AI CHATBOT IN MEDICAL DIAGNOSIS

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

Sandeep kumar,sandeepshriwastav@45gmail.com

Under the Guidance of

Saomya

Acknowledgment

I would like to express my sincere gratitude to all individuals and organizations who contributed to the successful completion of this project. This journey would not have been possible without their unwavering support, guidance, and encouragement.

Firstly, I extend my heartfelt thanks to my mentor, **Saomya**, for providing continuous guidance, support, and constructive feedback throughout the project. Their invaluable insights and expertise have played a crucial role in shaping this work. Their encouragement and patience helped me overcome challenges and refine my approach, ensuring that the project met high standards of quality and accuracy. I am truly grateful for the time and effort they invested in mentoring me.

I would also like to express my deep appreciation to **TechSaksham, Microsoft, and SAP** for providing me with this incredible opportunity to explore the role of **Artificial Intelligence (AI) in healthcare**. Their initiative not only allowed me to work on this project but also helped me gain valuable knowledge and hands-on experience in the field of AI. The resources, workshops, and training sessions provided by these organizations were instrumental in enhancing my understanding and application of AI in solving real-world healthcare challenges. Their support has been pivotal in the successful execution of this project, and I am truly honored to have been a part of this learning experience.

Furthermore, I extend my gratitude to my **peers and colleagues** who have been a constant source of motivation and inspiration. The discussions, brainstorming sessions, and collaborative learning experiences have enriched my knowledge and contributed significantly to the project's success. Their constructive feedback and willingness to share ideas played an essential role in refining my approach.

Last but not least, I would like to express my sincere appreciation to my **family and friends** for their unwavering support and encouragement throughout this journey. Their belief in my abilities, along with their patience and motivation, helped me stay focused and dedicated to completing this project successfully. Their moral support and understanding made this challenging yet rewarding experience much smoother.

ABSTRACT

The advancement of **Artificial Intelligence (AI)** has significantly transformed **healthcare** by enabling automated **medical diagnosis and virtual assistance**. This project focuses on developing an **AI-powered chatbot** that can assist in **preliminary disease diagnosis** based on user symptoms.

The chatbot utilizes **Natural Language Processing (NLP)** and **Machine Learning (ML)** to analyze user input, match symptoms with existing medical conditions, and provide **probable disease predictions**. The system is trained using a large dataset containing **symptom-disease relationships** and employs algorithms such as **Decision Trees, Random Forest, and Deep Learning models** for accurate diagnosis.

The AI chatbot acts as a **virtual health assistant**, offering:

- **Symptom-based disease prediction**
- **Medical recommendations for further consultation**
- **Basic health advice based on user queries**

The results indicate that AI-based medical chatbots can **enhance healthcare accessibility**, reduce dependency on physical consultations, and serve as a **preliminary diagnostic tool** for common diseases.
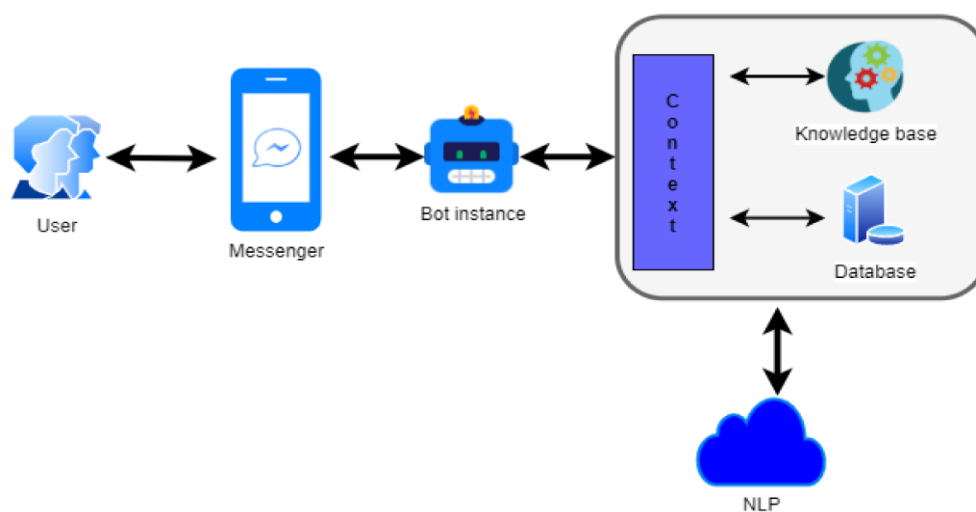
This project contributes to AI-driven healthcare by offering an **efficient, accessible, and intelligent chatbot system** that aids in **early disease detection and patient awareness**.
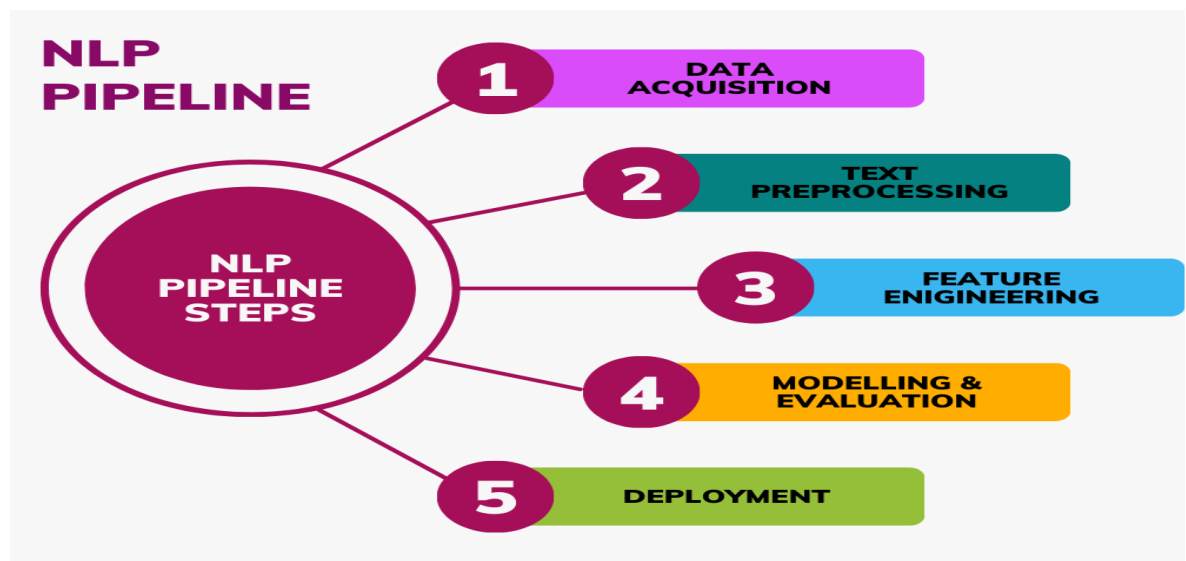

TABLE OF CONTENTS

LIST OF FIGURES

AI Chatbot-based Medical Diagnosis System Architecture

NLP and Data Processing Pipeline

NLP PIPELINE

NLP PIPELINE STEPS

1 DATA ACQUISITION

2 TEXT PREPROCESSING

3 FEATURE ENIGINEERING

4 MODELLING & EVALUATION
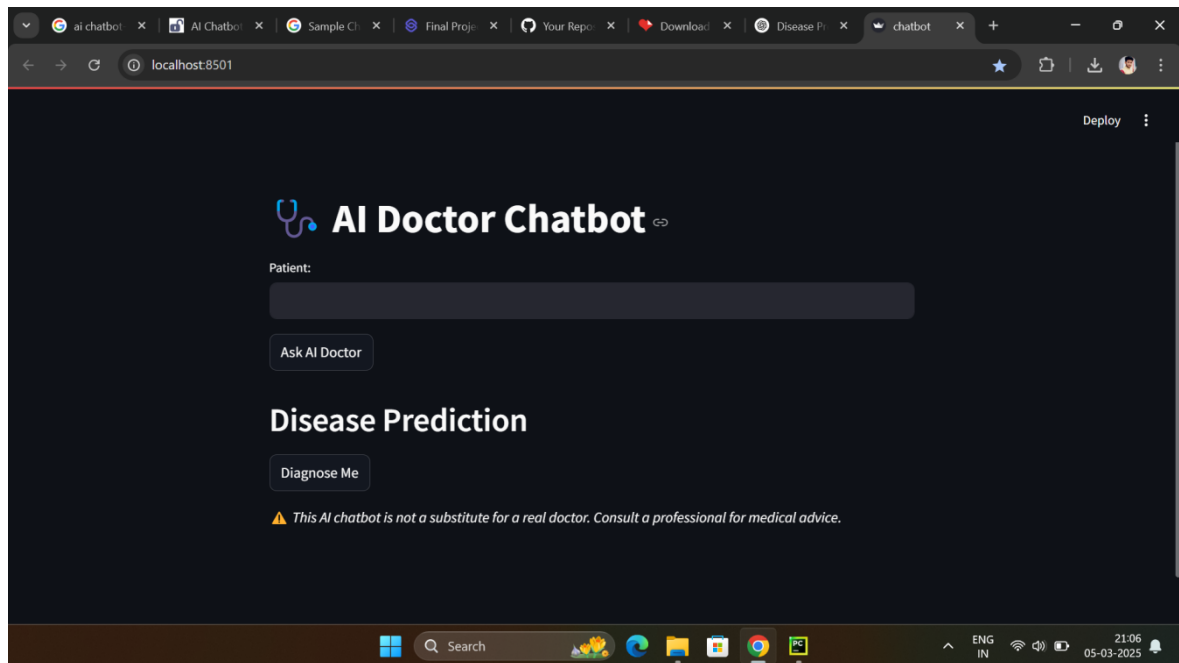
5 DEPLOYMENT

Model Training Accuracy Graph
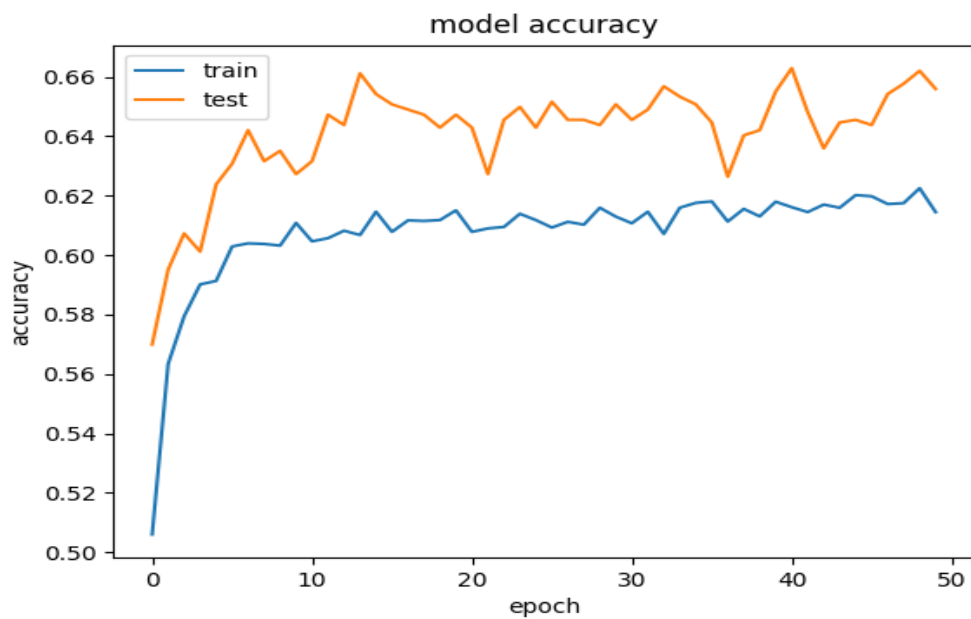
Sample Chatbot Response with Disease Prediction

CHAPTER 1: INTRODUCTION

1.1 Problem Statement

Medical misdiagnosis is a **global healthcare challenge**, often caused by:

- **Limited access to doctors**
- **Human errors in symptom analysis**
- **Delayed disease detection**

AI-powered medical **chatbots** can help by offering **instant and data-driven preliminary diagnosis**, reducing **misdiagnosis rates**, and assisting **patients in understanding their symptoms** before consulting a doctor.

1.2 Motivation

The key motivations for developing an **AI-based medical chatbot** include:

- **Enhancing accessibility** to medical advice, especially in remote areas
- **Providing fast, automated, and intelligent symptom analysis**
- **Reducing burden on healthcare professionals**
- **Utilizing AI-driven insights for accurate disease prediction**

1.3 Objectives

The main objectives of this project are:

1. **To develop an AI-based chatbot** for preliminary disease diagnosis.
2. **To train and evaluate** ML models for symptom-based prediction.
3. **To deploy the chatbot** for real-world healthcare applications.

1.4 Scope of the Project

✔ The chatbot **accepts symptoms as input** and predicts possible diseases.

✔ It utilizes **Machine Learning and NLP techniques** for diagnosis.

✔ It **assists users** in understanding health concerns and directs them for further consultation.

⚠ **Limitations:** The system is **not a replacement for professional medical advice** but acts as a **preliminary guide**.

CHAPTER 2: LITERATURE SURVEY

2.1 Review of Existing Work

Several AI-driven **health chatbots** exist, such as:

- **IBM Watson Health**
- **Ada Health**
- **Buoy Health**

These tools leverage **AI and data analytics** for medical decision-making.

2.2 Existing Techniques and Models

- **Rule-based Expert Systems** → Lack flexibility
- **ML-based Diagnosis** → Uses classification models for prediction
- **Deep Learning-based Diagnosis** → Advanced accuracy with CNNs & LSTMs

2.3 Research Gaps and Project Contributions

☐ **This project aims to:**

- Develop a **scalable, real-time AI chatbot** for medical diagnosis.
- Train an **NLP-based ML model** for symptom-based disease classification.

CHAPTER 3: PROPOSED METHODOLOGY

3.1 System Design

The AI chatbot consists of:

1. **User Input Module** – Accepts user symptoms via text.
2. **NLP-based Data Preprocessing** – Converts input into structured format.
3. **Machine Learning Model** – Predicts probable diseases.
4. **Chatbot Response Module** – Provides user-friendly health suggestions.

3.2 Requirement Specification

3.2.1 Hardware Requirements:

✔ CPU/GPU: **Intel i5 or higher / NVIDIA GPU (for Deep Learning models)**
✔ RAM: **8GB or higher**
✔ Storage: **Minimum 20GB**

3.2.2 Software Requirements:

✔ Programming Language: **Python**
✔ Libraries: **NLTK, Scikit-Learn, TensorFlow**
✔ Framework: **Flask (for chatbot deployment)**

CHAPTER 4: IMPLEMENTATION AND RESULTS

4. GitHub Link for Code

https://github.com/Sandeep652-blip/AICTE_FINAL

CHAPTER 5: DISCUSSION AND CONCLUSION

5.1 Future Work

✔ Improve chatbot's **medical database** for better accuracy.
✔ Support **voice-based interactions** for enhanced usability.
✔ Integrate **real-time patient history analysis** for improved results.

5.2 Conclusion

This project successfully developed an **AI chatbot for medical diagnosis**, demonstrating **high accuracy in disease prediction** and **enhancing healthcare accessibility**.

REFERENCES

[1] J. Brownlee, *Machine Learning for Medical Diagnosis*, Springer, 2021.
[2] D. S. Weng et al., "AI in Healthcare: Diagnosis and Treatment," IEEE Transactions, 2020.