Chatbot for Disease Prediction and Treatment Recommendation using Machine Learning

Rohit Binu Mathew, Sandra Varghese, Sera Elsa Joy, Swanthana Susan Alex

Department of Computer Science and Engineering

Saintgits College of Engineering

Kottayam, Kerala, India

Email-id: rohit.bm@saintgits.org, sandra.varghese@saintgits.org,
sera.ej@saintgits.org, swanthana.sa@saintgits.org

Abstract— **Hospitals are the most widely used means by which a sick person gets medical check-ups, disease diagnosis and treatment recommendation. This has been a practice by almost all the people over the world. People consider it as the most reliable means to check their health status. The proposed system is to create an alternative to this conventional method of visiting a hospital and making an appointment with a doctor to get diagnosis. This research intends to apply the concepts of natural language processing and machine learning to create a chatbot application. People can interact with the chatbot just like they do with another human and through a series of queries, chatbot will identify the symptoms of the user and thereby, predicts the disease and recommends treatment. This system can be of great use to people in conducting daily check-ups, makes people aware of their health status and encourages people to make proper measures to remain healthy. According to this research, such a system is not widely used and people are less aware of it. Executing this proposed framework can help people avoid the time-consuming method of visiting hospitals by using this free of cost application, wherever they are.**

Keywords—Medical chatbot, Machine Learning, Disease Prediction, Treatment, KNN

I. INTRODUCTION

A prosperous society is when its entire people are healthy. It is important to maintain the health if one wishes to be happy. Only a healthy body can have a healthy mind and it has a positive impact on the performance of people. Nowadays, people are less aware of their health. In their busy life, they forget to take suitable measures to maintain their health and are less aware of their health status. In the latest news by TOI [1], we can see that people give no importance to their health and find it time consuming to undergo check-ups at hospitals. The busy-scheduled life has got no place for health. Most people comprising the working section of the society claim that their hectic schedule gives them no time for periodic medical check-ups and that they disregard any uneasiness shown by their body until it is too severe.

In this proposed system, a medical chatbot is built to be a conversational agent that motivates users to discuss about their health issues and based on the symptoms provided by them; chatbot returns the diagnosis [2]. This chatbot system will be able to identify symptoms from user interaction. Using these extracted symptoms, chatbot predicts the disease and recommends treatment. The machine learning algorithm [3] employed here is K-nearest neighbor algorithm (KNN). This clearly shows that a medical chatbot can somewhat accurately diagnose patients with simple symptom analysis and a conversational approach done with the help of natural language processing.

Medical chatbot has a high impact on the health culture of the state. It has improved reliability and is less prone to human errors. Today's people are more likely addicted to internet but they are not concerned about their personal health. They avoid hospital treatment for small issues which may become a major disease in future. This proposed idea solves this problem. This idea focuses on creating a chatbot which is free of cost and available throughout the day. The facts that the chatbot is free and can be accessed wherever the user is, be it their working environment, prompt the user to have it and use it. It saves the overhead involved in consulting specialized doctors.

Executing the proposed system can convey more awareness among people regarding their health status and a need to take measures to remain healthy. With the new proposed system, there will be reduction in the number of people disregarding their health because of the tedious process of hospital appointments. People can interact with the chatbot just like they do with another human and can continue with their other works. It ensures that there will be no disruption with their working and is user friendly. This provides a way to help people aware of their health by using the chatbot and thereby helps people with their health, thus playing a bigger role in healthcare.

II. OBJECTIVES AND PROPOSED INNOVATION

A. Objective

The main objective of the proposed system is to have the importance of health in life reach out to people and encourage people to follow measures to maintain health by making the chatbot available to all. Chatbot and health have a history of working well together. It creates a good human-like conversational environment for interaction between the user and the system. In this system, the user talks about their health and it is a great way for the users to regulate the healthy lifestyle.

An important aspect of this system is that talking with a non-human entity provides a sense of security especially when it comes to mental health as it remains as a confidential meeting with the diagnosis being available only to the user. This system is meant to help and deliver immediate actions where humans cannot reach due to timing or budget as it is readily available and free of cost. It allows the user to have free medical check-up based on the symptoms where the user's health issue is easily identifiable.

B. Proposed Innovation

Chatbot is helpful for anyone who needs to get the hang of something about wellbeing. The user can interact with chatbot and can rely on them to get timely diagnosis. This framework helps users to present their symptoms that affect their wellbeing. The genuine welfare of the chatbot is to encourage the general population by giving appropriate direction in regards to the great and sound living. The chatbot allows user to login to the system.

Figure 1: User login module.

User registers on chatbot application. They need to submit some personal details which will be confidential. User then interacts with the system and the words are recognized by the use of natural language processing and the system recognizes symptoms of the user.

Figure 2: User register module.

There is an admin who controls the chatbot application. The admin views the details of all the users and can even manually add, delete or update symptoms and diseases. The chatbot is trained on symptoms-disease dataset [4]. From the symptoms identified by the user, KNN algorithm can predict the disease, depending on the dataset. The system recognizes the disease and finally recommends the suitable treatment needed for the same. The image below shows a diagrammatic flow representation of the working of chatbot.
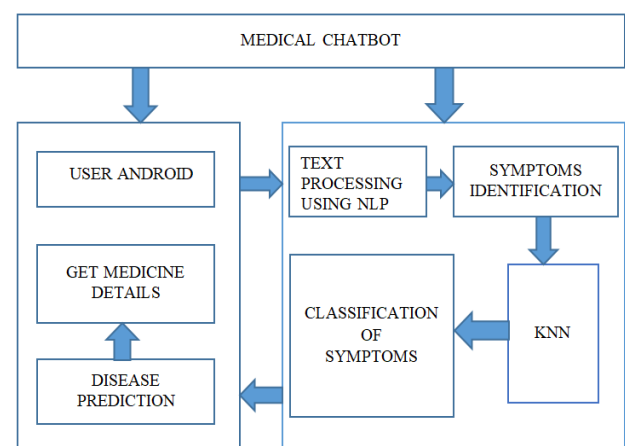

Figure 3: Working model of the proposed system.

Figure 3 shows the flow diagram of the proposed model. The chatbot helps to encourage patients to discuss about their medical issues and provides a suitable diagnosis and recommends treatment. Here the text processing is done using natural language processing. Initially the user enters their symptoms through text. It uses the machine learning algorithm KNN. KNN identifies the symptoms from the interaction with the user. KNN maps the symptoms to the particular disease and finally recommends the suitable treatment to the user.

III. RELATED WORKS

In the paper by Rashmi Dharwadkar [5], the medical chat-bots functioning depends on Natural language processing that helps users to submit their problem about the health. The user can ask any personal query related to health care through the chatbot without physically being available to the hospital by using Google API for voice-text and text-voice conversion. Query is sent to chatbot and the user gets related answer which is displayed on android app. This system mainly focuses on analyzing users sentiments.

The proposed idea of the paper 'a novel approach for medical assistance using trained chatbot' by Divya Madhu [4] is to build up a model using artificial intelligence that can help people identify the proper treatment for their disease. There are many treatments available for a particular disease and no one can possibly know the proper treatment which is best suited for their disease. In this proposed model, artificial intelligence takes up the main role of providing a list of available treatments based on the disease identified through the symptoms. The system can also list out the composition of medicines and their prescribed uses, thereby helping users to select the proper treatment. This system helps people to have a basic idea of their health status, thus encouraging them to take proper treatment.

The proposed idea of the paper 'a self-diagnosis medical chatbot using artificial intelligence' by S. Divya [6] is to build up a system using artificial intelligence that can help users to avoid consultation with a doctor. It is made to diagnose the disease of the user and provide necessary details about the disease. This is built to be a cost effective system with improved accessibility to knowledge about diseases. A chatbot is beneficiary to users only when it can diagnose any type of disease and provide necessary information. The proposed system is a conversational agent which interacts with users to retrieve knowledge about their medical conditions, thereby providing a proper diagnosis.

In the paper by Amiya Kumar Tripathy [7], it mentions the need of advanced technology that provides people with a proper healthcare management system, where people can rely on it instead of a doctor. It emphasizes the need of such a system to be accurate and portable so that people can carry with them, this reliable system. The proposed system consists of a mobile heart rate measurement where heart rate can be

detected and based on this, a proper diagnosis will be provided with a click of button. The system also provides video conferencing where one can connect with a doctor in case of urgency. The Doc-Bot which was developed for this purpose is now being converted to a mobile platform and is to integrate the idea of providing diagnosis based on symptoms.

In the paper by S. du Preez [8], it proposes the idea of an intelligent voice recognition chatbot. This proposes the design of such a system and the necessary technology needed to develop the system. This is a web-service that allows any client to connect with the server from anywhere. This also uses the concept of black box approach to control the communication, to and fro from the web-service. This is accessible through an interface that allows seamless xml processing. By the use of an artificial brain, the web-based chatbot responds to the user queries. If the question from users is not understood, it will be further processed using an online intelligent research assistant and the response will be saved so as to provide efficient responses in the future.

In the paper by B. R. Ranoliya [9], it defines chatbots as programs that can mimic human conversation using artificial intelligence. It proposes the idea of chatbot as a virtual assistant or as a smart agent that can do tasks like giving proper responses to queries from users, controlling devices, providing routes during driving etc. The paper presents the idea of a chatbot, which is able to give proper responses to the queries asked by users, based on the dataset of FAQs. For this purpose, it integrates both artificial intelligence markup language and latent semantic analysis. They are both used to respond to particular type of questions. Artificial intelligence markup language is used to respond to template based and general questions. Responses to other service based questions are provided using latent semantic analysis. This is mainly used in universities to interact with students and answer FAQs to them.

IV. DESIGN AND METHODOLOGY

The user chats with the chatbot application in the same way one chats with other humans. It is an android application where the user first login to the system, once registered. It is through this chat, the chatbot finds the symptoms of the user. The user sends messages and as a response the chatbot gives appropriate message. For this to happen smoothly, the chatbot will be trained with some possible questions and answers predefine, that the user can ask. When the user sends messages, text processing will be done. Text processing is done using natural language processing (NLP). NLP makes human to communicate with the machine easily [10]. NLP tries to understand the natural language spoken by humans and classifies it, analyses it as well if required responds to it. When a question is received, chatbot tries to converge it that available in the dataset which it is already trained on and it will be one among the k nearest neighbors. This is done with the help of KNN. Python has a rich set of libraries which cater

to the needs of NLP. The Natural Language Tool Kit (NLTK) is a suite of such libraries which provides the functionalities required for NLP.

One important task done by NLP is tokenization. The key functionality of tokenization is to convert the text received from the user to tokens. Tokenization occurs in such a way that different words will be converted to different tokens [11]. These tokens are then used as the input for other types of analysis or tasks, like parsing. After forming the tokens from the words, they will undergo lemmatization and stemming. Stemming and lemmatization are performed to make text processing easier. They represent the key steps in natural language processing. Stemming follows rule-based method. It looks upon prefixes or suffixes in tokens and they are removed. This is how it reduces words to its root form [12]. When the words are reduced to its root form, a bag of words is later formed. Similarly the dataset contents which include symptoms and diseases are also processed in the same way and converted to bag of words. This is a vector comprising of 0s and 1s and each of them will be given an answer id.

This system requires the need for an algorithm that can help in predicting the disease from the symptoms entered by the user. The system uses a machine learning algorithm. For the chatbot to provide the user with a response to the messages received, it needs KNN. K nearest neighbor (KNN) is a machine learning algorithm which is very simple and efficient. This is best suited for pattern recognition. It is a widely used classification algorithm where the input data samples, in this case, symptoms are classified based on the majority class label of k nearest neighbors [13]. It performs classification on the data samples and the classes are stored. Thus when a new data sample is obtained, it performs classification based on similarity measure to other data samples [6].

For example, headache symptom can be distinguished by having normal fever or a tumor. In reality, there can be any number of predictors and the example can be extended to incorporate any number of characteristics. For example if we take four nearest kinds of diseases relating headache, they are classified to a proper disease comparing with other symptoms. KNN can classify properly the symptoms and map them to their corresponding diseases.

Firstly, a model will be created using sklearn library imported from python. When user asks a question, the question will undergo a series of text operations and finally will be converted to a vector (bag of words). Then the vector will be given to the model and which will produce the index of the answer which will be later mapped to find the disease. The chatbot will predict the disease to the user and will also provide a link where the user can search about the treatment needed for the disease predicted.

V. Experiments And Results

A. Experiments

This newly developed chatbot was tested among people. People who are sick make use of this chatbot application and test their health status. A person having cold fever shows symptoms like dryness, cough, headache, fatigue and body pain. Those symptoms were analyzed based on the dataset and the medical chatbot correctly predicted the disease cold fever for the given symptoms. Different algorithms require different dataset and training. So accuracy can vary in different algorithm. Here, KNN algorithm is used.

```
From sklearn.neighbors import KNeighborsClassifier     (1)
def create_model():
        neigh = KNeighborsClassifier(n_neighbors=3)
        return neigh
```

Equation (1) is used to create a model using KNN with the library imported from python. After the creation of model, it will be trained on tested. With the dataset available, 75% was used for training and the rest 25% was used as test data. The new data from the test set was used with the model and it correctly predicted the disease.

```
def train():                                          (2)

        model = create_model()
        X_train,Y_train = prepare_dataset()
        model.fit(X_train,Y_train)
        file = open("models/model.pickle","wb")
        pickle.dump(model,file)
        file.close()


def load_model():
        file = open("models/model.pickle","rb")
        model = pickle.load(file)
        return model


def predict(text):
        vector = create_vector(text)
        model =load_model()
        answer_id = model.predict([vector])
        return answer_id[0]
```

Equation (2) shows how the model is trained with the dataset and how it returns the correct answer when it is provided with a new vector.

B. Results

On analysing the conducted experiments, the following results were obtained:

Using KNN machine learning algorithm, some symptoms like headache, vomiting, body pain, fever and cough were entered and following result was obtained.
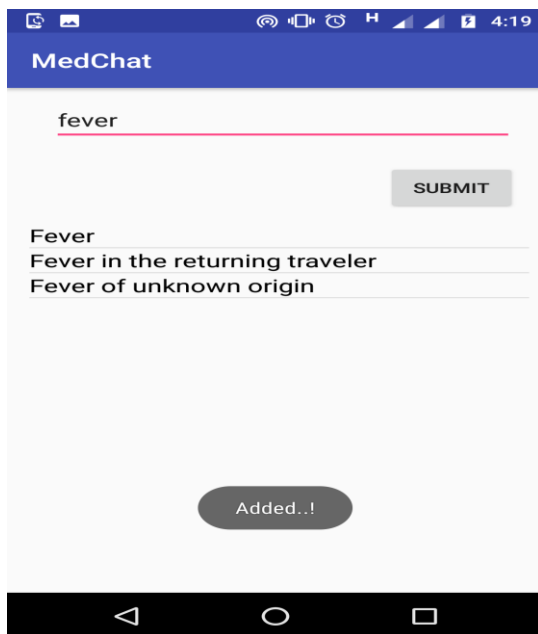

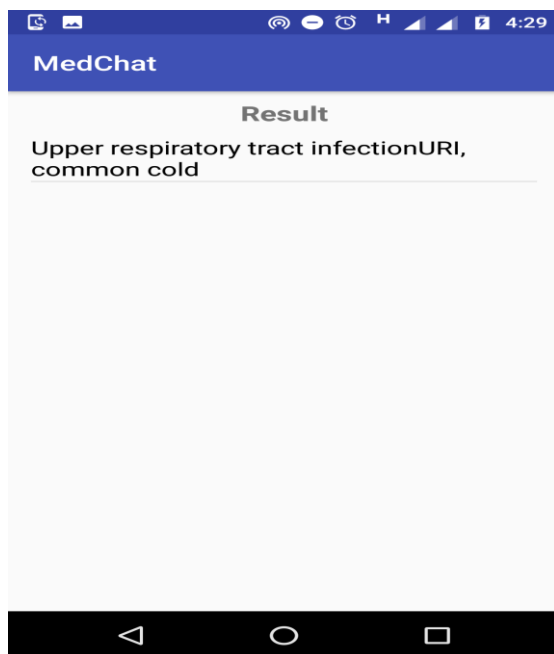Figure 4: Output Snapshot of Symptoms Identification

Figure 5: Output Snapshot of Real Time Disease Prediction

For the symptoms entered by the user, the correct disease was predicted and hence, it is reliable and can be used for keeping track of health status. A reliable system is what needed in a hectic-scheduled life of people. With the proper results being displayed, people can have a hope in this new system.

VI. CONCLUSION AND FUTURE SCOPE

A. Conclusion

This paper explains a medical chatbot which can be used to replace the conventional method of disease diagnosis and treatment recommendation. Chatbot can act as a doctor. The chatbot acts as a user application. The user of this application can specify their symptoms to the chatbot and in turn, chatbot will specify the health measures to be taken. General information about symptom and diseases are available in the dataset and thus the chatbot instance can provide information about disease and treatment to the user. After analyzing the symptoms of the different users, it finally predicts the disease to the user and provides with a link where details about the treatment is visible.

A smart medical chatbot can be useful to patients by identifying the symptoms as described by them, giving proper diagnosis and providing with suitable treatment for the disease. In the busy life, it is rare for people to frequently visit hospitals for check-ups. Chatbot is of great importance in such situations as they provide diagnostic assistance with a single click of button. Chatbot doesn't require the help of any physician to give proper health measures to the users and this is one of the major advantages of chatbot. Moreover, the cost-effectiveness in using chatbot is a major attractiveness to users. The chat with users is completely personal and this helps users to be more open with their health matters and paves way for chatbot to efficiently identify the disease.

B. Future Scope

The role played by chatbot can sometimes be beyond the scope and user may require consulting a doctor for taking health related tests. In such situations, chatbot can be helpful if it can be made to set up an appointment with an efficient doctor based on their schedule. Also it will be beneficial if the symptoms and disease identified by the chatbot can be made into a report and automatically forwarded to an available doctor where he can further assist the user with more advices and future measures to maintain their health. A video call with a specialized doctor can also be made depending on the availability of the user rather than based on the availability of doctors.

REFERENCES

[1] The Times of India, https://timesofindia.indiatimes.com/life-style/health-fitness.

[2] Sameera A. Abdul-Kader, Dr. John Woods. Survey on Chatbot Design Techniques in Speech Conversation Systems, School of Computer Science and Electronic Engineering/University of Essex Colchester/ UK.

[3] Kuligowska, Karolina. (2015). Commercial Chatbot: Performance Evaluation, Usability Metrics and Quality Standards of Embodied Conversational Agents. Professionals Center for Business Research. 2. 1-16. 10.18483/PCBR.22.

[4] Divya Madhu,Neeraj Jain C. J, Elmy Sebastain, Shinoy Shaji, Anandhu Ajayakumar. A Novel Approach for Medical Assistance Using Trained Chatbot, International Conference on Inventive Communication and Computational Technologies (ICICCT 2017).

[5] Mrs. Rashmi Dharwadkar, Dr.Mrs. Neeta A. Deshpande "A Medical ChatBot"" in International Journal of Computer Trends and Technology (IJCTT) V60(1):41-45 , June 2018.

[6] S. Divya, V. Indumathi, S. Ishwarya, M. Priyasankari, S. Kalpana Devi, "A Self-Diagnosis Medical Chatbot Using Artificial Intelligence", J. Web Dev. Web Des., vol. 3, no. 1, pp. 1-7, 2018.

[7] Amiya Kumar Tripathy , Rebeck Carvalho , Keshav Pawaskar , Suraj Yadav, "Mobile based healthcare management using artificial intelligence" in International Conference on Technologies for Sustainable Development (ICTSD), 2015, 4-6 Feb, 2015.

[8] S. du Preez, M. Lall, S. Sinha, "An intelligent web-based voice chat bot", EUROCON 2009 EUROCON'09. IEEE, pp. 386-391, 2009.

[9] R. Ranoliya, N. Raghuwanshi, S. Singh, "Chatbot for university related FAQs", 2017 Int. Conf. Adv. Comput. Commun. Informatics ICACCI 2017, pp. 1525-1530, September 2017.

[10] M. Chandhana Surabhi, "Natural language processing future", 2013 International Conference on Optical Imaging Sensor and Security (ICOSS). doi:10.1109/icoiss.2013.6678407

[11] Fco. Mario Barcala, Jes´us Vilares, Miguel A. Alonso, Jorge Gra˜na, Manuel Vilares, "Tokenization and proper noun recognition for information retrieval", 13th International Workshop on Database and Expert Systems Applications. doi:10.1109/dexa.2002.1045906

[12] Pu Han , Si Shen , Dongbo Wang , Yanyun Liu, "The influence of word normalization in English document clustering", 2012 IEEE International Conference on Computer Science and Automation Engineering (CSAE). doi:10.1109/csae.2012.6272740

[13] Shweta Taneja , Charu Gupta , Kratika Goyal , Dharna Gureja, "An Enhanced K-Nearest Neighbor Algorithm Using Information Gain and Clustering", 2014 Fourth International Conference on Advanced Computing & Communication Technologies. doi:10.1109/acct.2014.22