

Healthcare Chatbot using Natural Language Processing

Papiya Mahajan¹, Rinku Wankhade², Anup Jawade³, Pragati Dange⁴, Aishwarya Bhoge⁵

^{1,2,3,4,5}Student's Dept. of Computer Science & Engineering, Dhamangaon Education Society's College of Engineering and Technology, Dhamangaon Rly., Maharashtra, India.

Abstract - To start a good life healthcare is very important. But it is very difficult to consult the doctor if any health issues. The proposed idea is to create a healthcare chatbot using Natural Language Processing technique it is the part of Artificial Intelligence that can diagnose the disease and provide basic. To reduce the healthcare costs and improve accessibility to medical knowledge the Healthcare chatbot is built. Some chatbots acts as a medical reference books, which helps the patient know more about their disease and helps to improve their health. The user can achieve the benefit of a healthcare chatbot only when it can diagnose all kind of disease and provide necessary information. The system provides text or voice assistance, that means user can use his own convenient language, Bot will provides which type of disease based on the user symptoms, and provides doctor and also provides food suggestion that means which type of food you have to take. Thus, people will have an idea about their health and have the right protection.. Chatbots are programs that work on Machine Learning (ML) as well as Artificial Intelligence (AI) Natural Language Processing (NLP) techniques such as NLTK for Python can be applied to analyses speech, and intelligent responses can be found by designing an engine to provide appropriate human like responses.

Key Words: Chatbot, Natural Language Processing (NLP), Natural Language Toolkit (NLTK), Machine Learning (ML), Artificial Intelligence (AI).

1. INTRODUCTION

Now a days, health care is extremely necessary in our life. Today's people are busy with their works reception, workplace works and additional addicted to web. They are not involved regarding their health. So they avoid to travel in hospitals for little issues. It may become a significant drawback.

So, we will offer a thought is to make a health care chatbot system using AI that may identification the illness and supply basic information regarding the illness before consulting a doctor. Which helps the patients apprehend additional regarding their illness and improves their health. User can do the all reasonably illness information. The system application uses question and answer protocol within the style of chatbot to answer user queries. The response to the question is replied supported the user question. The significant keywords are fetched from the sentence and answer to those sentences. If match is

discovered or vital answer are given or similar answers are displayed can identification which sort of illness you have got supported user symptoms and additionally offers doctor details of explicit illness. It may cut back their health problems by victimization this application system. The system is developed to scale back the tending price and time of the users because it isn't potential for the users to go to the doctors or consultants once in real time required.

2. LITERATURE SURVEY

Simon Hoermann [1] discuss this proof for the practicability and effectiveness of online one-on-one psychological state interventions that use text-based synchronous chat. Synchronous written conversations are getting well-liked as Web-based psychological state interventions. This review is predicated on associate analysis of individual synchronous Web-based chat technologies. Several of the prevailing systems have live chats through texts and a few limitation like there's no instant response given to the patients they need to attend for consultants acknowledgement for an extended time. A number of the processes could charge quantity to measure chat or telecom communication. However, the difficulty of those technologies are cost effective in clinical practice remains a thought for future analysis studies.

Saurav Kumar Mishra [2] says that the chatbot will act as a virtual doctor and makes possible for the patient to interact with virtual doctor. Natural language processing and pattern matching algorithm for the development of this chatbot. It is developed using the python Language. Based on the survey given it is found that the no of correct answer given by the chatbot is 80% and incorrect/ambiguous answer given is 20%. From this survey of chatbot and analysis of result suggested that this software can be used for teaching and as a virtual doctor for awareness and primary care.

Divya Madhu [3] proposed an idea in which the AI can predict the diseases based on the symptoms and give the list of available treatments If a person's body is analyzed periodically, it is possible to predict any possible problem even before they start to cause any damage to the body. Some Challenges are research and implementation costs, and government regulations for the successful implementation of personalized medicine, they are not mentioned in the paper.

Hameedullah Kazi [4] describes the development of a chatbot for medical students, that is based on the open source AIML based Chatterbean. The AIML based chatbot is customized to convert natural language queries into relevant SQL queries. A total of 97 question samples were collected and then those questions were divided into categories depending on the type of question. According to the number of questions in each category the resultant categories were ranked. Questions were based on queries, where 47% are of posed questions

3. EXISTING SYSTEM

Many of the existing systems have live chats through texts and some limitation such as there is no instant response given to the patients they have to wait for experts acknowledgement for a long time. Some of the processes may charge amount to live chat or telephony communication. However, the issue of these technologies are cost effective in clinical practice remains a consideration for future research studies.

Here the studies are based on to recognize emotions classification using AI methods. The studies train emotions classification models from a lot of labelled data based on RNN, deep learning, convolutional neural network. Linguistic interaction is most important in counselling using NLP and NLG to understand dialogues of users. Here the multi-modal approach is used of emotion-recognition. They have collected corpuses to learn semantic information of words and represent as vector using the word vector, synonym knowledge of lexical are collected. [1]

In this paper a voice recognition chat-bot is developed, if the questions are not understood asked to the bot is further processed using the third party expert-system. The web-bots are created as text-based web-friends, an entertainer for the user. Here they focused on the improved system if the system is not only text-based but also voice-based trained. Here the voice recognition requires a 2 part process of capturing and analysis of an input signal. Server response recognition data retrieval and information output. The server used here is SOAP based on black box approach. The use of expert system allows unlimited and autonomous intelligence improvements. [2]

This chatbot aims to make a conversation between human and machine. Here the system stores the knowledge database to identify the sentence and making a decision to answer the question. The input sentence will get the similarity score of input sentences using bigram. The chatbot knowledge is stored in RDBMS. [3]

The chatbot implemented using pattern comparison in which the order of the sentence is recognized and saved response pattern. Here the author describes the implementation of the chatbot Operating system, software,

programming language, and database. How results input and output is stored. Here the input is taken using text () function and other punctuation is removed using trim () function and random () function is used to choose a response from the database. The chatbot is used for an entertainment purpose. [4]

Here they use n-gram technique for extracting the words from the sentences. Here n-gram is used for comparison and deduction of the input with case data using Moro phonemes and phonemes as the deciding parameter. Probability analysis for the closest match is performed. The final expression is redirected through an expert system. [5]

The chatbot developed here for healthcare purposes for the android application. The user sends the text message or voice message using Google API. Here the user gets only related answer from the chatbot. SVM algorithm is used to classify the dataset. Here the Porter algorithm is used to discard unwanted words like suffixes or prefixes. [6]

The different documents served in web, the content is checked by tagging the dataset using n-gram based low dimensional demonstration, TF-IDF matrix that generates S, U, and V and finally multiplying the 3 matrices cosine similarity is calculated. [7]

Here the chatbot is created for the customer service that functions as public health service. The application uses N-gram, TF-IDF and cosine similarity. The knowledge base is created for storing the question and answer. The application clearly shows extracted the keyword from the question ad by using unigram, bigram, and trigram which helps in fast answering. [8]

3.1 Disadvantages in Existing system

- It takes more time to response to the user question
- Pay some charges to perform live chat

4. PROPOSED SYSTEM

In our proposed system the user can chat with the bot regarding the query through voice or text. The system uses an expert system to answer the queries. User can also view the available doctors for that particular disease. This system can be used by the multiple users to get the counselling sessions online. The data of the chatbot stored in the database in the form of pattern-template. Bot will provide analgesics and food suggestions that means which food you have to take based on the disease.

4.1 Advantages in proposed system

- Reducing health care cost
- Save the user time

- Don't go to hospital for even any small problem

5. SYSTEM ARCHITECTURE

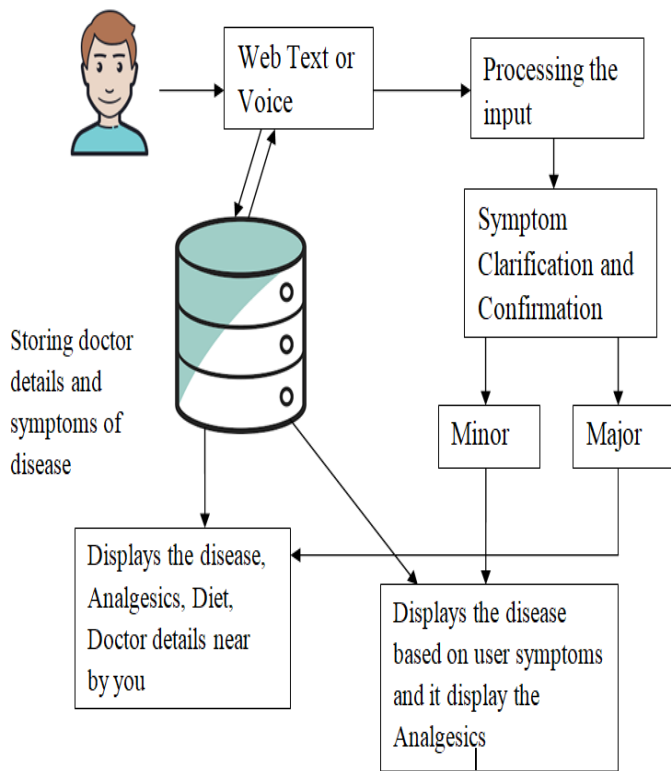


Fig-1: system architecture

The above Figure proceeds with the user can start their conversation with the chatbot like user friendly and it will be stored in the database for future reference. The chatbot will clarify the users symptoms with serious of questions and the symptom conformation will be done. The disease will be categorized as minor and major disease. Chatbot will reply whether it's a major or minor disease. If it's a major one user will be suggested with the doctor details near by you for further treatment and display the analgesics and also provides food suggestions that means which food you have to take more to recover the disease. The chatbot user interface can chat with like user friendly.by using chatbot don't go to hospitals for even small problems.

5.1 Dataflow Diagram

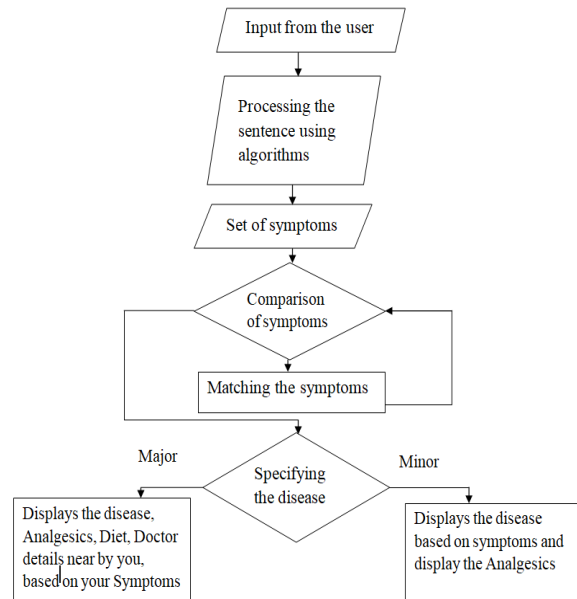


Fig-2: Data flow diagram of chatbot

The chatbot will take the input from the user and then processing the input by using algorithms. Bot will apply the algorithms on whatever the user give the input to the bot.it will understand the input by using algorithms, set of symptoms in the database. The chatbot will clarify the users symptoms with serious of questions and the symptom conformation will be done. The disease will be categorized as minor and major disease. Chatbot will reply whether it's a major or minor disease. If it's a major one user will be suggested with the doctor details near by you for further treatment and display the analgesics and also provides food suggestions that means which food you have to take more to recover the disease

6. ALGORITHMS

We are using three algorithms to implement health care chatbot system.

1. N-gram Algorithm
2. TF-IDF(Term frequency-inverse data frequency)
3. Cosine similarity algorithm

6.1 N-gram Algorithm

N-Grams are way to help machines understand a word in the content to get a better understanding of word. N-gram is a neighboring sequence of n-items from a given sample of text. N-items means we can have two items, three items and so on. So, it is a contiguous sequence of some items.it helped to predicting the next words in a sentence. Things will be characters, words, sentences. When n is 2 then we

can call it as bigrams and n is 3 then we can call it as trigrams. Based on sentence we can change the value of 'n'.

6.2 TF-IDF

6.2.1 Term frequency (tf)

Generally, when building a model with the goal of understanding text, you'll see all of stop words being removed. Another approach is to score the relative importance of words using TF-IDF. The number of times a word seems during a document divided by the entire number of words within the document. Each document has its own term frequency.

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{i,j}}$$

6.2.2 Inverse Data Frequency (IDF)

The log of the number of documents divided by the number of documents that contain the word. Inverse data frequency determines the weight of unique words across all documents in the corpus.

$$idf(w) = \log\left(\frac{N}{df_t}\right)$$

6.3 Cosine similarity Algorithm

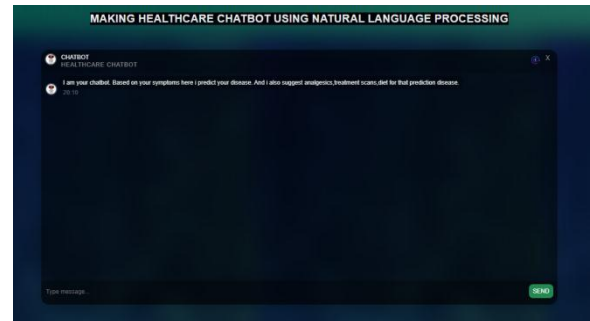
Cosine similarity finds a similarity between two non-zero vectors of an inner product space that measures the cosine of the angle between them. This technique is also used to measure cohesion among clusters within the field of information (data) mining.

Cosine similarity = $\frac{A \cdot B}{|A| |B|}$.

Cosine distance is nothing however getting distance between two vectors in n dimension area. Distance represent however words associated with one another.

7. SCREENSHOTS

Chatbot user interface:



Communication page:



8. CONCLUSION

Chatbot is great tool for conversation language between human and machine. The application is developed for obtaining a fast response from the bot which implies with none delay it provides the correct result to the user. It's ended that, the usage of chatbot is user friendly and might be utilized by someone who is aware of the way to sort in their own language. Chatbot provides personalised diagnosis supported symptoms.

9. FUTURE SCOPE

The future era is that the era of messaging app as a result of people spend longer time in messaging app than the other apps. The implementation of personalized drugs would with success save several lives and build a medical awareness among the people. No matter how far people are, they will have this medical voice communication. The sole demand they have easy desktop or smartphone with active web association. The economical of chatbot will be improved by adding a lot of combination of words and increasing the use of database information so of the medical chatbot may handle all type of diseases.

10. REFERENCES

- [1] K. Oh, D. Lee, B. Ko and H. Choi, "A Chatbot for Psychiatric Counseling in Mental Healthcare Service Based on Emotional Dialogue Analysis and Sentence

- Generation," 2017 18th IEEE International Conference on Mobile Data Management (MDM), Daejeon, 2017, pp. 371-375. doi: 10.1109/MDM.2017.64
- [2] Du Preez, S.J. & Lall, Manoj & Sinha, S. (2009). An intelligent web-based voice chat bot. 386 - 391.10.1109/EURCON.2009.5167660
- [3] Bayu Setiaji, Ferry Wahyu Wibowo, "Chatbot Using a Knowledge in Database: Human-to-Machine Conversation Modeling", Intelligent Systems Modelling and Simulation (ISMS) 2016 7th International Conference on, pp. 72-77, 2016.
- [4] Dahiya, Menal. (2017). A Tool of Conversation: Chatbot. INTERNATIONAL JOURNAL OF COMPUTER SCIENCES AND ENGINEERING. 5. 158-161.2017.
- [5] C.P. Shabariram, V. Srinath, C.S. Indhuja, Vidhya (2017). Ratatta: Chatbot Application Using Expert System, International Journal of Advanced Research in Computer Science and Software Engineering, 2017
- [6] Mrs Rashmi Dharwadkar¹, Dr.Mrs. Neeta A. Deshpande, A Medical ChatBot, International Journal of Computer Trends and Technology (IJCTT) - Volume 60 Issue 1-June 2018
- [7] N-gram Accuracy Analysis in the Method of Chatbot Response, International Journal of Engineering & Technology. (2018)

BIOGRAPHIES



Papiya P. Mahajan- Btech Student
Dept. CSE ,DES's College of Engineering and Technology ,
Dhamangaon Rly. Maharashtra,
India.
Domain : Machine Learning



Rinku B. Wankhade-Btech Student
Dept. CSE, DES's College of Engineering and Technology ,
Dhamangaon Rly. Maharashtra,
India.
Domain : Machine Learning



Anup A. Jawade-Btech Student
Dept. CSE, DES's College of Engineering and Technology ,
Dhamangaon Rly. Maharashtra,
India.
Domain : Machine Learning



Pragati P. Dange - Btech Student
Dept. CSE, DES's College of Engineering and Technology ,
Dhamangaon Rly. Maharashtra,
India.
Domain : Machine Learning

**Aishwarya S. Bhoge -Btech**

Student

Dept. CSE, DES's College of
Engineering and Technology ,
Dhamangaon Rly. Maharashtra,
India.

Domain : Machine Learning