

Medical AI Assistant: Symptom Analysis and Treatment Suggestions

Abstract— This paper presents the design and development of a Medical AI Assistant that leverages transformer-based large language models to provide disease predictions and treatment plan suggestions based on user-provided symptoms and medical history. The system integrates Gradio for a user-friendly interface and IBM's Granite-3.2-2B-Instruct model for natural language understanding. The assistant emphasizes that its outputs are for informational purposes only and encourages consultation with healthcare professionals for accurate diagnosis and treatment.

Keywords— Medical AI, Natural Language Processing, Gradio, Healthcare, Transformer Models

I. INTRODUCTION

The integration of artificial intelligence (AI) in healthcare has shown promising advancements in patient support, diagnostics, and treatment planning. However, ensuring accessibility and usability remains a key challenge. This project introduces a Medical AI Assistant built using IBM's Granite-3.2-2B-Instruct model and deployed with a Gradio interface. The system enables users to input symptoms or medical conditions and receive possible conditions and general treatment suggestions. Importantly, the tool is designed as an informational resource and does not replace professional medical consultation.

II. METHODOLOGY

The system architecture consists of two main modules: disease prediction and treatment plan generation. Disease prediction accepts symptom descriptions and generates possible conditions. Treatment planning accepts structured patient data, including age, gender, and medical history, to provide personalized recommendations. The backend leverages Hugging Face's Transformers library for tokenization and model inference. Gradio provides an interactive web-based interface with separate tabs for each functionality.

III. RESULTS AND DISCUSSION

The Medical AI Assistant successfully provides coherent and contextually relevant disease predictions and treatment suggestions. However, limitations exist in terms of medical accuracy, as the model is not trained as a certified medical diagnostic system. Ethical considerations such as patient privacy, misinformation, and overreliance on AI-generated outputs are discussed. The system includes disclaimers to encourage professional consultation.

IV. CONCLUSION AND FUTURE WORK

This paper demonstrates the feasibility of deploying an AI-powered medical assistant using transformer models and Gradio. Future work includes integrating more domain-specific medical datasets, enhancing explainability of predictions, and incorporating multilingual support. Additionally, compliance with healthcare data standards such as HIPAA and GDPR will be explored.

REFERENCES

H. Touvron et al., “LLaMA: Open and Efficient Foundation Language Models,” arXiv preprint arXiv:2302.13971, 2023.

T. Wolf et al., “Transformers: State-of-the-Art Natural Language Processing,” in Proc. EMNLP, 2020, pp. 38–45.

IBM Research, “Granite Models,” IBM, 2024. [Online]. Available: <https://huggingface.co/ibm-granite>