



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 10    **Issue:** XII    **Month of publication:** December 2022

**DOI:** <https://doi.org/10.22214/ijraset.2022.48322>

[www.ijraset.com](http://www.ijraset.com)

Call:  08813907089

E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)

# A Deep Learning Enabled Chatbot Approach for Self-Diagnosis

Soumiki Chattopadhyay<sup>1</sup>, Souti Chattopadhyay<sup>2</sup>

<sup>1</sup>Department of Computer Science and Engineering, Institute of Engineering and Management (IEM), Kolkata, India

<sup>2</sup>Department of Computer Science, University of Southern California, Los Angeles, California

**Abstract:** *Healthcare is an essential need that is often inaccessible or unavailable. With rising populations and deadly viruses, the load on healthcare workers is the highest, while the ratio of healthcare workers per person has hit an all time low. This extreme demand for assistance in the face of unavailability has opened up avenues for computer assisted healthcare. In this paper, we propose an approach to use advancements in computer assisted healthcare in two directions. We employ machine learning algorithms to predict the disease correctly. We used Random Forest and XGBoost as this system's core classifiers that gave an accuracy of 100% and 99.0% correspondingly. The chatbot is used to trigger the appropriate function for the disease prediction concluding from the conversation with the user. It has an accuracy of 97.22%. Overall our chatbot has the potential to provide immediate assistance to those who are in need, and in turn, reduce the burden on healthcare workers.*

**Keywords:** *Machine Learning, Deep Learning, Healthcare, Chatbot*

## I. INTRODUCTION

Medical care is one of the most essential services for people, and governments take on responsibility to provide accessible and affordable healthcare to their people. However, governments around the world fail to provide functional healthcare access due to varying reasons like costly medical supply and care provider systems in the United States [1], or the lack of resources to provide for a huge population in India. While high costs of medical care are driven by the economic strategies and policies, lack of resources and shortage of healthcare is a global problem across all developing nations [2]. WHO estimates a projected shortfall of 10 million health workers by 2030, mostly in low- and lower-middle income countries [3]. A 2011 study estimated that India has roughly 20 health workers per 10,000 population [4] [5]. The World Health Organization (WHO) has said India faces an "enormous challenge" to train enough health personnel to meet the needs of a burgeoning, ageing population [6]. With the recent COVID 19 pandemic, even affluent countries like the US are facing healthcare worker shortage [7].

The lack of personnel creates a crisis on an upward trend in the lack of accessibility and affordability of healthcare. A study reports that in rural areas in India, only 37% of people were able to access in-patient facilities within a 5 km distance, and 68% were able to access out-patient facilities [8]. Less than half of Africa's citizens (52%) – some 615 million people – have access to the healthcare they need [9]. Due to the lack of accessibility, private sectors often carry the burden to provide healthcare to exponentially increasing the cost of healthcare, with healthcare expenditure as high as 70% in Africa and 75% in India coming from the private sector [4, 10]. While the world is focused on improving national and international policies, supply chains, and training ground forces to cope with the ever increasing demand for healthcare, especially, in the face one of the biggest pandemics faced by mankind, the gap between demand and accessibility of healthcare seems unlikely to be resolved by simply personnel or policies. The WHO reports that the demand for physicians is projected to grow in the US by 17% percent by 2025 [11], and the Indian Ministry of Health and Family Welfare puts the same estimate for India to grow by nearly 50% by 2025 [12]. This has caused governments and industry to examine computer mediation in healthcare, investigating how digital procurement and localisation can be used to improve access and make healthcare more affordable [13].

One such computer mediated approach is using computers to assist with diagnosis, popularly, in the form of medical chatbots. These chatbots have advantages that most automated systems bring like unlimited availability, learn, and disseminate critical information faster than people having to consult primary care or search for valid medical articles [14]. Chatbots are popularly used to provide automated assistance with booking appointments, building a community who share similar health experiences, or collect patient data [15]. However, most chatbots don't perform diagnosis, identifying the health issue from the patients' conversations. When chatbots have the ability to diagnose and identify what kind of health support the patient needs, they can make an impact by contributing to the healthcare workforce and making healthcare more accessible.

In this paper, we implement a diagnostic assistant that can predict specific diseases the patients are suffering from by asking for information, and then provide additional assistance or temporary relief. Our chatbot uses Deep Learning to learn with time and answer more detailed questions and provide personalised assistance. Our chatbot uses ML algorithms to predict heart and kidney diseases with 99.02% and 100% accuracy. And if the patient is diagnosed with a disease, the chatbot follows up with potential assistance. We posit, our chatbot will be useful in scenarios where people want a confirmed diagnosis to identify which specialist or doctors to contact. Since the chatbot uses minimal computation power, and can be hosted from a server, the chatbot will improve accessibility to healthcare or alleviate immediate concerns. Eventually, it will lower the cost associated with healthcare.

## II. BACKGROUND AND RELATED WORK

Our approach marries two threads of advancements in computer aided healthcare—predicting diseases accurately from health data, and using chatbots to provide medical assistance. In this section, we talk about how our approach builds upon the existing work and adds functionality to make a practically usable chatbot.

### A. Predicting Diseases using Machine Learning

Researchers have explored various Machine Learning models and their powers to predict diseases based on existing medical data. The UCI Machine learning repository [16] serves as the benchmark dataset to explore and evaluate algorithms. In recent years, authors have compiled the performances of the algorithms by various other researchers to understand the need and scope for improvement [17]. Other authors have explicitly compared the performances of various algorithms to predict heart diseases [18]. They found that researchers most frequently used Support Vector Machine (SVM). Some papers have also built hybrid algorithms to improve the performance on heart disease prediction [19]. Various papers have documented the performance of algorithms on different classes of diseases like heart, breast cancer etc. available from the repository [20].

Machine Learning algorithms are well documented to perform good prediction of diseases, however, the algorithms alone are not usable by the general users. The predictors are used by expert healthcare officials or for various stimulated drug trials. In this paper, we propose an approach where the disease predictor algorithms are made available directly to the users through a chatbot interface

### B. Chatbots in Healthcare

Chatbots have been implemented in various capacities in the field of healthcare. Some of them are trained and able to answer questions like “how to treat a cut?” “how to cure a fever?” developed the application using the N-gram, TF-IDF for extracting the keyword from the user query [21]. Each keyword is weighed down to obtain the proper answer for the query. The authors also implemented a web interface developed for the users, to the input query. The application uses n-gram for text compression using bigram and trigram for faster execution of the query. The N-gram, TF-IDF, and cosine similarity to convey the answers to the users. Other authors like Prayitno et. al. [22] builds a chatbot using Cosine Similarity, ID3 Decision Tree, and Natural Language Processing. It is designed to make therapeutic suggestions depending on the disease the user describes. The cosine similarity is used to find the similarities between the query words or the questions that are asked by the user and the documents, then return the answers of the document with the highest similarity. The chatbot can successfully diagnose user illness with approximately 87% accuracy.

Some authors have used fully trained neural networks as the core for chatbots. For instance, Reshmanth et. al. [23] built a chatbot based on natural language processing and trained using RNN. The accuracy of the chatbot is 100% and is programmed to answer questions like "how to treat dog bite?", "how to treat fever?", "how to handle the cold?". The dataset comprises these general problems and their solutions. Finally, they used Flask to create a UI for the chatbot. The large number of chatbots have been surveyed and studied by some authors to understand the use cases for the chatbots. Detailed study of various research papers related to chatbots and analyse the different tools, algorithms, software and platforms that are utilised. The study of the papers included comparison of their chatbots, the methodologies they used, advantages, disadvantages, dataset, domain, accuracy, and whether the chatbot is voice or text activated [24].

Authors have also worked to make the chatbots multilingual, for example, some papers used five classification algorithms SVM, KNN, Decision Tree, MNB, Random Forest Classifier for disease prediction [25]. The Random Forest Classifier produces the best results and gives an accuracy of 98.43%. The chatbot uses TF-IDF and cosine similarity to respond to users and also supports speech to text and text to speech conversion so that the user can also communicate using voice in three languages: English, Hindi, and Gujarati to make it suitable for use in rural areas.

In our approach, we devised a chatbot that will diagnose the disease of the user through conversational flow and utilise databases through the forms filled up by the user to detect what type of disease the user has, if any. Further, it can also provide aid for temporary relief. The chatbot is trained using LSTM with an accuracy of 97.22%, and the disease detection is performed using XGBoost or Random Forest. It utilises the concept of natural language processing to parse and understand the intents file which was created by wielding various resources.

### III. METHODOLOGY

Our approach marries the two threads of using machine learning to predict diseases from users' symptoms and using a chatbot to identify symptoms. We explain the design choices of these components, and then describe how we combined the two components to work together as a system.

#### A. Using ML Algorithms to Predict Diseases

We used data from the UCI machine learning repository [16]. For this paper, we will demonstrate the approach with two datasets, heart disease dataset and chronic kidney disease dataset. Both the datasets are multivariate and designed to classify the diseases with various metabolic measurements. To make our approach more generalizable, the two datasets were merged and passed on to the data preprocessing stage. To broaden the approach, datasets for different diseases can be merged at this stage.

##### 1) Data Processing

The data preprocessing stage took place within two functions: Heart and Kidney.

- a) *Heart Function:* When the chatbot detects the user has a possibility of heart disease (See Section 3.2), the chatbot asks the patient to fill up the heart disease form. The form collects the variable measures used for the classification and passes the form data to the heart function.
- b) *Kidney Function:* If the chatbot detects that the user might have kidney disease (See Section 3.2) from the conversation flow, it prompts the user to fill up the kidney form. Then sends the form data to the kidney function which only works with the kidney classification related variable values from the merged dataset. Since the kidney dataset has categorical values and null values, we used label encoder to encode the categorical values to allow us to perform numerical operations. Also drop the rows with null values.

##### 2) Model Training and Prediction

The processed data then invoked the algorithms, and the best performing algorithm for the specific data is then used to predict the disease on the new patient data. The model training and prediction also occurs within the specific disease functions.

- a) *Heart Disease Prediction:* The heart function only subsets the heart disease related data features from the merged dataset and uses the best performing model for heart data, XGBoost model. Then uses the trained model to predict whether the user has heart disease. The results of the prediction are then shown to the user in the "result" portion of the form. The result is a dichotomous, and the form returns the result of 0 or 1.
- b) *Kidney Disease Prediction:* The Kidney function uses random forest to train the model, and then uses the trained model on the user provided form data to predict whether the user has CKD (Chronic Kidney Disease). The prediction result is dichotomous, and the model returns 2 or 3 to the form for the user to see.

#### B. Building the Chatbot

Chatbots are virtual conversational agents that stimulate conversation with the users. There are mainly three types of chatbots: rule-based, learning enabled, and hybrid. Rule-based chatbots have predefined conversational paths and do not respond accurately if the user says anything outside the predefined path. Deep learning enables chatbots to improve themselves over time with more training and learns to mimic the human brain and stimulate human-like conversations. Hybrid chatbots are a combination of these two.

Rule based chatbots follow a predefined conversational path. If the user says something beyond the predefined path, the chatbot can't attach semantic meanings to the user conversations. This makes the user interaction with the chatbot not conversational. To avoid this limitation, in this paper, we are using a Deep Learning trained chatbot approach to build an intelligent conversational agent which, over time and training, learns the patterns of human conversations and eventually mimics the way humans talk. This allows the chatbot to provide meaningful responses even if a user asks something beyond the scope of the conversation.

Since Deep Learning chatbots are powered by Natural Language Processing, rather than using keywords to answer, they try to understand the intent of the user. The hidden layers in the neural network captures the underlying trends and improves with training. This allows the chatbot to handle variations in user conversations.

In Figure 1, we demonstrate the components and workflow of the chatbot. First, we compiled a set of user intents for interacting with the chatbot. The tags, responses, and patterns were stored in separate lists correspondingly. Then, we applied Tokenization on patterns or user inputs to split them into smaller, understandable units. Table 1 provides some examples of intents from our database.

Table 1. Intents-based data

| Tag                      | Patterns in User Conversation   | Responses   |
|--------------------------|---|---|
| greeting                 | “just going to say hi”, “heya”, “hello”, “hi”, “hey”, “hi”, “need help”                             | “Hello! This is Doctor Bot, how can I help you?”  |
| pain                     | “i am having chest pain”, “pain”, “chest pain”, “heart pain”, “heart ache”, “pain in my chest”      | “Alright. What is your age?”  |
| chestpainshortnessbreath | “Chest pain while resting”, “Sudden chest pain”, “Shortness of breath”, “Fast or Shallow breathing” | “Alright. An ideal BMI is in the 18.5 to 24.9 range for adults. What is your BMI?”                              |
| bmi                      | “i dont know”, “my bmi is 24.6”, “my bmi is 36.0”, “57.1”, “20.2”, “dont know”, “dont know bmi”     | “Okay. Ideal blood pressure is considered to be between 90/60mmHg and 120/80mmHg. What is your blood pressure?” |

Table 1. User intents in the conversations. The Tag is automatically generated, then the patterns are detected based on the tag, and corresponding responses from the chatbot.

Next, we performed Padding to convert all inputs to the same size and store them in the features data. The corresponding list of tags was the label data that the model has to learn to predict. These training data (the features dataset and the training dataset) were then passed to the Deep Learning (LSTM) model. The model used softmax activation function, sparse categorical cross entropy for loss function, and Adam optimizer to perform the training. The model was trained for a total of 320 epochs and resulted in an accuracy of 97.22%.

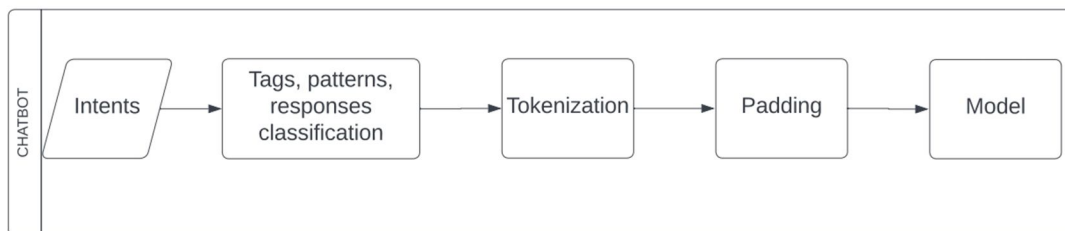


Figure 1. Chatbot architecture demonstrating the steps taken within the chatbot to build the conversation model with users.

C. Tool Design: Merging the Disease Predictor and Chatbot

The chatbot uses deep learning while the disease prediction is done using machine learning models we implemented through functions. The user starts having a conversation with the chatbot and there are two ways the conversation could flow. If the user mentions the pain around the chest area, the chatbot asks more questions related to heart disease symptoms. However, if the user mentions pain around the stomach area, the chatbot asks more questions based on symptoms a kidney disease patient could have. Once the chatbot is assured, it suggests the user to fill the kidney disease form or heart disease form accordingly. Figure 2 demonstrates the entire workflow of the system.

When the user fills out the form, the data is sent to the corresponding algorithm for kidney or heart disease. The algorithm then uses the trained model to predict whether the user has a disease. The heart disease prediction form returns a value of 0 or 1, and the kidney disease form returns 2 or 3. After the prediction value is returned, the chatbot asks the user for the value, and interprets what the value means in terms of the disease and provides assistance with tips for immediate relief until the user can consult a specialist (e.g. when the user gets the value 0, the chatbot responds "you don't have a high probability of a disease, but if would you like to receive further assistance about temporary relief?").

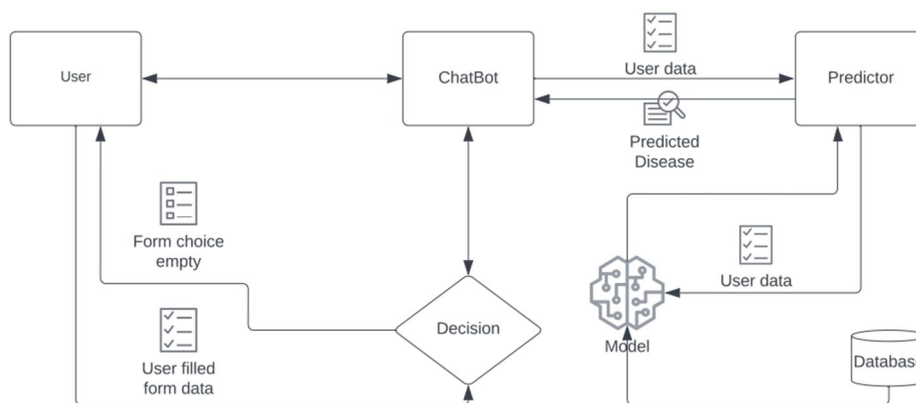


Figure 2. System architecture. The left part of the figure shows the component exchange and interaction of the chatbot and user. The right side shows the interaction between the chatbot and the disease predictor.

#### IV. RESULTS

The evaluation of our approach is based on the performance of the two components of the system—the predicting algorithms and the chatbot.

##### A. Evaluating Prediction Performance

First, we ran 8 classifications algorithms with the two datasets separately to determine the optimal model. The accuracy table for the heart disease dataset is Figure 3.b. It shows Decision Tree, Random Forest and XGBoost gave the highest accuracies. To determine which is the best among these three models, we further performed cross validation with 10 folds. XGBoost gave the highest accuracy of 99.02%.

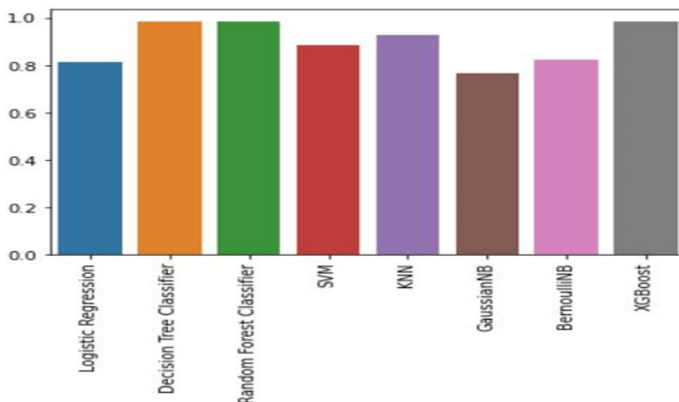


Figure 3.a. Bar Graph representing accuracies of different models evaluated on heart data. XGBoost shows the best accuracy among the models after performing cross validation.

|   | Models                   | Accuracy |
|---|--------------------------|----------|
| 0 | Logistic Regression      | 0.814634 |
| 1 | Decision Tree Classifier | 0.985366 |
| 2 | Random Forest Classifier | 0.985366 |
| 3 | SVM                      | 0.882927 |
| 4 | KNN                      | 0.926829 |
| 5 | GaussianNB               | 0.765854 |
| 6 | BernoulliNB              | 0.824390 |
| 7 | XGBoost                  | 0.985366 |

Figure 3.b. Accuracies of different models evaluated on heart data

Similarly, for the kidney disease prediction, we ran the same models to find out the optimal model. According to Figure 4.b., Random Forest, GaussianNB, BernoulliNB, and XGBoost have an accuracy of 100%. We applied hyper-parameter tuning to these models to improve models' predictive accuracy. Although the other models' accuracies reduced to some extent, Random Forest stayed consistent with a 100%. This could be possible due to overfitting. We conducted a post-hoc inspection to overrule the possibility of overfitting, and additionally, triangulated with other findings who report that 100% accuracy is a valid outcome when using Random forest [26]. Random forest is an ensemble algorithm. It combines multiple trees and each tree is grown to the greatest extent, there is no pruning, and it decides on majority vote rule. So, it is quite common for Random forest to get a 100% accuracy.

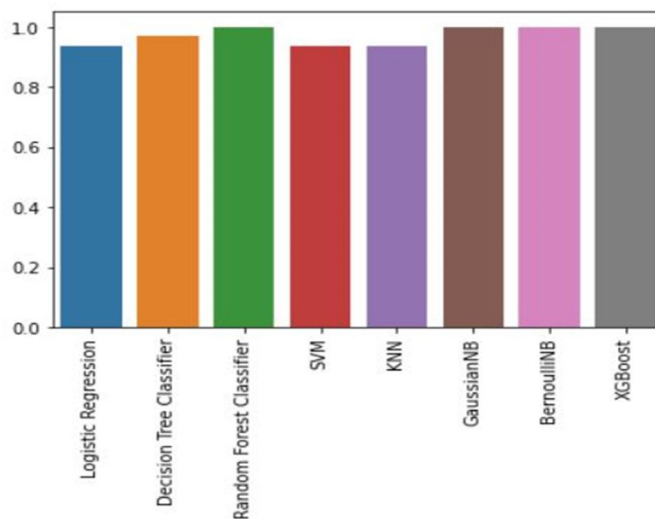


Figure 4.a. Bar Graph representing accuracies of different models evaluated on kidney data. Random Forest shows the best accuracy among the models.

|   | Models                   | Accuracy |
|---|--------------------------|----------|
| 0 | Logistic Regression      | 0.93750  |
| 1 | Decision Tree Classifier | 0.96875  |
| 2 | Random Forest Classifier | 1.00000  |
| 3 | SVM                      | 0.93750  |
| 4 | KNN                      | 0.93750  |
| 5 | GaussianNB               | 1.00000  |
| 6 | BernoulliNB              | 1.00000  |
| 7 | XGBoost                  | 1.00000  |

Figure 4.b. Accuracies of different models evaluated on kidney data.

### *B. Evaluating Chatbot Performance*

The goal of the chatbot was to engage the user with conversation that will lead to the diagnosis of the disease by matching symptoms that the user mentioned. Hence, the chatbot needed to be adaptive to various natural language conversations. This is suitable for an LSTM model, as LSTM thrives on sequential input like natural language where the meaning of a word depends on the ones that preceded it.

The model consisted of an Input layer, an Embedding Layer, LSTM with ten hidden layers, a Flatten layer, and a Dense layer with Softmax activation. The loss function used was sparse categorical cross entropy, and Adam optimizer to perform the training. The best stable accuracy was obtained when trained for 320 epochs and where the model reached a maximum accuracy of 97.22% with a learning rate lambda of 0.001.

## **V. DISCUSSION AND IMPLICATIONS**

The performance of the components of the system indicate that the approach is usable and applicable to reduce the time to reach a diagnosis and the pressure on healthcare workers. Using the two underlying functionalities, we can allow users to get immediate feedback. With the vast amount of articles available to people to search their symptoms, the searches often lead to a spiral of gradually worse diseases that the articles relate with the symptoms. Users looking for a possible correlation for a symptom like mild chest pain can get articles that diagnose it with heartburn and indigestion, while the other articles would associate the same symptoms to blockage of vessels near the heart, or cardiac arrest. However, with no one to chat with and no way of getting an initial confirmation of the disease, the extreme range in the diagnosis often puts users in panic going down a spiral of information search leading to more and more serious diagnosis.

Our approach can be used to intervene in this process, saving patients from feeling drowned in information and spooked by rumoured diagnosis, by providing immediate confirmation on whether certain serious diseases like heart failure or chronic kidney disease occurs. By making this diagnostic confirmation more accessible and available to the users, the chatbot can also save them time from worrying about a serious disease they don't have and focus on what might be the true cause of their symptoms.

Additionally, the prediction is often the first step in the healthcare process. Predicting a possible disease requires patients to go through testing their blood and measuring biophysical data. The process of getting the measurements and tests itself requires multiple visits to various doctors who then refer the patients to the clinics. Using our approach, users will be receiving a form with the names of various measures and tests they need to perform. This approach saves the time to initiate the diagnosis and also reduces the burden on healthcare workers which in turn makes healthcare more available to critical patient needs.

## **VI. FUTURE WORK**

For future work, we target to build a web application with the same chatbot integrated on a user-friendly and visually enticing platform. This paper only works with two diseases so we will increase the types of diseases that can be diagnosed by adding more data. This would need us to further add models that can also accommodate other disease data. In the future, we also plan to train the chatbot on a larger corpus of conversational cues, and enable the chatbot to be trained with LLMs like GPT3, that can parse existing information on the web and learn to explain and respond to humans.

### *A. Deploying The Chatbot To Increase Healthcare Accessibility*

We plan to develop a full stack web application with the chatbot integrated. It will require customers to make accounts, and store their data and information in a database for future references. The database will also have patients' history that will include past diagnosis, medical records, medicines prescribed for the chatbot to utilise. It will be secure as there is no third-party intervention. To make the system available to everyone, the entire application will be deployed via a cloud service.

### *B. Enhancing Chatbot Functionality To Connect To Medical Assistance Networks*

If the user feels the need to take medicines even after being diagnosed of not having any disease, the user can use the chatbot to get prescriptions or refill prescriptions, the chatbot can identify if the desired medicine is safe and adequate for the diagnosis and prescribe it to the patient.

We also wish to utilise location services and integrate with care provider network data, so the chatbot can identify closest doctor/clinic/specialist and eventually schedule appointments by sharing patient's information and initial diagnosis result from the chatbot. If a user needs urgent medical attention while in a foreign region, this will be incredibly helpful.



### C. Integrating Chatbot with Automated Scheduling Systems

Lastly, we target to function the chatbot to schedule automated appointments for tests, check-ups, and other facilities. The chatbot can be modified to have access to a database containing information about general clinics and pharmacies within the various networks, the doctors affiliated with those clinics, and their scheduling systems. Using this data, the chatbot can be integrated with a scheduling assistant module to allow users to directly book appointments with clinics and pharmacies based on the diagnosis. This module can also be refined to make distinction between low acuity medical needs like ear infection or cavity, and more serious and immediate needs like detecting lung cancer or renal failure. The chatbot can then schedule initial visits and surgeries for serious disease requiring such procedures.

## VII. CONCLUSION

Healthcare is an important resource that is often not available or accessible to users. Especially since the pandemic, the number of healthcare workers have drastically reduced making it difficult to access immediate care for millions of people. In this paper, we proposed an approach to merge the two capabilities of machine learning algorithms to help with this problem. We use ML to predict the presence of diseases in a user, and use a DL trained chatbot to converse with the patient to detect cues that can trigger the prediction algorithm. Our algorithms performed with high accuracy, and overall we envision that our approach can help users get immediate access to healthcare and in turn reduce the burden on healthcare workers.

## VIII. ACKNOWLEDGEMENTS

We would like to thank the members of the ACE Lab for their feedback that helped shape the results.

## REFERENCES

- [1] Robert H. S. (2021). Is our healthcare system broken? Harvard Health Publishing. <https://www.health.harvard.edu/blog/is-our-healthcare-system-broken-202107132542>
- [2] Naicker, S., Plange-Rhule, J., Tutt, R. C., & Eastwood, J. B. (2009). Shortage of healthcare workers in developing countries--Africa. *Ethnicity & disease*, 19(1 Suppl 1), S1-64
- [3] Health workforce. WHO. [https://www.who.int/health-topics/health-workforce#tab=tab\\_1](https://www.who.int/health-topics/health-workforce#tab=tab_1)
- [4] Kasthuri, A. (2018). Challenges to healthcare in India-The five A's. *Indian journal of community medicine: official publication of Indian Association of Preventive & Social Medicine*, 43(3), 141.
- [5] Rao, M., Rao, K. D., Kumar, A. S., Chatterjee, M., & Sundararaman, T. (2011). Human resources for health in India. *The Lancet*, 377(9765), 587-598.
- [6] Figueras, J., McKee, M., Cain, J., Lessof, S., & World Health Organization. (2004). *Health systems in transition: learning from experience*.
- [7] "A public health crisis: Staffing shortages in health care: Usc mph," Oct 2022. [Online]. Available: <https://mphdegree.usc.edu/blog/staffing-shortages-in-health-care/>
- [8] Krishna, A., & Ananthpur, K. (2013). Globalization, distance and disease: spatial health disparities in rural India. *Millennial Asia*, 4(1), 3-25.
- [9] K. Cullinan, "Universal health coverage: Only half of africans have access to health care," <https://healthpolicy-watch.news/only-half-of-africans-have-access-to-health-care/>, accessed: 2022-12-08.
- [10] Natasha Sunderji, (2022). "What's the key to affordable care in Africa?," <https://www.accenture.com/us-en/blogs/life-sciences/key-to-affordable-care-in-africa>.
- [11] Sheldon, G. F., Ricketts, T. C., Charles, A., King, J., Fraher, E. P., & Meyer, A. (2008). The global health workforce shortage: role of surgeons and other providers. *Advances in surgery*, 42, 63-85.
- [12] Sarwal, R., Prasad, U., Gopal, K. M., Kalal, S., Kaur, D., Kumar, A., ... & Sharma, J. (2021). *Investment Opportunities in India's Healthcare Sector*.
- [13] Baur, A., Yew, H., & Xin, M. (2021). *The future of healthcare in Asia: digital health ecosystems*. McKinsey & Company.
- [14] Engati Team. (2022). "How are intelligent healthcare chatbots being used? [New Uses for 2022]."<https://www.engati.com/blog/chatbots-for-healthcare#:~:text=Healthcare%20Chatbots%20are%20conversationalists%20that,generation%2C%20result%20analysis%2C%20etc>.
- [15] Cem Dilmegani. (2022). Top 6 Use Cases & Examples of Chatbots in Healthcare. AI Multiple. <https://research.aimultiple.com/chatbot-healthcare/>
- [16] UCI Machine Learning Repository. (2021). Center for Machine Learning and Intelligent Systems. <https://archive.ics.uci.edu/ml/index.php>.
- [17] Soni, J., Ansari, U., Sharma, D., & Soni, S. (2011). Predictive data mining for medical diagnosis: An overview of heart disease prediction. *International Journal of Computer Applications*, 17(8), 43-48.
- [18] Uddin, S., Khan, A., Hossain, M. E., & Moni, M. A. (2019). Comparing different supervised machine learning algorithms for disease prediction. *BMC medical informatics and decision making*, 19(1), 1-16.
- [19] Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. *IEEE access*, 7, 81542-81554.
- [20] Kohli, P. S., & Arora, S. (2018, December). Application of machine learning in disease prediction. In 2018 4th International conference on computing communication and automation (ICCCA) (pp. 1-4). IEEE.
- [21] Athota, L., Shukla, V. K., Pandey, N., & Rana, A. (2020, June). Chatbot for healthcare system using artificial intelligence. In 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO) (pp. 619-622). IEEE.
- [22] Prayitno, P. I., Leksono, R. P. P., Chai, F., Aldy, R., & Budiharto, W. (2021, October). Health Chatbot Using Natural Language Processing for Disease Prediction and Treatment. In 2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI) (Vol. 1, pp. 62-67). IEEE.

- [23] Reshmanth, P., Chowdary, P. S., Yogitha, R., & Aishwarya, R. (2022, April). Deployment of Medibot in Medical Field. In 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) (pp. 325-329). IEEE.
- [24] Swain, S., Naik, S., Mhalsekar, A., Gaonkar, H., Kale, D., & Aswale, S. (2022, April). Healthcare Chatbot System: A Survey. In 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM) (pp. 75-80). IEEE.
- [25] Badlani, S., Aditya, T., Dave, M., & Chaudhari, S. (2021, May). Multilingual Healthcare Chatbot Using Machine Learning. In 2021 2nd International Conference for Emerging Technology (INCET) (pp. 1-6). IEEE.
- [26] Santos, I. H. F., Machado, M. M., Russo, E. E., Manguinho, D. M., Almeida, V. T., Wo, R. C., ... & Silva, E. (2015, October). Big data analytics for predictive maintenance modelling: Challenges and opportunities. In OTC Brasil. Offshore Technology Conference.

## AUTHORS



Soumiki Chattopadhyay is a student of Institute of Engineering and Management, currently in her final year of B.Tech in Computer Science and Engineering.



Souti Chattopadhyay is a professor of computer science at the University of Southern California. She works at the intersection of human-computer interaction & software engineering. Her work has been awarded by IEEE, and other prominent research bodies.



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)