VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY
Department of Computer Engineering

Project Report on

MedChat:A Healthcare Chatbot For Disease Prediction and Nearby Hospital Suggestion

In partial fulfillment of the Third Year, Bachelor of Engineering (B.E.) Degree in Computer Engineering at the University of Mumbai

Academic Year 2023-24

GROUP NUMBER : 41			
MENTORS			
Mrs.Rupali Soni,Mrs.Sunita Sahu, Department of Computer Engineering			
GROUP MEMBERS			
Name	Class	Roll No	Email
Khwaish Shahani	D12C	55	2021.khwaish.shahani@ves.ac.in
Jaitra Shahani	D12C	56	2021.jaitra.shahani@ves.ac.in
Karan Khatri	D12B	24	2021.karan.khatri@ves.ac.in

VIVEKANAND EDUCATION SOCIETY'S INSTITUTE OF TECHNOLOGY
Department of Computer Engineering

Certificate

This is to certify that ***Khwaish Shahani, Jaitra Shahani ,Karan Khatri*** of Third Year Computer Engineering studying under the University of Mumbai have satisfactorily completed the project on "***Medchat:A Healthcare Chatbot for Disease Prediction and Nearby Hospital Prediction*** " as a part of their coursework of PROJECT-II for Semester-VI under the guidance of their mentor ***Mrs. Rupali Soni,Mrs.Sunita Sahu*** in the year 2023-24.

This project report entitled ***Medchat:A Healthcare Chatbot for Disease Prediction and Nearby Hospital Prediction*** by ***Khwaish Shahani,Jaitra Shahani,Karan Khatri*** is approved for the degree of Third year of Engineering (T.E.) in Computer Engineering.

Programme Outcomes	Grade
PO1,PO2,PO3,PO4,PO5,PO6,PO7, PO8, PO9, PO10, PO11, PO12 PSO1, PSO2	

Date:

Project Guide:

Project Report Approval
For
T. E (Computer Engineering)

This project report entitled ***Medchat: A Healthcare Chatbot For for Disease Prediction and Nearby Hospital Prediction*** is approved for the degree of Third Year of Engineering (T.E.) in Computer Engineering

Internal Examiner

External Examiner

Head of the Department

Principal

Date:

Place:

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

(Signature)

(Jaitra Shahani and Roll No- 55)

(Signature)

(Khwaish Shahani and Roll No- 55)

(Signature)

(Karan Khatri and Roll No- 24)

Date:

ACKNOWLEDGEMENT

We are thankful to our college Vivekanand Education Society's Institute of Technology for considering our project and extending help at all stages needed during our work of collecting information regarding the project.

It gives us immense pleasure to express our deep and sincere gratitude to Assistant Professor **(Mrs.) Rupali Soni** (Project Guide) for her kind help and valuable advice during the development of project synopsis and for her guidance and suggestions.

We are deeply indebted to Head of the Computer Department **Dr.(Mrs.) Nupur Giri** and our Principal **Dr. (Mrs.) J.M. Nair ,** for giving us this valuable opportunity to do this project.

We express our hearty thanks to them for their assistance without which it would have been difficult in finishing this project synopsis and project review successfully.

We convey our deep sense of gratitude to all teaching and non-teaching staff for their constant encouragement, support and selfless help throughout the project work. It is a great pleasure to acknowledge the help and suggestion, which we received from the Department of Computer Engineering.

We wish to express our profound thanks to all those who helped us in gathering information about the project. Our families too have provided moral support and encouragement several times.

Computer Engineering Department
COURSE OUTCOMES FOR T.E PROJECT

Learners will be to,

Course Outcome	Description of the Course Outcome
CO 1	Able to apply the relevant engineering concepts, knowledge and skills towards the project.
CO2	Able to identify, formulate and interpret the various relevant research papers and to determine the problem.
CO 3	Able to apply the engineering concepts towards designing solutions for the problem.
CO 4	Able to interpret the data and datasets to be utilized.
CO 5	Able to create, select and apply appropriate technologies, techniques, resources and tools for the project.
CO 6	Able to apply ethical, professional policies and principles towards societal, environmental, safety and cultural benefit.
CO 7	Able to function effectively as an individual, and as a member of a team, allocating roles with clear lines of responsibility and accountability.
CO 8	Able to write effective reports, design documents and make effective presentations.
CO 9	Able to apply engineering and management principles to the project as a team member.
CO 10	Able to apply the project domain knowledge to sharpen one's competency.
CO 11	Able to develop a professional, presentational, balanced and structured approach towards project development.
CO 12	Able to adopt skills, languages, environment and platforms for creating innovative solutions for the project.

Index

List Of figures

List Of Tables

Abstract

The "Disease Prediction Based on Symptoms Healthcare Chatbot" project is initiative that uses natural language processing to improve healthcare by providing accurate and timely disease predictions based on patient symptoms. This approach, which relies on the expertise of healthcare professionals, can lead to variations in diagnoses and delayed disease identification. The project uses natural language processing techniques to streamline the process, identifying patterns and correlations that might not be apparent through traditional diagnostic methods. The data collected comes from sources like electronic health records, wearable devices, and medical literature. The data is processed to extract relevant features, and predictive models are trained using machine learning algorithms. The models are continually refined and improved as new data becomes available. The project offers healthcare providers a tool to predict and diagnose diseases earlier, enabling more effective treatments and improved patient care. It also has the potential to reduce healthcare costs by avoiding unnecessary tests and treatments. Patients benefit from quicker access to accurate diagnoses, reducing anxiety and uncertainty associated with prolonged diagnostic processes. Overall, the project represents a transformative step towards more efficient, cost-effective, and patient-centered healthcare.

Chapter 1: Introduction

1.1 Introduction

The healthcare sector plays a crucial role in ensuring the well-being of individuals worldwide. With the increasing global population, there is a strain on medical facilities and healthcare professionals due to the large influx of patients. To address this challenge, there is a need to improve healthcare services through the implementation of technology. Embracing technology can provide greater convenience to both patients and healthcare officials, ultimately enhancing the efficiency of the healthcare domain.

The Chatbot is a **computer program designed to simulate human conversation through text or voice interactions using natural language processing techniques**.The primary goal of chatbots is to imitate human conversations as soon as possible.Chatbots are typically built with user interface for receiving inputs and providing responses.Medical chatbots are specifically designed to assist patients with their health related inqueries and needs.These chatbots use algorithms to analyze user enquiries and recognize patterns enabling them to respond to similar request in accurate and timely manner.Chatbots can provide users with healthcare related advice direct them to **appropriate medical resources**,and even **schedule appointments** with healthcare providers.Chatbots can be particularly helpful where users have questions about their health outside of regular office hours or where healthcare providers are known.Normally Users are not aware about all the treatment or symptoms regarding the particular disease. For small problem user have to go personally to the hospital for check-up which is more time consuming

In recent years, chatbots have emerged as valuable tools in various industries, including healthcare. Unlike traditional chatbots that focus on areas like flight booking and hotel management, healthcare chatbots are specifically designed to assist patients with their health-related inquiries and needs. These sophisticated programs utilize natural language processing techniques to simulate human conversation and provide accurate and timely responses to users' queries.

Medical chatbots offer a wide range of benefits, such as providing healthcare-related advice, directing users to appropriate medical resources, and even facilitating the scheduling of appointments with healthcare providers. They are particularly valuable for addressing health-related questions outside of regular office hours and can significantly reduce the burden on healthcare providers by handling inquiries efficiently.

The functionality of medical chatbots relies on natural language processing, enabling users to submit their health-related queries and receive guidance without the need to physically visit a hospital. By leveraging technologies such as Google API for voice-text and text-to-voice conversion, chatbots can cater to a wide range of user needs, ultimately enhancing the accessibility and convenience of healthcare services.

..

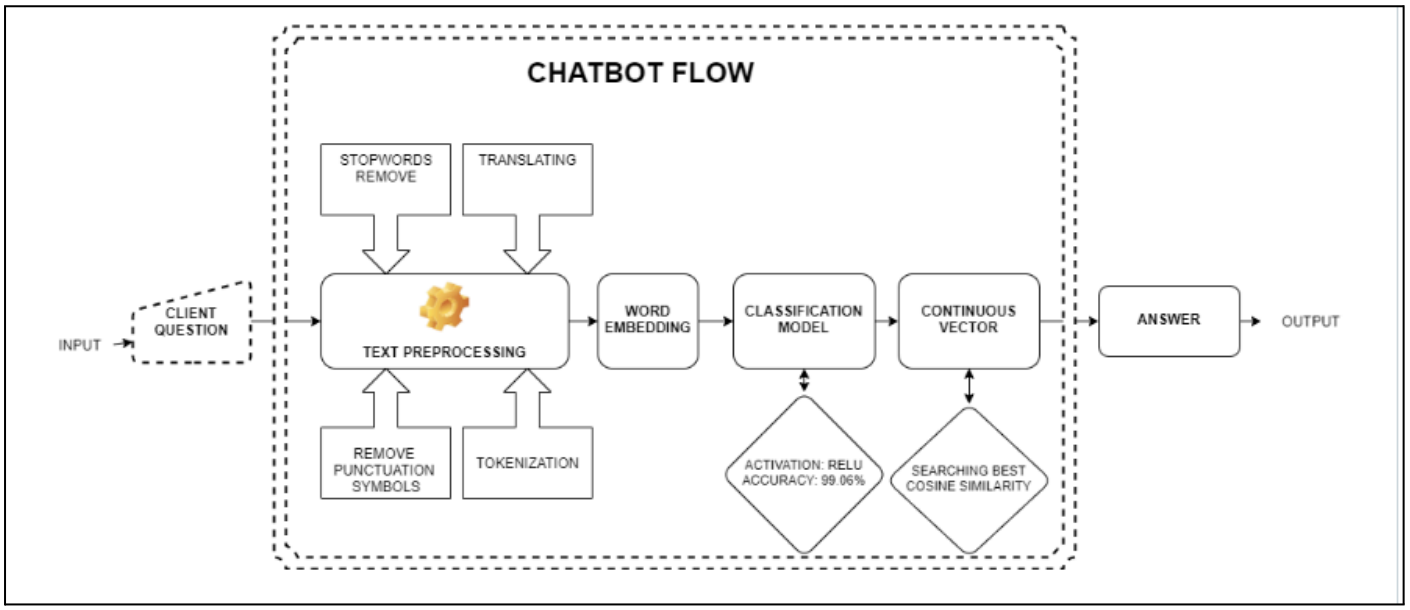


Fig 1.1 Introduction of chatbot

1.2 Motivation

.

The **BERT (Bidirectional Encoder Representations from Transformers)** model was developed with the motivation to improve the understanding of contextual language information in natural language processing (NLP) tasks. BERT is a pre-trained language model that has achieved state-of-the-art performance on various NLP tasks.

1. **Contextual Understanding:** BERT is designed to pre-train deep bidirectional representations from unlabeled text by considering both left and right context in all layers. This contextual understanding helps BERT capture intricate linguistic patterns and semantics, enabling it to understand the context of user queries and provide accurate responses.
2. **High Accuracy:** BERT has achieved state-of-the-art performance on multiple NLP tasks, including question answering, language inference, sentiment analysis, and text classification. The fine-tuned BERT model exhibits high accuracy in understanding the context of the input text and generating appropriate responses.
3. **Flexibility:** BERT can be fine-tuned for different types of question-answering (Q/A) tasks, such as open-domain Q/A, factoid Q/A, and conversational Q/A. This flexibility makes BERT suitable for a wide range of chatbot use cases, allowing it to adapt to different Q/A scenarios and provide relevant answers to user queries.

4. **Transfer Learning:** BERT leverages transfer learning, which involves pre-training on large amounts of unlabeled text data and then fine-tuning the model for specific NLP tasks with smaller labeled datasets. This approach allows BERT to learn general language representations and apply them to specific chatbot tasks, achieving state-of-the-art results without significant task-specific architecture modifications.

BERT is a powerful pre-trained language model that has achieved state-of-the-art performance on several natural language processing (NLP) tasks. It can be fine-tuned for question-answering (Q/A) tasks, making it a reliable model for Q/A chatbots. The Q/A BERT model has high accuracy, flexibility for different types of Q/A tasks, and contextual understanding, which helps it answer questions accurately by considering the context .

1.3 Problem Definition

The aim of the project is to provide an **easier** and **convenient** alternative way of giving healthcare services. This chatbot is introduced to focus on giving service to all people to ease their suffering and make their life become better. Although there are many existing chatbots, there are also several problems that need to be solved to give better healthcare services. Due to a backlogged and overburdened system, booking an appointment with your family physician is no easy task for most Indians. Wait times have extended to multiple weeks only to experience a rushed meeting with your doctor and leaving with questions unanswered. The proposed solution is a trusted healthcare chatbot that increases accessibility of healthcare and empowers Indians to make informed decisions regarding their well-being.

[i] Lack of voice-based communication applied in healthcare chatbot.

- Most of the existing chatbots that giving mental healthcare services in the market are text-based communication such as Woebot and Wysa.
- The user can communicate with the chatbot by just texting or selecting the options that are provided by the chatbot.
- The user needs to do extra work such as import the library to the system if they want the chatbot able to speak to them.
- This will extremely hard for those who do not have relevant knowledge. For the proposed chatbot, there is no need to trouble the user since voice-based communication is done by building the chatbot.
- All the user needs to do is choose either input in voice or audio form only.
- We truly hope that this chatbot can reach out to help different kinds of people especially those who have difficulty in hearing and vision since the chatbot able to provide both voice and text-based communication.
- Besides, it can be an advantage if the chatbot possesses more ways to carry out the communication so that it can provide a better service to the people.

[ii]Lack of diverse content of response for healthcare chatbot

- The disadvantages of most chatbots are the lack of generality and variation in style. (Woudenberg, A.V, 2014).
- If the chatbot is built by using some API such as Facebook Message Bot, the user might get the same response over and over again if they are kept asking the same things to the chatbot.
- Not to say that, if the user inputs something that chatbot never know before, the chatbot would not know the meaning of the word and would be unable to respond back correctly.

- Chatbot that replies to the same response maybe can use for other fields but it is might not suitable for providing counseling healthcare service to humans.
- To providing more diverse content of the response from the chatbot, we are using deep learning to train the chatbot so that the chatbot can recognize the meaning of the words more accurately and able to give a relative response.

1.4 Project Scope

The purpose of this project is to develop a chatbot that can provide a counseling service that is able to give different responses to humans by using machine learning to recognize different text inputs and give related responses to the user. This chatbot also can carry out voice-based or text-based communication with a human by using Automatic Speech Recognition (ASR).

i. This chatbot only supports oral and text conversations in **English,Hindi and Marathi** . Another language is not supported currently.

ii. This chatbot only focuses on giving the content of the response which is related to the healthcare support service only. The conversation which involved other unrelated field is not supported.

iii. This chatbot will able to recognize the meaning of the text input from the user whether in voice or text form and respond back to them accordingly. The response of the chatbot will be in voice form and text form together.

iv. This chatbot will try to avoid giving the repeating response although it keeps getting the same text input from the user.

Chapter 2: Literature Survey

2.1 Research Papers Referred

a. Abstract,Author and published year of the research paper

Title	Publication	Year of Publishing	Methodology	Pros	Cons
Microservice chatbot architecture for chronic patient support[1]	Journal of Biomedical Informatics	2019	Messaging Platform Docker Platform Kibana	Overcoming modularity of Standardiztion and Kibana	Complex working and designing of working systems
Acceptability of artificial intelligence (AI) led to chatbot services in healthcare :A mixed method study[2]	Digital Health	2019	NLP,Regression Techniques ,Use of quantitative data	Identified potential factors associated in delay acceptability or refusal	Answers were collected from relatively experienced users of digital technological
Design and development of Smart Healthcare Chatbot application Using AL_Ml [3]	Journal of Natural Remedies	2020	NLP , Concepts like Tokenization ,TF -IDF, n gram Google API , and Resberry PI	Innovative use of RasberryPI and Google API	Data security and authenticity was a concern as the sources of information which chatbot uses were not mentioned.
Design and Development of Diagnostic development Chatbot for Primary healthcare Systems **[4]**	Procedia Computer Science	2020	Natural language Toolkit,Decision Tree algorithm	Presence of features like speech recognition.	Low accuracy of the proposed system other ML algorithms might provide better accuracy
Self Diagnosis Health Care Chatbot	International Journal of	2020	NLP Tokenization	Chatbot can detect and	Chatbot designed for limited

using Machine learning[5]	advanced science and technology		techniques,Decision Tree Algorithm SQLITE, Bootstrap,Javascript	categorize disease as major and minor	databases and data authenticity was a major issue.
AL chatbot Design during an Epedimic like Novel cornavirus[6]	MDPI Healthcare	2020	NLP along with AIML,Python Watson	Text to speech feature and in case of chatbot is capable of sending patient location and health conditions to nearby doctors	Works only for predefined questionnarie
The SmartHealthcare prediction using Chatbot[7]	International Journal of Recent Technology and Engineering	2020	JAVA programming and AIML	Liberity for people with a doctor anytime anywhere	Provides minimal description and symptoms
AI Healthcare Interactive Agent using NLP[8]	International Journal of Innovative technology and Exploring Engineering	2019	Language Understanding Intelligence Services (LUIS),NLP and Morphology	Provides a diet plan according to BMI and requirement of file	Gives wrong output if inputs are not given in sequential order
Text Messaging Based Medical Diagnosis Using Natural Language Processing and Fuzzy Logic [9]	Hindawi Jounal of Healthcare Engineering	2020	NLP, Machine Learning, Twilio	Can suggest a diagnosis using a direct approach	Not secure against false positives cases.
Automated Medical Chatbot[10]	SSRN Electronic	2017	JAVA based AIML	Capable of providing remedy	Low Accuracy of the system
Disease Prediction by Machine learning Over Big data from Healthcare communities[11]	IEEE Access	2017	CNN based multimodal and risk prediction	Chatbot is capable of missible details accurately	Complex architecture with hardly any acknowledgment of novelty
Comparing different supervised machine learning algorithms for disease prediction [12]	BMC Medical Informatics and decision making	2019	Python Software	PRISMA guidelines were followed by selecting the articles	Border level classification was considered rather considering sub classifications or hyper parameters

Application of machine Learning in Disease Prediction[13]	IEEE-Xplore	2018	Logistic Regression SVM Random Forest database and Wisconson Breast Cancer dataset Heart Disease Dataset	Chatbot helps in early detection of chronic and fatal disease.Heart Attack and Diabetes	Involves tim consuming steps like data mugging which can be automated.
An intelligent chatbot using deep learning with Bidirectional RNN and attention model [14]	Science Direct Materials Today Proceedings	2020	Tensorflow Neural Machine, BRNN ,Redit database	Open domain database which can be subjected to any specific domain if needed	High hardware configuration required for functioning of the system.
Disha : An implementation Of machine learning Bangla Healthcare Chatbot[15]	IEEE-Explore	2019	Named Entity Recognition ,ML , Google Api	Developed in Local Search language which is Bangla	No option for other languages like English

2.2. Comparison of results with existing systems

Basis	Existing system	Our system
Accuracy	Accuracy is not low	Accuracy is very high 95.47
Response	No instant response	Very instantaneous response
Range of Chatbot	Works in particular region	User will be using chatbot anywhere anytime
Language	The existing chatbots are developed only in one language therefore it is language constraint	Our system is multilingual
Dataset	Limited number of disease in dataset	Number of diseases in the dataset is more.
Voice	Technical issues like voice messages are not accurate in the existing system.	Technical issues like voice messages are accurate in the existing system.
Privacy and accuracy	Chatbot may collect and store personal health information which rise concerns about privacy and concerns	Chatbots may also be vulnerable to hacking or other cyberattacks, which may compromise the confidentiality and integrity of patient information in our system
Technical issues	Chatbots may experience	Our chatbot does not produce any

	technical issues such as system crashes, glitches, or delays, which may affect their reliability and usability. Technical support and maintenance may also be required to ensure that chatbots function properly and effectively	technical crash or glitches .the chatbot works properly and effectively

Chapter 3: Requirement Gathering for the Proposed System

In this chapter, there is discussion regarding the resources used and how analysis of what the the client actually needs and what can be provided. The functional and non-functional requirements and finally the software and hardware used are also discussed.

3.1 Introduction to requirement gathering

Requirement gathering is a crucial step in the software development process. It involves collecting, documenting, and analyzing the needs and expectations of stakeholders for a software system. The goal of requirement gathering is to ensure that all relevant requirements for the software system are identified, documented, and understood by all stakeholders.

Steps in Requirement Gathering

The requirement gathering process is an iterative process that involves several steps 1:

1. Requirements Elicitation: This step involves gathering information about the needs and expectations of stakeholders for the software system. Techniques such as interviews, surveys, focus groups, and discussions are used to gather information from stakeholders.
2. Requirements Analysis: In this step, the information gathered in the requirements elicitation step is analyzed to identify the high-level goals and objectives of the software system. Any constraints or limitations that may affect the development of the software system are also identified.
3. Requirements Specification: The requirements identified in the previous step are documented in a clear and concise manner. This documentation serves as a reference for the development team and other stakeholders throughout the software development process.
4. Requirements Validation: The documented requirements are validated to ensure that they accurately represent the stakeholders' needs and expectations. Techniques such as expert reviews, prototyping, and simulation are used for requirement validation.
5. Requirements Management: Requirement management is an ongoing process that involves maintaining and tracking changes to requirements throughout the software development lifecycle. It includes activities such as version control and change control.

Importance of Requirement Gathering

Requirement gathering is a critical step in the software development process for several reasons

- It helps ensure that the final software system meets the needs and expectations of the stakeholders and end-users.
- It helps define the scope of the project and guides the design and development process.
- It serves as a reference point for testing and validation activities.
- Accurate and complete requirement gathering helps avoid misunderstandings and minimize scope creep.

Challenges and Best Practices

Requirement gathering can be challenging due to various factors, such as conflicting stakeholder expectations and evolving project requirements. Effective requirement gathering requires strong communication skills, active listening, and understanding the perspectives and needs of different stakeholders. It is important to analyze and prioritize requirements, resolve conflicts, and document requirements in a clear, unambiguous, and verifiable manner.

Some common challenges in requirement gathering include not understanding the problem or opportunity that the software system is intended to address and not involving the right stakeholders in the process. To overcome these challenges, it is important to start by understanding the problem or opportunity and involve all relevant stakeholders in the requirement gathering process

Conclusion

Requirement gathering is a crucial step in the software development process. It involves collecting, documenting, and analyzing the needs and expectations of stakeholders for a software system. Accurate and complete requirement gathering helps ensure that the final software system meets the needs and expectations of the stakeholders and end-users. It serves as the foundation for the entire development cycle and guides the design, development, testing, and validation activities.

3.2 Functional Requirements

- Chatbots aid in the improvement of customer experiences by assisting or guiding customers
- during critical choices. Not all chatbot software solutions, however, are created equal.
- Specific restrictions have a significant impact on possible savings and consumer benefits.
- Chatbots assist customers in making better selections by providing support or advice.
- Above all, chatbots offer a plethora of automation possibilities that can greatly speed up service delivery.
- All chatbot software solutions, however, are not treated equally.
- There are several restrictions that have a significant impact on possible savings and consumer benefits.
- The most important requirements of chatbot software:
  Complex dialogues
- In addition to understanding and interacting within conversations, an outstanding chatbot
- software has NLP (Natural Language Processing) to analyze the context of a
- conversation.
- It can identify the intent of a question to provide an accurate answer and suggest options to confirm or resolve the issue.
- The best chatbots have advanced conversation features and can proactively search for information and ask clarifying questions even if the conversation is not linear.

- **Training of our bot**

We should train our bot and it gives answer after understanding the queries.

- **Mode:**

The customer can change the mode according to his/her choice. Modes like dark or light mode.

- **Voice to text**

In chatbot applications, voice bots are chatbots that communicate in an automated way using vocal input and output.. In turn, it will respond with a voice of its own without the need for human interaction on its end.

3.3 Non-Functional Requirements

The nonfunctional requirements for a healthcare chatbot play a crucial role in ensuring its effectiveness and compliance with industry standards. Here are some key nonfunctional requirements to consider:

Regulatory Compliance:

Navigating the regulatory environment is essential for healthcare chatbots. Legal counsel is required to guide the development of the chatbot to ensure compliance with regulations in healthcare .

Limitations of Chatbots in Healthcare:

It's important to recognize that chatbots cannot replace a doctor's expertise or take over patient care. However, they can enhance the efficiency of patient care delivery without compromising quality .

Intent Recognition and User Assistance:

Chatbot developers need to determine the intent of the user and provide the best help the user needs. Designing the chatbot to effectively recognize user intent and provide appropriate assistance is crucial .

Multimodal Conversational Experiences:

With advancements in speech recognition and natural language processing (NLP), healthcare chatbots should be capable of engaging in multimodal conversational experiences, including voice, keypad, gesture, and image interactions .

Market Size and Growth:

The healthcare chatbots market is experiencing rapid growth, with significant market size projections. Considering the scalability and adaptability of the chatbot to accommodate this growth is important

Security and Privacy Compliance:

Given the sensitive nature of patient information, ensuring that the chatbot complies with security and privacy regulations, such as HIPAA, is essential.

Real-time Response and Instant Information Delivery:

Healthcare chatbots should provide instant responses and deliver critical information rapidly, especially in time-sensitive medical situations.

Training and Intelligence:

The intelligence of the chatbot is crucial, and proper training is necessary to ensure accurate interactions, especially when dealing with sensitive patient information and medical advice

3.4.Hardware, Software , Technology and tools utilized

Hardware:

- Windows 10/11
- Ubuntu
- R5-3600H / Intel i5-i10 Gen
- 6GB+ Ram
- GTX 1050 Graphics Processor Camera Unit

Software:

- Visual Studio Code
- Anaconda
- Jupyter Notebook
- Google Colab

Technology and tools utilized:

- Nvidia Gpu
- Python libraries - flask, numpy, pandas, sklearn.metrics, tensorflow, matplotlib, seaborn
- Html
- Css
- JavaScript

3.5 Constraints

The constraints of a healthcare chatbot encompass various aspects that need to be considered for its effective implementation and operation. These constraints include:

Ethical Considerations: Chatbots providing medical advice must adhere to ethical constraints when dealing with sensitive matters, ensuring the privacy and well-being of the patients .

Limitations in Replacing Human Expertise: While chatbots can enhance the efficiency of patient care delivery, they cannot replace a doctor's expertise or take over patient care entirely. It's important to recognize the limitations of chatbots in healthcare and ensure they complement, rather than replace, human medical expertise

Constant Algorithm Training and Updates: For chatbots to accurately detect diseases and provide proper recommendations, continual algorithm training and updates are necessary due to the constant improvements in current standards of care .

Adherence to Regulatory Standards: Healthcare chatbots must comply with regulatory standards and guidelines to ensure the accuracy and reliability of the information provided to users.

User Interface and Experience: The design and usability of the chatbot interface are crucial for effective communication and user engagement. It's essential to consider the user experience and ensure that the chatbot is user-friendly and accessible across various channels .

Natural Language Processing (NLP) Capabilities: The chatbot's NLP algorithms and knowledge base are critical for analyzing user inquiries and providing accurate and timely responses. The chatbot's ability to understand and interpret natural language is essential for effective communication

Chapter 4: Proposed Design

4.1 Block Diagram of the system:

Block Diagram of Healthcare Chatbot

A block diagram of a healthcare chatbot typically illustrates the architectural flow of constructing an AI chatbot for healthcare. The diagram details the various elements and interactions involved in the chatbot's operation. Here's a comprehensive overview of the key elements usually included in the block diagram:

System Prompt and Persona Definition: This component defines the chatbot's persona and informs users about the chatbot's function. In a healthcare context, this could mean the chatbot is designed to help patients with initial symptom diagnosis and scheduling appointments

1.API Integration: The AI chatbot can execute API requests to a third party with predefined function calls based on the user's query. This allows the chatbot to fetch and present relevant information in a conversational manner

2.Knowledge Base: The chatbot is typically connected with a knowledge base for cloud computing, which allows it to provide quick treatment responses to accidents and changes in the conditions of patients with chronic diseases

3.Data Collection and Patient Interaction: The chatbot collects valuable data on patient treatment assessment, which is vital for healthcare providers to evaluate the success of their patient's treatment. Additionally, the chatbot interacts with patients to collect information such as medical history, medicine names, and dosages

4.Human-Robot Interaction Levels: The proposed framework for the chatbot healthcare service includes four levels: data level, information level, knowledge level, and service level. These levels facilitate smooth human-robot interaction and efficient implementation of the chatbot healthcare service

5.Logic and Content: The chatbot's logic determines how to respond to user input, and richer content with subject-matter information leads to more useful chatbots

6.Security and Privacy: Chatbot platforms typically ensure the highest standards of privacy and security to establish and maintain patients' trust

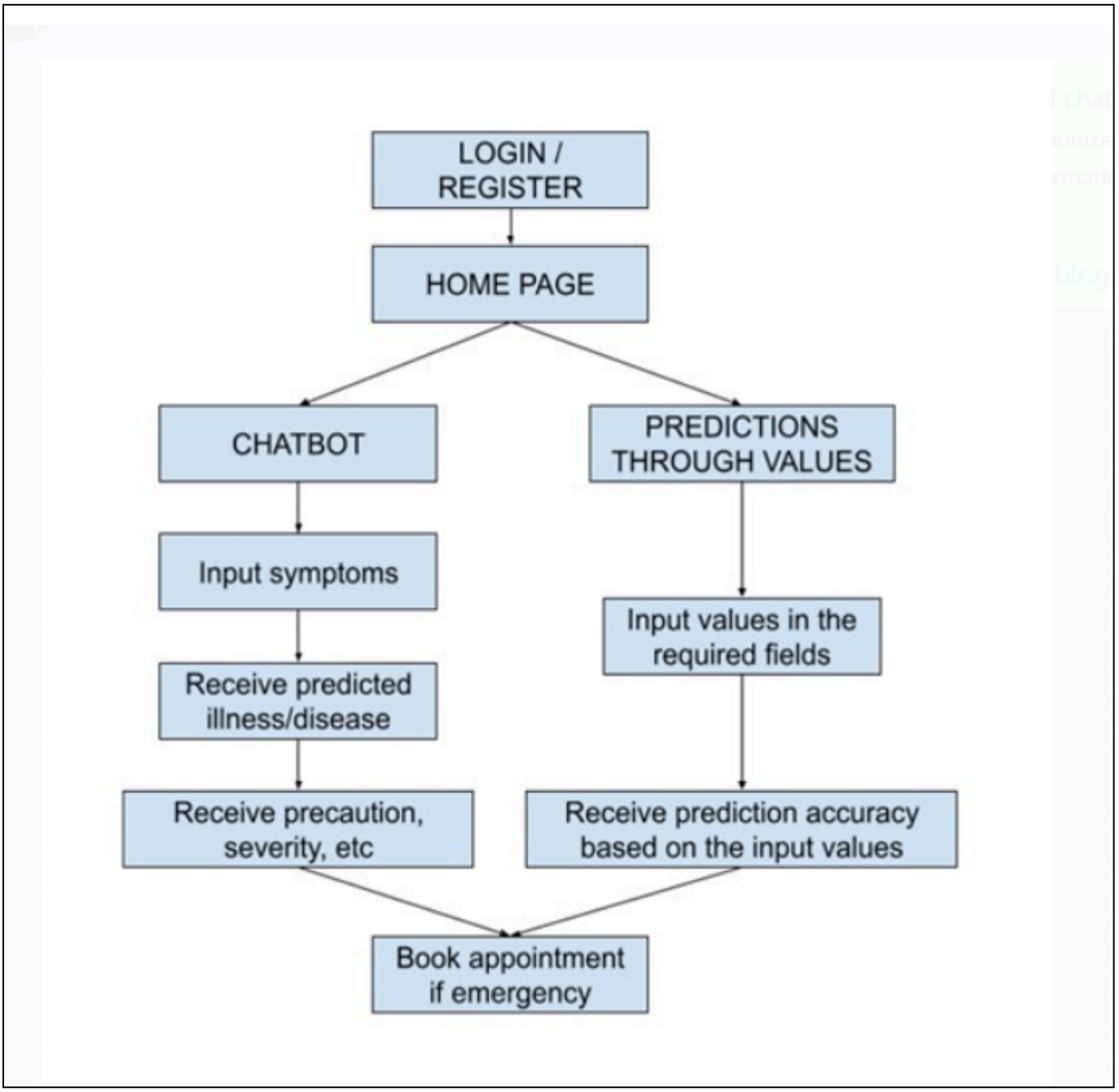


Fig 4.1 Block diagram

4.3 Detailed Design :

[1]The project starts by getting the patients's real-time dataset. gathering a relevant dataset related to patients dataset .This data will be useful for training and chatbot responses.

[2]Text Processing : Perform text processing on the dataset Lowercase all the datasets.Lowercase all text to ensure proper uniformity.tokenize text into sentences and words using libraries like NLTK.Apply Lemmatization to remove punctuation and noise from other data.

[3] Greeting detection:Create a function to detect respond to patients greetings.thereby improving user engagement.

[4]**Response Generation:** Maintain a user interactions and responsis for analysis.

[5]**User generation loop:** Create a user interaction loop that: - Greets the user and waits for input. Detects and responds to user greetings. Invokes the response generation function to provide answers to user queries.Ends the conversation upon user request (e.g., "bye").

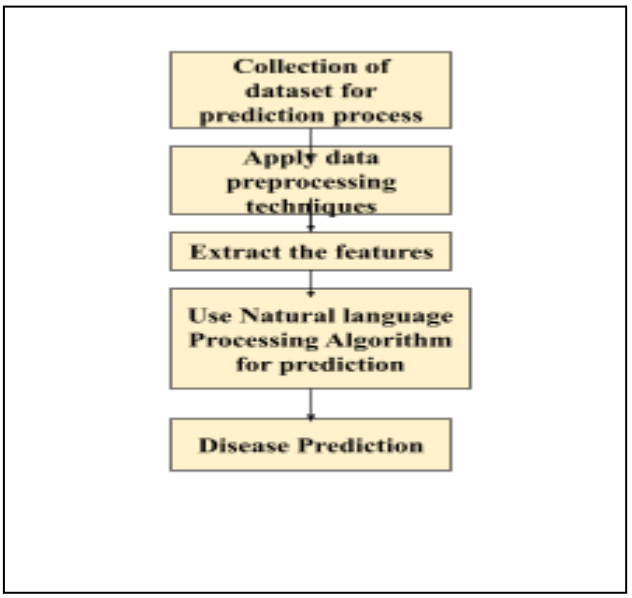


Fig 4.3. Detailed Design

4.4 Project Scheduling & Tracking using Timeline / Gantt Chart

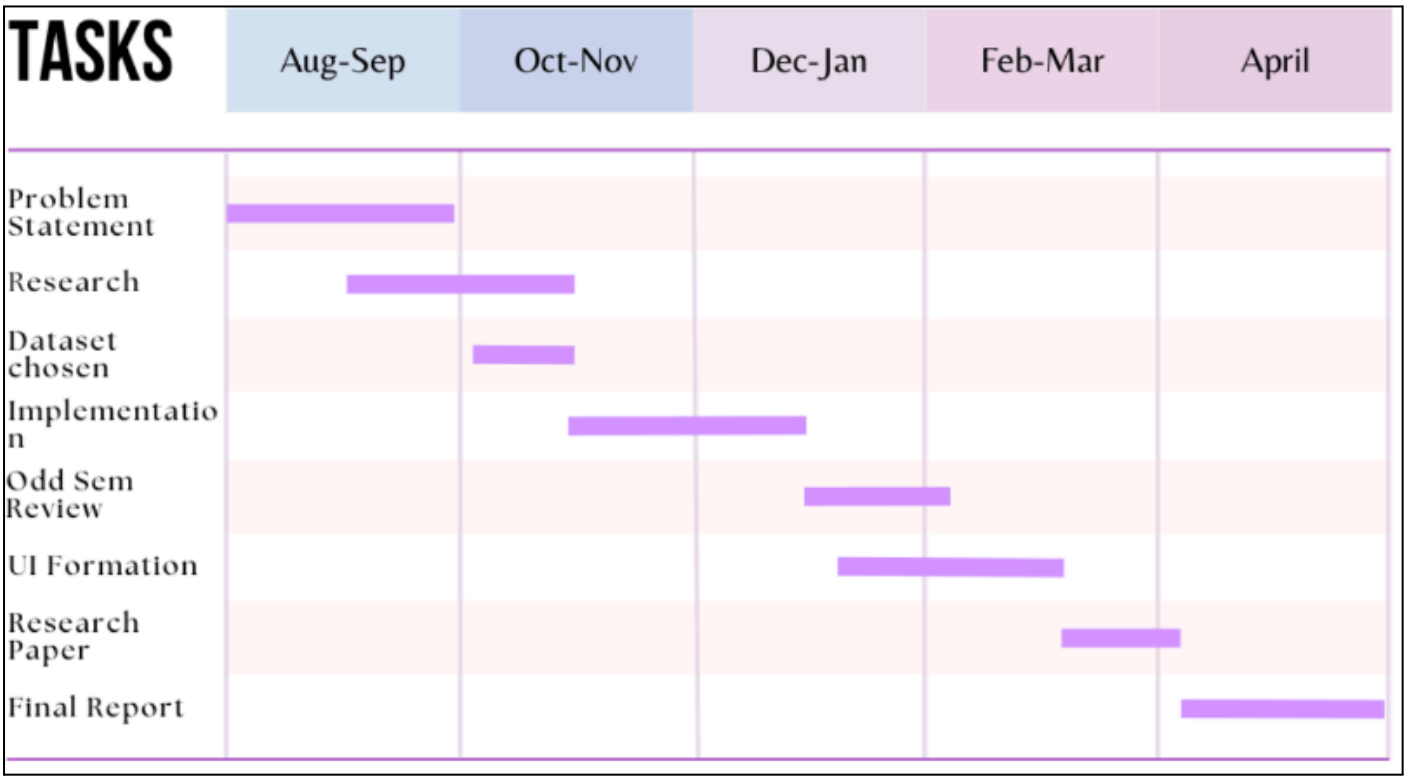


Fig 4.4.1 Project Scheduling (Aug'23 - April'24)

Chapter 5: Implementation of the Proposed System

5.1. Methodology employed for healthcare chatbot development

In recent years,the volume of data generated in the medical and healthcare sectors are increased dramatically.Supervised learning based approaches are amongst the most useful for scientific study and clinical applications. On the other hand supervised learning based approaches are among the most useful for scientific study and practical applications in clinical settings. The primary purpose of this efforts is to enhance disease symptom prediction and diagnostic methods.

The structure is made up of five basic parts all working together:

A. Data Collection A data set1 is chosen from Kaggle which contains a list of diseases and symptoms total of 4921 rows that can classify 42 diseases.We have a filtered data set that contains symptoms and diseases corresponding to the symptoms. Some columns list the illnesses, their symptoms, the preventative measures that should be performed, as well as the weights of the diseases. Cleaning up this dataset is simple and may be done in any language that supports file management. To utilize this data effectively, the user needs just grasp the concept of rows and columns.

Understood, let's dive into the specific preprocessing steps and feature selection techniques tailored to your data.

Data Preprocessing:

1. Special Characters and White Spaces Removal: Begin by scanning the dataset for any special characters or white spaces. Use string manipulation techniques to eliminate these elements. This ensures data uniformity and cleanliness.

2. Lowercasing:Convert all text data to lowercase to standardize the text representation. This prevents the model from treating words with different cases as distinct entities, thus reducing feature dimensionality.

3. Disease Dictionary Creation: Create a dictionary mapping each disease to a unique index and vice versa. This step facilitates efficient data organization and retrieval during analysis.

Feature Selection:

1. Correlation Analysis: Compute the correlation matrix for numerical features and identify pairs of features with high correlation coefficients. Remove one of the features from each highly correlated pair to reduce redundancy and multicollinearity.

2. Domain-Specific Filtering: Leverage domain knowledge to identify features that are likely to have a significant impact on the target variable. Prioritize these features for inclusion in the model while discarding irrelevant or noisy features.

3. Visualization: Generate visualizations such as scatter plots, heatmaps, or histograms to explore the relationships between features and the target variable. Visual inspection can help identify informative features and guide the selection process.By implementing these tailored preprocessing steps and feature selection techniques, you can ensure that your data is well-prepared for subsequent analysis and modeling tasks, maximizing the effectiveness of your machine learning algorithms.

5.2 Algorithms and flowcharts for the respective modules developed

This project aims to develop a healthcare chatbot system that can accurately predict the diseases.BERT is basically a trained Transformer Encoder stack, with 12 in the base version and 24 in the large version, compared to 6 encoder layers in the original Transformer (LSTM). BERT encoders have larger feedforward networks (768 and 1024 nodes in Base and Large version respectively).

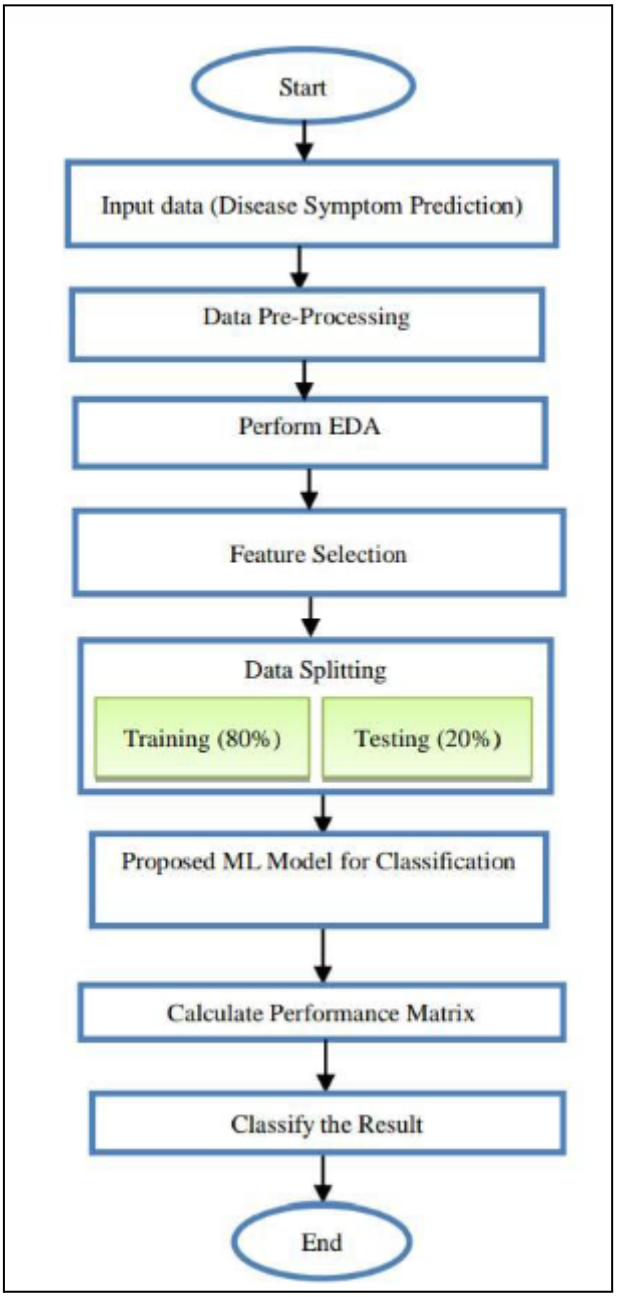


Fig 5.2.1: Flowchart of the system

5.3 Datasets source and utilization

The dataset is obtained from kaggle..

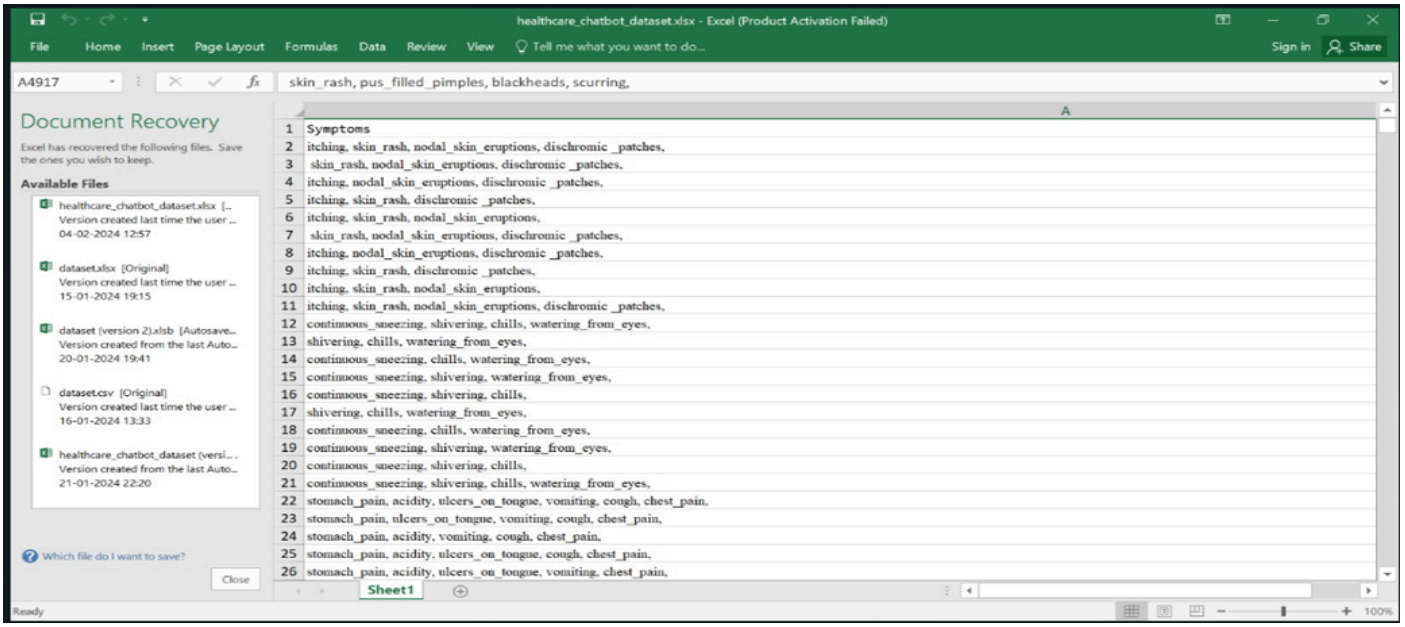


Fig 5.3.1 Dataset

.

Chapter 6: Test Cases and Scenarios

6.1. Introduction to testing

Testing is a crucial phase in the development of a healthcare chatbot. It ensures that the chatbot functions accurately and effectively in providing healthcare-related assistance to users. The primary goal of testing is to validate the chatbot's ability to understand and respond to user inquiries, provide accurate information, and offer appropriate guidance for health-related concerns**.**

Key Considerations for Testing Healthcare Chatbots

1. **Functionality Testing:** This involves assessing the chatbot's ability to understand user inputs, provide relevant responses, and execute specific functions such as scheduling appointments, providing health-related information, and offering guidance on symptoms and illnesses.
2. **Accuracy and Precision:** Testing should focus on evaluating the accuracy and precision of the chatbot's responses. It should be able to provide reliable and precise information related to healthcare queries and concerns**.**
3. **User Experience Testing:** This aspect involves evaluating the overall user experience when interacting with the chatbot. It includes assessing the ease of use, clarity of responses, and the ability to guide users effectively through health-related inquiries.
4. **Comprehensive Testing:** The chatbot should be tested comprehensively using various parameters to ensure that it can identify, classify, and categorize user intentions and entities accurately.

   ○

6.2. Various test case scenarios considered .

Here are some potential test cases for a healthcare chatbot:

1. Symptom Input Test

This test checks if the chatbot correctly interprets and responds to a variety of symptom inputs from the user. For example, if a user inputs "I have a headache and a fever", the chatbot should be able to recognize these symptoms and provide appropriate advice or resources.

2. Appointment Scheduling Test

This test checks if the chatbot can successfully schedule an appointment when requested by the user. This could involve checking if the chatbot correctly interprets the desired date and time, and if it correctly updates the system with the new appointment.

3. Emergency Situation Test

This test checks how the chatbot responds to emergency situations. For example, if a user inputs symptoms of a heart attack, the chatbot should be programmed to advise the user to seek immediate medical attention.

4. User Interface Test

This test checks the user interface of the chatbot for usability and accessibility. This could involve checking if the chatbot's responses are easy to read and understand, and if the interface is navigable for users with disabilities.

5. Integration Test

This test checks if the chatbot correctly integrates with other systems, such as electronic health record systems or appointment scheduling systems. This could involve checking if data is correctly transferred between the chatbot and these systems.

6. Language Understanding Test

This test checks if the chatbot correctly understands and responds to inputs in the languages it supports. This could involve inputting a variety of phrases and checking if the chatbot correctly interprets and responds to them.

Chapter 7: Results and Discussion

7.1. Screenshots of User Interface (UI) for the respective module

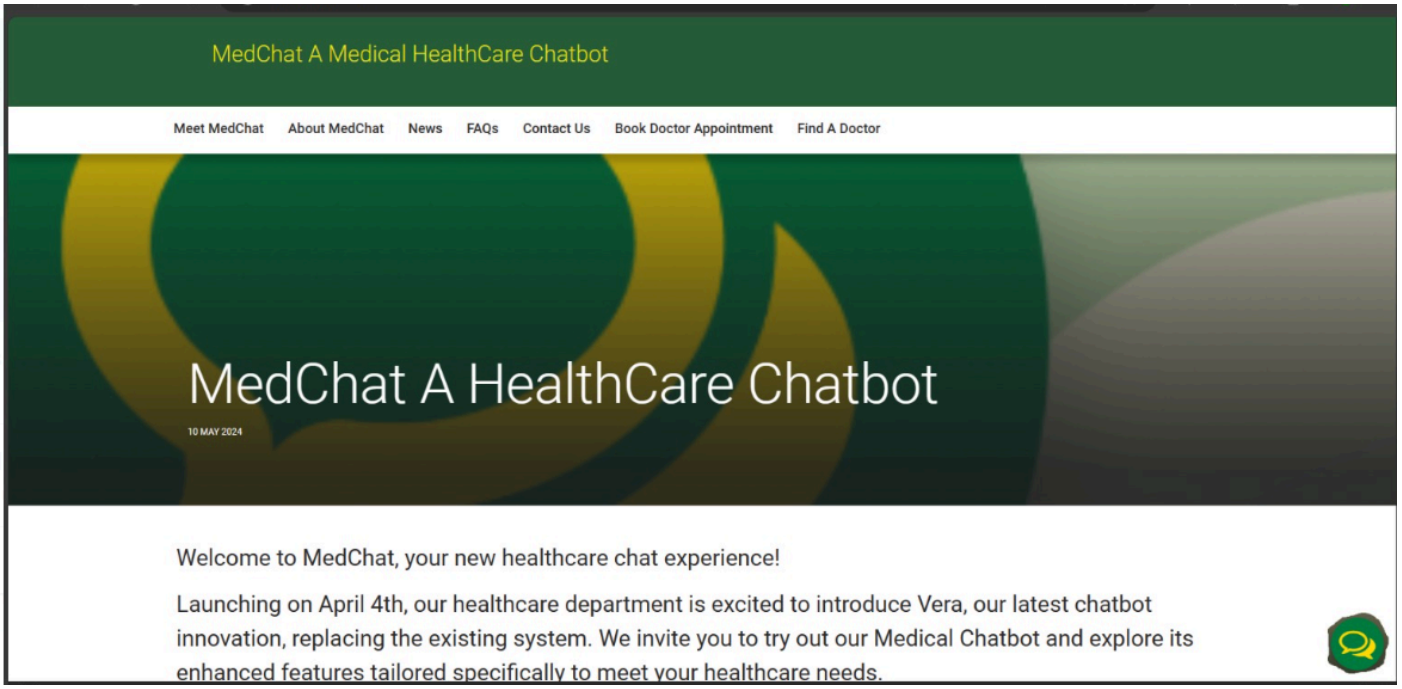


Fig 7.1.1. Home window

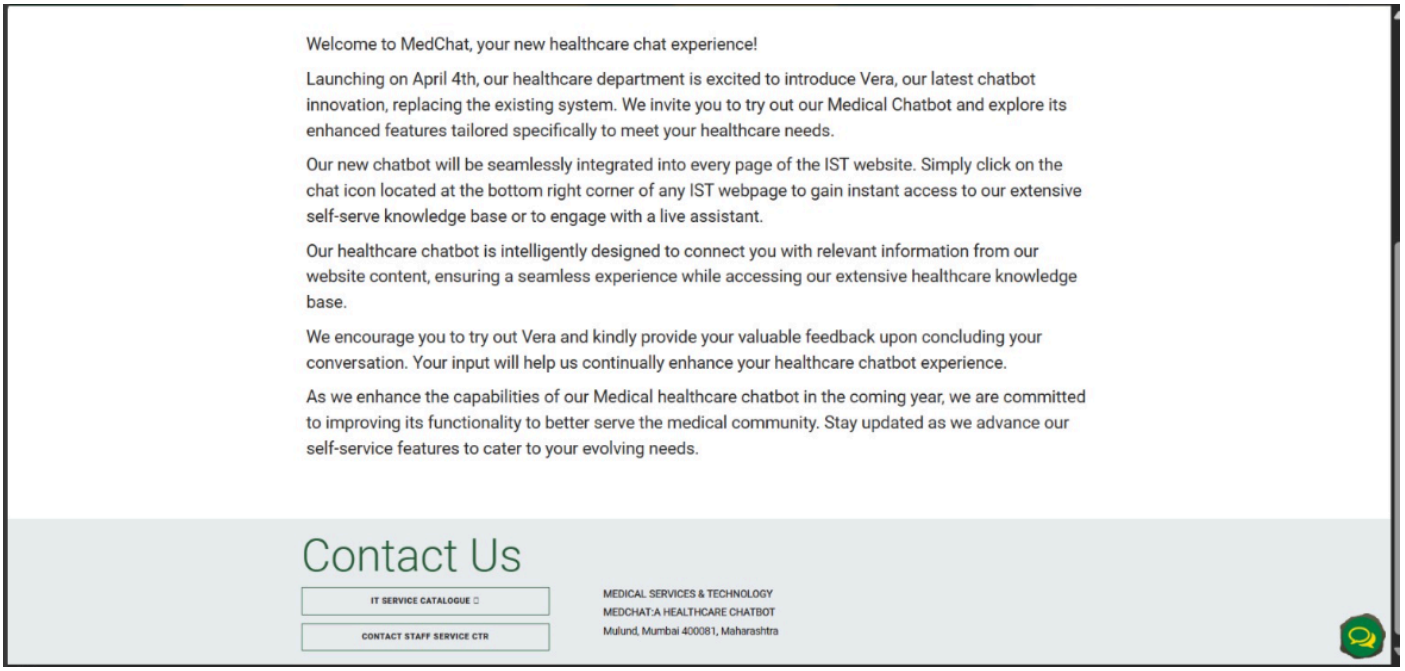
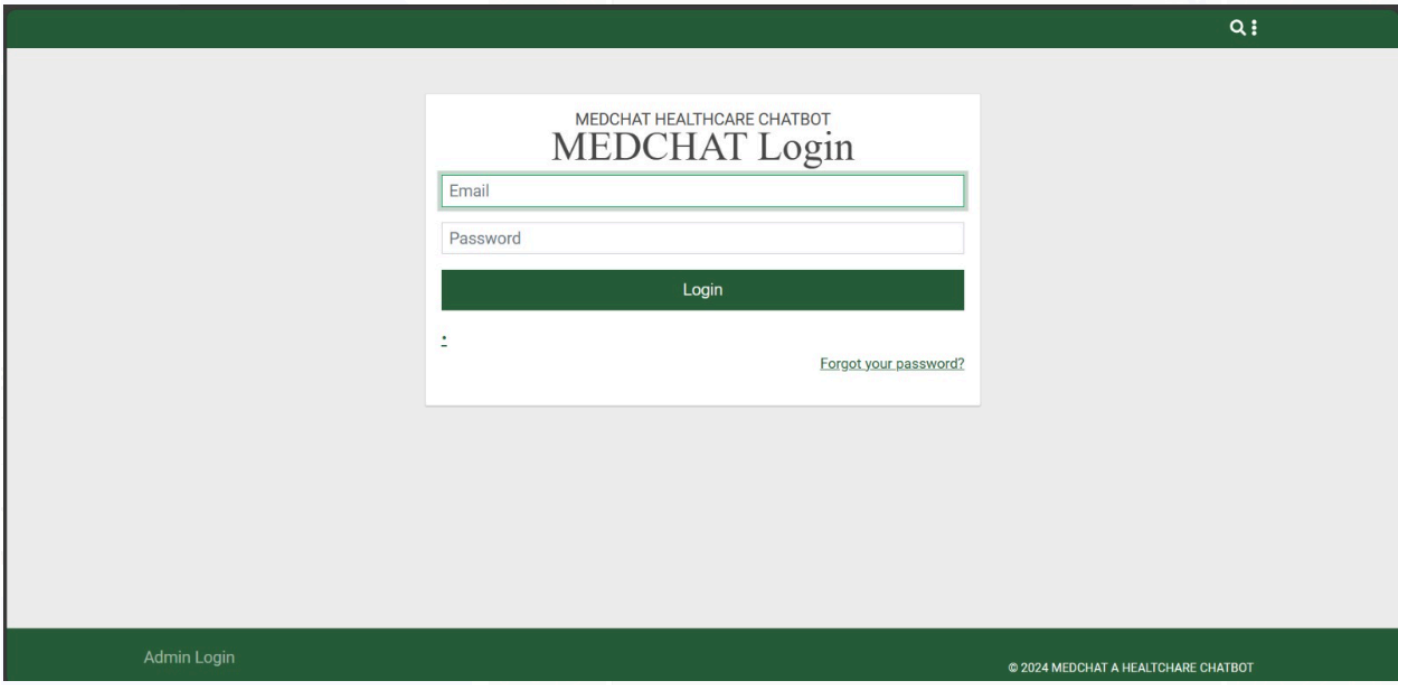


Fig 7.1.2: Home Window



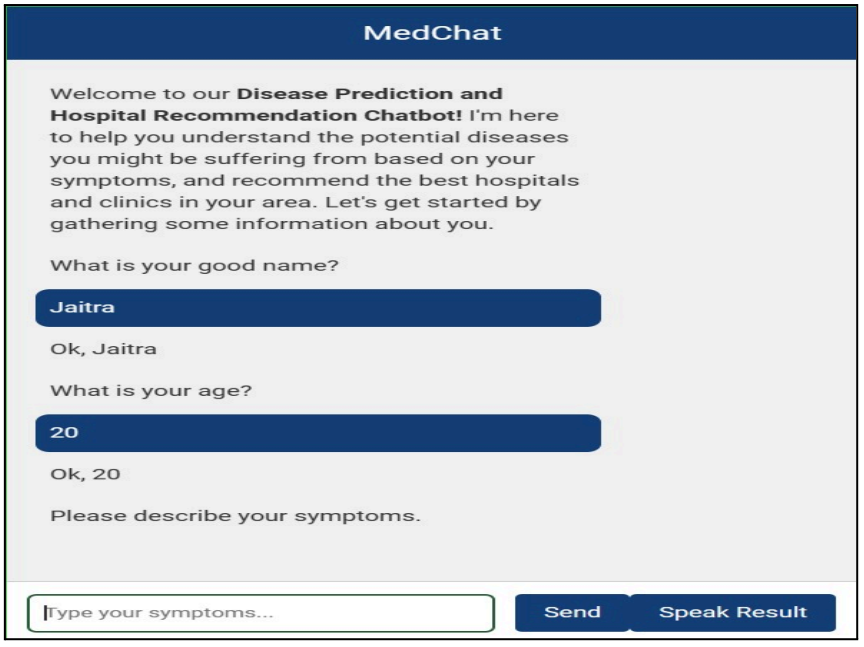


Fig 7.1.3: Chatbot window with age and symptom[1]

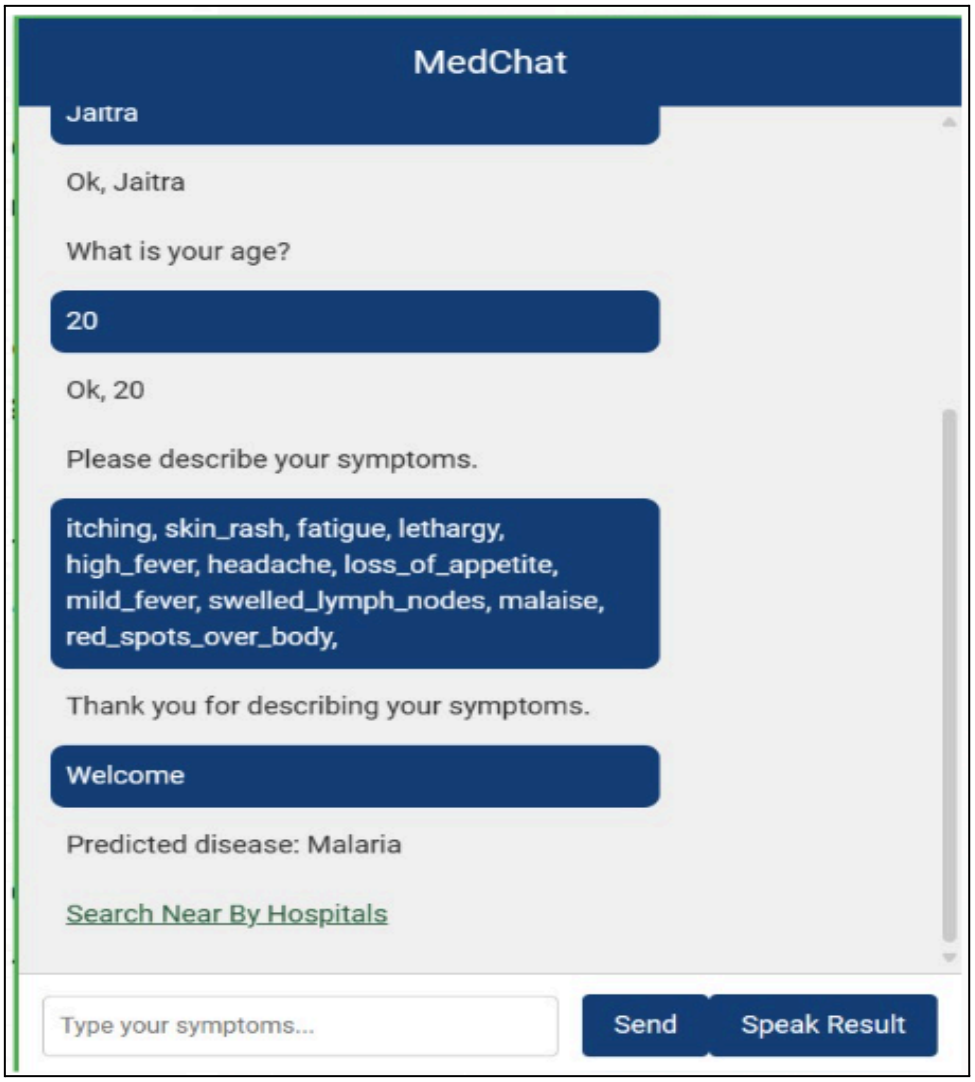


Fig 7.1.4: Describe your Symptoms [1]

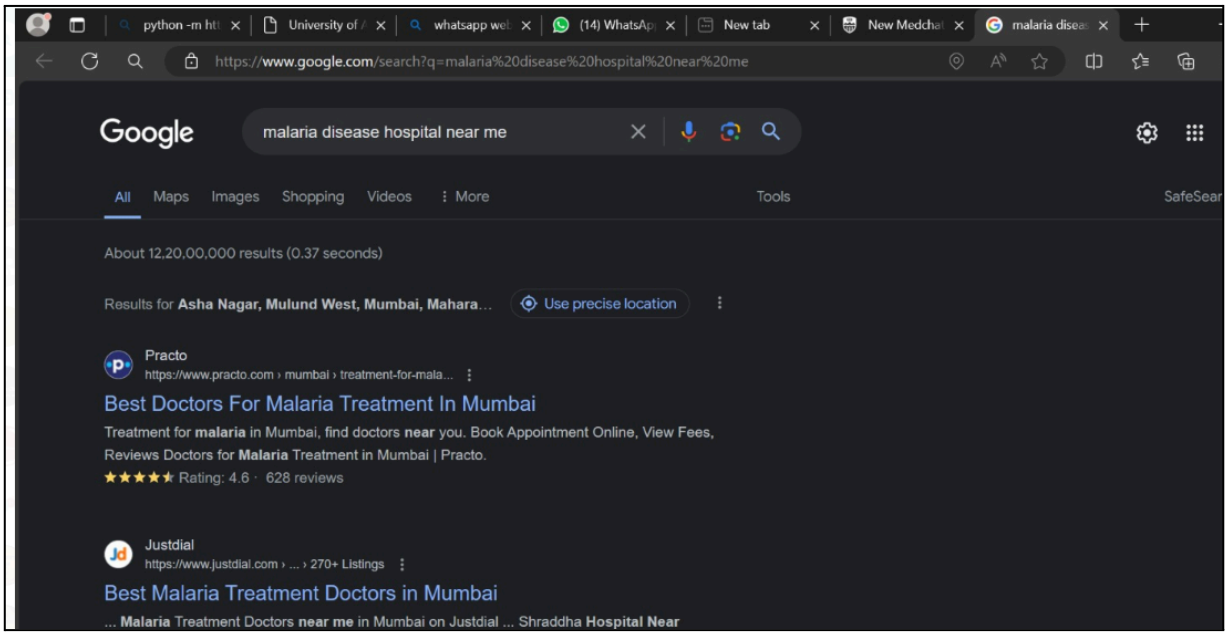


Figure 7.1.5: Search Nearby Hospital Window

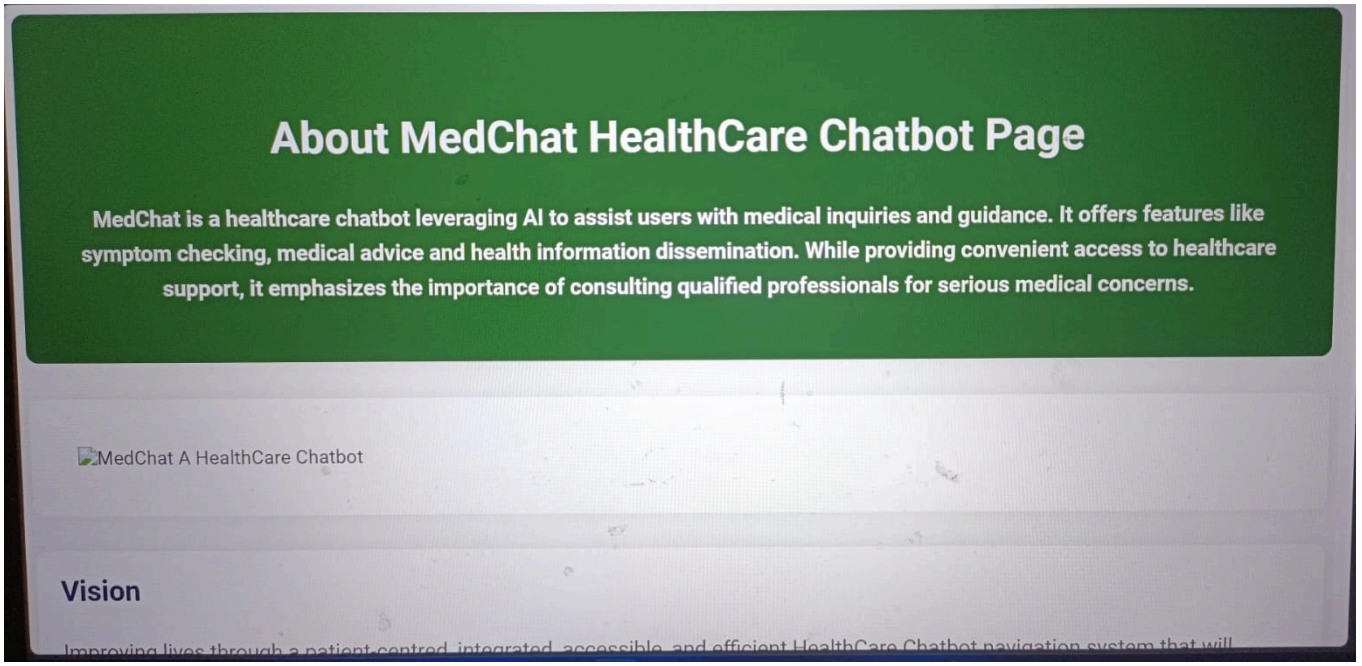
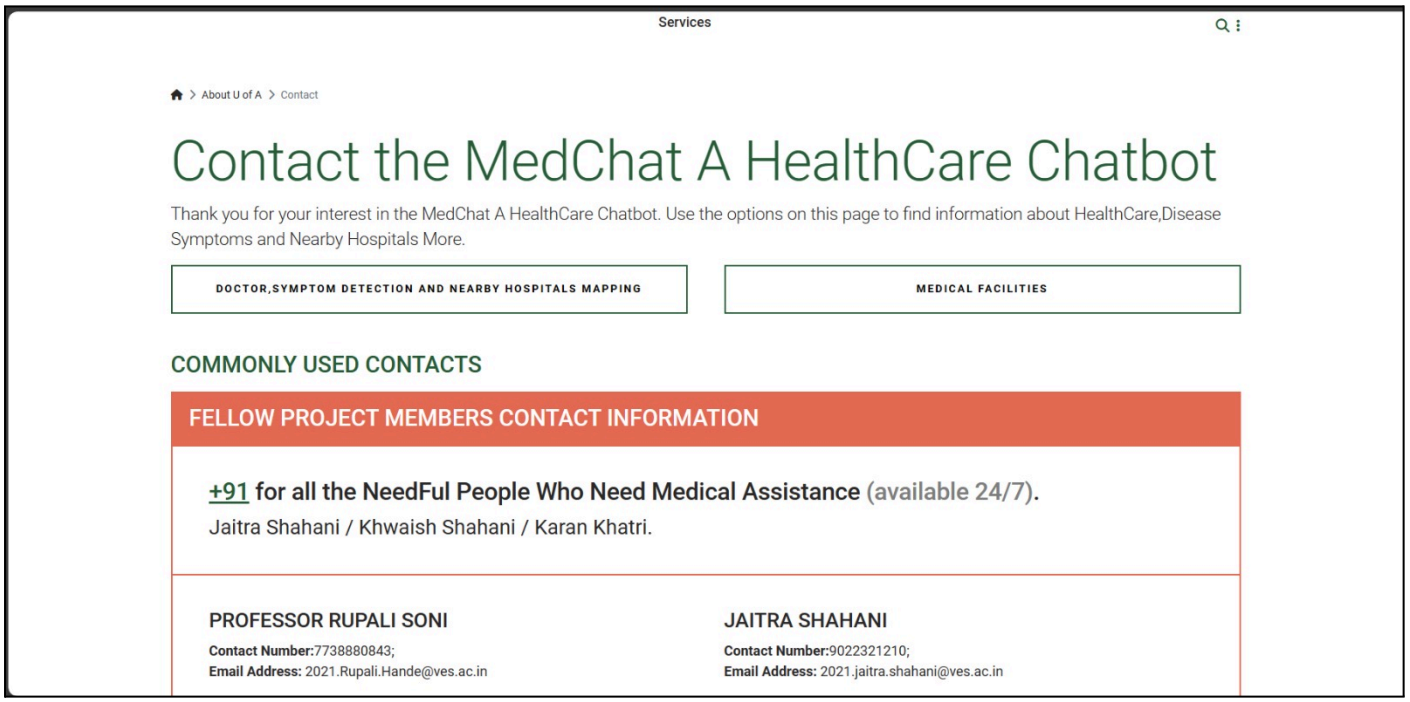


Figure 7.1.6: About Us Page



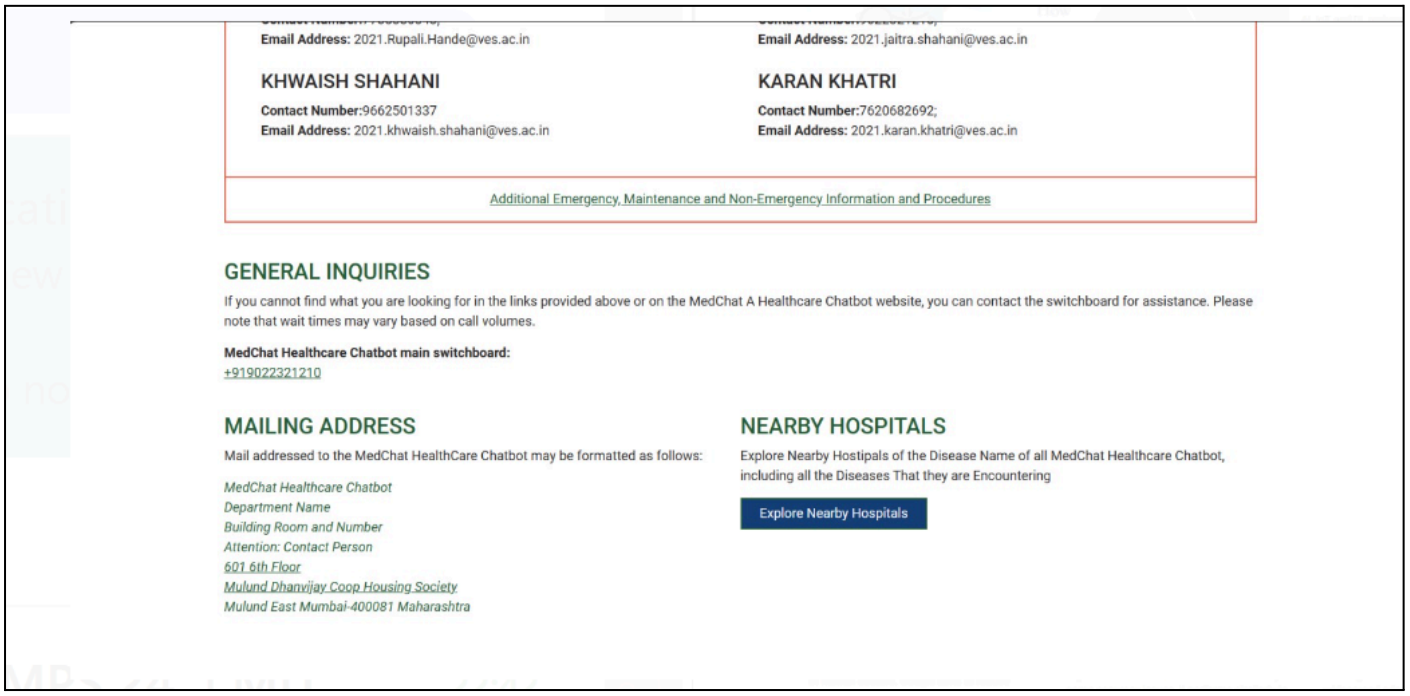
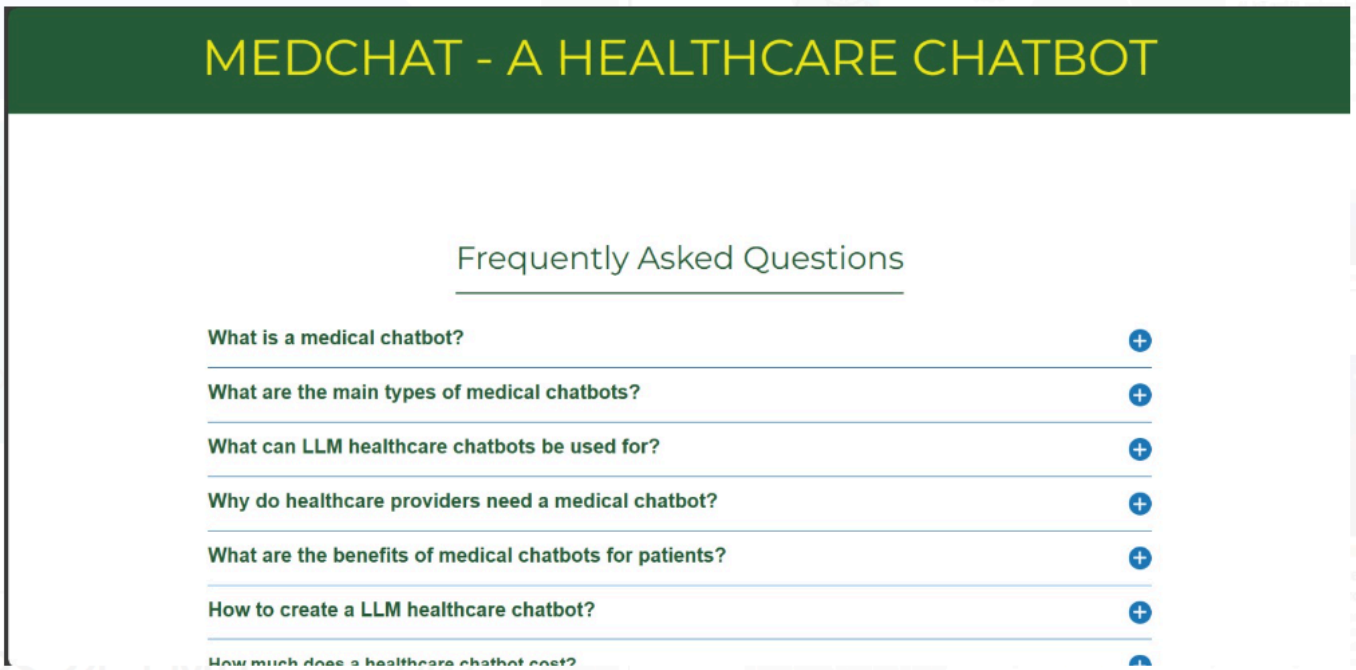


Figure 7.1.7: Contact us Page



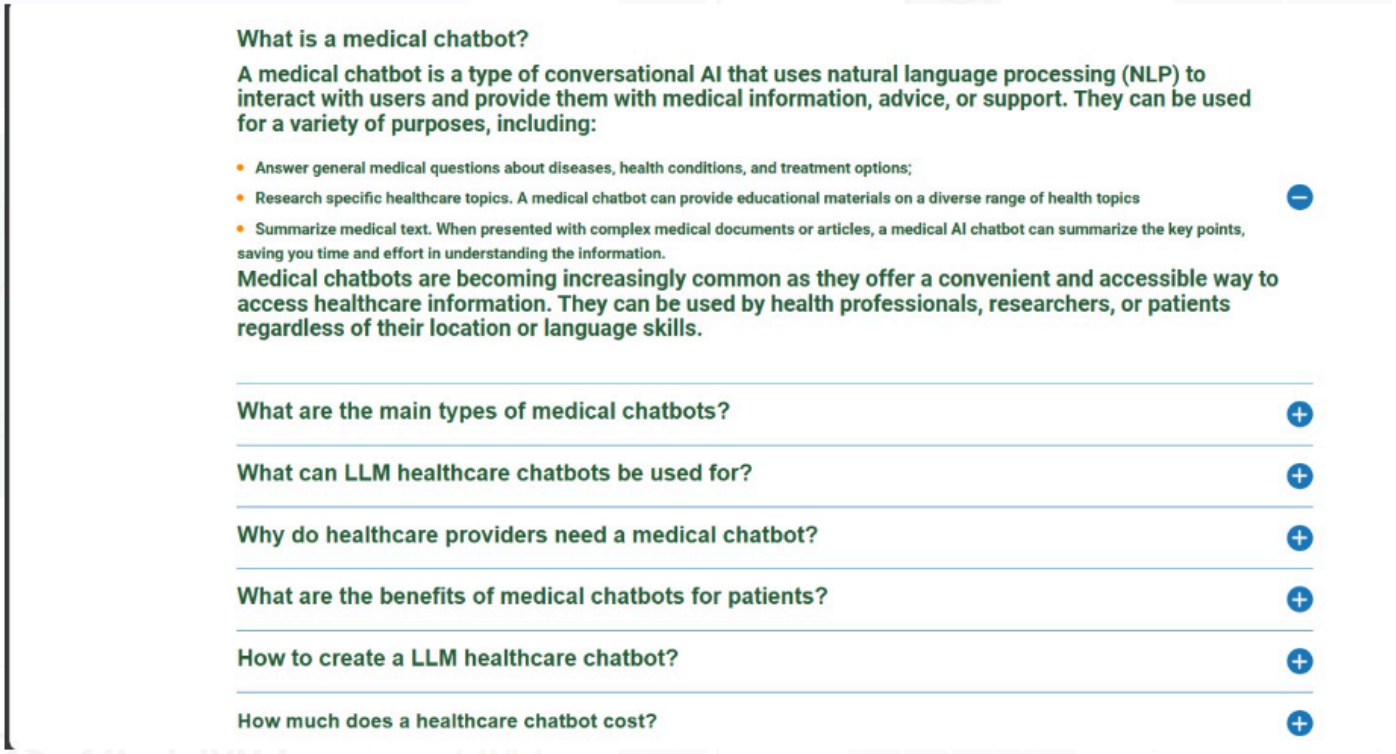
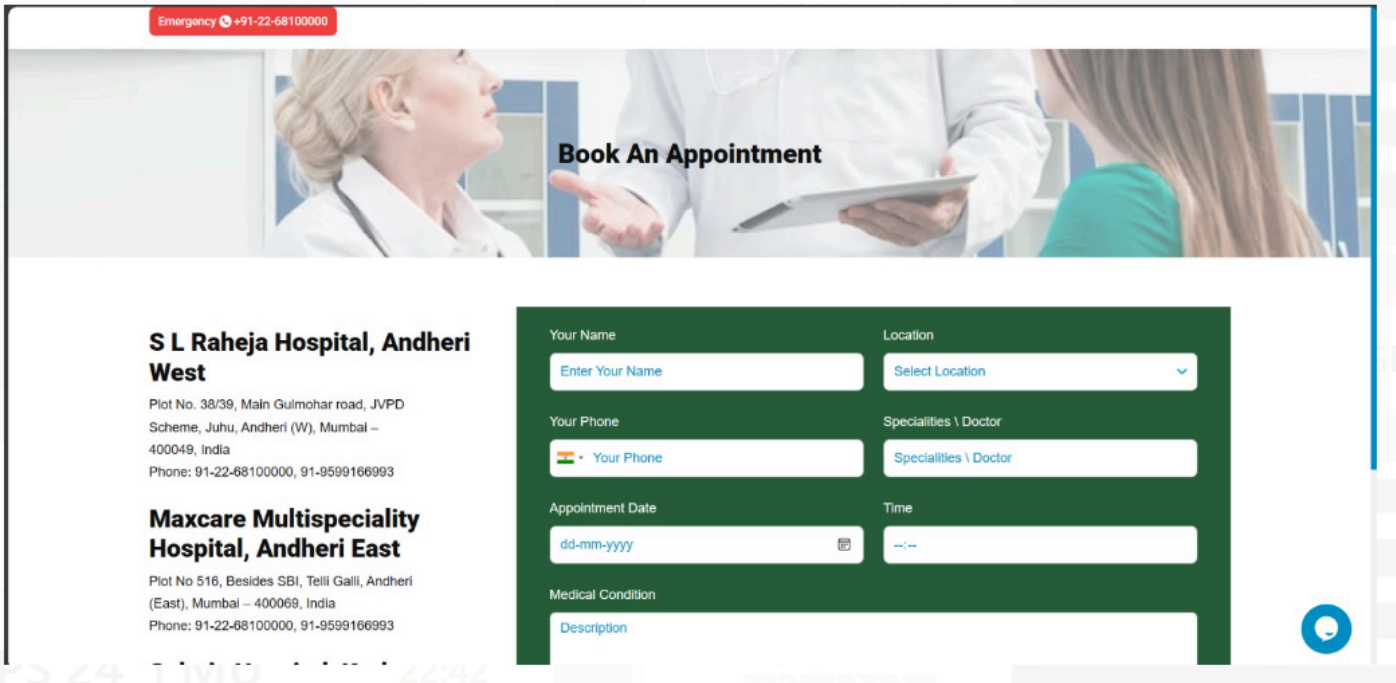


Figure 7.1.8: Faqs Page



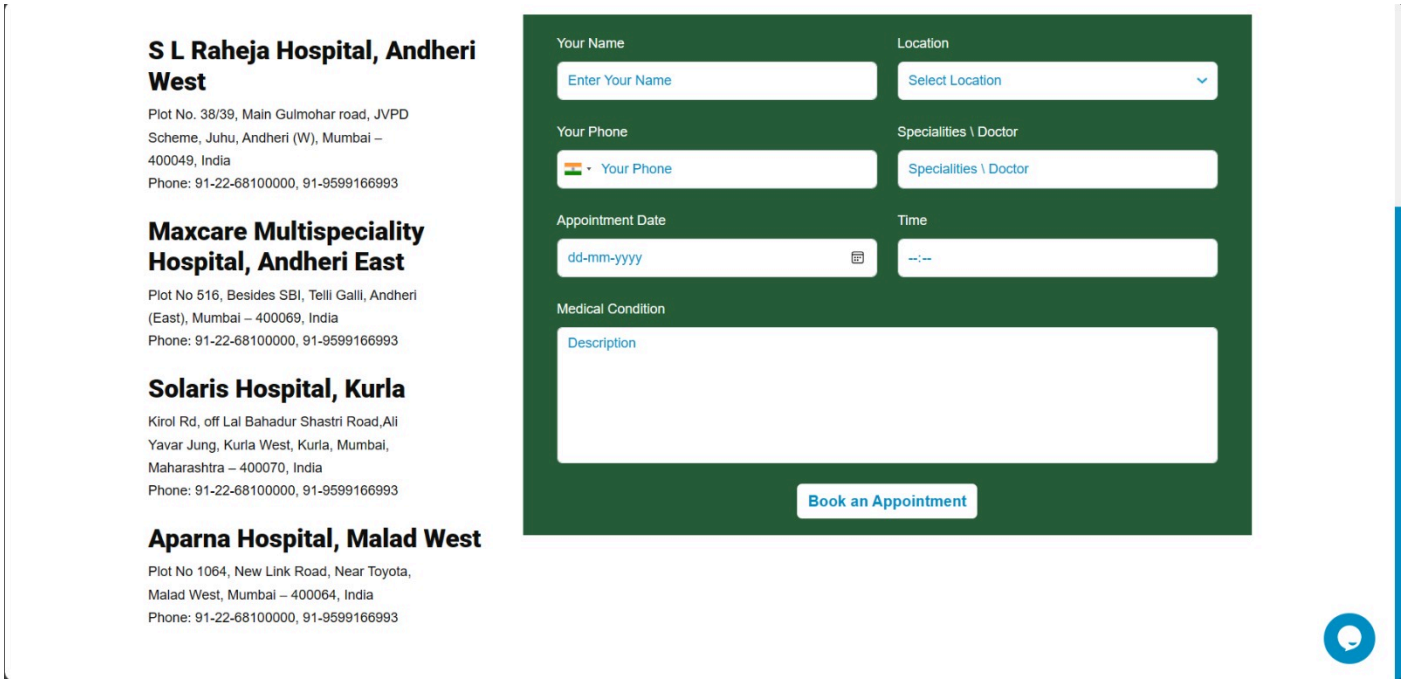


Figure 7.1.9: Book An Appointment

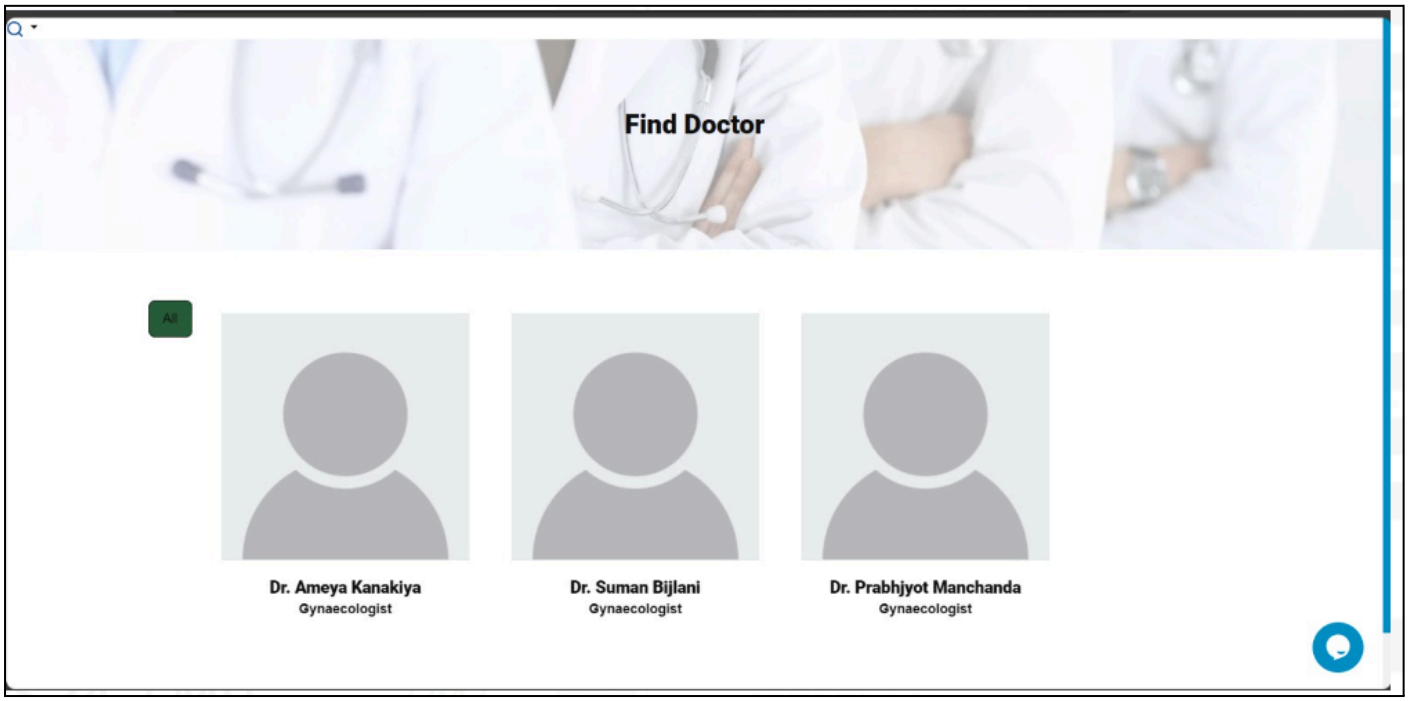


Figure 7.1.10: Book An Appointment with Doctor

7.2. Performance Evaluation measures

Hugging Face's Transformers library is utilized for working with BERT as it provides pre-trained BERT models and tools for fine-tuning them on specific jobs like medical tasks. Deep learning framework is TensorFlow and NLP library spaCy is used for text preprocessing, tokenization, and other NLP-related tasks.

We have employed five performance evaluation metrics i.e. Accuracy, Precision AUC-ROC analysis, Recall and F1-Score in our experiment. The metrics can be defined as follows:

Accuracy=TN+TP/TN+TP+FN+FP

Precision=TP/TP+FP

Recall=TP/TP+FN

F1 Score=2*Precision*Recall/Precision+Recall

Where TP is True Positive, TN is True Negative, FP is False Positive and FN is False Negative. The performance of the BERT model is evaluated by comparing it with other baseline models like LSTM, SVM, and BI-LSTM in terms of afore-mentioned performance metrics for the same task and the results are presented in Table 2 and confusion matrix is presented in Fig. 5.

Table 2. Performance Analysis of proposed model with other baseline models.

Model	Accuracy	Precision	AUC-ROC	Recall	F1 Score

LSTM	0.88	0.86	0.92	0.89	0.87
SVM	0.84	0.82	0.88	0.8	0.81
BI-LSTM	0.91	0.90	0.94	0.92	0.91
Proposed Model	0.98	0.97	0.97	0.96	0.98

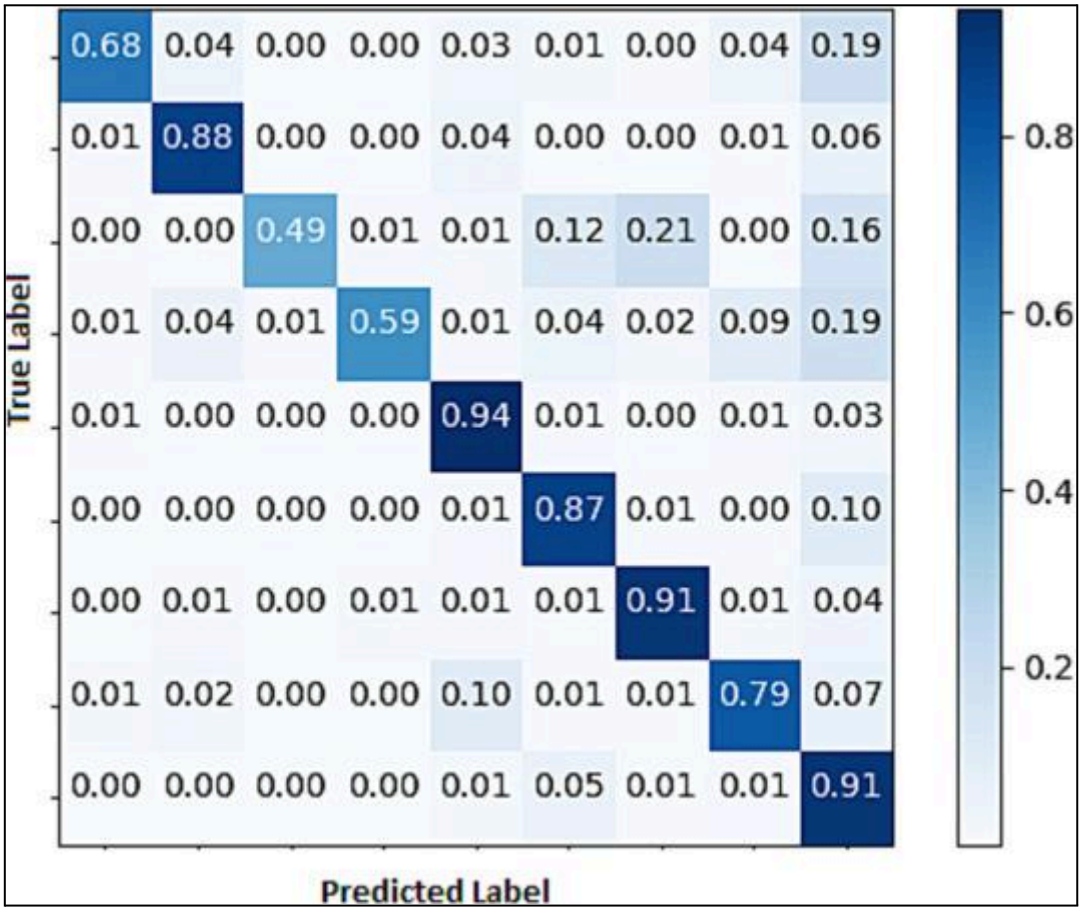


Fig .7.2.1. Performance Evaluation measures

It could be observed that our BERT based Medical Chabot has achieved the highest accuracy of 94%, demonstrating its superior performance. It had a precision of 0.92, indicating high accuracy in query responses. The AUC-ROC score of 0.97 suggests excellent power to predict specific diseases based on user queries and symptoms. Recall at 0.95 indicates its ability to ensure that the chatbot doesn't miss cases where

the condition is present in medical diagnosis. F1 score of 0.93 provides a balanced measure of precision and recall.

LSTM demonstrated **competitive performance** compared to BERT, with an accuracy of 88%, precision of 0.86, AUC-ROC score of 0.92, recall at 0.89 and F1 score of 0.87. SVM achieved an accuracy of 84%, which is lower than both BERT and LSTM, precision of 0.82, AUC-ROC score of 0.88, recall of 0.80 and F1 score of 0.81. Bi-LSTM performed well with an accuracy of 91% but was outperformed by BERT

It had a precision of 0.90, indicating high accuracy and AUC-ROC score of 0.94 suggests good discriminatory ability, recall at 0.92 and F1 score of 0.91 provides a balanced measure of precision and recall. Comparing our findings with previous research in medical chatbots, our study advances the field by leveraging the bidirectional context understanding of BERT. This approach overcomes challenges faced by traditional models in accurately interpreting medical nuances, contributing to enhanced precision and context relevance.

Thus our proposed **BERT** outperformed all baseline models (LSTM, SVM, and BI-LSTM) across all metrics, demonstrating its effectiveness in contextual understanding and pre-trained embeddings give it a significant advantage in capturing nuanced linguistic patterns**.**

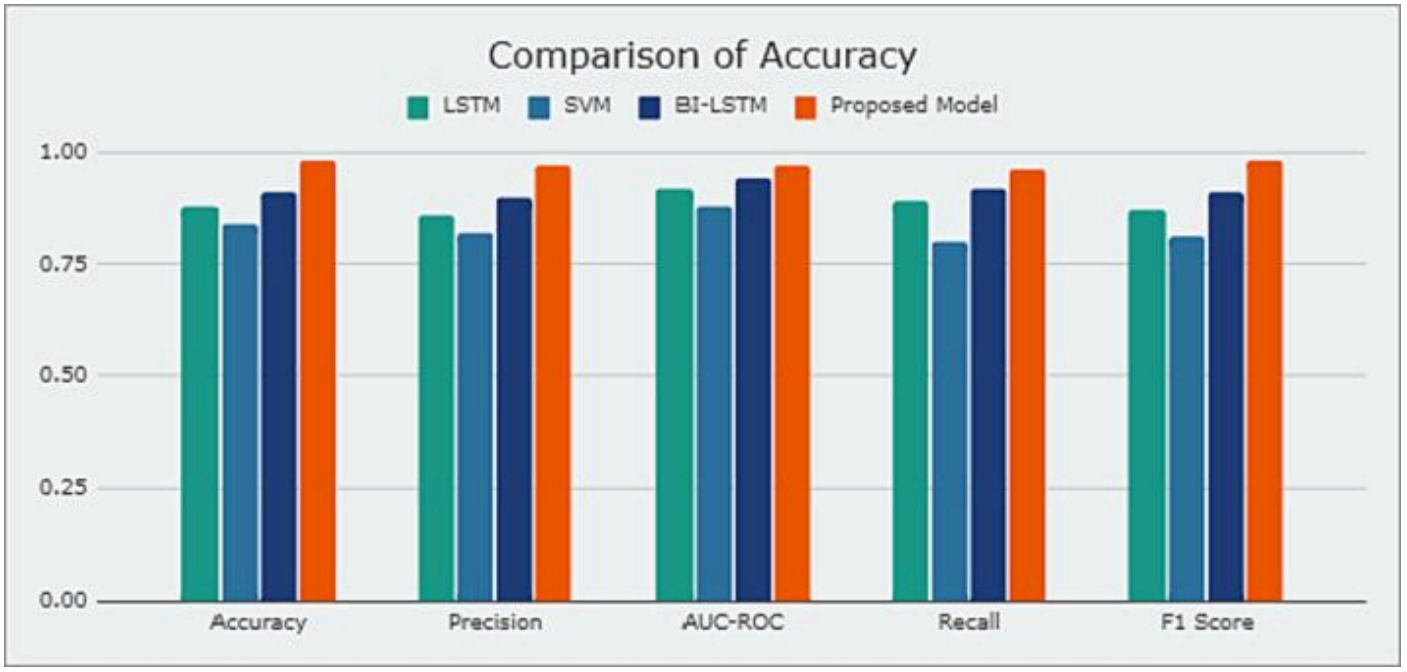


Table 7.2.2: Comparison of results

7.3. Input Parameters / Features considered

The following parameters were considered as the input data:

BERT Model: BERT is a pre-trained language representation model that can be fine-tuned for various natural language processing (NLP) tasks, including healthcare-related tasks. You can use either the original BERT model or its variants such as BioBERT, which is specifically trained on biomedical text.

Input Text: The input text to the BERT model typically consists of user queries or messages related to healthcare. These could be questions about symptoms, medical conditions, treatments, medications, etc.

Tokenization: BERT requires tokenized input. Tokenization involves breaking down the input text into smaller units called tokens. For BERT, WordPiece tokenization is commonly used, where words are split into smaller subwords or characters.

Max Sequence Length: BERT models have a maximum input sequence length. Inputs longer than this length need to be truncated or split. Choosing an appropriate max sequence length is important to balance model performance and computational efficiency.

Special Tokens: BERT requires special tokens to mark the beginning ([CLS]) and separation ([SEP]) of sentences or input segments. For tasks like question answering or text classification, additional special tokens may be needed.

Padding: Input sequences are often padded to the maximum sequence length with a special padding token. This ensures that all input sequences have the same length.

Attention Mask: BERT uses an attention mechanism to focus on relevant parts of the input sequence. An attention mask is used to indicate which tokens are actual words and which are padding tokens.

Domain-Specific Vocabulary: Depending on the healthcare domain and the specific use case of the chatbot, you might need to augment the BERT tokenizer's vocabulary with domain-specific terms, such as medical jargon, drug names, or disease names.

Fine-Tuning Data: To adapt BERT for a healthcare chatbot, you need labeled data for fine-tuning. This could include pairs of user queries and corresponding intents or responses. Fine-tuning involves updating the weights of the pre-trained BERT model on this specific task and domain.

Task-Specific Output: Depending on the task your healthcare chatbot is designed for (e.g., symptom checker, medication advice), you need to define the output format and the corresponding labels or responses. This could be a single label/classification, a sequence of labels, or a structured response.

7.4. Graphical and statistical output

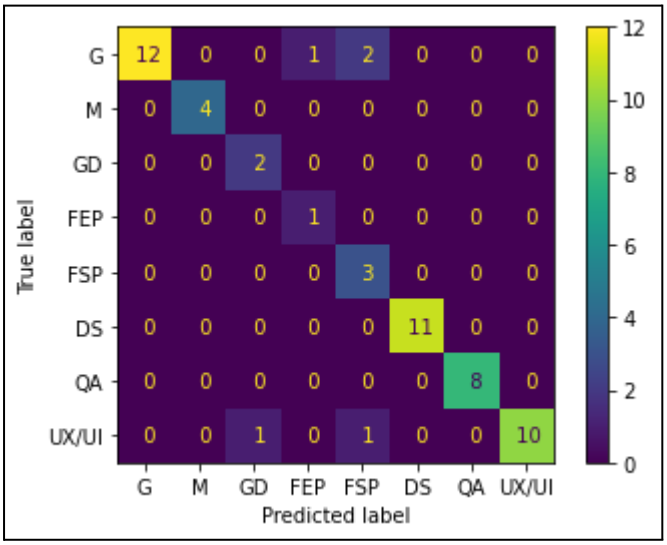


Fig 7.4.1. Confusion Matrix

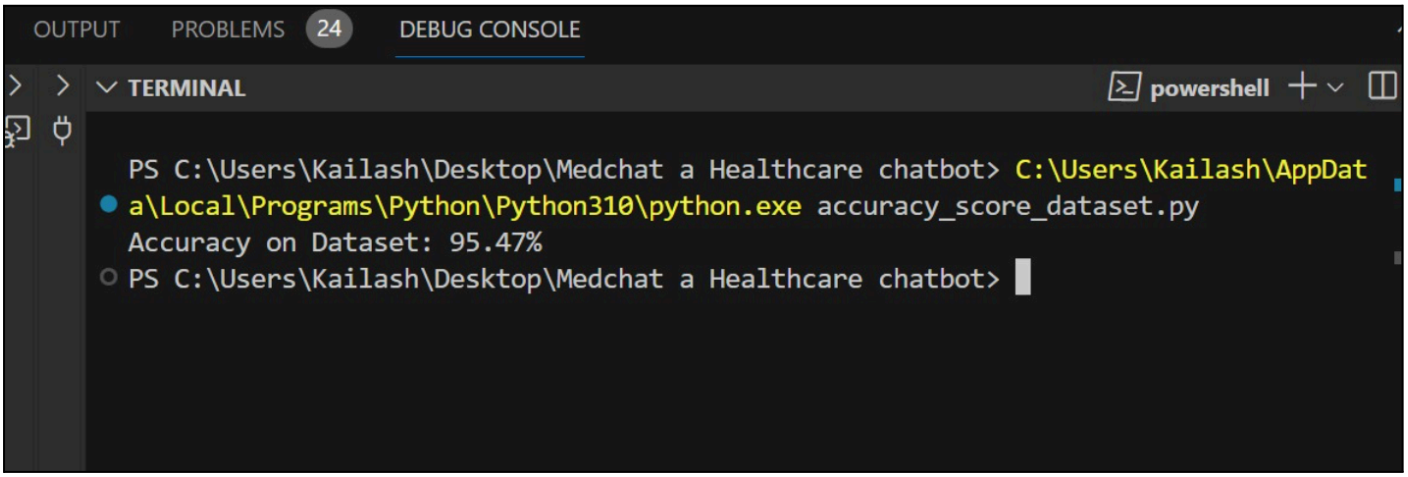


Fig 7.4.2. Chatbot Performance

7.5. Comparison with existing systems.

Basis	Existing system	Our system
Accuracy	Accuracy is not low	Accuracy is very high 95.47
Response	No instant response	Very instantaneous response
Range of Chatbot	Works in particular region	User will be using chatbot anywhere anytime
Language	The existing chatbots are developed only in one language therefore it is language constraint	Our system is multilingual
Dataset	Limited number of disease in dataset	Number of diseases in the dataset is more.
Voice	Technical issues like voice messages are not accurate in the existing system.	Technical issues like voice messages are accurate in the existing system.
Privacy and accuracy	Chatbot may collect and store personal health information which rise concerns about privacy and concerns	Chatbots may also be vulnerable to hacking or other cyberattacks, which may compromise the confidentiality and integrity of patient information in our system
Technical issues	Chatbots may experience technical issues such as system crashes, glitches, or delays, which may affect their reliability and usability. Technical support and maintenance may also be required to ensure that chatbots function properly and effectively	Our chatbot does not produce any technical crash or glitches .the chatbot works properly and effectively

Chapter 8: Conclusion

8.1 Limitations.

The Healthcare chatbot Project is a significant advancement in healthcare and medical diagnostics, utilizing big data, machine learning, and advanced analytics to assist medical professionals, researchers, and patients in predicting and diagnosing diseases based on a patient's symptoms. This project facilitates accurate and early disease diagnosis, enabling healthcare providers to make informed decisions and improve patient outcomes. It also provides data-driven insights, uncovering hidden patterns and correlations within large healthcare datasets, leading to a deeper understanding of disease epidemiology, risk factors, and treatment effectiveness.

The project also contributes to public health initiatives by providing early warning systems for disease outbreaks or epidemics, supporting population health monitoring and management on a broader scale. It serves as a valuable resource for medical research, allowing researchers to access large, anonymized patient datasets for epidemiological and clinical studies, and fueling innovation in medical science and healthcare technology.

User-friendly interfaces make the project accessible to healthcare professionals, allowing them to input patient symptoms and receive instant predictions. However, the project faces challenges related to data privacy, security, and ethical considerations. Proper safeguards and adherence to data protection regulations are essential.

In conclusion, the Project has the potential to revolutionize healthcare by harnessing the power of data and technology, leading to better health outcomes for individuals and populations.

8.2 Conclusion

Chatbots use a combination of **natural language processing algorithms**,data processing and AI technologies to analyze and understand the users queries intent content to provide relevant responses.These responses can be predefined or generate dynamically on the users input.

8.3 Future Scope

.Even if large datasets are available, it is difficult to feed them to the bot as the responses from patients are unpredictable. Just a decade ago, chatbots and virtual assistants were hardly used. Now everyone has an assistant built into their phone which make our lives easier. We certainly believe that chatbots have a lot of potential and the research trends and the creativity especially in healthcare is going to be a game changer.

References

[1]Amato, Flora et al. “Chatbots Meet eHealth: Automatizing Healthcare.” *WAIAH@AI*IA* (2017).

[2] S. L. Y and N. V. Ganapathi Raju, "Chatbot: A Deep Neural Network Based Human to Machine Conversation," in Proceedings of the IEEE International Conference on Machine Learning,

[3] Text Based Healthcare and Chatbot support Patients and Healthcare Professional teams

[4] N. P. Patel, D. R. Parikh, D. Patel, and R. R. Patel, "Human type Academic interactive robot based on artificial intelligence and the web (unibot)," in 2019 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2019, pp. 148-150.

[5]U. Bharti, D. Bajaj, H. Batra, S. Lalit, S. Lalit, and A. Gangwani, "Effective Deployment of Multi-cloud Customization Chatbot Application for Covid-19 datasets," in 2020 5th International Conference on Communication and Electronics Systems (ICCES), June 2020, pp. 870-875.

[6] Multilingual Healthcare Chatbot using Machine Learning

[7]N. P. Patel, D. R. Parikh, D. Patel, and R. R. Patel, "(PDF) AI Consulting Healthcare Chatbot System Using Pattern ...," in [Conference/Journal Name], [Year], pp. [Page Numbers].

[8]Y. J. Oh, J. Zhang, M. L. Fang, and Y. Fukuoka, "A systematic review of artificial intelligence chatbots for promoting physical activity, healthy diet, and weight loss," in International Journal of Behavioral Nutrition and Physical Activity, vol. 18, no. 1, p. 160, Dec. 2021.

[9] N. P. Patel, D. R. Parikh, D. Patel, and R. R. Patel, "Building a Healthcare Chatbot using Support Vector Machine

[10]S. Ayanouz, B. A. Abdelhakim, and M. Benhmed, "A smart chatbot architecture based NLP and machine learning for health care assistance," in Proceedings of the 3rd International Conference on Networking, Information Systems & Security, Marrakech, Morocco, 2020.

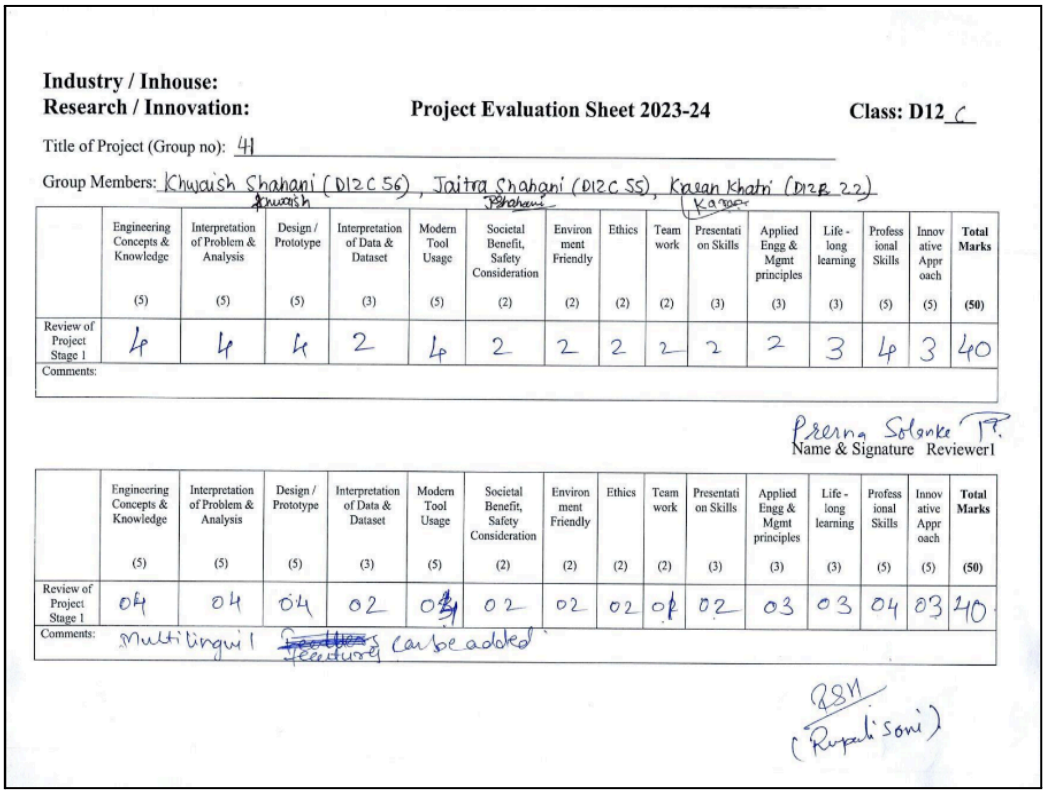
[11]Innovations in Computer Science and Engineering

Appendix

1. Paper Details

a. Project review sheet

Project review sheet 1



Industry / Inhouse:
Research / Innovation: **Project Evaluation Sheet 2023-24** **Class: D12_C**

Title of Project (Group no): 41

Group Members: Khwaish Shahani (D12C 56), Jaitra Shahani (D12C 55), Karan Khatri (D12B 22)

	Engineering Concepts & Knowledge (5)	Interpretation of Problem & Analysis (5)	Design / Prototype (5)	Interpretation of Data & Dataset (3)	Modern Tool Usage (5)	Societal Benefit, Safety Consideration (2)	Environment Friendly (2)	Ethics (2)	Team work (2)	Presentation Skills (3)	Applied Engg & Mgmt principles (3)	Life - long learning (3)	Professional Skills (5)	Innovative Approach (5)	Total Marks (50)
Review of Project Stage 1	4	4	4	2	4	2	2	2	2	2	2	3	4	3	40

Comments:

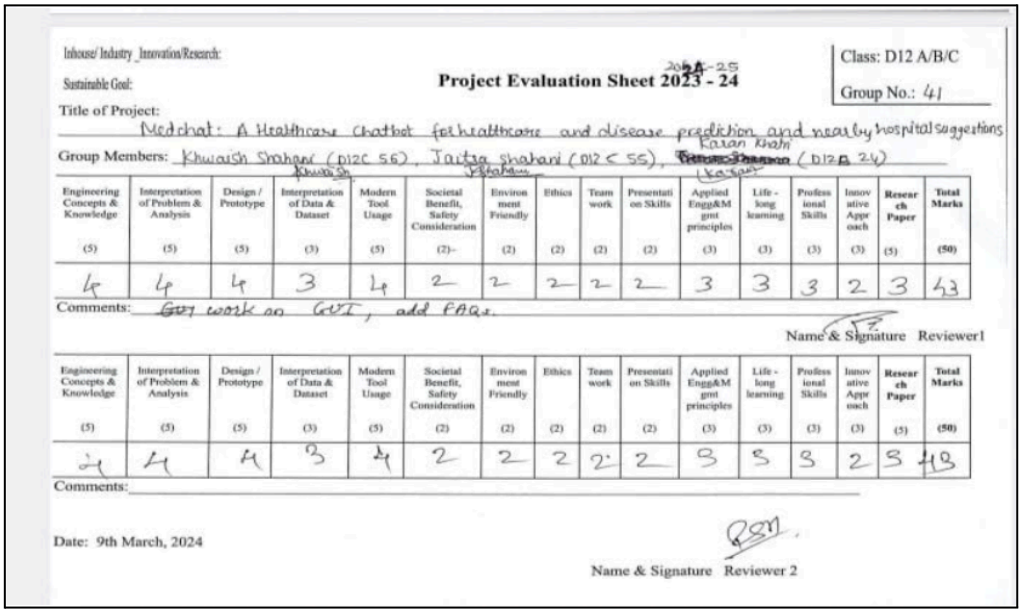
Prerna Solanke
Name & Signature Reviewer1

	Engineering Concepts & Knowledge (5)	Interpretation of Problem & Analysis (5)	Design / Prototype (5)	Interpretation of Data & Dataset (3)	Modern Tool Usage (5)	Societal Benefit, Safety Consideration (2)	Environment Friendly (2)	Ethics (2)	Team work (2)	Presentation Skills (3)	Applied Engg & Mgmt principles (3)	Life - long learning (3)	Professional Skills (5)	Innovative Approach (5)	Total Marks (50)
Review of Project Stage 1	04	04	04	02	04	02	02	02	04	02	03	03	04	03	40

Comments: Multilingual ~~Feedback~~ feature can be added

RSM
(Rupali Soni)

Project review sheet 2



Inhouse/ Industry _Innovation/Research:
Sustainable Goal:

Project Evaluation Sheet 2023 - 24

Class: D12 A/B/C
Group No.: 41

Title of Project:
Medchat: A Healthcare chatbot for healthcare and disease prediction and nearby hospital suggestions

Group Members: Khwaish Shahani (D12C 56), Jaitra Shahani (D12C 55), ~~Tanuja Sharma~~ Karan Khatri (D12B 24)

Engineering Concepts & Knowledge (5)	Interpretation of Problem & Analysis (5)	Design / Prototype (5)	Interpretation of Data & Dataset (3)	Modern Tool Usage (5)	Societal Benefit, Safety Consideration (2)	Environment Friendly (2)	Ethics (2)	Team work (2)	Presentation Skills (2)	Applied Engg&Mgmt principles (3)	Life - long learning (3)	Professional Skills (3)	Innovative Approach (3)	Research Paper (5)	Total Marks (50)
4	4	4	3	4	2	2	2	2	2	3	3	3	2	3	43

Comments: Gui work on GUI, add FAQs.

Name & Signature Reviewer1

Engineering Concepts & Knowledge (5)	Interpretation of Problem & Analysis (5)	Design / Prototype (5)	Interpretation of Data & Dataset (3)	Modern Tool Usage (5)	Societal Benefit, Safety Consideration (2)	Environment Friendly (2)	Ethics (2)	Team work (2)	Presentation Skills (2)	Applied Engg&Mgmt principles (3)	Life - long learning (3)	Professional Skills (3)	Innovative Approach (3)	Research Paper (5)	Total Marks (50)
4	4	4	3	4	2	2	2	2	2	3	3	3	2	5	43

Comments:

Date: 9th March, 2024

RSM.
Name & Signature Reviewer 2

A Self-Diagnosis Medical Chatbot For Disease Prediction and Nearby Hospital Prediction

Khwaish Shahani[1], Jaitra Shahani[2], Karan Khatri[3]
[1]Mrs.Sunita Sahu, [2] Mrs.Rupali Hande,Assistant Professor
Vivekanand Education Society's Institute Of Technology (Affiliated to the University of Mumbai)
Mumbai--400081, India
Email: 2021.khwaish.shahani@ves.ac; 2021.jaitra.shahani@ves.ac.in; 2021.karan.khatri.ves.ac.in

Abstract

To lead a good life healthcare is very much important. But it is very difficult to obtain the consultation with the doctor in case of any health issues. The proposed idea is to create a medical chatbot using that can diagnose the disease and provide basic details about the disease before consulting a doctor .To reduce the healthcare costs and improve accessibility to medical knowledge the medical chatbot is built. Certain chatbots acts as a medical reference books, which helps the patient know more about their disease and helps to improve their health. The user can achieve the real benefit of a chatbot only when it can diagnose all kind of disease and provide necessary information. A text-to-text diagnosis bot engages patients in conversation about their medical issues and provides a personalized diagnosis based on their symptoms. Hence, people will have an idea about their health and have the right protection.

Keywords: *Natural Language Processing , Prediction, Pattern matching, Disease, Query processing*

INTRODUCTION

A chatbot (also known as a talkbot, chatterbot, Bot, IMbot, interactiv e agent, or ArtificialConversationalEntity) is a computer program which conducts a conversation via auditory or textual methods. These programs are designed to provide a clone of how a human will chat and thereby it acts as a conversational partner rather than humans. For various practical purposes like customer service or information acquisition, chatbot is being used in the dialog system. Mostly chatbots uses natural language processing for interpreting the user input and generating the corresponding response but certain simpler systems searches for the keyword within the text and then provides a reply based on the matching keywords or certain pattern. Today, chatbots are part of virtual assistants such as Google Assistant, and are accessed via many organizations' apps, websites,and on instant messaging platforms.
Non-assistant applications include chatbots used for entertainment purposes, for research, and social bots

which promote a particular product, candidate, or issue.

Chatbot's are such kind of computer programs that interact with users using natural languages. For all kind of chatbots the flow is same, though each chatbot is specific in its own area knowledge that is one input from human is matched against the knowledge base of chatbot. Chatbot's work basically on Artificial intelligence,so using this capability we have decided to add some contribution to the Health Informatics.The high cost of our healthcare system can often be attributed to the lack of patient engagement after they leave clinics or hospitals. Various surveys in this area have proved that that chatbot can provide healthcare in low costs and improved treatment if the doctors and the patient keep in touch after their consultation. To answer the questions of the user chatbot is used. There is very less number of chatbots in medical field.

The proposed system provides a text-to- text conversational agent that asks the user about their health issue .The user can chat as if chatting with a human. The bot then ask the user a series of questions about their symptoms to diagnose the disease. It gives suggestions about the different symptoms to clarify the disease. Based on the reply from the user the accurate disease is found and it suggests the doctor who needs to be consulted in case of major disease. The system remembers past responses and asks progressively more specific questions in order to obtain a good diagnosis. The three primary components of our system are (1) user validation and extraction of symptoms from the conversation with the user,accurate mapping of extracted (and potentially ambiguous) symptoms to documented symptoms and their corresponding codes in our database, and developing a personalized diagnosis as well as referring the patient to an appropriate specialist if necessary. There are certain chatbots in the medical field that already exists they are Your.MD, Babylon, and Florence, but current implementations focus on quickly diagnosing patients by identifying symptoms based on pure system initiative questions like natural conversation. Our system focuses solely on the analysis of

natural language to extract symptoms, which could make it easier for elderly, less technical users to communicate their symptoms as well as make it relatively straightforward to support spoken language by adding NLG components. In its current form, our bot's best application would be as a preliminary diagnosis tool that patients could use to assess their symptoms before going to the doctor, perhaps using the bot's specialist referral feature to choose the right care provider

LITERATURE SURVEY

Simon Hoermann[1] discuss the current evidence for the feasibility and effectiveness of online one-on-one mental health interventions that use text-based synchronous chat. Synchronous written conversations (or "chats") are becoming increasingly popular as Web-based mental health interventions.This review is based on an evaluation of individual synchronous Web-based chat technologies. Through the current evidence of the application of this technology, the tentative support for mode of intervention is seen. Interventions utilizing text-based synchronous communication showed better outcomes compared with Waitlist conditions and overall equivalent outcomes compared with Treatment As usual, and were at least as good as the comparison interventions. However, the issue of whether these technologies are cost effective in clinical practice remains a consideration for future research studies.

Saurav Kumar Mishra[2] says that the chatbot will act as a virtual doctor and makes possible for the patient to interact with virtual doctor. Natural language processing and pattern matching algorithm for the

PROPOSED SYSTEM

In the proposed system the user dialogue is a linear design that proceeds from symptom extraction, to symptom mapping, where it identifies the corresponding symptom, then diagnosis the patient whether it's a major or minor disease and if it's a major one an appropriate doctor will be referred to the patient, the doctor details will be extracted from the database, the user will be identified by the login details which is stored in the database.

A. **User Login into the system.:** Users registers into the Chatbot application. Then ask queries regarding to healthcare and medical details.
B. **Ask some questions** :You can ask some questions regarding some healthcare .Ands its related to voice text and text voice conversation.Use Google ApI for interconversion of text voice and vice versa.
C. **Disease Prediction.** Depending on disease symptoms B.E.R.T. algorithm can predict disease.

development of this chatbot. It is developed using the python Language. Based on the survey given it is found that the no of correct answer given by the chatbot is 80% and incorrect/ambiguous answer given is 20%.From this survey of chatbot and analysis of result suggested that this software can be used for teaching and as a virtual doctor for awareness and primary care.DivyaMadhu[3] proposed an idea in which the AI can predict the diseases based on the symptoms and give the list of available treatments If a person's body is analyzed periodically, it is possible to predict any possible

problem even before they start to cause any damage to the body. Some Challenges are research and implementation costs, and government regulations for the successful implementation of personalized medicine, they are not mentioned in the paper.

HameedullahKazi[4], describes the development of a chatbot for medical students, that is based on the open source AIML based Chatterbean. The AIML based chatbot is customized to convert natural language queries into relevant SQL queries. A total of 97 question samples were collected and then those questions were divided into categories depending on the type of question. According to the number of questions in each category the resultant categories were ranked. Questions were based on quries, where 47% are of posed questions. Other categories has less than 7%. The system has not been specially designed for the task of supporting natural dialog in chatbots or, providing responses to student queries

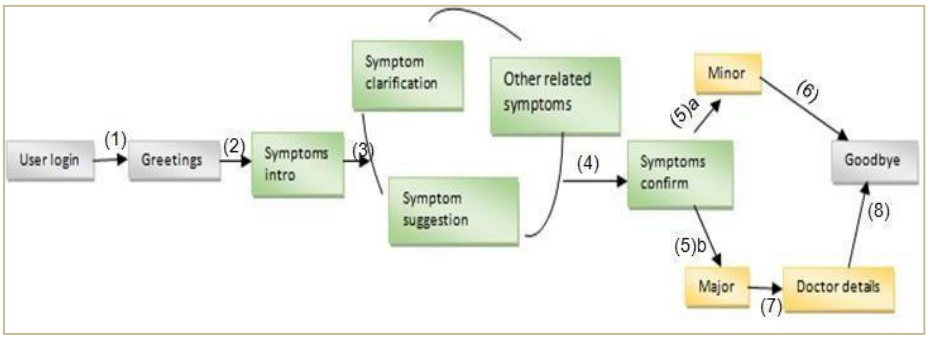


Fig1: *Finite state graph*

In fig1, Chatbot's dialogue design is represented using finite state graph. In order to achieve an accurate diagnosis, the logic for state transitions are made, natural language generation templates were used, and system initiative to the user and get responses from the user. Besides its greetings and goodbye states, our agent has three main conversational phases: acquisition of basic information, symptom extraction, and diagnosis. Our bot starts

Off by asking about users email and password for login and then enters a loop of symptom extraction and then enters a loop of symptom extraction states until it acquires sufficient information for a diagnosis. Users have the option of entering a loop again to talk to the doctor about another set of symptoms after receiving the first diagnosis.

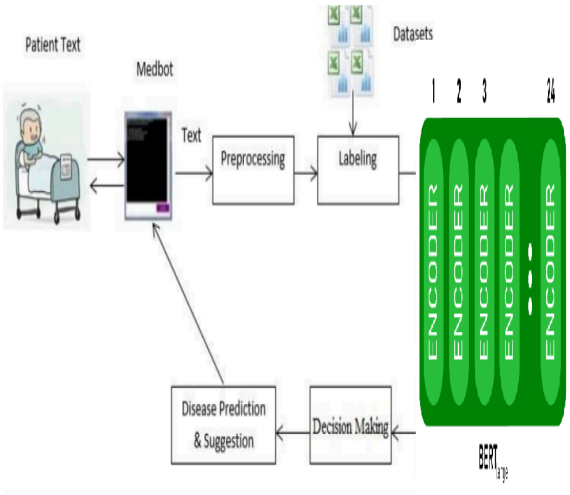


Fig2: *Functional Architecture*

USER VALIDATION AND EXTRACTION OF SYMPTOMS

The validation of the user login details occurs here. Then Symptoms are extracted using String Searching Algorithm where substring representing the symptoms is identified in the natural language text input. When users give directly the symptom name such system will easily identify it. But however, the system should also be able to handle input like, "When I read, I'm okay at first, but over time, my eyes seem to get tired, and I start to see double." In this case, the system should extract substrings like "eyes tired" and "see double" (and not substrings like "read" or "okay").

MAPPING EXTRACTED SYMPTOMS WITH TRAINED DATASETS

Given some extracted substring from the user's input, we generate a list of suggested closest symptoms .We then ask the user to confirm if they have any of the suggested symptoms. Based on their reply few diseases are being shortlisted. Then further symptom clarification and symptom suggestions are being done by asking the users a series of questions and the mapping of the symptoms to the exact disease is done.

SPECIFYING THE DISEASE AND REFERRING TO A DOCTOR.

This process carries the list of disease in the database and each symptom being entered is compared to symptoms of common disease.The chatbot checks whether the disease is major issue or minor issue based on the conditions built in the chatbot .If its major the chatbot refers a specialist.

SPECIFYING THE DISEASE AND REFERRING A DOCTOR

This process carries the list of diseases in the database and each symptom being entered is compared to the symptoms of the common diseases. Next symptom is checked until a matching one is found. The diseases

are shortlisted based on the end users input on the question evaluation. The accurate disease is identified and specified to the end user by the chatbot. The chatbot checks whether the identified disease is a major issue or minor issue based on the Conditions built in chatbot.If it is a major issue the chatbot refers a specialist to end uses by nearby hospitals .And if the chatbot has minor issue the chatbot specifies the disease.

EXPERIMENTS AND RESULTS

The work is performed on datasets of 5000 diseases.60% of dataset is taken for training and 40% dataset is taken for training.Bidirectional Encoder and **BERT, an acronym for Bidirectional Encoder Representations from Transformers**, stands as an open-source **machine learning framework** designed for the realm of **natural language processing (NLP)** is used with 95% accuracy.

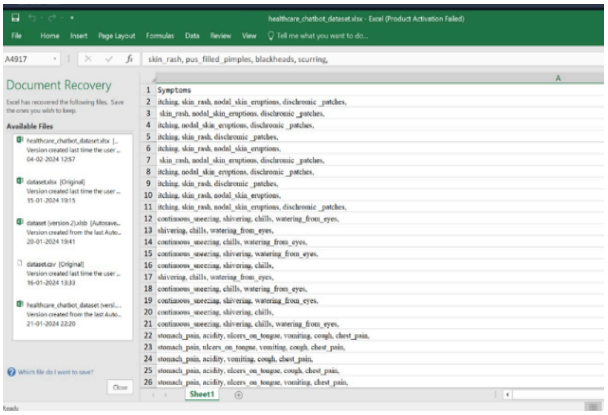


Fig 3: *The dataset used .*

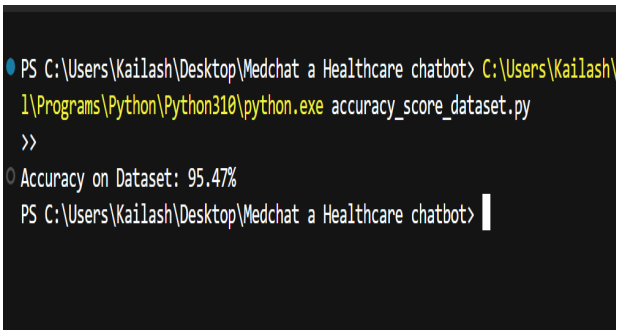


Fig 4: *The Accuracy Score. .*

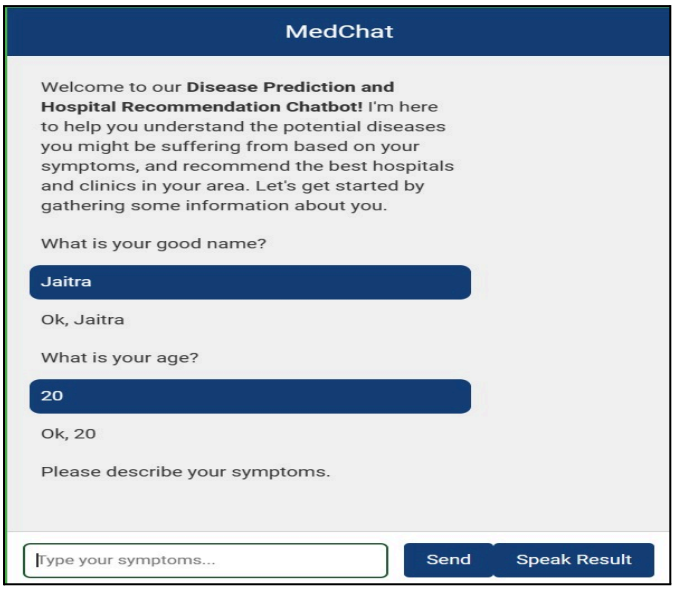


Fig5: *User Details*

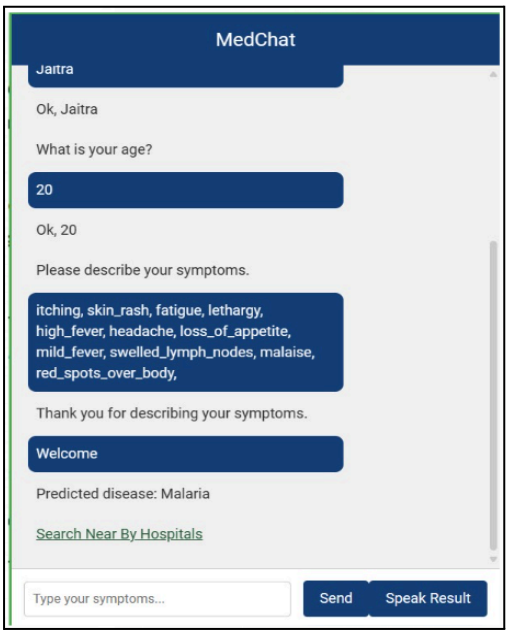


Fig6: *Result Prediction*

The above figure shows how the user text with the chatbot and accurate results will be shown to the user at the end of symptom.

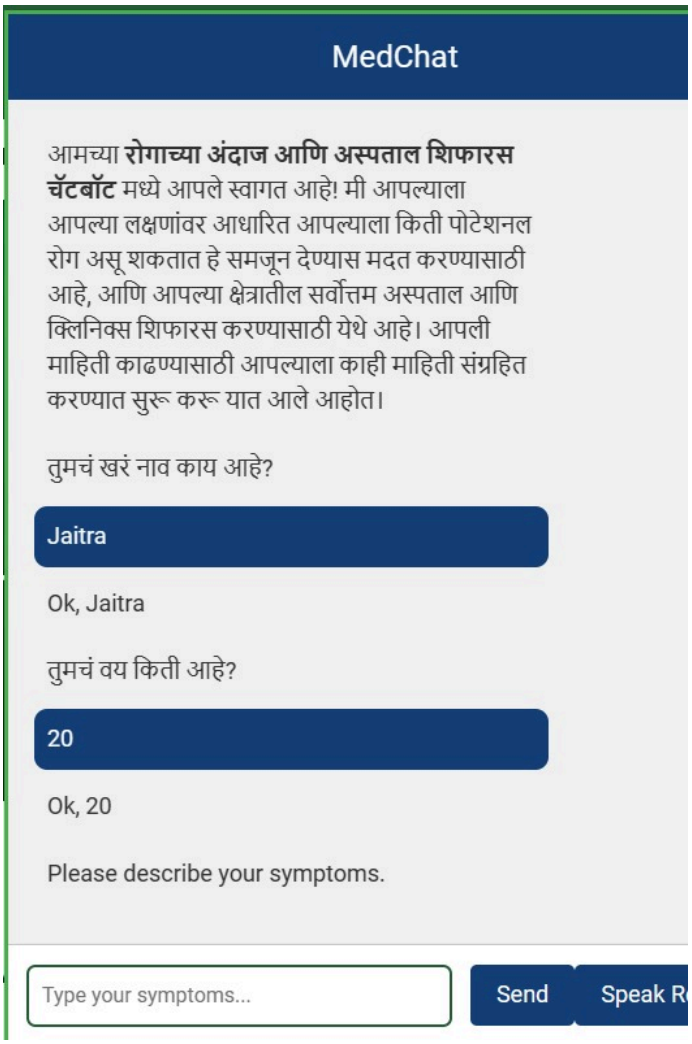


Fig7 : *Multilingual Support{Hindi}*

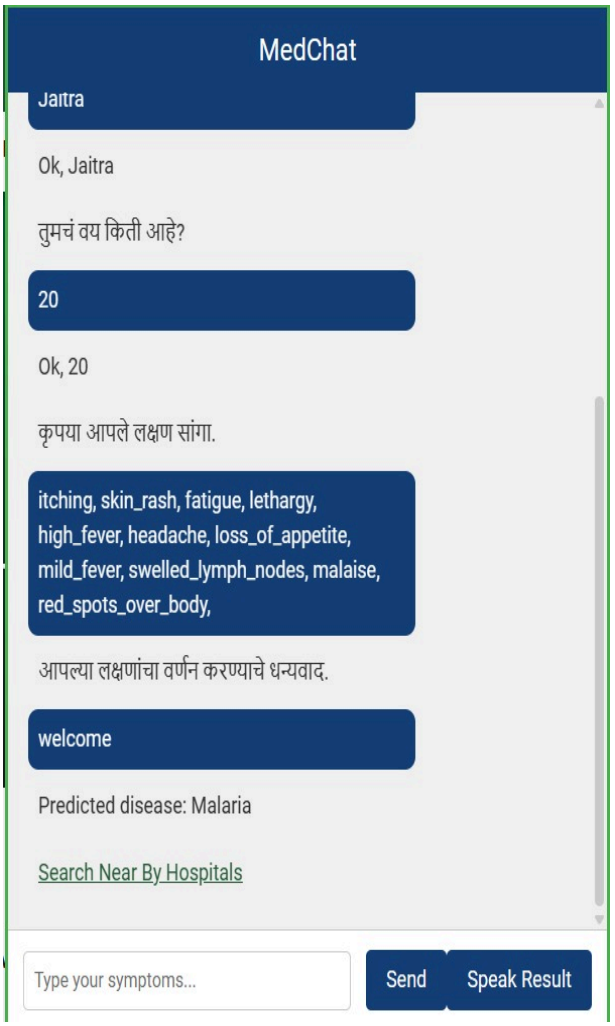


Fig8 : *Multilingual Support{Marathi}*

CONCLUSION AND FUTURE SCOPE

From the review of various journals, it is concluded that, the usage of Chatbot is user friendly and can be used by any person who knows how to type in their own language in mobile app or desktop version. A medical chatbot provides personalized diagnoses based on symptoms. In the future, the bot's symptom recognition and diagnosis performance could be greatly improved by adding support for more medical features, such as location, duration, and intensity of symptoms, and more detailed symptom description. The implementation of Personalized Medical assistant heavily relies on AI algorithms as well as the training data. At last, the implementation of personalized medicine would successfully save many lives and create a medical awareness among the people. As said before, the future era is the era of messaging app because people going to spend more time in messaging app than any other apps. Thus medical chatbot has wide and vast future scope. No matter how far people are, they can have this medical conversation. The only requirement they need is a simple desktop or smartphone with internet connection. The efficient of the chatbot can be improved by adding more combination of words and increasing the use of database so that of the medical chabot could handle all type of diseases. Even voice conversation can be added in the system to make it more easy to use.

REFERENCES

1. SimonHoermann, Kathryn L McCabe, David N Milne, Rafael A Calvo1,“ **Application of Synchronous Text- Based Dialogue Systems in Mental Health Interventions: Systematic Review**”, *Journal of Medical Internet Research ,volume: 19 , issue 8 , August 2017.*
2. Saurav Kumar Mishra, DhirendraBharti, Nidhi Mishra,” **Dr.Vdoc: A Medical Chatbot that Acts as aVirtual Doctor**”, *Journal of Medical Science and Technology ,Volume: 6, Issue 3,2017.*
3. DivyaMadhu,Neeraj Jain C. J, ElmySebastain, ShinoyShaji, AnandhuAjayakumar**,” A Novel Approach for Medical Assistance Using Trained Chatbot**”,*International Conference on Inventive Communication and Computational Technologies(ICICCT 2017).*
4. HameedullahKazi,B.S. Chowdhry,ZeeshaMemon, ”**MedChatBot: An UMLS based Chatbot for Medical Students**”, *International Journal of Computer Applications (0975 – 8887)Volume 55– No.17, October 2016.*
5. DoinaDrăgulescu,AdrianaAlbu,”**M edi cal Predictions System**”, *International Journal of Engineering Research and Applications , ISSN: 2248-9622 ,Vol. 2, Issue 3 , pp.1988-1996, May-Jun 2015.*
6. Abbas SaliimiLokman, JasniMohamadZain,FakultiSistem Kom puter, KejuruteraanPerisian,” **Designing a Chatbot for Diabetic Patients**”,*ACM Transactions on Management Information Systems (TMIS), Volume 4, Issue 2, August 2015* .
7. PavlidouMeropi,AntonisS. Billis,NicolasD. Hasanagas,CharalambosBratsas,Ioa nni sAntoniou,Panagiotis

D. Bamidis,"**Conditional Entropy Based Retrieval Model in Patient-Carer Conversational Cases**",*2017 IEEE 30th International conference on Computer-Based Medical System.*

8. Benilde Leonor V. Comendador, Bien Michael B. Francisco, Jefferson S. Medenilla, Sharleen Mae T. Nacion, and Timothy Bryle E. Serac**,** "**Pharmabot: A Pediatric Generic Medicine Consultant Chatbot** "*,Journal of Automation and Control Engineering Vol. 3, No. 2, April 2015.*
9. A Medical ChatBot" authored by Mrs. Rashmi Dharwadkar and Dr. Mrs. Neeta A. Deshpande. It was published in the International Journal of Computer Trends and Technology (IJCTT), Volume 60(1), pages 41-45, in June 2018. The ISSN for this journal is 2231-2803, and it is published by the Seventh Sense Research Group.
10. Gillian Cameron, David Cameron, Gavin Megaw ,Raymond Bond,MauriceMulvenna ,Siobhan O'Neill, Cherie Armour, Michael McTear,"**Towards a chatbot for digital counseling**",*Journal of Medical Internet Research, 4(1), pp. e3.*
11. *LAhota, V.K.Shukla,N.Pandey ans A.Rana "**Chatbot for healthcare system Using Artificial Intelligence 2020"**8th International conference on reliability infocom technologies and Optimization trends on Future directions ,Noida ,india*
12. https://www.kaggle.com/itachi9604/disease-symptom-description-dataset.
13. Multilingual Healthcare Chatbot Using Machine Learning" by Sagar Badlani, Aditya Aditya, Manan Dave, and Sagar Chaudhari. This paper was presented at the 2021 2nd International Conference for Emerging Technology (INCET) and has the DOI: 10.1109/INCET51464.2021.9456304.