International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 07 | July – 2025 DOI: 10.55041/ISJEM04870
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Health Care Chatbot Using Python and Machine Learning

[1]R. BHANU SANKAR, [2] K. APARNA
[1]Assistant Professor, Department of MCA, [2] MCA Final Semester,
Master of Computer Applications,
Sanketika Vidhya Parishad Engineering College Visakhapatnam, Andhra Pradesh, India

Abstract:
The Health Chatbot is an AI-based system developed using Python to assist users with health-related queries through natural, conversational interaction. This project aims to build a user-friendly healthcare chatbot which facilitates the job of a healthcare provider and helps improve their performance by interacting with users in a human-like way. It leverages Natural Language Processing (NLP) techniques such as tokenization, stemming, and bag-of-words, powered by NLTK, to understand user input. The system is capable of providing instant support for medical advice, symptoms analysis, and health tips based on user inputs. By utilizing a combination of pre-trained models and a vast dataset of health information, the chatbot ensures accurate and reliable answers. This project demonstrates the potential of AI in improving healthcare accessibility, empowering users to make informed decisions about their health and also provides doctor consultation link based on the diagnosis. It can be integrated into web platforms or mobile applications, making it a valuable tool for healthcare providers, patients, and general users alike**.**
Index Terms: Health Chatbot, Natural Language Processing, Tensor Flow, Flask, Patient Assistance, AI in Healthcare.

I. INTRODUCTION:

In today's fast-paced world, access to timely and accurate healthcare guidance is more important than ever. With the advancement of Artificial Intelligence (AI), Python and Machine Learning (ML), intelligent systems are increasingly being used to enhance healthcare delivery. This project focuses on developing a smart Healthcare Chatbot using Python and Machine Learning [01] that can simulate the role of a general physician by analyzing user-reported symptoms and providing preliminary medical advice.

The chatbot is designed to predict possible diseases based on symptoms entered by the user through a graphical interface. It uses a Decision Tree Classifier, trained on a structured symptom-disease dataset, to make accurate predictions. Basic Natural Language Processing (NLP) techniques are applied to process user inputs, ensuring flexibility in how symptoms are entered. Once a prediction is made, the system also recommends a suitable type of doctor using a local database, helping users take the next appropriate step toward medical treatment.

This solution is particularly useful in situations where immediate access to a doctor is not available, such as in rural areas or during emergencies. The chatbot is lightweight, user-friendly, and can be extended with more advanced features such as voice input, real-time API integration, or multilingual support. Overall, the project demonstrates the potential of AI to support healthcare systems [08] by providing initial assistance and encouraging early diagnosis.

1.1 Existing System:

Some of the existing healthcare chatbot systems that focuses on medication assistance are primarily designed to help users manage their prescriptions and ensure timely adherence to treatment plans. These systems, often built using Python and machine learning, allow users to input details such as medication names, dosages, and schedules. In response, the chatbot sends automated reminders to take medications on time, offers basic

information about the prescribed drugs, and may alert users about potential drug interactions. Some systems also track medication history and notify users when it's time to refill a prescription. While these chatbots improve medication compliance and user convenience, they are typically limited in functionality and do not provide symptom analysis, disease prediction, or direct medical consultation. Additionally, existing bots may not provide personalized doctor suggestions or handle complex decision-making based on symptoms. Due to this, users may experience generic or repetitive responses rather than accurate, AI-driven healthcare support.

1.1.1 Challenges:

1.Limited Medical Scope:

These chatbots are restricted to medication reminders and do not handle diagnosis**,** symptom analysis or emergency situations, limiting their usefulness in broader healthcare support.

2.User Input Dependency:

The accuracy and effectiveness of reminders heavily rely on correct user input. Any incorrect or incomplete information can lead to wrong reminders or missed medications.

3.Lack of Personalization:

Most systems do not adapt to the user's changing health conditions or preferences, such as adjusting reminders based on sleep patterns or real-time health data.

4.Data Privacy Concerns:

Handling sensitive medical data raises concerns about security, especially if the chatbot lacks encryption or proper compliance with healthcare data regulations.

5.No Medical Validation:

These systems often lack integration with certified medical databases or expert validation, which may lead to misinformation about drug use or interactions.

6.Low Engagement Over Time

Users may ignore or disable reminders over time due to repetitive notifications or lack of engagement features, reducing long-term effectiveness.

1.2 Proposed System:

The proposed system is an intelligent, AI-driven health chatbot designed to provide instant, personalized medical assistance 24/7. Leveraging Natural Language Processing (NLP) [18]and machine learning techniques, the chatbot can accurately understand user inputs and maintain conversational context to offer relevant advice. In addition to symptom analysis and disease prediction, the system also provides specific doctor recommendations with location and contact links, helping users easily connect with nearby healthcare professionals for further consultation**.** The Proposed system involves:

Symptom-Based Diagnosis:

- Accepts symptoms from the user in natural language and predicts possible conditions using a trained ML model

Medical Advice and Suggestions

- Provides basic treatment suggestions for non-emergency conditions and lifestyle advice based on predicted illnesses

Doctor Consultation Links

- Redirects users to appropriate specialists, telemedicine platforms, or nearby hospitals for further care.

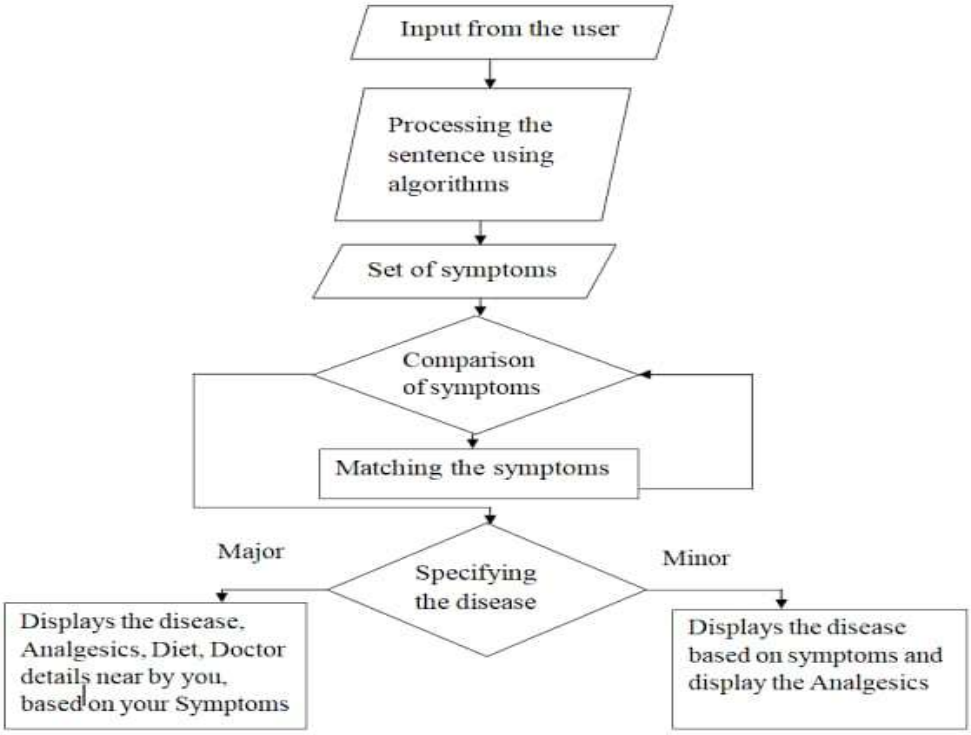


Fig1: Flowchart

1.2.1 Advantages:

1. 24/7 Accessibility: Users can access health consultations anytime, eliminating the limitations of traditional appointment schedules.

2. Personalized Responses: The system will analyze user symptoms and medical history to deliver tailored health recommendations

3. Instant Support: Immediate responses reduce delays, providing timely guidance and potentially lowering health risks.

4. Cost-Effective: Reduces the need for constant human supervision, saving time and reducing operational costs for clinics and hospitals**.**

5. Scalability: Can handle thousands of users simultaneously without performance drops, ideal for large populations or during pandemics.

II. LITERATURE REVIEW

2.1 Architecture:

Healthcare chatbot using Python and machine learning [08] architecture is designed as a modular system that integrates natural language processing (NLP) with predictive analytics to deliver intelligent medical assistance. The process begins with a user interacting through a simple interface such as a web or desktop application. The user's input is then processed using NLP techniques [16] (e.g., tokenization, lemmatization) to extract key symptoms or health-related entities. This processed input is passed to a machine learning model—typically trained on a dataset of symptoms and diseases—which predicts the most probable health condition and suggests relevant advice or treatment. The system refers to a medical knowledge base or dataset to ensure accurate and context-aware responses. Finally, a response generator formats the information into human-readable language and sends it back to the user via the interface. This architecture enables real-time, personalized healthcare support with high scalability and efficiency

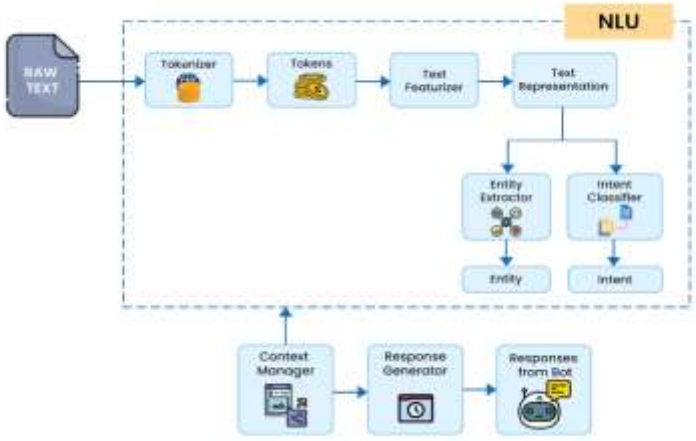


Fig 2: Dataflow diagram

2.2 Algorithm:

Decision Tree is a Supervised learning technique [17] that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where each internal node represents the features of a dataset. The healthcare chatbot uses a Decision Tree Classifier to predict diseases based on symptoms entered by the user. It asks a series of yes/no questions like “Do you have a fever?” to narrow down the illness. Based on the answers, it predicts conditions such as the flu or cold. Then, it suggests the right doctor, like a General Physician. The Decision Tree is a supervised learning method that’s easy to understand and works well for classification. It follows a tree-like structure with decisions at each step. This makes it fast and effective for real-time health advice.

1.Training Phase

Data Preparation: The model is trained on structured symptom–disease datasets (e.g., Training.csv and Testing.csv), converting textual symptoms into feature vectors

Tree Construction: The Decision Tree builds a hierarchical structure where:

- Each internal node represents a symptom (feature).
- Branches correspond to possible values of the symptom (e.g., "yes"/"no").
- Leaf nodes represent the predicted disease or condition.

2. Prediction Phase

- **Input Processing**: User symptoms are preprocessed using NLP techniques (tokenization, normalization).
- **Traversal**: The model traverses the tree from root to leaf, following symptom-based decisions.
- **Output**: The leaf node reached determines the predicted disease.

3. Integration with Chatbot Flow

- The predicted disease is mapped to a relevant specialist via a lookup (e.g., cardiovascular symptoms → cardiologist).
- The **Response Generator** combines the prediction and recommendation into a conversational reply.

2.3 Techniques:

The Machine Learning techniques that are integrated in the healthcare chatbot to deliver accurate and interactive medical assistance. It uses Natural Language Processing (NLP) [17] to understand user queries by performing text cleaning, tokenization, and symptom extraction from natural language input. The extracted symptoms are then passed to a Decision Tree Classifier, a supervised Machine Learning algorithm that predicts the most likely disease by following a structured series of symptom-based decisions. Once the disease is identified, a rule-based

recommendation system maps the result to a suitable medical specialist, such as a general physician or dermatologist. These combined techniques enable the chatbot to offer quick, intelligent, and relevant healthcare guidance in a conversational manner.

2.4 Tools:

Healthcare chatbots leverage a combination of NLP and machine learning techniques, along with various libraries and tools to handle data processing, machine learning, and user interaction. The chatbot interface is built using lightweight web framework while the machine learning model is trained using healthcare-related datasets.

Python – Core programming language for the entire project.
Scikit-learn – Used to build and train the Decision Tree Classifier.
Pandas – For reading and managing symptom-disease datasets.
NumPy – For efficient numerical computations.
NLTK – Used for text preprocessing and NLP tasks.
Tkinter– To develop the chatbot's user interface.
Jupyter Notebook – For developing, testing, and presenting code.

2.5 Methods:

The proposed healthcare chatbot system follows a modular and step-by-step approach, combining data preprocessing, machine learning, and rule-based logic to deliver disease predictions and doctor recommendations. The following methods were used in the design and development of the system:

Data Preprocessing: A symptom-disease dataset was cleaned and converted into a binary format, where each symptom is marked as present or absent.
Model Training: A Random Forest Classifier was trained using Scikit-learn to predict diseases based on symptoms. The model was evaluated for accuracy using a test split.
User Input Handling: The chatbot accepts symptoms via a command-line interface, processes them, and converts them into a machine-readable format.
Disease Prediction: The trained model predicts the most likely disease based on the user's symptoms.
Doctor Recommendation: A local JSON file maps diseases to recommended doctors and links, which are displayed after prediction.
Output Display: The chatbot provides the disease name, doctor suggestion, and consultation link in a clear format.

III. METHODOLOGY:

3.1 Input:

The Chatbot system begins by loading a structured symptom-disease dataset using the Pandas library. This dataset contains a list of symptoms as features and corresponding diseases as target labels. Once the data is loaded and the machine learning model is trained, the user is prompted to log in or register through a simple login page. After authentication, the user can enter symptoms through a graphical user interface (GUI) in plain text—for example, "fever," "cough," or "headache." These inputs are then cleaned, tokenized, and standardized using basic Natural Language Processing (NLP) techniques to match them with known symptom terms from the dataset. The matched symptoms are converted into a binary feature vector that aligns with the trained model's input structure. This vector is then passed to the Decision Tree Classifier, which predicts the most likely disease. Based on the predicted condition, the system recommends a suitable doctor type by referencing a predefined mapping stored in a JSON file or dictionary. The final output includes the disease name and doctor recommendation [07], helping users understand their condition and seek appropriate medical advice.

ISSN: 2583-6129
DOI: 10.55041/ISJEM04870

Loading the dataset:

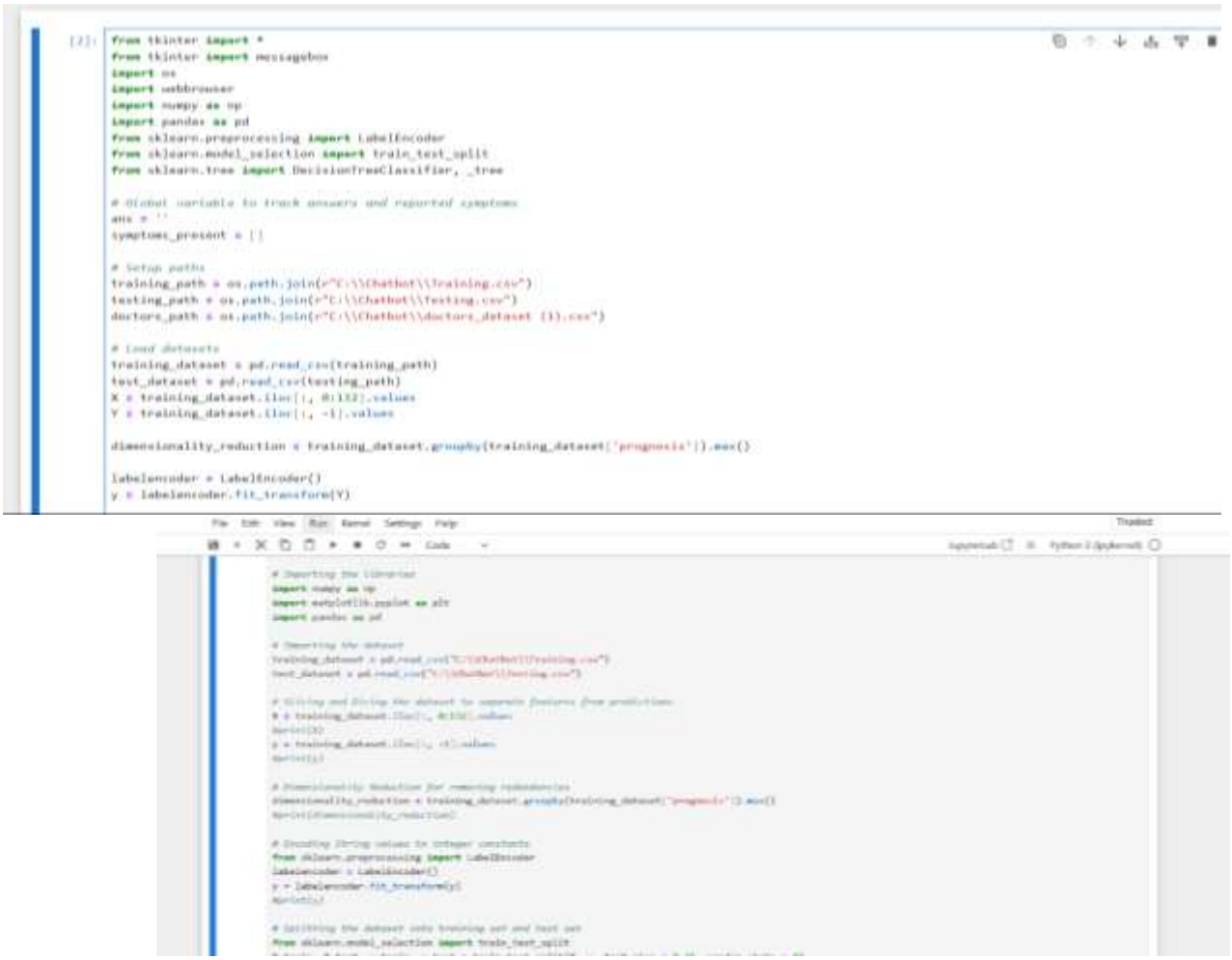


Fig 2: Health Console

3.2 Method of the process:

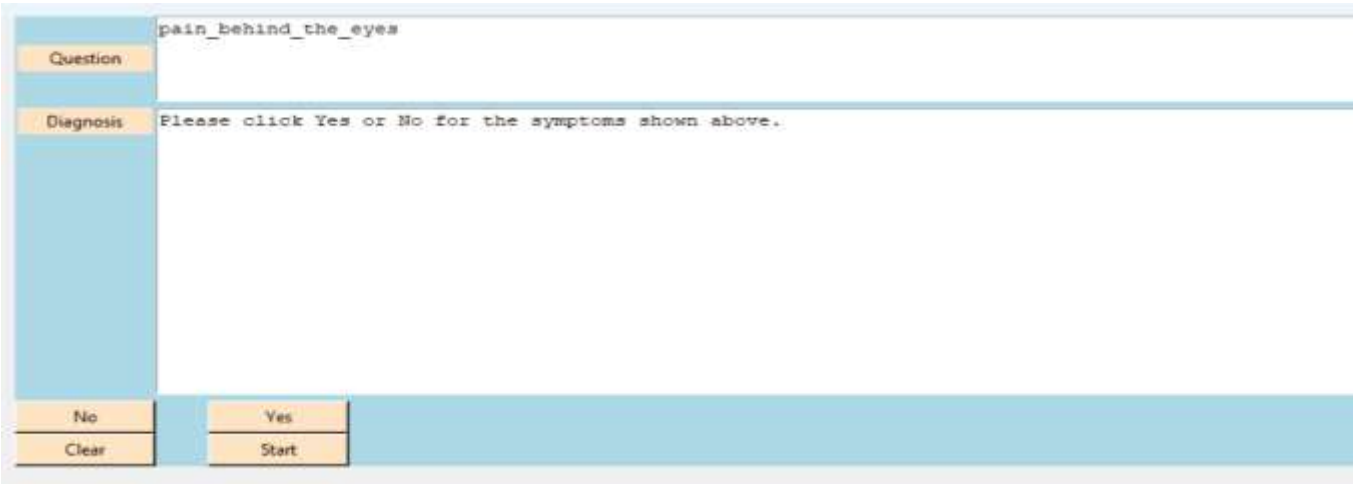
The method used in the Healthcare Chatbot project begins with collecting a structured dataset containing various symptoms and their associated diseases. The data is cleaned by handling missing values, removing duplicates, and standardizing the symptom entries. Users interact with the chatbot through a GUI by entering their symptoms in natural language, which are then processed using basic NLP techniques. These cleaned and processed symptoms are passed to a trained Decision Tree Classifier, which follows a path of logical decisions to predict the most likely disease. Finally, the chatbot displays the diagnosis, confidence score, and suggests a relevant doctor type along with a consultation link.

3.3 Output:

The output of the Healthcare Chatbot is a predicted disease based on the user's symptoms, along with a recommended type of doctor. After analyzing the input through the trained Decision Tree model, the chatbot displays the diagnosis clearly in the GUI. It also provides a short message or response suggesting the next steps, such as consulting a medical professional. This output helps users quickly understand their possible condition and seek appropriate care.

ISSN: 2583-6129
DOI: 10.55041/ISJEM04870

Fig 1: Login page



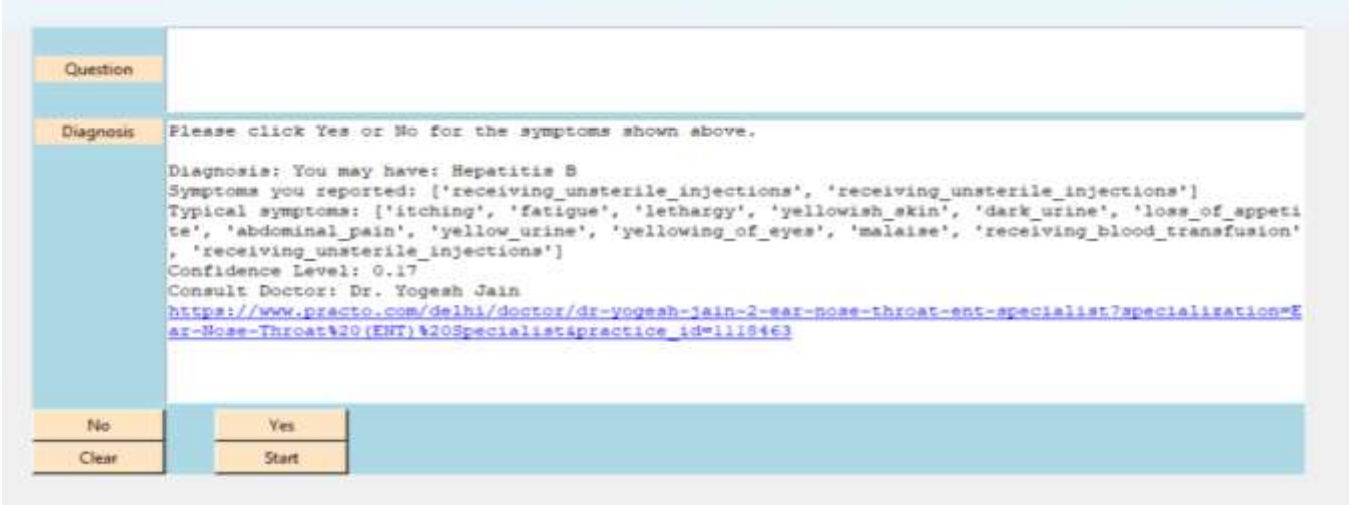


Fig 3: Diagnosis Result and Doctor Recommendation

IV. RESULTS:

The healthcare chatbot system was successfully developed and tested using a Decision Tree Classifier trained on a structured symptom-disease dataset. The Chatbot delivers reliable results by accurately predicting diseases based on user-reported symptoms. It allows users to confirm symptoms through a Yes/No interface, which is then processed by a trained Decision Tree model [19]. Once the diagnosis is complete, the chatbot displays the identified disease, symptoms matched, and a confidence score indicating the model's certainty. The overall response time for symptom analysis and result generation was under two seconds, ensuring smooth user interaction. It also suggests a specialist doctor with a clickable consultation link. The interface is simple, responsive, and provides real-time interaction, making it user-friendly for basic medical guidance.

V. DISCUSSIONS:

The healthcare chatbot developed in this project demonstrates the practical application of machine learning in preliminary medical diagnosis. By leveraging a Decision Tree Classifier and a structured symptom-disease dataset, the system is able to predict potential illnesses based on user-reported symptoms with good accuracy. The integration of basic Natural Language Processing (NLP) [05] ensures that even simple, unstructured text inputs from users can be correctly interpreted and mapped to known medical symptoms. One of the key strengths of the system is its simplicity and accessibility. Additionally, the chatbot offers not only disease prediction but also doctor recommendations, which helps bridge the gap between diagnosis and treatment guidance.

VI. CONCLUSION:

The Healthcare Chatbot project successfully demonstrates the integration of machine learning and natural language processing for basic medical assistance. The system applies Natural Language Processing to interpret user inputs, and a Decision Tree Classifier to make accurate disease predictions. It also recommends a relevant doctor type [07], enhancing its practical usefulness. The chatbot provides a user-friendly interface and guides users toward appropriate healthcare options. It offers a quick, accessible solution for initial symptom analysis. Although the system has limitations in handling rare diseases or vague symptom descriptions, it lays a solid foundation for further improvements. With future enhancements such as voice input, severity analysis, and integration with real-time medical data, the chatbot has the potential to become a valuable tool in preliminary medical consultation and awareness. The project also highlights the importance of clean datasets and model training.

VII. FUTURE SCOPE:

The current version of the healthcare chatbot is a basic yet functional system that allows users to input symptoms through a GUI and receive possible disease predictions along with relevant doctor recommendations. It provides fast responses, and uses a Decision Tree Classifier trained on a symptom-disease dataset. The interface is simple and effective, making it suitable for initial medical assistance in both urban and rural settings**.**
However, there are several promising areas where the system can be improved and expanded:

- **Voice-Based Input**: Integration of speech recognition to allow symptom entry via voice commands.
- **Multilingual Support**: Adding support for regional languages to improve accessibility.
- **Real-Time API Integration**: Connecting with hospital systems and health APIs for live doctor availability and appointment booking**.**
- **Mobile App Development**: Deploying the chatbot as an Android or iOS app for on-the-go access.
- **Electronic Health Record (EHR) Integration**: Personalizing responses based on user's past medical history.

VIII. ACKNOWLEDGEMENT:

Rampilli Bhanu Sankar working as Assistant Professor in Master of Computers Applications (MCA) at Sanketika Vidhya Parishad Engineering College, Vishakhapatnam, Andhra Pradesh, Accredited by NAAC. With over 2 years of experience in Masters of Computer Applications (MCA), he has published a paper in Journal of Emerging Technologies and Innovative Research (JETIR) and he is a member in IAENG. His area of expertise includes C, Data Structures, Java Programming, Python Programming.

Karipe Aparna is pursuing her final semester MCA in Sanketika Vidya Parishad Engineering College, accredited with A grade by NAAC, affiliated by Andhra University and approved by AICTE. With interest in Artificial intelligence. K. Aparna has taken up her PG project on “HEALTHCARE CHATBOT USING PYHTON AND MACHINE LEARNING” and published the paper in connection to the project under the guidance of R. Bhanu Sankar, Assistant Professor, SVPEC.

IX. REFERENCES:

[01] Contextual chatbot for Healthcare purposes

https://ieeexplore.ieee.org/abstract/document/9210351

[02] Multilingual Healthcare Chatbot Using Machine Learning

https://ieeexplore.ieee.org/abstract/document/9456304/

[03] Machine learning based healthcare chatbot

https://ieeexplore.ieee.org/abstract/document/9823901/

[04] A medical chatbot using machine learning and natural language understanding

https://link.springer.com/article/10.1007/s11042-022-12820-4

[05] AI conversational chatbot for primary healthcare diagnosis using natural language processing and deep learning

https://link.springer.com/chapter/10.1007/978-981-16-2543-5_22

[06] A healthcare system using machine learning techniques for disease prediction with chatbot assistance

https://www.scienceopen.com/hosted-document?doi=10.14293/PR2199.000474.v1

[07] Chatbot for Disease Prediction and Treatment Recommendation using Machine Learning

https://ieeexplore.ieee.org/abstract/document/8862707/

[08] Meena Locharla, M. Aravind, P. Prem Kumar AI Healthcare Bot System using Python

https://share.google/jUHCR4OX9bmop7CyC

[9] Sawad, A. B., et al. A Systematic Review on Healthcare AI Conversational Agents for Chronic Conditions, **Sensors (MDPI)**, 2022.
https://doi.org/10.3390/s22072625

[10] G. Kurup and S. D. Shetty, "AI Conversational Chatbot for Primary Healthcare Diagnosis Using Natural Language Processing and Deep Learning," ResearchGate, 2022.
https://www.researchgate.net/publication/357162231_AI_Conversational_Chatbot_for_Primary_Healthcare_Diagnosis_Using_Natural_Language_Processing_and_Deep_Learning.

[11] J.-H. Chen *et al.*, "Online Textual Symptomatic Assessment Chatbot Based on Q&A Weighted Scoring for Female Breast Cancer Pre-screening," Applied sciences, vol. 11, no. 11, pp. 5079–5079, May 2021, Doi : https://doi.org/10.3390/app11115079.

[12] Jim Elliot Christopher James et al., "Natural Language Processing based Human Assistive Health Conversational Agent for Multi-Users," *2021* Second International Conference on Electronics and Sustainable Communication Systems (ICESC), Aug. 2021, doi: https://doi.org/10.1109/icesc51422.2021.9532913.

[13] A. Softic, Jasmina Barakovic Husic, A. Softic, and S. Barakovic, "Health Chatbot: Design, Implementation, Acceptance and Usage Motivation," Mar. 2021, doi: https://doi.org/10.1109/infoteh51037.2021.9400693.

[14] T.-H. Hwang, J. Lee, S.-M. Hyun, and K. Lee, "Implementation of interactive healthcare advisor model using chatbot and visualisation," Oct. 2020, doi: https://doi.org/10.1109/ictc49870.2020.9289621.

[15] R. Rajan, G. Amudhavalli. Python Chatbot for Healthcare, ResearchGate, 2024. https://www.researchgate.net/publication/379690974_Python_Chatbot_for_Healthcare

[16] "AI Based Healthcare Chatbot using Natural Language Processing and Pattern Matching," © 2023 IJNRD , vol. 8, no. 5, p. 793, 2023. https://www.ijnrd.org/papers/IJNRD2305598.pdf

[17] Decision Tree in Machine Learning - GeeksforGeeks https://www.geeksforgeeks.org/machine-learning/decision-tree-introduction-example/

[18] Natural Language Processing

https://www.ibm.com/think/topics/natural-language-processing

[19] Vinod Kumar Shukla; Nitin Pandey; Ajay Rana Chatbot for Healthcare System Using AI
https://ieeexplore.ieee.org/abstract/document/9197833

[20] A literature review on chatbots in healthcare domain

https://www.viit.ac.in/images/Quality/NAAC/NAAC-Data/Criteria-3/Research-Journal-Paper/JP-AY-2019-20-PART-1.pdf

[21] Design and development of diagnostic Chatbot for supporting primary healthcare systems

https://www.sciencedirect.com/science/article/pii/S1877050920306499

[22] AI enabled NLP based text to text medical chatbot

https://ieeexplore.ieee.org/abstract/document/10117966/