BACHELOR OF TECHNOLOGY (B.TECH)

MINOR PROJECT REPORT

INTELLIGENCE CONVERSATIONAL HEALTHCARE CHATBOT

SUBMITTED BY

YASH SHARMA
OUI118BCS017
2018-2022
COMPUTER SCIENCE & ENGINEERING (Modular)

SUPERVISED BY

Miss. OJASVEE KANERIA

ASSISTANT PROFESSOR

Department of Computer Science & Engineering



ORIENTAL UNIVERSITY, INDORE MADHYA PRADESH, INDIA`

DECLARATION

I declare that this written submission represents my idea in my own words and where

others' ideas or words have been included, i have adequately cited and referenced the

original sources. I also declare that i have adhered to all principles of academic

honesty and integrity and have not misrepresented or fabricated or falsified any

idea/data/fact/source in my submission. I understand that any violation of the above

will be cause for disciplinary action by the Institute and can also evoke penal action

from the sources which have not been properly cited or from whom proper

permission has not been taken when needed.

Name of Student and Roll number: Yash Sharma, (OUI118BCS017)

Signature

Date: 1st June 2021

I

ACKNOWLEDGEMENT

No project is ever complete without the guidance of those expert who have already traded this past before and hence become master of it and as a result, our leader. So, we would like to take this opportunity to take all those individuals who have helped me in visualizing this project.

I would take this opportunity to thank my project coordinator **Prof. Ojasvee Kaneria** and other faculty members for their guidance in selecting this project and also for providing timely assistance to my query and guidance of this project.

I am very grateful to my Head of the Department **Dr. Dinesh Jain** for extending his help directly and indirectly through various channels in my project work.

I am really thankful to all my Professors from Department Of Computer Science and Engineering, Oriental University Indore for their valuable advices during the designing of the project. Their contributions have been valuable in so many ways that i find it difficult to acknowledge all of them individually.

Thanking You.



CERTIFICATE

This is to certify that the project report entitled "Intelligence Conversational Healthcare Chatbot" submitted to the Department of Computer Science and Engineering, Oriental University Indore (M.P.), in partial fulfillment for the award of the degree of Bachelor of Technology in Computer Science and Engineering, is a record of bona fide work carried out by Mr. Yash Sharma, Roll No. OUI118BCS017, under my supervision and guidance. All help received by him from various sources have been duly acknowledged. No part of this report has been submitted elsewhere for award of any other degree.

Ms. Ojasvee Kaneria (Assistant Professor) Supervisor

Place: Indore
Date: 1st June

Index

Sr.No.	Topic	Page No.
	Declaration	I
	Acknowledgement	II
	Certificate	III
	Index	1-3
	List of Figures	4
	List of Tables	5
	Abstract	6
1	Chapter 1 - Introduction	7-8
	1.1 Introduction	7
	1.2 Purpose and Scope	7
	1.3 Problem Statement	8
2	Chapter 2 – Project Analysis	9-10
	2.1 Review of Literature	9
	2.2 Project Timeline	9
	2.3 Dataset details	10
	2.5 Methodology Used	10

3	Chapter 3 - Project Design	11-14
	3.1 Block Diagram	.11
	3.2 Data Flow Diagram	12
	3.3 Use Case Diagram	13
	3.4 Sequence Diagram	14
4	Chapter 4 - Implementation	15-22
	4.1 Project Implementation Technology	15
	4.1.1 Hardware Requirement	15
	4.1.2 Software Requirement	
	4.2 Experimental Setup	16
	4.3 Coding (Max 5-7 pages)	17-21
	4.4 Testing	22
5	Chapter 5 - Result	23-25
	5.1Snapshot of Result	
	• Snapshot	23-24
	 Analysis of result 	25

6	Chapter 6 - Advantage and Disadvantages of Model	
	6.1 Advantages	26
	6.2 Disadvantages	27
7	Chapter 7 - Conclusion & Future Scope	
	7.1 Conclusion	28
	7.2 Future Scope	
8	Chapter 8 - References	29

List of Figure

Figure No.	Name of Figure	Page No.
3.1	Block Diagram	11
3.2	Data Flow Diagram	12
3.3	Use Case Diagram	13
3.4	Sequence Diagram	14

List of Table

Table No.	Name of Table	Page No.
	Project Timeline	1-3

ABSTRACT

Chatbots are intelligent negotiators with which users can hold conversations, usually via text or voice. Over the past few years, chatbots became popular in businesses focused on client service especially in retail banking domains. Despite an increasing interest of chatbots in healthcare, clear information on the way to design them as intelligent conversational healthcare assistants has been scarce. The primary aim of Conversational Healthcare Chatbot is to help you better visualize the presentation of mined data (information). It deals with all the health care issues which may benefit stakeholders who belong to the health care field. Health professionals have limited resources and are not completely able to personally monitor and support patients in their lifestyles during this pandemic time. Against this background and due to the increasing number of self-service channels and digital health interventions, this paper presents a conventional methodology for designing and implementing a conversational healthcare Chatbot as an intelligent assistant especially for those who want to possess contact-free assistance from a healthcare professional. The methodology is constructed upon first-order logic predicates which can be utilized in different commercially available tools and focus on two phases: knowledge abstraction and modeling, and conversation flow. As the main results of this research, we propose mathematical definitions to model conversation elements, reasoning processes, and conflict resolution to formalize the methodology and make it framework-independent.

Introduction

1.1 Introduction:

Chatbots are the substitute of humans, which communicate using text or voice medium and users obtain reply through AI enabled interface where user can have the sense of interacting with human. Chatbots are utilized in applications like ecommerce customer service, call centers and Internet gaming.

Bots are often programmed to reply an equivalent way whenever, to reply 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 hospitals, helping them discover specialists, booking their appointments, and getting them access to the right treatment. An ML(machine learning) model has got to be created wherein we could give any text input and on the ground 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 analyze outputs of numerous models and suggest the most effective model which will 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 so as to learn more from the training set. This healthcare chat bot system will help healthcare system to provide 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 trying to find.

1.2 Purpose and Scope:

Almost everyone kept on hold while operators connect you to a customer care executive. On an average people spend around 7 minutes until they are assigned to a person. Gone are the frustrating days of waiting in a queue for the next available operative. They are replacing live chat and other forms of slowercontact methods such as emails and phone calls. Since chat bots are basically virtual robots they never gettired and continue to obey your command. They will continue to operate every day throughout the year without requiring taking a break.

1.3 Problem Statement:

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 ecommerce 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 itmust 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.

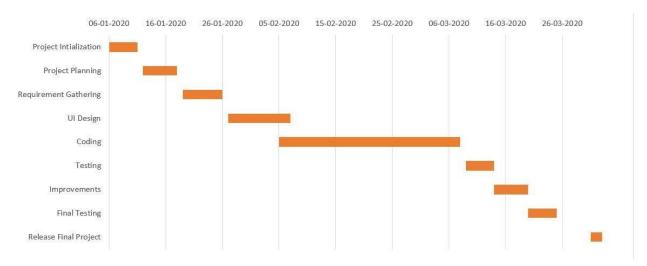
In any case, the utilization of artificial intelligence in an industry where individuals' lives could be in question still starts misgivings in individuals. It brings up issues about whether the task mentioned above ought to be assigned to human staff. This healthcare chat bot system will help hospitals 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.

Project Analysis

2.1 Review of Literature:

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 health. They avoid going to the hospital for small problems which may become a major disease in the 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. This system aims to replicate a person's discussion.

2.2 Project Timeline:



2.3 Dataset Details:

Dataset contains description of different types of diseases. There are different sets of differenttypes of diseases. These sets consist of descriptions of a single disease with different doctors, hospitals, etc.

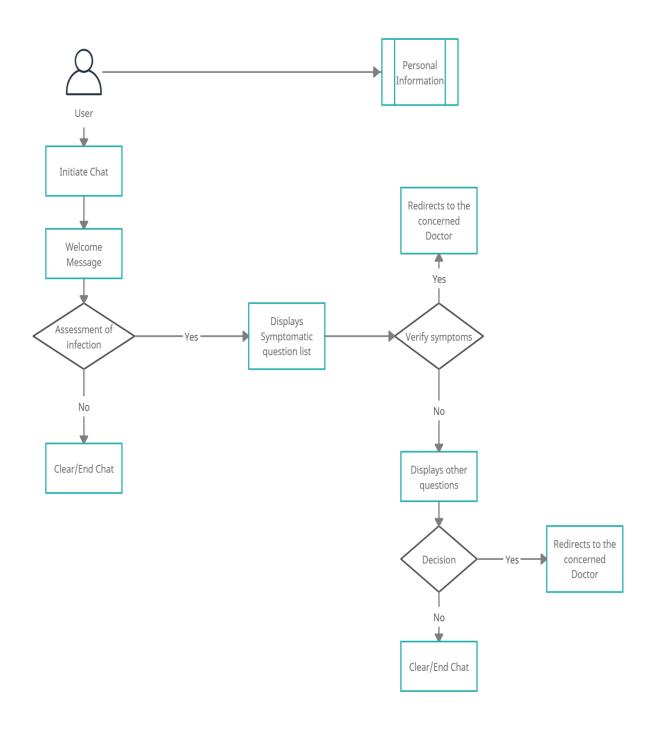
A dataset has been created by recording sequences from over 133 numbers of diseases anddoctors and hospitals.

2.4 Methodology Used:

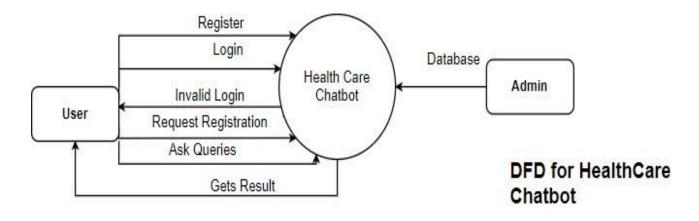
The Health-Care Chat Bot System should be written in Python, GUI links, and a simple, accessible network API. The system must provide a capacity for parallel operation and system design should not introduce scalability issues about the number of surface computers, tablets, or displays connected at any one time. The end system should also allow for seamless recovery, without data loss, from individual device failure. There must be a strong audit chain with all system actions logged. While interfaces are worth noting that this system is likely to conform to what is available. With that in mind, the most adaptable and portable technologies should be used for the implementation. The system has criticality in so far as it is a live system. If the system is down, then customers must not notice or notice that the system recovers quickly (seconds). The system must be reliable enough to run; crash and glitch-free more or less indefinitely, or facilitate error recovery strong enough such that glitches are never revealed to its end-users.

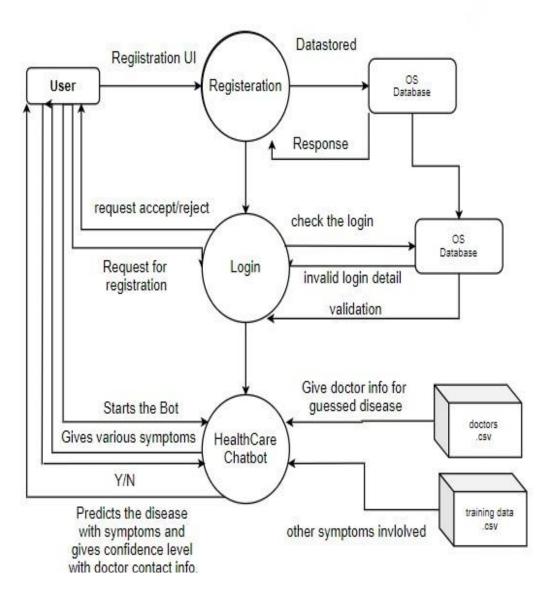
Project Design

3.1 Block Diagram:

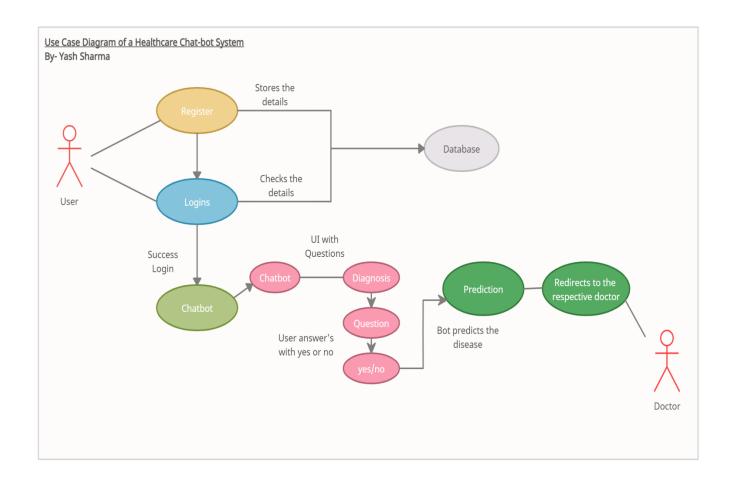


3.2 Data Flow Diagram:

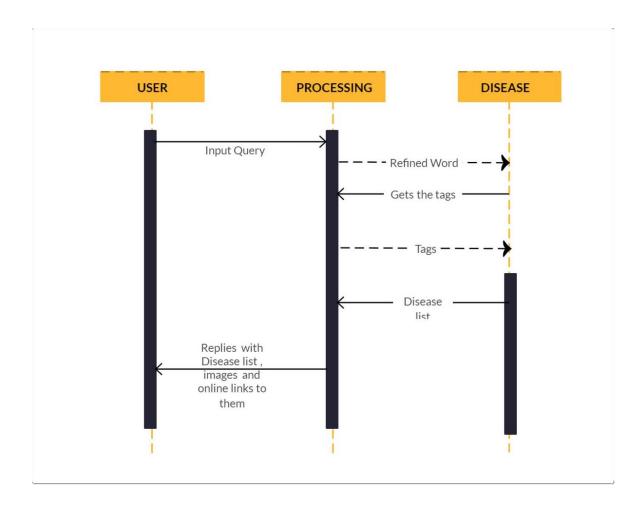




3.3 Use Case Diagram:



3.4 Sequence Diagram:



Implementation

4.1 Project Implementation Technology:

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 wideas 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.

4.1.1 Hardware Requirement:

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

4.1.2 Software Requirements:

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

4.2 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.

4.3 Coding:

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
# Importing the dataset
training_dataset = pd.read_csv('Training.csv')
test_dataset = pd.read_csv('Testing.csv')
X = training_dataset.iloc[:, 0:132].values
#print(X)
y = training_dataset.iloc[:, -1].values
#print(y)
dimensionality_reduction = training_dataset.groupby(training_dataset['prognosis']).max()
#print(dimensionality_reduction)
from sklearn.preprocessing import LabelEncoder
labelencoder = LabelEncoder()
y = labelencoder.fit_transform(y)
#print(y)
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.25, random_state = 0)
from sklearn.tree import DecisionTreeClassifier
```

```
classifier = DecisionTreeClassifier()
classifier.fit(X_train, y_train)
         = training_dataset.columns
  cols
         = cols[:-1]
  cols
  importances = classifier.feature_importances_
  indices
                   np.argsort(importances)[::-1]
  features = cols
  from sklearn.tree import _tree
  def execute_bot():
    print("Please reply with yes/Yes or no/No for the following symptoms")
    def print_disease(node):
       #print(node)
       node = node[0]
       #print(len(node))
       val = node.nonzero()
       #print(val)
       disease = labelencoder.inverse_transform(val[0])
       return disease
    def tree_to_code(tree, feature_names):
       tree_ = tree.tree_
       #print(tree_)
       feature_name = [
         feature_names[i] if i != _tree.TREE_UNDEFINED else "undefined!"
         for i in tree_.feature
```

```
]
#print("def tree({}):".format(", ".join(feature_names)))
symptoms_present = []
def recurse(node, depth):indent
  = " " * depth
  if tree_.feature[node] != _tree.TREE_UNDEFINED:name =
     feature_name[node]
     threshold = tree_.threshold[node]
     print(name + "?")
     ans = input()
     ans = ans.lower()if ans
     == 'yes':
       val = 1else:
       val = 0
    if val <= threshold: recurse(tree_.children_left[node],
       depth + 1
     else:
       symptoms_present.append(name)
       recurse(tree .children right[node], depth + 1)
  else:
     present_disease = print_disease(tree_.value[node])print(
    "You may have " + present disease ) print()
     red_cols = dimensionality_reduction.columns
         symptoms_given =
red_cols[dimensionality_reduction.loc[present_disease].values[0].nonzero()]
    print("symptoms present " + str(list(symptoms_present)))
```

```
print()
print("symptoms given " + str(list(symptoms_given)) )print()
confidence_level = (1.0*len(symptoms_present))/len(symptoms_given)
print("confidence level is " + str(confidence_level))
print()
print('The model suggests:')print()
row = doctors[doctors['disease'] == present_disease[0]]print('Consult',
str(row['name'].values))
print()
print('Visit', str(row['link'].values))
#print(present_disease[0])
recurse(0, 1)
tree_to_code(classifier,cols)
doc_dataset = pd.read_csv('doctors_dataset.csv', names = ['Name', 'Description'])
diseases = dimensionality_reduction.index
diseases = pd.DataFrame(diseases)
doctors = pd.DataFrame()
doctors['name'] = np.nan
doctors['link'] = np.nan
```

```
doctors['disease'] = np.nan
doctors['disease'] = diseases['prognosis']
doctors['name'] = doc_dataset['Name']
doctors['link'] = doc_dataset['Description']
record = doctors[doctors['disease'] == 'AIDS']
record['name']
record['link']
execute_bot()
```

4.4 Testing:

Without a well-thought testing effort, the project will undoubtedly fail overall and will impact the entire operational performance of the solution. With a poorly tested solution, the support and maintenance cost will escalate exponentially, and the reliability of the solution will be poor.

Therefore, project managers need to realize that the testing effort is a necessity, not merely as an ad hoc task that is the last hurdle before deployment.

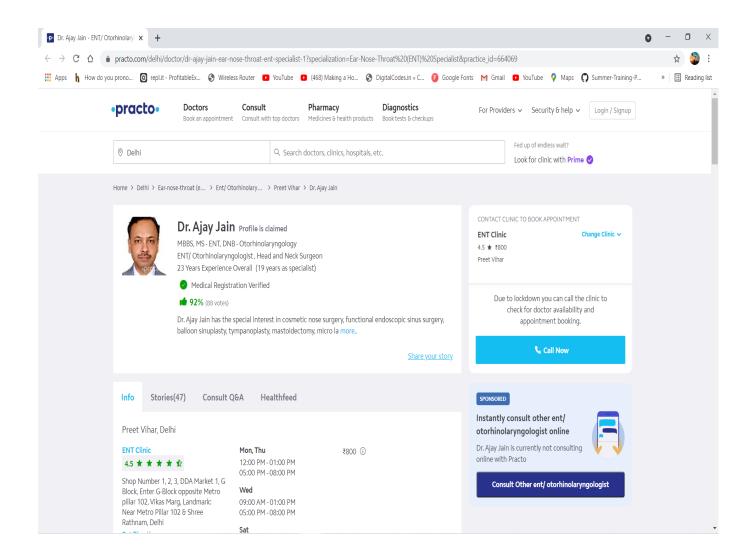
The project manager should pay specific attention to developing a complete testing plan and schedule. At this stage, the project manager should have realized that this effort would have to be accommodated within the project budget, as many of the testing resources will be designing, testing, and validating the solution throughout the entire project life cycle—and this consumes work-hours and resources.

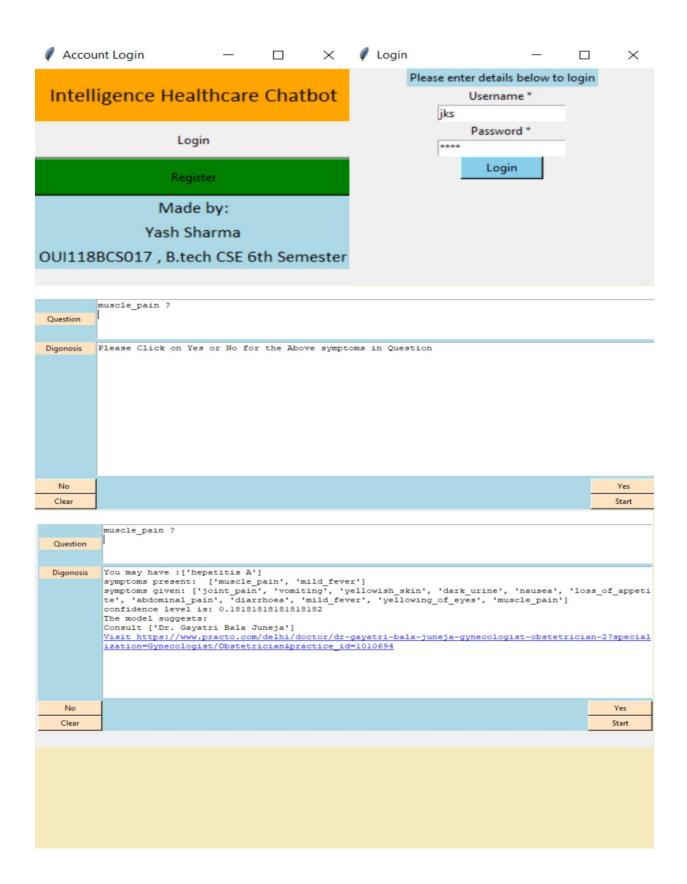
The testing effort begins at the initial project phase (i.e. preparing test plans) and continues throughout until the closure phase.

Result

5.1 Snapshot of Result:

- Snapshot
- Analysis of Result





• Analysis of Result:

A wide range of positive and negative perspectives was reported on the utilization of health care chatbots, including the importance to patients for managing their health and also the benefits on physical, psychological, and behavioral health outcomes. More consistent agreement occurred about administrative benefits associated with chatbots; many physicians believed that chatbots would be most beneficial for scheduling doctor appointments (78%, 78/100), locating health clinics (76%, 76/100), or providing medication information (71%, 71/100). Conversely, many physicians believed that chatbots cannot effectively care for all of the patients' needs (76%, 76/100), cannot display human emotion (72%, 72/100), and cannot provide detailed diagnosis and treatment because of not knowing all of the personal factors associated with the patient (71%, 71/100). Many physicians also stated that health care chatbots could be a risk to patients if they self-diagnose too often (714%, 74/100) and do not accurately understand the diagnoses (74%, 74/100).

Advantage and Disadvantages of Model

6.1 Advantages:

1. Omni-capable

 The chat bot converses seamlessly across multiple digital channels and retains data and and and and and and and are seamless experience. In best cases, even passing that information to a liveagent if needed.

2. Free to Explore

• The chat bot can reach, consume, and process vast amounts of data—both structured and unstructured—to surface insights from any source—to gather relevant data to solve customer issues quickly.

3. Autonomous Reasoning

• The chat bot can perform complex reasoning without human intervention. For example, a great Service chat bot should be able to infer solutions based on relevant case histories.

4. Pre-Trained

• The chat bot is pre-trained to understand brand-specific or industry-specific knowledge and terms. Even better, it's pre-configured to resolve common customer requests of a particular industry.

5. Register/Log-in

• To access this chat bot and individual needs to register and then use the registration ID tolog in to access the features.

6. User Interface

• A user friendly interface which is engaging and easy to access.

6.2 Disadvantages:

- Complex Interface Chatbots are often seen to be complicated and require a lot of time to understand user's requirement. It is also the poor processing which is not able to filter results in time that can annoy people.
- Inability to Understand Due to fixed programs, chatbots can be stuck if an unsaved query is presented in front of them. This can lead to customer dissatisfaction and result in loss. It is also the multiple messaging that can be taxing for users and deteriorate the overall experience on the website.
- Time-Consuming Chatbots are installed with the motive to speed-up the response and improve customer interaction. However, due to limited data-availability and time required for self-updating, this process appears more time-taking and expensive. Therefore, in place of attending several customers at a time, Chatbots appear confused about how to communicate with people.
- Zero decision-making Chat bots are known for being infamous because of their
 inability to make decisions. A similar situation has landed big companies like Microsoft
 etc. in trouble when their chat bot went on making a racist rant. Therefore, it is critical to
 ensure proper programing of your chat bot to prevent any such incident which can
 hamper your brand.
- Poor Memory Chat bots are not able to memorize the past conversation which forces the user to type the same thing again & again. This can be cumbersome for the customer and annoy them because of the effort required. Thus, it is important to be careful while designing chatbots and make sure that the program is able to comprehenduser queries and respond accordingly.

Conclusion & Future Scope

7.1 Conclusion:

Thus, we can conclude that this system is giving accurate results. As I am using a large dataset which will ensure better performance. I build up this system which is useful for people to detect the disease by looking at the symptoms and clicking a yes/no button. This system will not harm any patient as it will first show the question list to the user and as per the question selected by the user, it will just detect the symptoms and refer a doctor to the user with a consultation booking link.

7.2 Future Scope:

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

- https://en.wikipedia.org/wiki/Chatbot https://en.wikipedia.org/wiki/Disease
- https://data-flair.training/blogs/python-chatbot-project/
- https://www.jmir.org/2019/4/e12887/
- https://www.youtube.com/playlist?list=PLQVvvaa0QuDdc2k5dwtDTyT9aCja0on8j