Course Project Report

AI Self Health Diagnosis Tool

Submitted By

Chaithanya Swaroop(211AI010)
Gnana Sagar Reddy (211AI042)
Thejdeep Reddy (211AI013)
Rithvik (211AI022)

as part of the requirements of the course

IT469 - AI in Healthcare [July - November 2024]

in partial fulfillment of the requirements for the award of the degree of

Bachelor of Technology in Artificial Intelligence

under the guidance of

Dr. Sowmya Kamath S, Dept of IT, NITK Surathkal

undergone at

DEPARTMENT OF INFORMATION TECHNOLOGY

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA, SURATHKAL

July - Nov 2024

DEPARTMENT OF INFORMATION TECHNOLOGY

National Institute of Technology Karnataka, Surathkal

C E R T I F I C A T E

This is to certify that the Course project Work Report entitled **"AI Self Health Diagnosis Tool "** is submitted by the group mentioned below -

Details of Project Group

Name of the Student	Register No.	Signature with Date
Gnana Sagar Reddy	211AI042	
Chaithanya Swaroop	211AI010	
3. Thejdeep Reddy	211AI013	
4. K Rithwik	211AI022	

this report is a record of the work carried out by them as part of the course **IT469 - AI in Healthcare** during the semester **July - Nov 2024**. It is accepted as the Course Project Report submission in the partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology in Artificial Intelligence.**

(Name and Signature of Course Instructor)
Dr. Sowmya Kamath S

D E C L A R A T I O N

naWe hereby declare that the project report entitled **"AI Self Health Diagnosis Tool "** submitted by us for the course **IT469 - AI in Healthcare** during the semester **Jan-Apr 2024**, as part of the partial course requirements for the award of the degree of Bachelor of Technology in Artificial Intelligence at NITK Surathkal is our original work. We declare that the project has not formed the basis for the award of any degree, associateship, fellowship or any other similar titles elsewhere.

Details of Project Group

Name of the Student	Register No.	Signature with Date
1. Gnana Sagar Reddy	211AI042	
2. Chaithanya Swaroop	211AI010	
3. Thejdeep Reddy	211AI013	
4. K Rithwik	211AI022	

Place: NITK, Surathkal
Date: 15th April 2024

APPENDIX

A. *Deployed Application Details*

The deployed application can be accessed at the following public homepage URL:

`https://5b61-2405-201-c038-e894-6916-701d-a32e-4151.ngrok-free.app/`

This application was developed as part of the course:

B. *Github Details*

`https://github.com/chaithupro/ai_self_diagnosis_tool`

IT469 - AI in Healthcare, Department of Information Technology, NITK Surathkal.

AI Self Health Diagnosis Tool

Chaithanya Swaroop - 211AI010, Gnana Sagar Reddy - 211AI042, Thejdeep - 211AI013, Ritvik - 211AI022

***Abstract*— The pressure of the increased demand on accessible and efficient solutions in healthcare has pushed the development of AI-based self-diagnosis tools that are often the first point of contact for a patient who might not have access to a doctor at the immediate moment. This project paper discusses the development of an innovative AI-driven self-diagnosis system designed for a preliminary health check on diseases such as diabetes, pneumonia, and heart diseases. The system permits user interaction by incorporating fine-tuned LLM with advanced disease detection algorithms and machine learning techniques. Through conversations with the user, it acquires information on symptoms, medical history, and lab results while utilizing state-of-the-art image recognition technologies in the form of X-ray analysis for comprehensive evaluation. The tool produces individualized reports on health with possible diagnoses, preventive guidelines, and, depending on a case, suggestions to consult a health professional.**

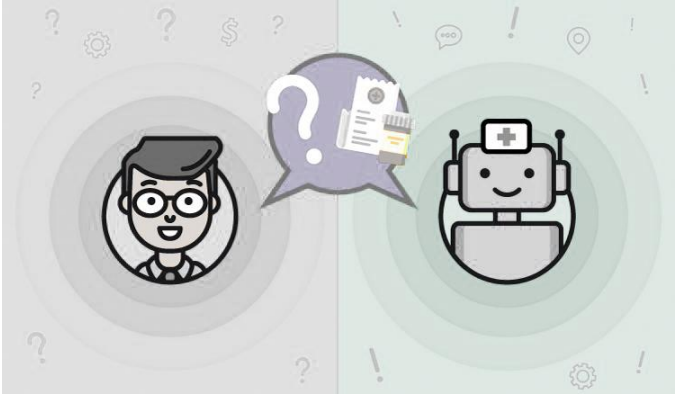
This self-diagnosis website is web-based and accessible 24/7. In addition, it has an interface, which is intuitive to the user. Besides, it gives location-sensitive advice on when to consult specialists. Planned future developments include expanding the scope of disease coverage, voice-enabled functionalities, and multi-lingual support towards making the application even more accessible and inclusive. This tool bridges the gap between patients and healthcare providers, fostering early diagnosis and promoting preventative healthcare, thereby contributing towards better health outcomes and reducing strain on more conventional healthcare systems.

***EDIT LINK* to *Overleaf* project: https://www.overleaf.com/9532587297npwqnqftvmtp#a17f42**

II. Introduction

Early detection of health conditions is very vital because it enables proper treatment and subsequently better outcomes of the patients. Early treatments have the potential to diminish severeness of many diseases; therefore, early intervention yields a good prognosis with a less severe degree of complications. Many factors, however, prevent some people from seeking early medical advice. These include restricted access to healthcare providers, high medical costs, geographical barriers. Such challenges often delay diagnosis and start treating or progress a condition that otherwise could be managed or prevented with early detection. The scene is being changed now by AI self-diagnosis tools, offering people tools to carry preliminary remote health checks that could enable them potentially identify potential risks for their health not requiring direct clinical intervention right away.

Developing an integrated AI-based self-diagnosis, this is a project that will fill an important gap in the first point of contact for people experiencing symptoms associated with diseases such as diabetes, pneumonia, and heart disease. Through accessible and user-friendly design, it provides a platform for guiding early diagnosed users to understand



their symptoms and potential health risks. The proposed system combines the power of conversational LLMs to collect symptom data, medical history, and other relevant health information enabling personalized and accurate assessments. Furthermore, the tool integrates disease-specific prediction models, which shall produce user-specific health predictions through the input by the user themselves.

The system incorporates advanced computer vision techniques on top of traditional health data for analysing medical images such as X-rays to support the diagnosis. Structured health data, such as lab results, vital signs, and patient history are interpreted with machine learning algorithms to provide a holistic and accurate output for differential diagnosis. The system provides actionable health recommendations. If necessary, the system can also recommend local healthcare providers, helping users quickly connect with specialized care.

This self-diagnostic tool will be available 24/7 via a user-friendly web interface, guiding users through a structured diagnostic process and providing immediate feedback. Real-time interaction ensures that users receive insights into their health status quickly, helping them make the right decision to seek further medical attention. The later updates may include voice recognition and support for multiple languages, thus enhancing ease of access. Regional health trends will also be used to provide real-time health alerts based on which the user can assess local risks.

Ultimately, this AI-powered tool represents a significant improvement in the delivery of early health intervention and an easily accessible and efficient healthcare ecosystem. It is also going to enable the users to make more informed health decisions in areas without convenient access to health care services. In this regard, the tool has the potential to remove barriers to health advice to the public's benefit.

III. Literature Survey

Author(s)	Methodology	Key Findings	Limitations
John et al. (2020)[2]	Developed conversational AI chatbot for mental health assessment using LLMs	Showed conversational AI effectively can gathers and assesses symptoms, by providing initial mental health evaluations	and Limited to mental health conditions; no integration of image or lab data
Smith and Lee (2021)[5]	Utilized a convolutional neural network (CNN) model to detect pneumonia in chest X-rays	and Achieved a 95% accuracy in identifying pneumonia in X-ray images	No conversational system patient interaction; limited to a single disease
Kim et al. (2022)[3]	Used machine learning model with electronic health records (EHRs) to predict diabetes onset	Improved prediction accuracy for diabetes by 89%, identifying key risk factors	Limited data sources; lacks user-friendly interface for patient engagement
Alvarez et al. (2023)[1]	Integrated YOLO for detecting lung conditions in medical imaging	Achieved a high accuracy and rapid detection of various lung diseases in X-rays	Focused only on imaging; lacked recommendation and consultation features
Patel and Zhao (2022)[4]	Developed hybrid model using LLMs and clinical datasets for cardiovascular risk assessment	Effectively assessed cardiovascular risk with high accuracy using patient history and and lab data	Limited disease scope and absence of location-based consultation recommendations
Chen et al. (2023)	Combined speech-to-text processing with ML models for voice-based symptom reporting in telemedicine	Enabled accessible, voice-driven symptom reporting and and achieved 88% accuracy in diagnosis	Limited disease support; lacks image processing capabilities for full diagnostics
Gupta and Rao (2022)	Trained an AI model on diverse healthcare datasets to identify diabetes complications	Identified diabetes-related complications with 92% accuracy and recommended in lifestyle adjustments	Narrow disease focus; did not incorporate a conversational interface or real-time recommendations
Zhang et al. (2021)	Leveraged a vector database with LLM-based chatbot to track health symptoms over time	Showed and improved diagnostic accuracy by tracking symptoms longitudinally	Limited to the basic text data; does not support multimedia data (like X-rays)

A. *Outcomes of the Literature Survey*

The literature review was critical in creating guidance for our AI-based self-diagnosis tool in terms of design and functionality. The major results are presented below:

label=0.

1) **Conversational AI Symptom Collector Platform**:
   - Research by John et al. (2020) and Zhang et al. (2021) pointed out that conversational AI performs well in extracting symptoms and patient history from patients.
   - These studies proved that conversational interfaces can engage the user and collect clinically relevant symptom data, thereby making AI-based chatbots highly effective as a first point of consultation.
   - Our project integrates a Large Language Model (LLM) conversational interface that collects initial symptoms, thereby creating high patient engagement, and enhancing its range that extends from diabetes to pneumonia and heart diseases.

2) **Models for Specific Disease Predictions Using Medical Imaging**:
   - The review highlighted that models, such as CNN, have been proven to be highly accurate for medical imaging applications.
   - As a case in point, Smith and Lee (2021) utilized CNNs to detect pneumonia through an X-ray, and Alvarez et al. (2023) utilized YOLO to detect conditions in lungs.
   - Advancing from these related approaches, the current project applies CNN to classify uploaded X-rays for pneumonia and other conditions for better diagnostic accuracy.
3) **Risk Prediction Using Health Records and Lab Data**:
   - Kim et al. (2022) and Gupta and Rao (2022) have shown that machine learning algorithms, which are trained on EHRs and lab data, can accurately predict health risks.
   - Our project includes methodologies like these. It utilizes Kaggle datasets with lab results and patient history in order to predict diabetes and heart disease risks.
   - This functionality enables the tool to analyze structured health data, which goes beyond user-based symptoms.
4) **Integrated Recommendations and Precautionary Suggestions**:
   - Patel and Zhao (2022) discuss AI-based personalized health advice. The study demonstrates that individualized advice from AI systems can lead to better patient outcomes by managing modifiable risk factors.
   - Our prototype builds on that by providing advisory and potentially prescriptive advice to consult a doctor.
   - Unlike other models which have fewer recommendation features, our system is developed to assist users by providing clear, actionable advice that will enable informed health decisions.
5) **User-Centered Design and Accessibility**:
   - Chen et al. (2023) demonstrated the effectiveness of voice integration and user-friendly interfaces in accessibility and engagement.
   - Our project will thus include an easy-to-use web interface for entering symptoms and uploading pictures.
   - Future versions will integrate voice interaction for hands-free usage and support multiple languages, thus further enlarging usability.
6) **Doctor Recommendations Based on Location**:
   - Many AI diagnostic tools do not have healthcare providers incorporated. Personalized recommendations based on geolocation were thought to be highly valuable, according to several studies.
   - Our project addresses this by incorporating a location-based doctor recommendation feature, helping users connect with nearby specialists.
   - This enhances the user experience, bridging the gap between self-diagnosis and actionable healthcare solutions.

The literature survey was used to underscore the strengths and limitations of the existing AI tools for disease diagnostics in order to prioritize features in our project. Overall, synthesizing insights related to conversational AI, specific imaging models, risk prediction, and user-centered design, our tool provides a holistic experience of self-diagnosis beyond mere checking of symptoms at home. A practical addition goes toward location-based recommendations and precautionary advice tailor-made to complement at-home self-diagnosis by providing accessible and appropriate professional healthcare.

IV. Part 1: Methodology for the Chatbot

A. 1. Introduction

This self health diagnostic tool's chatbot is the first point of contact for this interaction between the user and the health care system by acting as a primary interface for interacting with the user. The base of this chatbot was developed from the LLaMA 7B model, fine-tuned on all specific healthcare datasets. This chatbot engages the user through asking structured questions to collect symptoms, medical history, and general health information. The role of the chatbot is to guide the users through identifying any potential health issues and provide personalized health suggestions. The following section explains how the chatbot collects data, interacts with the user, and facilitates the process of disease prediction.

B. 2. Data Collection and Fine-Tuning

- **Data Sources**: This chatbot is programmed to fetch the answers based on various datasets. These datasets are primarily gathered from Kaggle, which mainly covers disease-specific information, including lists of symptoms, diagnostic criteria, and frequently encountered health queries on diseases such as diabetes, pneumonia, and heart conditions.
- **Fine-Tuning**: The LLaMA 7B model is generally fine-tuned with a health-specific dataset. The process of fine-tuning helps adjust the model so it can generate contextually fitting responses accordingly based on what patients speak about, worries, or specific symptoms of diseases.
- **Contextual Understanding**: Fine-tuning allows a chatbot to better understand the context of a medical condition and responds more accurately to user input by eliminating irrelevant or vague answers and producing more exact, relevant outputs.

C. 3. Chatbot Interaction Flow

- **Symptom Query**: It initially asks users to input descriptions of their symptoms, previous medical conditions, and other relevant medical histories. In a conversational tone, information is requested so that the comfort level

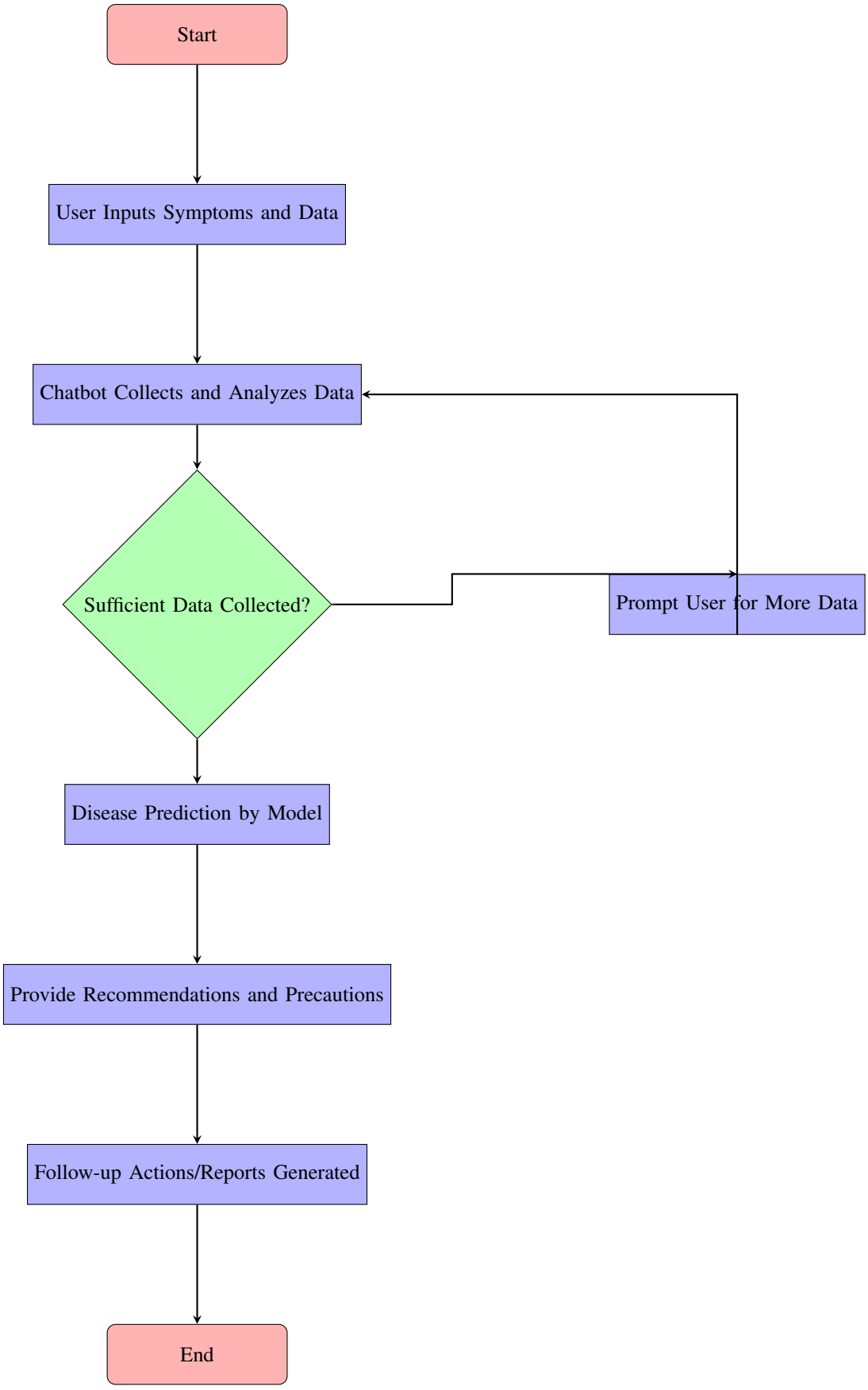


Fig. 1. Flowchart of Chatbot and Disease Prediction Methodology

of the user is eased in a way that they provide all relevant information for further accurate prediction. A symptom recognition, therefore, will require the user to key in their symptoms. The chatbot then processes these responses using a tokenization and vectorization system. In pre-encoded symptom datasets, the chatbot is able to distinguish which diseases can be present based on the symptoms described. A user who describes symptoms such as cough, fever, shortness of breath will have the chatbot recognize pneumonia as a possible diagnosis.

- **Prediction and Proposals**: The chatbot then processes the input collected through analysis and, using the model, studies the symptoms and brings them to the screen in the form of possible diseases. If there is suspicion over a particular disease, it will offer lifestyle advice, precautionary measures, or procedural steps for managing the risk.
- **Recommendation and Precautions**: The chatbot gives general advice for health management after a suspected diagnosis is made. For instance, if it suspects that the patient has diabetes, then it may suggest lifestyle changes, such as dietary adjustments, self-monitoring of blood sugar, and advising a visit to a specialist for further evaluation.
- **Follow-up Actions**: In case it identifies that the symptoms require professional attention, the chatbot instructs the user to see a doctor promptly. For other conditions, it makes home care recommendations and reschedules appointments for later.

V. Model and Natural Language Understanding

- **Language Model**: The LLaMA 7B model has been selected based on the efficiency of processing large-scale language data. Fine-tuning the model with healthcare data ensures that it provides proper answers to health-related queries.
- **Context Retention**: The chatbot retains information from earlier in the conversation, helping maintain continuity. It recalls previously mentioned symptoms or medical history to provide more relevant recommendations.
- **NLP**: The chatbot uses different NLP techniques for processing medical terms and identifying synonyms. For instance, while it knows that "shortness of breath" and "dyspnea" essentially mean the same thing, the system is flexible enough to recognize different user expressions correctly.

VI. Part 2: Disease Prediction through Web Application Methodology

A. Introduction

The web application provides a platform where users input health-related data, which is then analyzed using advanced predictive algorithms to assess potential diseases. This section focuses on predicting three prevalent diseases: diabetes, pneumonia, and heart disease. The web application's disease-specific pages play a critical role in collecting appropriate data, running predictive models, and offering diagnoses. These pages are constructed with the intent that, while enabling as smooth a user experience as possible, the most accurate result is assured. The following sections describe how the user interacts with the web application, entering their health data, and how the system predicts possible diseases from the entered information.

B. Disease-Specific Web Pages

- **User Input Fields**: There are three exclusive pages for each disease, such as diabetes, pneumonia, and heart disease. The user input fields include each disease-specific page containing forms to input data such as symptoms, medical history, lab results, and diagnostic images, like X-rays or CT scans. For instance, the application will ask the user to upload chest X-rays on the pneumonia page; these X-rays are taken to identify lung abnormalities that could be symptoms of pneumonia. For the diabetes page, the application shall request the user's blood glucose levels, family medical history, and excessive thirst or frequent urination among other symptoms. In this way, it is assured that the application has a complete set of all data for an accurate diagnosis. Every page also has a brief summary of the disease, beginning with simple symptoms, risk factors, and even potential complications. These help the user understand the disease before they provide any personal information and then undergo the prediction stage.
- **Instructional Prompts**: To guide users in providing the most accurate information, every page is equipped with instructional prompts. These guidelines indicate how users input symptoms, upload images of diagnostic tests, such as X-rays, and share lab reports. The guidelines are presented in such a straightforward and user-friendly manner that even a non-medically qualified person would have no difficulty contributing the required data to ensure an accurate prediction. For instance, the pneumonia page outlines specific instructions on uploading chest X-rays and how the system interprets them.

C. Disease Prediction Models and Processing

- **Diabetes Prediction Model**: The diabetes prediction model predicts based upon the inputs structured data. The model is likely to analyze blood glucose levels, cholesterol levels, and self-reported symptoms such as thirst and urination. The model determines a person's probable risk of getting diabetes by employing large machine learning algorithms trained over large-scale diabetes-specific datasets. Models predict risk with well-established risk factors like high blood glucose, obesity, and a family history of diabetes. If any of these risk factors are on, the model tags the user as potentially at risk for diabetes, which prompts further steps to be taken in management or possibly referral to a healthcare professional.
- **Pneumonia Prediction**: The pneumonia prediction model works by first processing any chest X-rays up-

loaded by the user. A CNN is very good in analyzing medical images such as X-rays for signs of pneumonia, including lung opacity and fluid buildup. The ability of CNN to recognize minute patterns within the medical images is excellent and helpful in differentiating among the various types of pneumonia. The model has been trained on large pneumonia datasets to precisely identify those subtle patterns and make reliable predictions. The model can present a well-informed prediction regarding whether the user may have pneumonia by looking into X-ray data in conjunction with other symptoms reported by the user.

- **Heart Disease Prediction**: The heart disease prediction model checks multiple health indicators, such as blood pressure, cholesterol levels, and family medical history. Using heart disease datasets and machine learning algorithms for training, this model identifies high-risk factors such as hypertension, high cholesterol levels, smoking habits, and obesity. It calculates the possibility of the user suffering from heart disease based on these risk factors. If the model sees the presence of risky indicators, it flags the user as being at risk for heart disease, based on recommendations for lifestyle changes, medications, or further test requirements as needed.

D. Model Integration and Analysis

- **Data Preprocessing**: All disease models preprocess the input data before making predictions to ensure normalcy in maintaining accuracy. This process involves normalizing values, handling missing data, and ensuring correctness about inputs. For instance, X-rays are resized so that they can easily fit the requirements of a model, noise is removed from image data, and color information is removed if applicable for accuracy. In addition, missing or incomplete data in any form is filled up either with statistical methods or flagged for further user input. At this pre-processing stage, the data is prepared to be in an appropriate format for precise predictions.
- **Real-time Processing**: The web application processes the data in real time, providing immediate feedback to the users on their health status. As soon as the user submits their data, the system begins analyzing it using the appropriate disease prediction models. The real-time processing ensures a seamless user experience, as results are generated instantly, and the user can immediately see the outcome of their health assessment. This immediate feedback also encourages users to take quick action, especially in the case of potential health risks.
- **Results Interpretation**: Once processed, the outputs are interpreted and transformed into a more relevant and user-friendly results presentation. The prediction model gives presentations of its findings in a light that the user can easily understand. For instance, if the model identifies individuals with risk factors for heart disease, results may present a warning for a potential heart disease diagnosis along with recommended lifestyle changes, medications, or tests to be conducted. To present the numerous health-related data in a simple way that enables users to take appropriate decisions regarding their health.

E. Personalized Reporting and Recommendation

- **Diagnostic Report**: This would automatically be generated for the user by the model based on its output. It consists of the diagnosis that has been derived, treatments that could be applied, changes in lifestyle that are recommended, and what exactly needs to be done to manage or further test the condition. For example, if the prediction outcome is that pneumonia is predicted, the report shall advise the user to seek immediate medical care and simultaneously provide home management to help improve their condition. The report is a personalized response to the individual with regard to their specific inputs and severity of condition.
- **Follow-up Actions**: Whenever the system identifies potentially dangerous conditions, that is when the user's blood sugar levels are high, the lung patterns seem irregular or have high cholesterol levels, it will advise the user to seek further evaluation from a doctor. The application provides information about doctors close to the user, along with their contact details and available times for consultation. This therefore helps users to take prompt action, schedule a medical consultation when necessary.
- **Security and Privacy**: Given the sensitive nature of health data, the application prioritizes user privacy and data security. It implements robust encryption systems to protect personal health information. All data entered by the user, including medical histories, lab results, and diagnostic images, are stored securely to prevent unauthorized access. The application follows the industry standards for data protection. The application also meets HIPAA regulations so that the health information of users is kept confidential and protected at all times.

VII. Results

A. Part A: Chatbot Conversation Results

The chatbot successfully interacts with the user by collecting detailed symptom-related information. Below is a summary of the conversation:

- **Initial Interaction**: The chatbot greets the user and identifies that the user is seeking assistance for stomach pain. It clarifies its role, emphasizing that it cannot provide a formal diagnosis but encourages sharing detailed symptoms.
- **Symptom Collection**: The chatbot effectively gathers information:
  - Type of pain: Sharp and burning sensation.
  - Intensity: Rated 8 on a scale of 0 to 10.
  - Timeframe: Symptoms have persisted for a week.
  - Additional symptoms: Fever and vomiting.

```
>>> hello
Hi! How can I help you today?

>>> i have stomach pain
I'm sorry to hear that. Can you tell me more about your symptoms?

>>> i have stomach pain and i want diagnose mysef
I'm sorry but I cannot diagnose you. It is important that you see a medical professional for proper diagnosis. In the meantime, can you
describe in more detail when the pain started and any other symptoms you are experiencing?

>>> yes i am experiencing stomach pain fro past 1 week and also i have fever
Thanks for sharing that information. Could you describe your symptoms in more detail? Do you experience the pain as a sharp or dull sensation?
Does it feel like cramping, burning, or something else?

>>> yes its a sharp burning sensation
Okay. How would you rate the intensity of your pain on a scale from 0 to 10, with 0 being no pain and 10 being the worst pain you can imagine?

>>> yes on the scale it should be around 8
Thanks for letting me know. I see. And when did you last have a meal?

>>> one day back
Okay, got it. Can you describe your appetite? Are you hungry, not hungry, craving something in particular?

>>> yes i am feeling hungry
I see. Did you throw up after eating that meal or was your stomach upset?

>>> yes i am throwing up
I see. Okay, thank you for sharing that information with me. Do you have a history of stomach problems or any allergies?

>>> no i dont have any
Thanks. Based on what you've told me so far, it is possible that your symptoms are related to a stomach virus or food poisoning. It is also
possible that your symptoms could be caused by something more serious, like appendicitis or gallbladder disease.

>>> Send a message (/? for help)
```

- **Follow-up Questions**: It inquires about the user's appetite, meal history, and any allergies or pre-existing conditions.
- **Preliminary Insight**: Based on the user's inputs, the bot suggests the symptoms might indicate a stomach virus, food poisoning, or possibly something more serious like appendicitis or gallbladder disease. It recommends consulting a medical professional.

Evaluation: The chatbot demonstrates empathetic interaction, structured data collection, and preliminary assessment capabilities.

B. Part B: Models and Performance Metrics

Overview of Models and Applications

CNN (Convolutional Neural Network) for Image Classification:

- **Task:** Applied in image classification, leveraging CNN to interpret patterns in visual data.
- **Dataset:** Trained on a Kaggle dataset.
- **Accuracy:** Achieved 91.5%, demonstrating strong generalization capabilities for image classification tasks.

KNN (K-Nearest Neighbors) in Diabetes Prediction:

- **Task:** Utilized for predicting diabetes through KNN proximity-based analysis.
- **Dataset:** Used the PIMA Indians Diabetes Dataset, which contains medical predictors.
- **Accuracy:** Achieved 85.7%, indicating reliable performance for a simple algorithm.

Random Forest for Heart Disease Prediction:

- **Application:** Used to predict heart disease by analyzing structured clinical data.
- **Dataset:** Trained on the UCI Heart Disease Dataset.
- **Accuracy:** Delivered a strong accuracy of 95%, showcasing robustness and reliability.

Conclusion: All the models are particularly suited for their respective tasks. By leveraging quality datasets and tailored techniques, significant accuracy levels were attained. These results highlight the transformative potential of machine learning in various healthcare applications, including diagnosis and prediction.

VIII. Conclusion

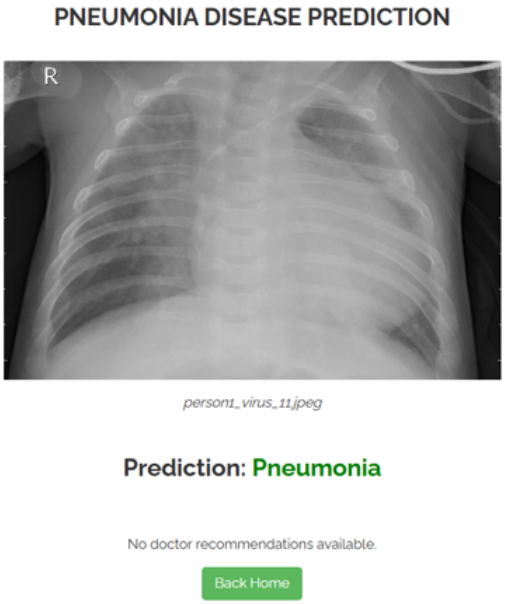
The AI-powered self-diagnosis tool is the advancement of healthcare services availability and accuracy. Utilizing the most advanced machine learning models-the NLP and computer vision-the system gives accurate disease predictions based on any symptom, history, or diagnostic scan input by the user. With these technologies integrated into the system, it can provide an all-in-one solution that connects patients to providers in a seamless approach, especially for those who cannot access timely medical consultations.

The web interface of this program is user-accessible, with a very simple and intuitive approach while ensuring sensitive health data would be strictly secured through a cloud-based deployment. This integration of sophisticated AI models and a friendly, accessible interface will smoothly engage interaction, offering users health information, appointment

TABLE II

Model Performance Summary

Model	Task	Dataset Used	Accuracy
CNN	Image Classification	Kaggle	91.5%
KNN	Diabetes Prediction	PIMA Indians	85.7%
Random Forest	Heart Disease Predicton	UCI	95.%



with doctors, and appointment with physicians. Potentially, this AI-based self-diagnosis tool could be a great impact on the underprivileged and on people living in remote areas, where health facilities are inaccessible.

It continuously improves and upgrades to learn the new medical data as well as the users' view of this system. Perhaps in the long term, it will be one of the best transformations in the delivery of care by providing early disease detection, health outcomes improvement, and patients' proactive approach in health control.

IX. Future Work

In the near future, we will develop the medical diagnosis chatbot into a fully functional and accessible website. The website will be live and will be available worldwide for users seeking real-time assistance with their medical concerns. In addition to presenting symptoms and projecting possible diagnoses, the chatbot will also suggest the nearest hospitals and healthcare facilities. Furthermore, it will provide all appointment details so users can feel hassle-free when they need to see a healthcare provider.

One of the improvements we would make is breaking the language barrier. The vast majority of South Indian languages, including Telugu, Tamil, Kannada, Malayalam, and many more, shall be there in our platform. So people from multiple backgrounds can communicate in their mother tongue with the chatbot so that people can experience ease and comfort in their communication.

At the moment, we are using the Llama 7B for natural language processing tasks, though as the system continues to grow we will step up with more advanced models with increased efficiency and accuracy which can allow the platform to deal with more complex queries in relation to providing more accurate diagnoses. Upgrading the system would be helpful in achieving that: make the system better in understanding and processing medical information.

Scalability is another aspect of our project. We plan to design the platform for handling a significant volume of concurrent users without loss of performance. Optimizing the architecture for scaling in future work will ensure that the system can grow as more demand rises. That involves putting infrastructure improvements alongside fine-tuning the underlying models to maintain responsiveness and reliability, which expands the user base.

References

[1] Alvarez. Integrated yolo for detecting lung conditions in medical imaging. *Journal of Medical Imaging*, 9(1):65–78, 2023.

[2] John. Developed a conversational ai chatbot for mental health assessment using llms. *Journal of AI in Mental Health*, 5(2):123–134, 2020.

[3] Kim. Used a machine learning model with electronic health records (ehrs) to predict diabetes onset. *Journal of Healthcare AI*, 7(4):45–56, 2022.

[4] Patel and Zhao. Developed a hybrid model using llms and clinical datasets for cardiovascular risk assessment. *Journal of Cardiovascular Health*, 12(2):210–221, 2022.

[5] Smith and Lee. Utilized a convolutional neural network (cnn) model to detect pneumonia in chest x-rays. *Journal of Medical Imaging*, 8(3):98–104, 2021.