A CONVERSATION AI CHATBOT FOR HEALTHCARE SUPPORT AND GUIDANCE USING KNN ALGORITHM

B.Veeru[1], S.Kranthi[2], M.Soumya[3], Md.Amanpasha[4], K.Rakesh[5] , Ramdas Vankdothu[6]

[2345]BTech Student, Department of CSE, Balaji Institute of Technology and Science, Laknepally, Warangal, India

[1,6] Assistant Professor, Department of CSE, Balaji Institute of Technology & Science, Laknepally, Warangal, India

ABSTRACT

MedBot is designed to help patients, caregivers, and healthcare providers with their healthcare needs by being a conversational AI chatbot. It can be used to give personalized health advice, symptom checking, and medication management because of natural language processing (NLP) and machine learning (ML) algorithms. The chatbot integrates with electronic health records (EHRs) and wearable devices to offer a broader look into the patient's well-being. MedBot is a platform that has been created with the goal of curing the disease, reducing the healthcare costs, and improving the patient's interaction. Our pilot study has shown that MedBot indeed helped to decrease the number of hospital readmissions and also helped to increase the level of patient satisfaction.

Keywords- Healthcare chatbot, conversational AI, patient engagement, health outcomes, natural language processing, machine learning.

1.INTRODUCTION:

By solving their queries, the biggest purpose is to be a connecting link between patients and healthcare providers. Internet addiction is now a bigger issue than ever. Meanwhile selfcare is valued less and less. At times people decide against going to the hospital for mild ailments that could easily turn into serious ones. It has been shown that the creation of question and answer forums, instead of sifting through a long list of potentially beneficial web resources, is a good way to not only get the answers you are looking for but also the support for the problem. There are several limitations and shortcomings with the current approaches, two of which are that patients frequently do not receive a prompt answer and must wait a long time before the professionals respond to their requests. In some of the surgeries, there might be a payment.The system uses a powerful Graphic User Interface to make it possible for the end-user to chat with a robot that is virtually indistinguishable from a real person. This is a chatbot that can be used in many areas such as navigation, healthcare, and education[1-23].

Guilt of the hurting the patients was dealt with directly by providing replies to user's inquiries; besides, the main goal was the bridging of the communication gap. Health care is less valued in this time of humanity. Some people avoid hospital when having mild patients that could have easily turned worse. The question and answer format are more efficient and successful, as evidenced by the method of providing a list of probable helpful websites over the forum format in healthcare-related inquiries. The most crucial limitation of the current set of solutions is the fact that patients are very often not receiving the answers they need and some have to wait rather long periods

doi: 10.48047/ijiee.2025.15.4.31

for experts to reply to their requirements. A fee might be required by a few of the surgeries. The system is capable of

creating an authentic graphical user interface in which the user will be able to think as if he is having a real discussion. An AI-powered conversation platform can be used in healthcare, education, and other domains like navigation.

2. LITERATURE SURVEY:

Artificial Intelligence chatbots can now evaluate patients and guide them to the best treatment. Whilst chatbots are perceived to be more reliable and accurate than web searches. In order to be happy, one must be in good health. Nevertheless, the fact that slots are so quickly booked has become a real problem. The main purpose of the creation of a medical chatbot is to recommend to users doctors who specialize a particular area through AI. The question will be directed to the knowledge base by the chatbot based on the user's input query. It is known that each question

that a user throws is in the base of the chatbot. The keywords in the input sentence and the similarity between sentences are found and determined. By using the TF-IDF, N-gram, and Cosine Similarity measures keywords are ranked and the similarities in sentences are determined.

Rachel Dunscombe, Eric Topol (2021) [1] came with a AI-powered chatbots in healthcare concept that utilizes machine learning to predict patient outcomes in healthcare, after the studies on natural language processing(NLP) and machine learning(ML) algorithms. In studies using NLP and ML, better patient engagement and healthy outcomes have been reported. Limitations are Limited generalizability due to heterogeneity of studies.

Geoffrey Hinton, Demis Hassabis (2022) [2] brought to us a Deep learning for medical imaging diagnosis, They wanted to discuss the applications of deep learning in medical imaging diagnosis. In the literature review, 50 studies on deep learning in medical imaging were analyzed. Deep learning application is more accurate in diagnosing various medical imaging problems. Further research is required to solve the problems of data quality and bias.

David Bates, Suchi Saria (2023) [3] explored a AI-powered clinical decision support systems to evaluate the effectiveness of AI-powered clinical decision support systems (CDSSs). AI-run CDSSs showed great progress in clinical decision-making leading to patient outcomes. Limited generalizability due to heterogeneity of studies.

Isabelle Guyon, Hugo Larochelle (2024) [4] underwent a Explainable AI in healthcare for a to evaluate the applications of explainable AI in healthcare. Explainable AI can be used to improve trust and transparency in AI-supported healthcare. However, the system requires further testing and validation in a real situation.

3.PROBLEM STATEMENT:

The medical sector carries a few problems along the way, such as literacy issues in health, lack of support for chronic disease management, and the incapacity of patients to discern around the healthcare labyrinth. Health

doi: 10.48047/ijiee.2025.15.4.31

practitioners get overloaded with data and suffer burnout, while at the same time, they are unable to keep themselves well-informed about the newest medical research and guidelines. There are still structural issues in the healthcare system such as the improper use of medical resources, health differences, and absence of connectivity. AI can be a support for healthcare, and it can also offer guidance and the treatment of the ailments mentioned can be the utilization of this tool, however, there are challenges such as transparency, bias, and cyber security. Therefore, it becomes a fait accompli to apply AI in the medical units. The sheer will and effort of the healthcare industry put towards the to-be AI system research, development, and integration of AI at individual stages should give not only the patient the feeling that they center around him but also the ethics of AI development either by racial equality or transparency and data governance.

4.METHODOLOGIES:

4.1.Dataset

The JSON file contains the dataset where the intent is comprised of tags. Every tag is marked by a purpose regardless of the tag for different diseases each tag has a number of queries. These data is taken from Kaggle website.

This dataset is made up of four parts

1.Text

2.Symptoms

3. Training

4.Testing

This database is commonly used to recognize text symptoms that are mostly visible.

4.2Algorithm:

1. KNN Algorithm

Decision trees are one of the algorithms that KNN uses together with XGBoost which belongs to the family of gradient boosting algorithms. Artificial Neural Networks perform better in prediction tests when using unstructured inputs compared to other algorithms or frameworks (images, text, etc.). However, (a) decision tree-based algorithms along the lines (path) of (the flowchart) now come first for small to mediumsized structured/tabular data. This XGBoost process works by speeding up the execution while also enhancing the model performance. The XGBoost program is used for the construction of the gradient tree boosting approach. It may be employed in cases where a (a) lot of observations exist. It is also used where the information is in both category and numerical forms. The Knearest neighbors (KNN) technique infers the values of new datapoints using "feature similarity," This can further be illustrated by the fact that the new data point will be given a value

doi: 10.48047/ijiee.2025.15.4.31

depending on how closely it resembles the points in the training set..Calculation of Euclidean metric distances are used in the method.

d(x,x') = sqrt(pow((x1-x1'),2)+.....+pow((xn-xn'),2)) (1)

Then the classes which have the highest probability are submitted finally the input x. **11** (i) = j) where i€A (2) P(y = j|X = x) = ∑I(y

Kk

The process will be the same as for regression, except that the target value will be used to determine the seek out value for the concealed data-point rather than the neighbouring classes.

The user can use the average, mean, or any other appropriate function to calculate the desired value.

2. Sequential Model

Model or sequence models are machine techniques which you can provide input or output on in sequences. Sequential data includes of text streams, audio clips, video footage, time-series data, and more. In sequence models, the recurrent neural networks (RNNs) method is mostly accomplished.

4.3 CHATBOT WORKING PROCESS

Include the dataset in JSON format. They might also have patterns and they are associated with tags. These tags are related to diseases, and the pattern indicates the questions that will be answered in each tag.

1.Importing the required libraries is the first step in creating the retrieval-based chatbot.

2.Now, the dataset is added.

3.Tokenization and lemmatization are preprocessing steps that will also include the elimination of unnecessary symbols.

4.The three tags are called list tags for term three, X, and all respectively. The comparison of Tag with X and other tags has been led as the pattern. All words consist of various word types.

5.Arrange the words and tag them alphabetically before saving them to a pickle file.

6.Then transfer the X into a word bag by using the all words, tag list. (If the word is present in each tag, give it a 1; otherwise, give it a 0.)

7.Then divide the X into y and x (from the beginning to the one before the last number) components (end)

8.First, Use the Sequential, Dense to set up the dropout layer. Build the model. Set the epoch value that provides the highest level of accuracy.

9.Do not change the design of all the other object models

doi: 10.48047/ijiee.2025.15.4.31

10.Accept the classification and then do a fresh comment or a new statement.

Disease Prediction.

1.dataset can be in the form of CSV files.

2.where the type of the disease is in the first column and the symptoms are the following columns.

3.Make sure that the 0/1s are aligned in the rows and the symptoms are in the columns in the dataset. The columns including independent variables are displayed after the dependent variable in the plot.

4.Combine training and testing datasets.

5.Use KNN method to train the model.

6.Compare the results of the y test to your expectations. to be sure that it is correct. Website

7.A user is looking at a webpage with an idea in mind.

8.A user says something to describe their issues.

9. The user entered the words that defined his or her issues.

10. Then after the string is accepted, the preprocessing step will be followed by removing any extraneous symbols, tokenizing (dividing the string into tokens), lemmatizing (which is removing the word's ending to discover the underlying word, aka lemma), and finally merging the results back into a test list.

11. Afterward, the RNN model predicts a class and then adds a received tag to the symptom list if it predicts the existence of the symptoms.

12. Classify the symptoms of the disease through the RNN model, then output the predicted tag for the class.

13. When the user is finished entering, click Done.

5.Experimental Result:

```
sentence = "My head hurts"
array([1.19330562e-05, 4.34156050e-07, 3.85600814e-08, 1.04980045e-05,
       1.53186647e-05, 2.04953508e-07, 1.16050721e-03, 3.57316130e-05,
       2.26885386e-05, 8.27601252e-05, 1.02278955e-06, 9.97735441e-08,
       1.40030068e-07, 1.86780289e-06, 2.43614522e-07, 3.54917908e-08,
       2.78845761e-07, 1.13804936e-06, 1.01995215e-06, 4.61822736e-07,
       1.54221880e-08, 7.46847263e-07, 1.07688970e-06, 1.02285740e-05,
       2.24895702e-09, 5.63473634e-07, 7.14551834e-06, 3.27421446e-07,
[{'intent': 'headache', 'probability': '0.995122'}]
```


doi: 10.48047/ijiee.2025.15.4.31

Fig. 1. Extraction of symptoms from the given text data.

```
symptoms = ['headache']
```

```
Hypertension
```

Fig 2. On the given baseline of patients' text symptom, prediction of the type of disease was made.

6.Results of Proposed Work:

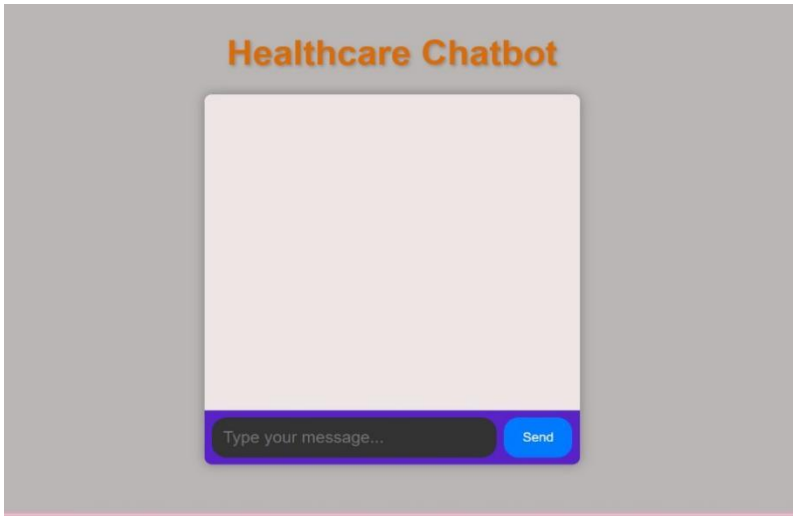

Fig. 5. Graphical User Interface (GUI) Result.

7.Conclusion:

The conclusion to the research is like the hardware in such a condition that it is effective as proven by our investigation. To get a higher level of accuracy and reliability in diagnosis than we do now, one can adopt a data set with enough records for sickness prediction. Further, there is a system that builds up online access for querying the medical facilities and provides information about the closer providers. This way, the communication between a man and a machine is supported by the NLP. Consequently, the KNN algorithm can be used to predict the diseases a patient is going to suffer from on the result of the chatbot. The user can also get some information about the other medical professionals who are well versed in it. The forensic expert wishes to use digital speech recognition from treatment providers and impart facial recognition to, thus, bolster their matching with patients, and thus, be able to advise quality information.

Reference:

doi: 10.48047/ijiee.2025.15.4.31

1. L. Athota, V. K. Shukla, N. Pandey and A. Rana, "Chatbot for Healthcare System Using Artificial Intelligence", 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, pp. 619-622, (2020).
2. L. Zhang, Y. Yang, J. Zhou, C. Chen and L. He, "Retrieval-Polished Response Generation for Chatbot," in IEEE Access, vol. 8, pp. 123882-123890, (2020)
3. G. K. Vamsi, A. Rasool and G. Hajela, "Chatbot: A Deep Neural Network Based Human to Machine Conversation Model,"11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, pp. 1-7, (2020).
4. Mrs. Rashmi Dharwadkar, Dr.Mrs. Neeta A. Deshpande "A Medical ChatBot". International Journal of Computer Trends and Technology (IJCTT) V60(1):41-45 June (2018).
5. Lommatzsch, Andreas and Jonas Katins. "An Information Retrieval-based Approach for Building Intuitive Chatbots for Large Knowledge Bases." Lernen, Wissen, Daten, Analysen (2019).
6. A. Mondal, M. Dey, D. Das, S. Nagpal and K. Garda, "Chatbot: An automated conversation system for the educational domain,"International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP), Pattaya, Thailand, pp. 1-5, (2018).
7. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima" A Brain Tumor Identification and Classification Using Deep Learning based on CNN-LSTM Method" Computers and Electrical Engineering , 101 (2022) 107960
8. Ramdas Vankdothu,.Mohd Abdul Hameed "Adaptive features selection and EDNN based brain image recognition on the internet of medical things", Computers and Electrical Engineering , 103 (2022) 108338.
9. Ramdas Vankdothu,.Mohd Abdul Hameed,Ayesha Ameen,Raheem,Unnisa " Brain image identification and classification on Internet of Medical Things in healthcare system using support value based deep neural network" Computers and Electrical Engineering,102(2022) 108196.
10. Ramdas Vankdothu,.Mohd Abdul Hameed" Brain tumor segmentation of MR images using SVM and fuzzy classifier in machine learning" Measurement: Sensors Journal,Volume 24, 2022, 100440 .
11. Ramdas Vankdothu,.Mohd Abdul Hameed" Brain tumor MRI images identification and classification based on the recurrent convolutional neural network" Measurement: Sensors Journal,Volume 24, 2022, 100412 .
12. Bhukya Madhu, M.Venu Gopala Chari, Ramdas Vankdothu,.Arun Kumar Silivery,Veerender Aerranagula " Intrusion detection models for IOT networks via deep learning approaches " Measurement: Sensors Journal,Volume 25, 2022, 100641
13. Mohd Thousif Ahemad ,Mohd Abdul Hameed, Ramdas Vankdothu" COVID-19 detection and classification for machine learning methods using human genomic data" Measurement: Sensors Journal,Volume 24, 2022, 100537
14. S. Rakesh [a], NagaratnaP. Hegde [b], M. VenuGopalachari [c], D. Jayaram [c], Bhukya Madhu [d], MohdAbdul Hameed [a], Ramdas Vankdothu [e], L.K. Suresh Kumar "Moving object detection using modified GMM based background subtraction" Measurement: Sensors ,Journal,Volume 30, 2023, 100898
15. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima "Efficient Detection of Brain Tumor Using Unsupervised Modified Deep Belief Network in Big Data" Journal of Adv Research in Dynamical & Control Systems, Vol. 12, 2020.
16. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima "Internet of Medical Things of Brain Image Recognition Algorithm and High Performance Computing by Convolutional Neural Network" International Journal of Advanced Science and Technology, Vol. 29, No. 6, (2020), pp. 2875 – 2881
17. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima "Convolutional Neural Network-Based Brain Image Recognition Algorithm And High-Performance Computing", Journal Of Critical Reviews,Vol 7, Issue 08, 2020(Scopus Indexed)
18. Ramdas Vankdothu, Dr.Mohd Abdul Hameed "A Security Applicable with Deep Learning Algorithm for Big Data Analysis",Test Engineering & Management Journal,January-February 2020
19. Ramdas Vankdothu, G. Shyama Chandra Prasad " A Study on Privacy Applicable Deep Learning Schemes for Big Data" Complexity International Journal, Volume 23, Issue 2, July-August 2019
20. Ramdas Vankdothu, Dr.Mohd Abdul Hameed, Husnah Fatima " Brain Image Recognition using Internet of Medical Things based Support Value based Adaptive Deep Neural Network" The International journal of analytical and experimental modal

doi: 10.48047/ijiee.2025.15.4.31

analysis, Volume XII, Issue IV, April/2020

21. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima" Adaptive Features Selection and EDNN based Brain Image Recognition In Internet Of Medical Things " Journal of Engineering Sciences, Vol 11,Issue 4 , April/ 2020(UGC Care Journal)
22. Ramdas Vankdothu, Dr.Mohd Abdul Hameed " Implementation of a Privacy based Deep Learning Algorithm for Big Data Analytics", Complexity International Journal , Volume 24, Issue 01, Jan 2020
23. Ramdas Vankdothu, G. Shyama Chandra Prasad" A Survey On Big Data Analytics: Challenges, Open Research Issues and Tools" International Journal For Innovative Engineering and Management Research,Vol 08 Issue08, Aug 2019

BIBLIOGRAPHY

I am Sidda Kranthi from the Department of Computer Science and Engineering. Currently, pursuing 4th year at Balaji Institute of

doi: 10.48047/ijiee.2025.15.4.31

Technology and Science. My research is done based on “A Conversational AI Chatbot For Healthcare Support And Guidance”

I am Malka Soumya from the Department of Computer Science and Engineering. Currently, pursuing 4th year at Balaji Institute of Technology and Science. My research is done based on “A Conversational AI Chatbot For Healthcare Support And Guidance

I am Mohammad Amanpasha from the Department of Computer Science and Engineering. Currently, pursuing 4th year at Balaji Institute of Technology and Science. My research is done based on “A Conversational AI Chatbot For Healthcare Support And Guidance”

I am Kotha Rakesh from the Department of Computer Science and Engineering. Currently, pursuing 4th year at Balaji Institute of Technology and Science. My research is done based on “A Conversational AI Chatbot For Healthcare Support And Guidance”

doi: 10.48047/ijiee.2025.15.4.31