

Disease Classification based on Symptoms using CHATBOT

H Shiva reddy¹, Mohammad Faraz², Darshan M jain³, Mani kanta G⁴, Asst. Prof. Prasad B S⁵

^{1,2,3,4}CMR Institute of Technology, Visvesvaraya Technological University, Bangalore, India.

⁵Asst. prof. Prasad B S, CMR Institute of Technology, Visvesvaraya Technological University, Bangalore, India.

Abstract - In today's world there are millions of diseases with various symptoms for each, no human can possibly know about all of these diseases and the treatments associated with them. So, the problem is that there isn't any place where anyone can have the details of the diseases or the medicines/treatments. What if there is a place where you can find your health problem just by entering symptoms or the current condition of the person. It will help us to deduce the problem and to verify the solution. The proposed idea is to create a system with artificial intelligence that can meet these requirements. The AI can classify the diseases based on the symptoms and give the list of available treatments. The System is a text-to-text diagnosis chat bot that will engage patients in conversation with their medical issues and provides a personalized diagnosis based on their symptoms and profile. Hence the people can have an idea about their health and can take the right action.

Key Words: diagnosis, chat bot, symptoms, disease, machine learning, disease classification.

1. INTRODUCTION

Hence there is a need of a system that helps people understand With the hustle and bustle of daily life backed-up with the fast paced development of modern world, People suffering from fever tend to assume it to be a common cold or something that will eventually reduce over time and completely neglect the fact that they may be suffering from something which may or may not be serious at the time but will eventually cause a problem in the near future. The current generation is in such a fast paced environment, that people hardly find the time to consult the doctor nor do they find time to look up the symptoms they are currently suffering from and hence are not capable of taking the appropriate action or treatment in time. There are also situations wherein people misunderstand a particular symptom to be something that isn't and take measures based on it on their own will and end up facing the consequences. their current condition or symptom, gives them the appropriate actions to be taken, by either providing them details of the medicines to be taken or to consult a doctor, without the risk of losing time on their side. We built a text-to-text conversational agent that diagnosis patients explaining their condition using natural language. The bot asks for relevant information, e.g., age and sex, and requests a list of symptoms. The system remembers past responses and asks progressively more specific questions in order to obtain a good diagnosis. The three primary components of our system are (1) identification and extraction of symptoms from the conversation with the user, (2) accurate mapping of

extracted (and potentially ambiguous) symptoms to documented symptoms and their corresponding codes in our database, and (3) developing a personalized diagnosis as well as referring the patient to an appropriate specialist if necessary.

2. RELATED WORK

Although several medical diagnosis chat bots already exist, including MD, Babylon, and Melody, current implementations focus on quickly diagnosing patients by identifying symptoms using tools such as radio buttons and pure system initiative questions instead of natural conversation. The systems, that currently exist do not provide an interface with the above mentioned necessities.

- Healthcare chat bots currently seem to be a mix of both patient-only (apps that help a patient track and make sense of health data) and patient-clinician applications (apps that connect the two groups, for diagnosis, treatment, etc.)
- The patient-clinician system offer user the capability of having a 'live chat', with an expert or a doctor, but lacks the speed with which the information is required nor the exact content.

There are a number of different paradigms for building dialog systems. A simpler text-to-text conversational paradigm is the finite state system, where states determine machine responses and handwritten logic interprets user input to choose the correct transition to the next state. In order to maximize our control of the logic with which our system identifies symptoms, our approach uses a finite state paradigm for managing the bot-patient dialogue. In order to properly identify and extract desired information from the user, we must have a good natural language understanding approach that accurately interprets information tokens and intents.

In terms of dealing with errors in speech recognition understanding, users preferred "progressive assistance" over repeated "I didn't understand" responses. This sort of feedback gave users the sense that the system was trying to understand what they were saying, avoiding the "brick wall" effect. We used elements of progressive assistance in our bot's user feedback templates, especially in the symptom gathering phase.

3. PROPOSED SYSTEM

3.1 Problem statement

The current generation is in such a fast paced environment, that people hardly find the time to consult the doctor nor do they find time to look up the symptoms they are currently suffering from and hence are not capable of taking the appropriate action or treatment in time. There are also situations wherein people misunderstand a particular symptom to be something that isn't and take measures based on it on their own will and end up facing the consequences.

Hence there is a need of a system that helps people understand their current condition or symptom, gives them the appropriate actions to be taken, by either providing them details of the medicines to be taken or to consult a doctor, without the risk of losing time on their side.

3.2 Proposed solution

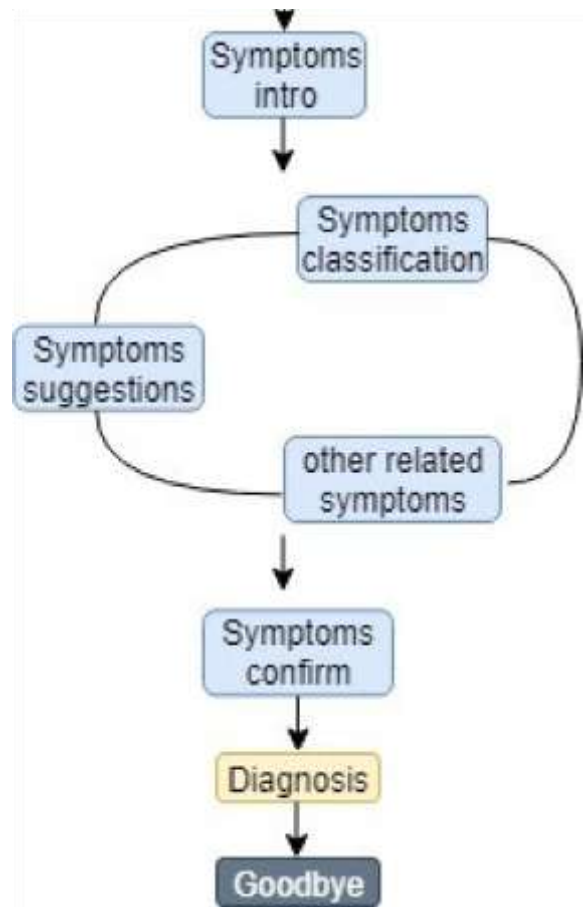
The proposed solution to tackle this problem, is to develop a chat bot, that will take in the symptoms from the user, uses machine learning, to come up with a classification of which type of fever it is, provide the user with the appropriate details of what he/she is suffering from, and will give them the appropriate actions to be taken.

4. SYSTEM DESIGN

Our chat-bot's user dialogue is a linear design that proceeds from symptom extraction, where we use spacy and nltk, and a logic layer to extract symptom-like entities, to symptom map-ping, where we identify corresponding symptom, using a text classification model, in our database, to diagnosis, where we use a trained decision tree classifier, to identify the most likely diagnosis based on the user's age, sex, and symptoms. Besides its greetings and goodbye states, our agent has three main conversational phases: acquisition of basic information, symptom extraction, and diagnosis. Our bot starts off by asking about the user's age and sex and then enters a loop of symptom extraction states until it acquires sufficient information for a diagnosis.

4.1 Symptom Extraction

Symptom extraction refers to the mechanisms used to identify potential symptom substrings in users' natural language text input. In cases where users simply list their symptoms, perhaps after some leading phrase (e.g. "I have a cough, fever, and nausea"), this is a relatively simple task. Our chat bot uses a set of text classification algorithms in order to fulfill this task. The first classification being the one which classifies a text as having symptoms or not. The second being the one which classifies the symptom in the text to a particular symptom. This helps to avoid user errors such as incorrect spellings or typing mistakes. Both algorithms use nltk package to stem the text input provided and are learned to classify the symptoms from this stemmed text using a text classifier.



4.2. Symptom Mapping

The process of identifying the given symptom with the one present in our database. This process uses one of the text classification algorithms again in order to carry out this task. The algorithm after classification, sends the symptom to a vectorizer which converts the text into binary value. This value along with the other converted symptoms is passed into the diagnosis phase.

4.3 Question Filterer

Our system, progresses further by asking a series of questions to the user. This is carried out by an algorithm that takes in a symptom which is provided by the user the user and the searches the database for all the disease and their corresponding symptoms, and then displays a set of symptoms from among those diseases. If a user provides no as the answer to the given set, it then moves to the next set of symptoms from the previously filtered disease list. If the disease is selected, the algorithm then filters the diseases with the current set of user input symptoms and the entire process is carried out in this manner.

4.4. Diagnosis

The system uses a pre trained decision tree model to classify the disease. The model is trained on a number of instances and the weights and model is saved and restored during classification to reduce load on memory. The classifier takes

in the symptom list gathered from the user through the chat bot, and runs the classification 10 times, and then presents the percentage of the diseases it has classified. This is done in order to give the user a broad perspective of what he/she maybe suffering from.

5. CONCLUSION

The proposed system has been successfully built, with an accuracy of 89% in classification of the disease. The system takes in input from the user, and uses the given symptoms for classification. In future, the bot can be integrated with a voice system that recognizes user's voice and converts it to text to extract the symptoms. The classifier can be further improved by using a hybrid algorithm..

6. REFERENCES

- [1] Monica Agarwal, Janette Cheng, Caelin Tran, "What's Up Doc?, A Medical Diagnosis Bot".
- [2] Rule based Mapping, spacy.io.
- [3] Entity Recognizer, spacy.io.
- [4] Decision Tree, tensorflow.org, medium.com.