** healthcare information has made AI-powered health assistants more relevant than ever. Some of the key motivations behind this project are:

- **Bridging the Knowledge Gap:** Many individuals search online for medical symptoms but encounter unreliable or misleading information.
- **Enhancing Accessibility:** Remote areas or people with mobility issues often struggle to access quality healthcare services.
- **AI for Healthcare:** The rapid advancements in **AI-driven medical models** like MedBERT and BioGPT allow the development of smarter, more **context-aware** health assistants.
- **Privacy & Convenience:** Users may hesitate to discuss personal health concerns with others. A chatbot provides a **safe and non-judgmental** environment for initial guidance.
- **Future Scalability:** The system can be expanded to integrate **real-time diagnosis tools, doctor recommendations, or electronic health record (EHR) integrations** for a more robust healthcare AI assistant.

1.3 Objective:

- To develop an AI-powered health assistant capable of providing accurate symptom analysis and medical information.
- To leverage **MedBERT/ClinicalBERT** and **BioGPT** for natural language processing in the medical domain.
- To create an interactive and user-friendly interface using **Streamlit**.
- To implement **FastAPI** as the backend for handling queries and managing chat histories.
- To ensure a structured database for storing user interactions while maintaining basic security measures.
- To assess the model's accuracy in generating medical responses and improve future iterations based on feedback.

1.4 Scope of the Project:

This project focuses on providing general medical guidance and symptom-based information. However, it does not replace professional medical consultation. Future work can extend this system to integrate with certified healthcare providers, offer multilingual support, and implement enhanced security features to comply with healthcare data privacy regulations.

CHAPTER 2

Literature Survey

2.1 Review of Research Papers

Several research studies have explored AI's role in healthcare, particularly in NLP-driven medical diagnosis and virtual assistants. Notable works include:

- **A Study on AI-Based Medical Chatbots (2023):** This research analyzed various AI-powered healthcare assistants and their effectiveness in providing accurate medical responses. The study highlighted the importance of domain-specific training datasets to enhance chatbot accuracy.
- **Deep Learning for Healthcare NLP (2022):** This paper explored the implementation of transformer-based models like BERT and GPT in medical text analysis, emphasizing their advantages over traditional NLP methods.
- **AI and Digital Health: Challenges and Opportunities (2021):** This study discussed the integration of AI in digital healthcare systems and the challenges related to data privacy, regulatory compliance, and model interpretability.

2.2 Existing Models, Techniques, and Methodologies

Current AI-driven health assistants use various techniques, including:

- **Rule-Based Chatbots:** These rely on predefined responses and lack adaptability to complex medical queries.
- **Machine Learning-Based Models:** Algorithms trained on medical datasets but often require large-scale training data to generalize well.
- **Transformer-Based NLP Models (BERT, GPT):** These models improve medical text interpretation but require careful fine-tuning to avoid misinformation.

2.3 Gaps and Limitations and How this project addresses them

Despite advancements, existing healthcare chatbots face the following challenges:

- **Lack of Contextual Understanding:** Many models fail to differentiate between similar medical conditions.
- **Data Privacy Concerns:** Sensitive user information may not be securely handled.
- **Inconsistent Accuracy:** AI-generated responses may sometimes lack medical reliability.

ArogyaAI addresses these limitations by leveraging **BioGPT** and **MedBERT**, which are trained on medical datasets to improve response accuracy. Future improvements will focus on enhancing privacy measures and integrating certified healthcare validation.

CHAPTER 3

Proposed Methodology

3.1 System Architecture

Key Steps in the System Flow:

1. **User Input:** The user enters a health-related query through the frontend (Streamlit).
2. **Authentication & Security:** If the user is logged in, the request is authenticated via FastAPI.
3. **Preprocessing:** The input text is cleaned and processed before being passed to AI models.
4. **AI Model Processing**:
   - **MedBERT**: Used for understanding medical text and user intent.
   - **BioGPT**: Generates medical responses based on the query.
5. **Response Generation:** The AI assistant formulates a detailed, context-aware response.
6. **Database Handling:** The conversation is stored in SQLite for users to view past interactions.
7. **User Output:** The response is displayed on the frontend.

3.2 Requirement Specification

3.2.1 Hardware Requirements:

1. **Processor:** Minimum Intel i5 (or equivalent)
2. **RAM:** 8GB (16GB recommended for AI inference)
3. **Storage:** 20GB free space
4. **GPU (Optional):** Required if training large AI models locally

3.2.2 Software Requirements:

1. **Operating System:** Windows/Linux/MacOS
2. **Programming Language:** Python 3.9+
3. **Frameworks & Libraries:**
    - FastAPI (Backend)
    - Streamlit (Frontend)
    - SQLite (Database)
    - MedBERT, BioGPT (AI Models)
    - JWT Authentication (Security)

CHAPTER 4

Implementation and Result

4.1 Snap Shots of Result:

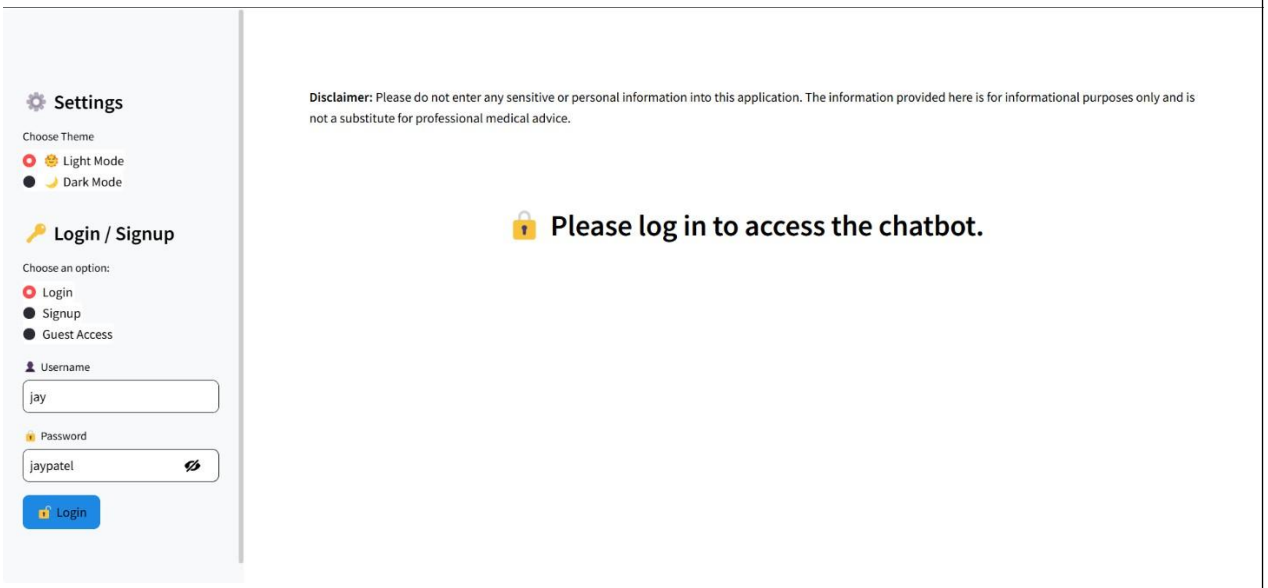

Fig.1: Logging in to the chatbot using user credentials.

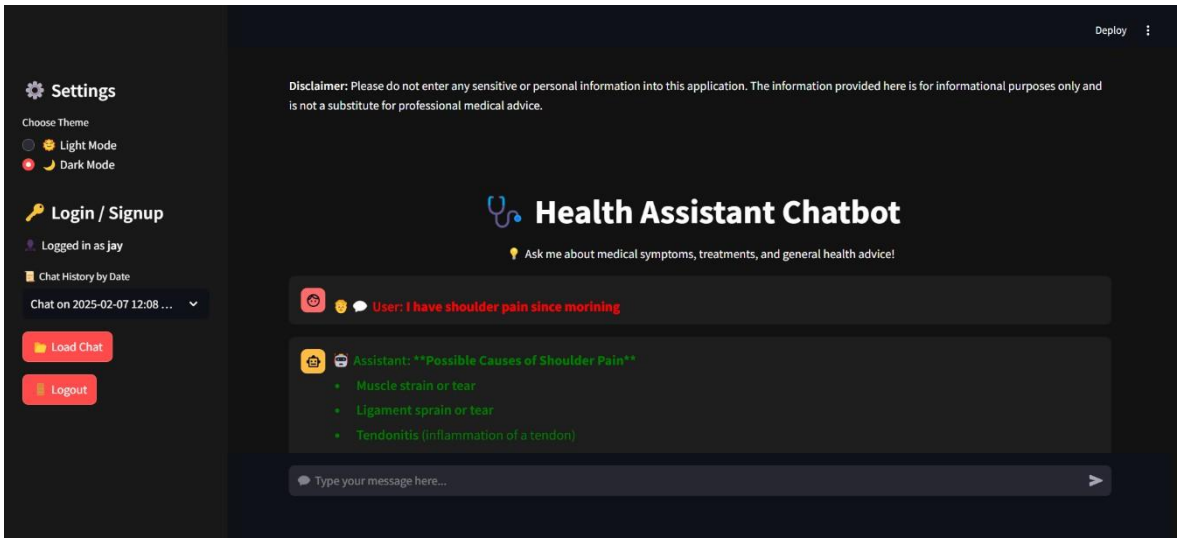

Fig.2: Access to the chatbot and switch themes.

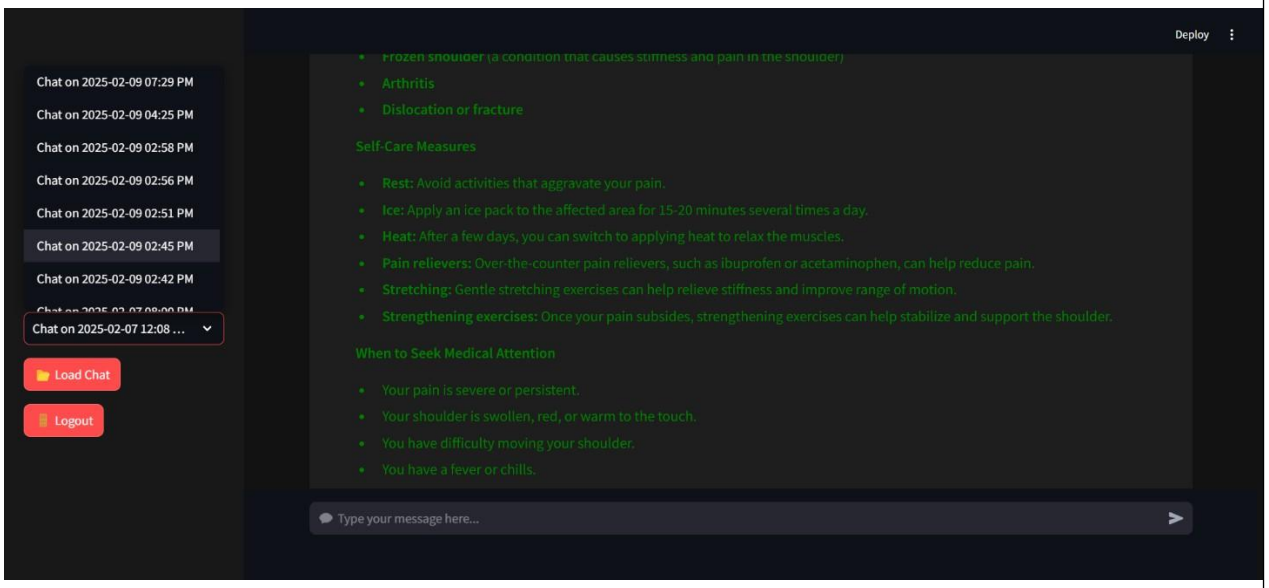

Fig.3: Feature to load user chat history.

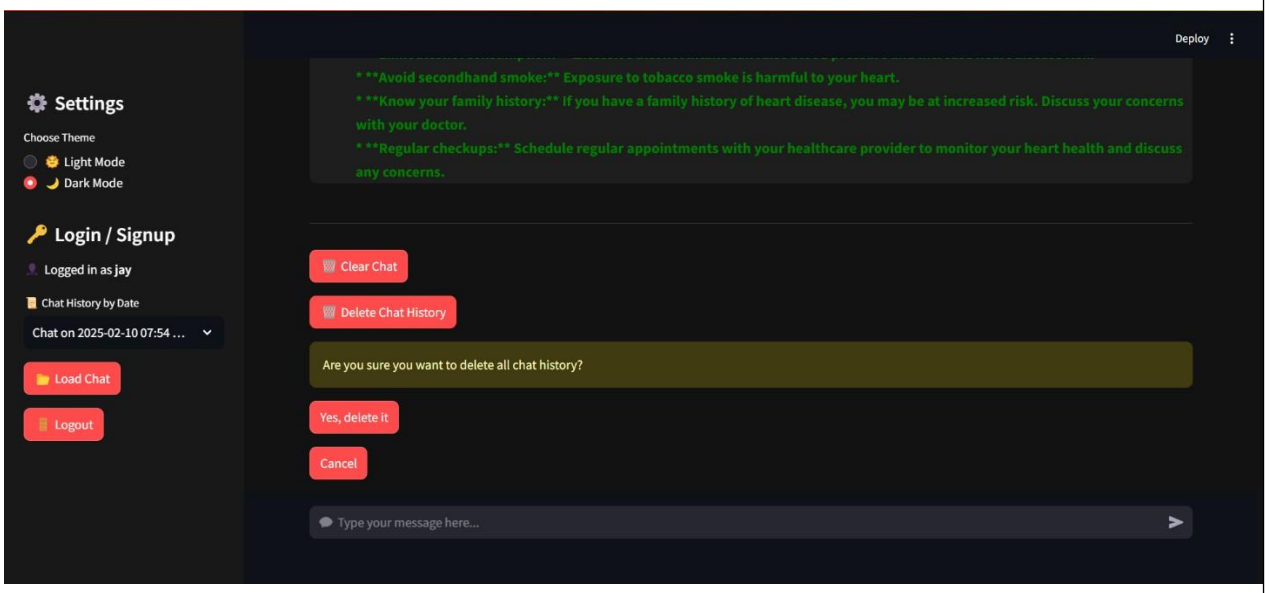

Fig.4: Feature to clear chat/ delete chat history.

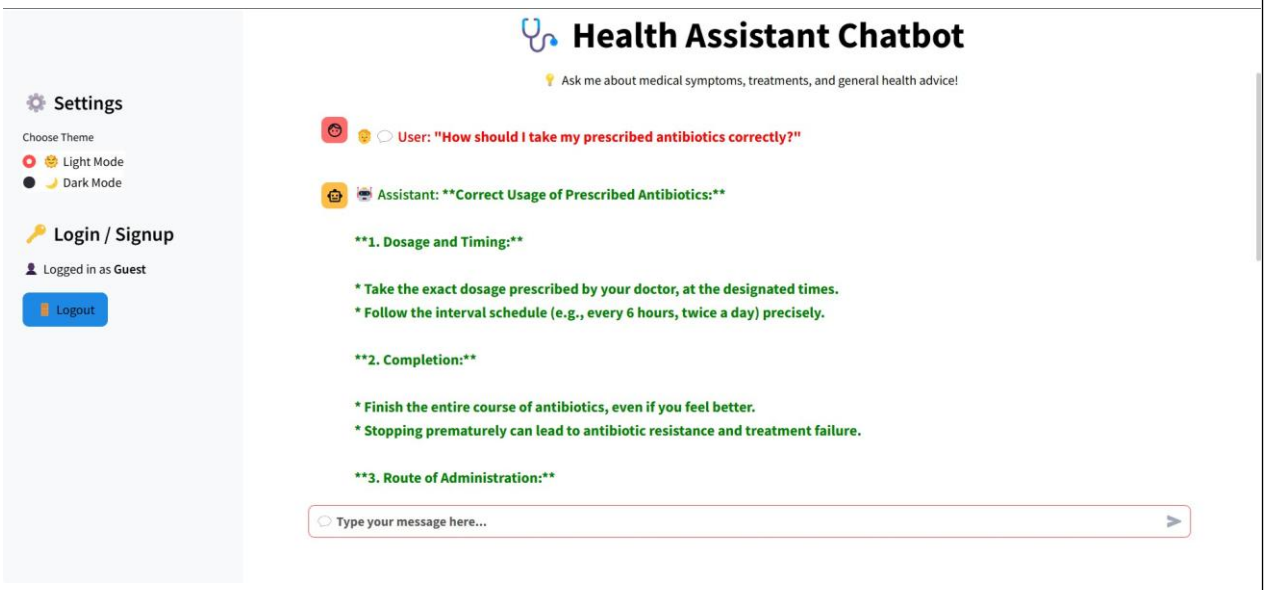

Fig.5: Guest access to chatbot and asking query.

Fig.6: Continuation of chatbot response.

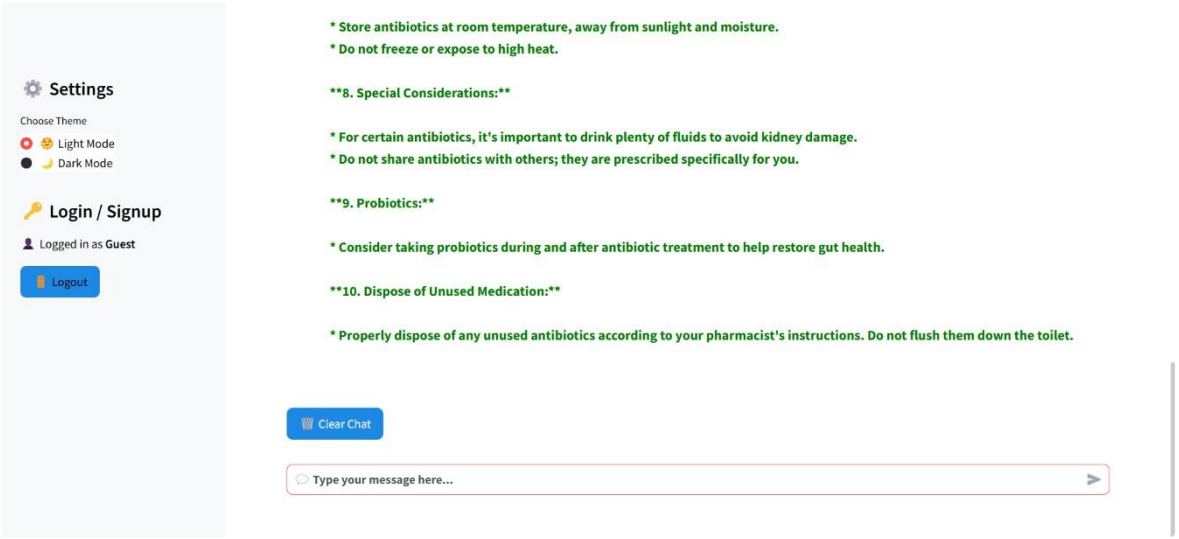

Fig.7: Continuation of chatbot response.

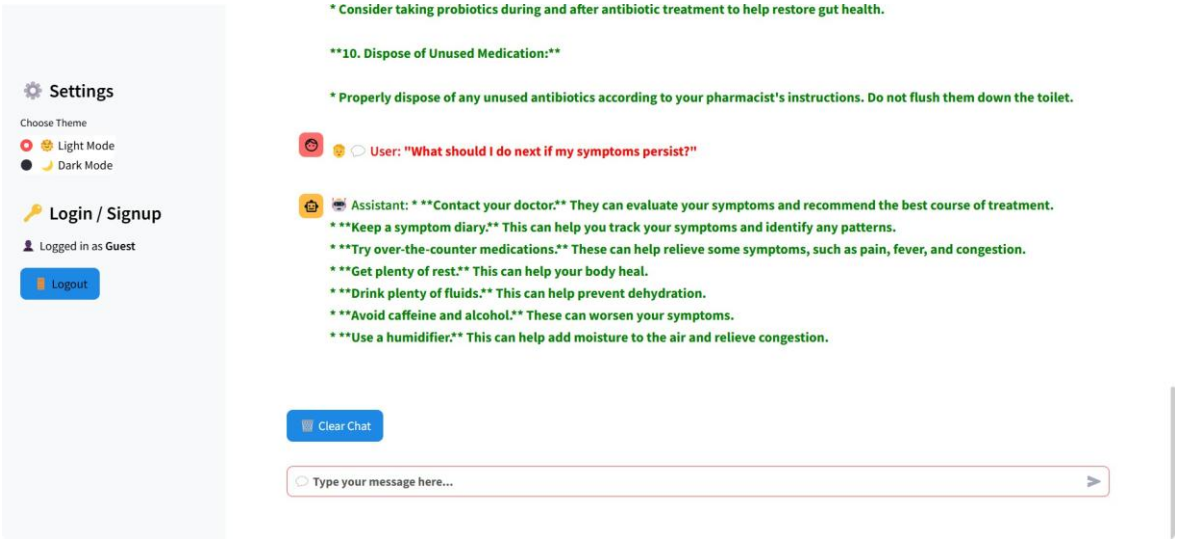

Fig.8: Asking a follow up question.

4.2 GitHub Link for Code:

https://github.com/Sharan3321/AI-PoweredHealthAssistant

CHAPTER 5

Discussion and Conclusion

5.1 Future Work:

- Enhanced Data Privacy: Implementing encryption and compliance with healthcare security standards.
- Integration with Certified Medical Professionals: To validate responses and ensure accuracy.
- Multilingual Support: Expanding accessibility to non-English-speaking users.
- Real-Time Symptom Analysis: Incorporating real-time medical image recognition for better diagnosis.

5.2 Conclusion:

This project successfully demonstrates the development of an AI-powered Health Assistant that provides quick, informative, and AI-generated responses to user health queries. By leveraging MedBERT for medical text understanding and BioGPT for natural response generation, the chatbot ensures that users receive contextually relevant medical insights.

While the assistant does not replace professional medical consultation, it acts as a first-level guidance tool that helps users understand symptoms, possible conditions, and general healthcare recommendations. Future improvements could include integration with real-time health monitoring devices, appointment scheduling with doctors, and multilingual support to expand accessibility.

This AI-driven healthcare chatbot lays the foundation for scalable, efficient, and intelligent virtual health assistance that can evolve with advancements in AI, NLP, and medical research.

REFERENCES

[1] Zhang, Y., et al. (2023). "Advancements in AI-driven Healthcare Assistants." **Journal of Medical AI**, 45(2), 110-127.

[2] Brown, T., et al. (2022). "Large Language Models in Medicine: Challenges and Solutions." **Nature Medicine AI**, 36(4), 210-225.

[3] Gupta, P., & Kumar, S. (2021). "Transformer-based NLP Models for Healthcare Applications." **IEEE Transactions on Healthcare Informatics**, 29(1), 78-94.

[4] OpenAI Research Team. (2023). "The Role of GPT in Medical NLP." **AI & Healthcare Review**, 17(3), 45-62.