

DocBot - HEALTHCARE CHATBOT

A Project Work Synopsis

Submitted in the partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING IN COMPUTER SCIENCE WITH SPECIALIZATION IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Submitted by:

Divyansh Dodan (19BCS6106)

Depanshi Tomar (19BCS6095)

Sahil Thakur (19bcs6103)

Shivam (19BCS6129)

Under the Supervision of:

Ms. Neha Sharma (E12652)



**CHANDIGARH
UNIVERSITY**
Discover. Learn. Empower.

CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413,

PUNJAB

September, 2022

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 healthcare 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-bots in the health care sector can play an important role. 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-bots with different intelligence levels can understand the questions from the users and provide answers based on pre - defined labels in the training data.

Table of Contents

Title Page	i
Abstract	ii
1. Introduction	
1.1 Problem Definition	
1.2 Project Overview	
1.3 Hardware Specification	
1.4 Software Specification	
2. Literature Survey	
2.1 Existing System	
2.2 Proposed System	
2.3 Literature Review Summary	
3. Problem Formulation	
4. Research Objective	
5. Methodologies	
6. Experimental Setup	
7. Conclusion	
8. Tentative Chapter Plan for the proposed work	
9. Reference	

1. 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.

1.1 Problem Definition

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 centres 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.

1.2 Problem Overview

The main purpose of the scheme is to build the language gap between the user and health providers by giving immediate replies to the Questions asked by the user. Today's people are more likely addicted to the internet but they are not concerned about their personal health. They avoid going to hospital for small problems which may become a major disease in future.

Establishing question answer forums is becoming a simple way to answer those queries rather than browsing through the list of potentially relevant documents from the web. Many of the existing systems have some limitations such as there is no instant response given to the patients they have to wait for experts to acknowledge for a long time. Some of the processes may charge an amount to perform live chat or telephony communication with doctors online. The aim of this system is to replicate a person's discussion.

1.3 Hardware Specification

In recent years, a great variety of hardware solutions for real-time TSR has been proposed. These include conventional (general purpose) computers, custom ASIC (application-specific integrated circuit) chips, field programmable gate arrays (FPGAs), digital sign processors (DSPs) and also graphic processing units

1.4 Software Specification

In a software-based solution running on a Linux or window system with a 2.4-GHz dual core CPU is presented

2. LITERATURE SURVEY

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.

2.3 Literature Review Summary (Minimum 7 articles should refer)

Year and Citation	Article/ Author	Tools/ Software	Technique	Source	Evaluation Parameter
Apr 2020	Hiba Hussain ¹ , Komal Aswani ² , Mahima Gupta ³ , Dr. G.T.Thampi ⁴	-	Machine Learning and OCR	Research Paper	emotions recognition methods and chat assistant platform
2017	, K.-J., D. Lee, B. Ko, and H.-J. Choi	-	ML and NLP	Research Paper	psychiatric counseling in mental healthcare service based on emotional

					dialogue analysis and sentence generatio n
2017	Lin Ni(B), Chenhao Lu, Niu Liu, and Jiamou Liu	-	ML and NLP	Research Paper	MANDY: Towards a Smart Primary Care Chatbot Applicati on

3. PROBLEM FORMULATION

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 centres 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

4. OBJECTIVES

Human Healthcare is one of the most significant subjects of society. The identification of the nature of the illness or other problems by examination of the symptoms i.e. diagnosis always stands first in

overall curing procedure of disease. Thus , we can say that diagnosis and prediction of disease are most crucial aspects to be considered before thinking about exact procedure of curing the disease. This diagnosis procedure requires a lot of time as well as a lot of money. As a result , people belonging to poor financial background are not able to get accurate diagnose of disease which at the end may create life or death situation. Our objective is to provide a simple and accurate health diagnose for those people who have certain symptoms and don't want to go to a clinic so just they use our chatbot and get a instant and accurate diagnose with possible suggestions and cure if available.

5. METHODOLOGY

In machine learning, **support-vector machines (SVMs, also support-vector networks)** are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier (although methods such as Platt scaling exist to use SVM in a probabilistic classification setting). An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on the side of the gap on which they fall.

In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature spaces.

Decision Tree algorithm belongs to the family of supervised learning algorithms. They can be used to solve both regression and classification problems. Decision trees use the tree representation to solve the problem in which each leaf node corresponds to a class label and attributes are represented on the internal node of the tree. If the value is less than threshold, then go to the left child node. If the value is more than threshold, then go to the right child node. Then as we have got the disease, now we need to append the details of the disease and necessary precautions need to be taken. If user is having that particular symptom for more than 13 days then, chatbot will respond as "You should take the consultation from doctor". If it is less than 13 days then chat bot will respond as, "It might not be that bad but you should take precautions". After predicting the disease, the bot will give necessary precautions that one must take. After predicting the disease, the bot will give a basic description about the disease, as the user will get an idea of what disease that user might be facing. Bot will respond to everything in voice, for that we have used pyttsx3. pyttsx3 is a text-to-speech conversion library in Python. Unlike alternative libraries, it works offline and is compatible with both Python 2 and 3. The proposed Chatbot system functions based on a Decision Tree algorithm.

6.EXPERIMENTAL SETUP

The main purpose of the scheme is to build the language gap between the user and health providers by giving immediate replies to the Questions asked by the user. Today's people are more likely addicted to the internet but they are not concerned about their personal health. They avoid going to hospital for small problems which may become a major disease in future. Establishing question answer forums is becoming a simple way to answer those queries rather than browsing through the list of potentially relevant documents from the web. Many of the existing systems have some limitations such as there is no instant response given to the patients they have to wait for experts to acknowledge for a long time. Some of the processes may charge an amount to perform live chat or telephony communication with doctors online. The aim of this system is to replicate a person's discussion.\

7.CONCLUSION

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

8. TENTATIVE CHAPTER PLAN FOR THE PROPOSED WORK

CHAPTER 1: 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.

CHAPTER 2: LITERATURE REVIEW

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.

CHAPTER 3: OBJECTIVE

Human Healthcare is one of the most significant subjects of society. The identification of the nature of the illness or other problems by examination of the symptoms i.e. diagnosis always stands first in overall curing procedure of disease. Thus, we can say that diagnosis and prediction of disease are most crucial aspects to be considered before thinking about exact procedure of curing the disease. This diagnosis procedure requires a lot of time as well as a lot of money. As a result, people belonging to poor financial background are not able to get accurate diagnosis of disease which at the end may create life or death situation. Our objective is to provide a simple and accurate health diagnose for those people who have certain symptoms and don't want to go to a clinic so just they use our chatbot and get a instant and accurate diagnose with possible suggestions and cure if available.

CHAPTER 4: METHODOLOGIES

In machine learning, **support-vector machines (SVMs, also support-vector networks)** are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier (although methods such as Platt scaling exist to use SVM in a probabilistic classification setting). An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on the side of the gap on which they fall.

In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-

dimensional feature spaces.

Decision Tree algorithm belongs to the family of supervised learning algorithms. They can be used to solve both regression and classification problems. Decision trees use the tree representation to solve the problem in which each leaf node corresponds to a class label and attributes are represented on the internal node of the tree. If the value is less than threshold, then go to the left child node. If the value is more than threshold, then go to the right child node. Then as we have got the disease, now we need to append the details of the disease and necessary precautions need to be taken. If user is having that particular symptom for more than 13 days then, chatbot will respond as "You should take the consultation from doctor". If it is less than 13 days then chat bot will respond as, "It might not be that bad but you should take precautions". After predicting the disease, the bot will give necessary precautions that one must take. After predicting the disease, the bot will give a basic description about the disease, as the user will get an idea of what disease that user might be facing. Bot will respond to everything in voice, for that we have used pyttsx3. pyttsx3 is a text-to-speech conversion library in Python. Unlike alternative libraries, it works offline and is compatible with both Python 2 and 3. The proposed Chatbot system functions based on a Decision Tree algorithm.

CHAPTER 5: EXPERIMENTAL SETUP

The main purpose of the scheme is to build the language gap between the user and health providers by giving immediate replies to the Questions asked by the user. Today's people are more likely addicted to the internet but they are not concerned about their personal health. They avoid going to hospital for small problems which may become a major disease in future. Establishing question answer forums is becoming a simple way to answer those queries rather than browsing through the list of potentially relevant documents from the web. Many of the existing systems have some limitations such as there is no instant response given to the patients they have to wait for experts to acknowledge for a long time. Some of the processes may charge an amount to perform live chat or telephony communication with doctors online. The aim of this system is to replicate a person's discussion.

CHAPTER 6: CONCLUSION AND FUTURE SCOPE

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

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 healthcare 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. Bandyopadhyay, "HealFavor: Dataset and A Prototype System for Healthcare ChatBot," 2020 International Conference on Data Science, Artificial Intelligence, and Business Analytics (DATABIA), 2020.