DocBot - Healthcare Chatbot

Shivam
Apex Institute of Technology,
Chandigarh University,
Gharuan, India
19BCS6129@cuchd.in

Divyansh Dodan
Apex Institute of Technology,
Chandigarh University,
Gharuan, India
19BCS6106@cuchd.in

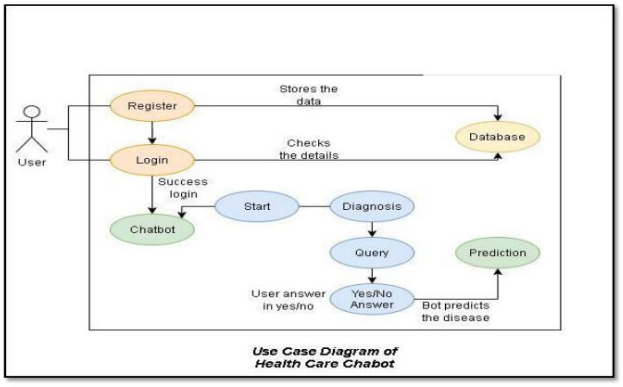
Sahil Thakur
Apex Institute of Technology,
Chandigarh University,
Gharuan, India
19BCS6103@cuchd.in

Depanshi Tomar
Apex Institute of Technology,
Chandigarh University,
Gharuan, India
19BCS6095@cuchd.in

Abstract -The main aim of project ―HEALTH CARE CHATBOT, is to help you better visualize the presentation of mined data (information). It deals with all the health care issues which will really benefit stakeholders in the health care space. Basically, chat-bot is a computer program that pretends chat with humans through natural language. On any platform like mobile, website and desktop application, this system can interact with the humans. While interacting with the human, chat-bot simulates as a human being. Human being only interacts with one human at a time, the chat-bot interacts and communicates with hundreds and thousands of persons simultaneously. It works and responds without considering how many persons are interacting and what time of the day and night it is. Chat-botsin the health care sector can play an importantrole. Chat-bot algorithms can be trained on massive healthcare data using disease symptoms, diagnostics, markers, and available treatments. Chat-bots can be updated continuously using Public datasets, such as COVID-19 for COVID-19, and Wisconsin Breast Cancer Diagnosis(WBCD). Conversational chat-botswith different intelligence levels can understand the questions from the users and provide answers based on pre - defined labels in the training data.

INTRODUCTION

Through chat bots one can communicate with text or voice interface and get reply through artificial intelligence. Typically, a chat bot will communicate with a real person. Chat bots are used in applications such as E-commerce customer service, call centers and Internet gaming. Chat bots are programs built to automatically engage with received messages. Chat bots can be programmed to respond the same way each time, to respond differently to messages containing certain keywords and even to use machine learning to adapt their responses to fit the situation. A developing number of hospitals, nursing homes, and even private centers, presently utilize online Chat bots for human services on their sites. These bots connect with potential patients visiting the site, helping them discover specialists, booking their appointments, and getting them access to the correct treatment. An ML model has to be created wherein we could give any text input and on the basis of training data it must analyze the symptoms. A Supervised Logistic Regression machine learning algorithm can be implemented to train the model with data sets containing various diseases CSV files. The goal is to compare outputs of various models and suggest the best model that can be used for symptoms in real-world inputs. Data set contains CSV file having all diseases compiled together. The logistic regression algorithm in ML allows us to process the data efficiently. The goal here is to model the underlying structure or distribution of the data in order to learn more from the training set. In any case, the utilization of artificial intelligence in an industry where individuals' lives could be in question, still starts misgivings in individuals. It brings up issues about whether the task mentioned above ought to be assigned to human staff. This healthcare chat bot system will help hospitals to provide healthcare support online 24 x 7, it answers deep as well as general questions. It also helps to generate leads and automatically delivers the information of leads to sales. By asking the questions in series it helps patients by guiding what exactly he/she is looking for.



Use Case Diagram of Health Care Chabot

BACKGROUND

This project is ML-based which is basically a subfield of artificial intelligence in which we use data and algorithmic approach to imitate human behavior.

Broadly Machine Learning is divided into two parts:

Supervised ML

Unsupervised ML

Supervised ML: Supervised learning is typically done in the context of classification when we want to map the input to output labels, or regression when we want to map the input to continuous output. Common algorithms in supervised learning include logistic regression, naive Bayes, support vector machines, artificial neural

networks, and random forests. In both regression and classification, the goal is to find specific relationships or structures in the input data that allow us to produce correct output data effectively. Note that “correct” output is determined entirely from the training data, so while we have a ground truth that our model will assume is true, it is not to say that data labels are always correct in real-world situations. Noisy, or incorrect, data labels will clearly reduce the effectiveness of your model. When conducting supervised learning, the main considerations are model complexity and the bias-variance tradeoff. Note that both of these are interrelated.

Unsupervised ML: The most common tasks within unsupervised learning are clustering, representation learning, and density estimation. In all of these cases, we wish to learn the inherent structure of our data without using explicitly-provided labels. Some common algorithms include k-means clustering, principal component analysis, and autoencoders. Since no labels are provided, there is no specific way to compare model performance in most unsupervised learning methods.

An application of counseling chatbot, which provides conversational service for mental health care based on emotions recognition methods and chat assistant platform. This application doesn’t consider the user's psychiatric status through continuous user monitoring.

2.1 Existing System-Many of the existing systems have chats through texts.Some limitations of such Chatbots are,there is no instant response given to the patient, they have to wait for experts acknowledgement for a long time. And also there are a limited number of diseases in the dataset.Technical issues like voice messages are not accurate in the existing system.

2.2 Proposed System-In our system the user can interact with the chatbot through text and chat bot will interact using voice and text manner. With respect to the users queries, the bot identifies the disease if user chatting with the chatbot. According to the diseases of the user, bot gives suggestions for the disease and also prescribe specialist doctors. This system can be used by multiple users at a time without any lagging.

METHODOLOGY

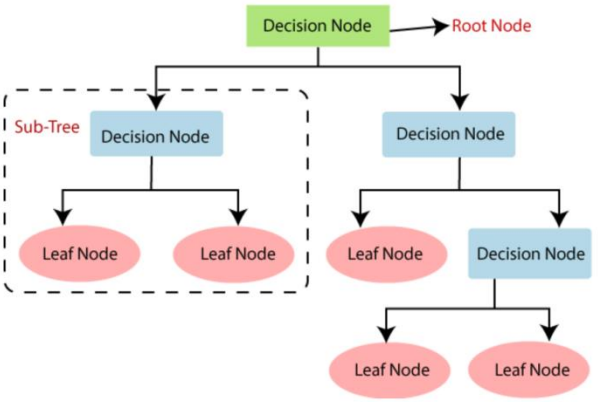
ALGORITHMS USED :

Decision Trees: Decision tree algorithms are a type of supervised machine learning algorithm that can be used for both classification and regression tasks. A decision tree is a hierarchical model that uses a tree-like structure to make decisions based on a series of rules and conditions.

The algorithm works by recursively splitting the data into subsets based on the features that provide the most information gain or decrease in impurity, such as entropy or Gini index. At each node of the tree, a feature is selected as the root node, and the data is split based on the value of the selected feature. The process is repeated at each subsequent node until a stopping criterion is met, such as reaching a maximum depth or having a minimum number of samples in a leaf node.

Once the decision tree is constructed, it can be used to make predictions by following the path of the tree that matches the values of the features in the new data point, ultimately reaching a leaf node that corresponds to the predicted class or value.

Some popular decision tree algorithms include ID3, C4.5, CART, and Random Forest. Decision trees are commonly used in various applications such as healthcare, finance, and marketing, where they can help make decisions based on complex data sets.

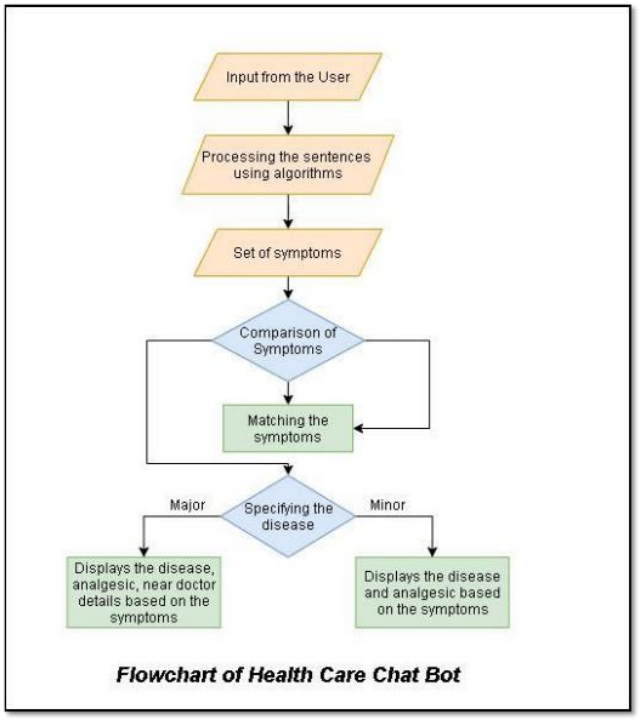


2) Label Encoding : Label encoding is a technique used in machine learning to convert categorical data into numerical data, which can be used as input to machine learning models. In label encoding, each unique category or label in a categorical feature is assigned a numerical value. The numerical values assigned to each label are typically integers that start from 0 and increase sequentially.

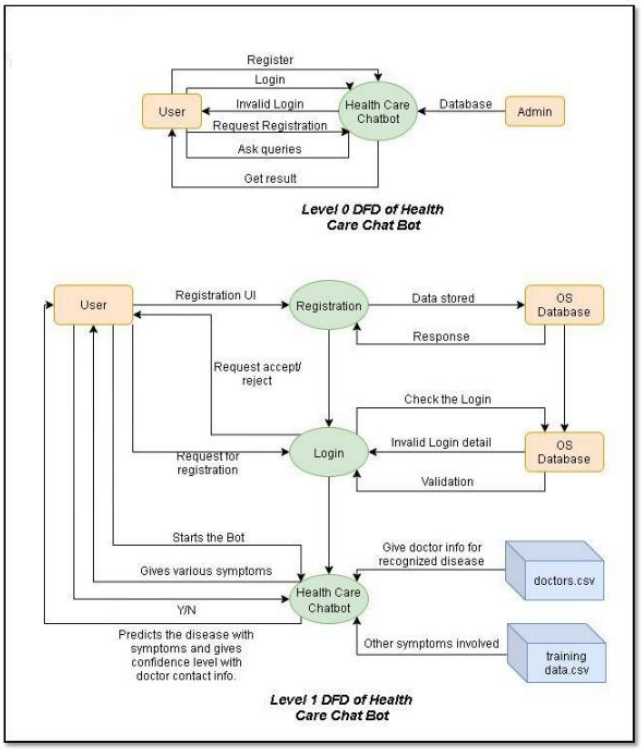
For example, suppose we have a categorical feature called "color" with three unique labels: "red", "green", and "blue". We could assign the numerical values 0, 1, and 2 to these labels, respectively, using label encoding.

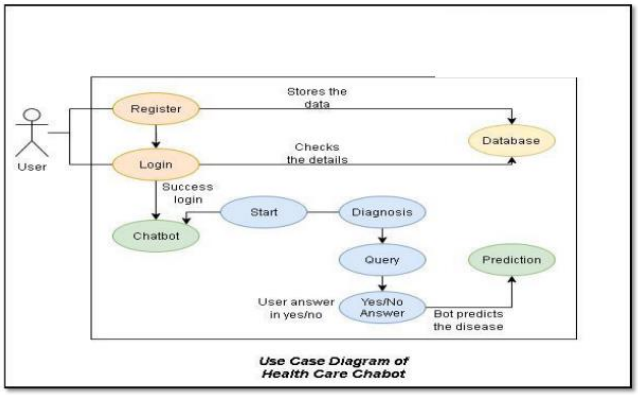
Label encoding is useful when working with machine learning algorithms that cannot handle categorical data directly. Many machine learning algorithms require input data to be in the form of numerical values, so label encoding can be used to convert categorical data into a format that can be used by these algorithms.

However, it is important to note that label encoding can introduce a false sense of ordering or hierarchy in categorical data, which can be misleading. For example, if we label encode clothing sizes as "small", "medium", and "large" as 0, 1, and 2, respectively, the resulting numerical values may suggest an ordering or hierarchy that does not actually exist. In such cases, one should use one-hot encoding or other techniques that preserve the categorical nature of the data.



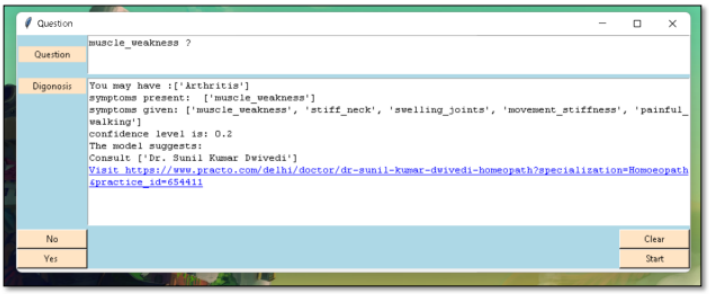
Flowchart of Health Care Chat Bot





RESULTS

Thus, we can conclude that this system giving the accurate result. As we are using large data set which will ensures the better performance. Thus we build up a system which is useful for people to detect the disease by typing symptoms



CONCLUSION AND FUTURE WORK

Chat bots are a thing of the future which is yet to uncover its potential but with its rising popularity and craze among companies, they are bound to stay here for long. Machine learning has changed the way companies were communicating with their customers. With new platforms to build various types of chat bots being introduced, it is of great excitement to witness the growth of a new domain in technology while surpassing the previous threshold.

REFERENCES

[1] Hiba Hussain1, Komal Aswani2, Mahima Gupta3, Dr. G.T.Thampi4,"Implementation of Disease Prediction Chatbot and Report Analyzer using the Concepts of NLP, Machine Learning and OCR,"IRJET,Apr 2020. [2] Oh, K.-J., D. Lee, B. Ko, and H.-J. Choi, A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation. In 2017 18th IEEE International Conference on Mobile Data Management (MDM). IEEE, 2017. [3] Kowatsch, T., M. Nißen, C.-H. I. Shih, D. Rüegger, D. Volland, A. Filler, F. Künzler, F. Barata, D. Büchter, B. Brogle, et al. (2017). Text-based healthcare chatbots supporting patient and health professional teams: preliminary results of a randomized controlled trial on childhood obesity. [4] Lin Ni(B), Chenhao Lu, Niu Liu, and Jiamou Liu," MANDY: Towards a Smart Primary Care Chatbot Application", SPRINGER,2017. [5] Divya, S., V. Indumathi, S. Ishwarya, M. Priyasankari, and S. K. Devi (2018). A self-diagnosis medical chatbot using artificial intelligence. Journal of Web Development and Web Designing, 3(1), 1–7. [6] Chung, K. and R. C. Park (2019). Chatbot-based heathcare service with a knowledge base for cloud computing. Cluster Computing, 22(1), 1925–1937. [7] Ahmed Fadil, Gianluca Schiavo,"Design for healthcare chatbot"Arxiv, 2019. [8] Beaudry, J., A. Consigli, C. Clark, and K. J. Robinson (2019). Getting ready for adult healthcare: Designing a chatbot to coach adolescents with special health needs through the transitions of care. Journal of pediatric nursing, 49, 85–91. [9] Kavitha, B. and C. R. Murthy (2019). Chatbot for healthcare system using artificial intelligence. [10] Kandpal, P., K. Jasnani, R. Raut, and S. Bhorge, Contextual chatbot for healthcare purposes (using deep learning). In 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4). IEEE, 2020. [11] A. F. Ur Rahman Khilji, S. R. Laskar, P. Pakray, R. A. Kadir, M. S. Lydia and S. Bandy-opadhyay, "HealFavor: Dataset and A Prototype System for Healthcare ChatBot," 2020 International Conference on Data Science, Artificial Intelligence, and Business Analytics (DATABIA), 2020.