HEALTHCARE-CHATBOT TO RECOGNIZE DISEASE USING SVM AND DT ALGORITHMS

A PROJECT REPORT

submitted by

SUGANYA S   210701270

SYLVIA V   210701279

TAEJASHWAR RB   210701280

in partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

in

COMPUTER SCIENCE AND ENGINEERING

RAJALAKSHMI ENGINEERING COLLEGE ANNA UNIVERSITY,

CHENNAI

MAY 2024

RAJALAKSHMI ENGINEERING COLLEGE, CHENNAI

BONAFIDE CERTIFICATE

Certified that this Project report titled "**HEALTHCARE-CHATBOT TO RECOGNIZE DISEASE USING SVM AND DT ALGORITHMS**" is the bonafide work of "**SUGANYA S - 2116210701270,SYLVIA V - 2116210701279,TAEJASHWAR RB -2116210701280**" who carried out the work under my supervision. Certified further that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE

Dr.V.KARTHICK,M.E,Ph.D

Associate Professor

Department of Computer Science

Rajalakshmi Engineering College

Chennai-602105.

Submitted to Project Viva-Voce Examination held on _____

Internal Examiner   **External Examiner**

ABSTRACT

The healthcare chatbot is engineered to aid users in diagnosing diseases by analyzing their symptoms and the duration of these symptoms. This system enhances accessibility, providing quicker and more informed healthcare decisions. It employs two sophisticated models: Decision Tree Classifier and Support Vector Machine (SVM). These models predict the most probable medical condition based on user-reported symptoms.The chatbot's efficiency stems from its ability to assess the severity of symptoms and retrieve detailed symptom descriptions. This capability enables the bot to suggest appropriate precautionary measures, which are derived from a structured dataset. The interaction with patients is significantly improved, as the bot collects symptoms through user interaction, utilizing regex-based pattern matching to ensure input accuracy.The decision-making process of the chatbot involves evaluating the severity and duration of the user's condition. Based on this assessment, it recommends whether the user should seek medical consultation or adopt preventive measures. This approach not only fosters patient involvement but also guides users towards timely and appropriate healthcare actions. By effectively integrating advanced machine learning models and user-friendly interfaces, the healthcare chatbot stands as a pivotal tool in modern digital healthcare, aiming to streamline the initial diagnostic process and enhance overall patient outcomes.

ACKNOWLEDGMENT

We thank the almighty god for the successful completion of the project. Our sincere thanks to our chairman **Mr. S. MEGANATHAN,B.E., F.I.E.,** for his sincere endeavor in educating us in his premier institution. We would like to express our deep gratitude and sincere thanks to our beloved Chairperson **Dr.(Mrs).THANGAM MEGANATHAN,Ph.D.,**for her enthusiastic motivation which inspired us a lot in completing this project and we extend our gratitude to our Vice Chairman **Mr.ABHAY SHANKAR MEGANATHAN,B.E., M.S.,** for providing us with the requisite infrastructure.We also express our sincere and warmful gratitude to our college Principal **Dr. S.N.MURUGESAN,M.E.,Ph.D.**,and **Dr. P. KUMAR,M.E.,Ph.D.**,**Director of Computing and Information Science** and **Head Of the Department of Computer Science and Engineering** and **Project coordinator Dr.V.KARTHICK,M.E.,Ph.D.**,for his encouragement and guiding us throughout project towards successful completion of project and to our parents, friends, all faculty members,supporting staff for direct and indirect involvement in successful completion of project for their encouragement and support.

SUGANYA S

SYLVIA V

TAEJASHWAR RB

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTIO

N

The employment of healthcare chatbots is seen in different areas contemporary to the healthcare field. Some benefits that is included to diagnose patients faster and triage patients. Therefore, through performing these tasks, chatbots assist in decreasing the load on healthcare personnel, increase involvement of patients and offer them appropriate medical aid on time.This work addresses the conceptualization and deployment of a healthcare chatbot, aimed at aiding clients to self-diagnose through the use of machine learning. The primary goals of the chatbot would be to give people an easy to use and always available option for their first step of an initial assessment of their health status and may benefit in helping identify diseases that are early on undiagnosed.The chatbot utilizes two main machine learning models: Decision Tree Classifier and another classification method, the Support Vector Machine (SVM). These models employ the consideration of symptoms that a user inputs with an aim of identifying the likely outcome. The system consists of several components, namely, the component for the evaluation of a symptom severity level, the component for searching for the description of a given symptom, and the component for recommendations of measures that are to be taken to prevent the development of the identified symptom, all of which involve local dataset.

1.1 PROBLEM STATEMENT

To create a healthcare chatbot that can identify diseases based on users' symptoms. This chatbot will use Support Vector Machine (SVM) and Decision Tree (DT) algorithms to analyze the symptoms and predict possible illnesses. By doing so, it will help users get a quick, initial diagnosis and decide if they need to see a doctor. This tool will improve access to health information and support early detection of diseases.By collecting information from different dataset to train the model to provide efficient outcome.

1.2 SCOPE OF THE WORK

The scope of work for the healthcare chatbot encompasses developing a robust platform capable of handling a diverse range of user inquiries related to health and wellness. This includes designing an intuitive user interface, implementing natural language processing capabilities for understanding and responding to user queries effectively, integrating databases of medical knowledge for accurate information retrieval, and ensuring compliance with relevant privacy and security regulations. Additionally, the scope extends to ongoing maintenance and updates to keep the chatbot aligned with the latest medical advancements and user needs.

1.3 AIM AND OBJECTIVES OF THE PROJECT

The aim of the healthcare chatbot is to provide accessible, accurate, and timely medical information and assistance to users, enhancing healthcare accessibility and promoting proactive health management. Its objectives include offering personalized health advice, symptom assessment, medication reminders, and facilitating communication with healthcare professionals, ultimately empowering individuals to make informed decisions about their health and well-being. The healthcare chatbot aims to revolutionize healthcare accessibility by providing users with accessible, accurate, and timely medical information and assistance. Its objectives include offering personalized health advice, symptom assessment, medication reminders, and facilitating communication with healthcare professionals. By leveraging technology to empower individuals, the chatbot seeks to promote proactive health management, ultimately leading to better-informed decisions about health .

1.4 EXISTING SYSTEM

Machine learning based healthcare chatbots are emerging as great tools to help patients that need constant assistance. Systems such as babylon Health employs NLP to interpret and respond to user inputs. The chatbot processes text inputs from users, recognizing symptoms and providing diagnoses based on medical knowledge base. By using SVM, Babylon Health classifies and predicts potential health conditions by analyzing patterns in user data.Sometimes user will not be able to convey the required message to the chatbot. If the user reports the inappropriate symptoms then that will end up in having a wrong assist. Users interact through a chat interface, describing their symptoms, medical history, and concerns. The system results in less accuracy and prediction due to ineffiency in diagnosis

1.5 PROPOSED SYSTEM:

The proposed system focuses to diagnostic models in a two-part algorithm using Decision Tree Classifier and Support Vector Machine (SVM) for better and more accurate identification of diseases based on the symptoms given by the user. This system also incorporates additional features including, evaluation of the symptoms' severity, descriptive assortment gathering data from various dataset CSV files for suggestions of precautionary measures that would improve the flow and user-friendliness of the system as well as to make it more inclusive. Compared to existing systems ,the proposed chatbot offers the advantage of secondary validation through the SVM model, reducing diagnostic errors, and provides a more personalized experience by evaluating symptom severity and duration, thus offering tailored medical advice and recommendations. Through this approach to implement chatbot, the chatbot model provides to produce more accuracy and suitable results.

CHAPTER 2

LITERATURE SURVEY

1]Machine Learning Based Healthcare Chatbot by <u>Rhythm Goel</u> Health care is very significant for you to live a well and healthy life.Unfortunately many people are not well-informed of all the treatments or symptoms of a certain illness. Young persons are attempting to be admitted to a laborious test in the hospital and it is very difficult to receive a doctor's appointment for ones. health problems, and managing phone calls. Problems will be solved by using Healthcare Chatbot with the help of appropriate guidance regarding healthy living. The idea is to create a medical chatbot using Neural Networks that can provide with the info and diagnose the disease and deliver basic information about the disease and when and where to consult a doctor. The effectiveness of medical chatbots rests on Natural language processing techniques that help users post their concerns about any disease and also their health. Any questions that is related to health care via chatbot can be asked by the user without being present at the clinic or hospital. This will help reduce the cost of healthcare and improve access to medical information through medical chat-bot. The program development plan is to analyze customer feelings.

2]Healthcare Chatbot System: A Survey by <u>Suprabha Swain</u>.A chatbot is a software that produces a computer model of conversation through an auditory or textual method between the user and the computer. Such bots use AI to understand user's questions and automate responses to them, simulating human conversation. Nowadays , every person utilize smartphones in their lives.Simply because, it has made life of a person simple by the touch of the screen, from extracting information from the internet, to providing personal assistance to entertainment, everything happens just by the click of a button. Today, Chatbots are being implemented in many fields such as education, customer relation, health, websites, help etc. for information retrieval, FAQs and help. Health care chatbots most of the time are created for medical screening; this is due to the enhancements in AI and approaches used in data mining. The scope

of this paper is to provide critical review of various research papers published in the domain of chatbots and find out the employed tools, algorithms, software and platforms.

3]A Smart Chatbot Architecture based Machine Learning for Health Care Assistance by Soufyane Ayanouz ,A chatbot or conversational agent is a software that can communicate with a human by using natural languageAmong the topics that play crucial roles in artificial intelligence and natural language processing, the conversation modeling is one of the most important ones. It has always been a big challenge to come up with a good chatbot as we see the evolution of Artificial intelligence. Though, chatbots are capable of accomplishing so many things, the role that is compulsory for them, is to comprehend the Human-to-Machine dialogues and, in return, to give the appropriate reaction. Earlier, basic statistical analyzes or manual scripts and procedures and rules were applied to build the structures of chatbots. End-to-end models and other related models have been replaced by these models in around the year 2015 due to the enhanced learning power.. Especially now, the encoder-decoder recurrent model is dominant in the modeling of conversations. This architecture is borrowed from the neural machine translation scenario where it was proving to be very effective. Many features and variations have been implemented that are discussed below, which have significantly improved the conversational nature of chatbots. For this purpose, extensive literature review has been conducted, which includes surveying recent literature in the field. We reviewed numerous articles from the extant literature published within the last five years focusing on the concept of chatbots. We then showcased other similar works studied on our topic, and the AI components required to develop an ICA as a deep learning model Lastly, we explained afunctional architecture which is the reason why we want to empower an intelligent chatbot for the health care support.

4]An Improved Chatbot for Medical Assistance using Machine Learning by Achuthan. S
In many scenarios, user may not be fully cognizant of all available treatment or signs of sickness. For small difficulties, the user must go straight to the hospital for a checkup, which takes much longer.Moreover, you shall be responsible for making and receiving the phone calls

concerning complaints. Such problem is very tricky to solve medical chatbot was used to rectify such an issue by offering the right information for the case. There is power in the technological advancement of appropriate machine learning algorithms that require interactive health assistance to the public with its significant classification performance. Hence the support vector machine (SVM) algorithm is used to predict the status of the health and by using the Google API we can convert the speech to text and text to voice. The input will be sent to a chatbot which will respond with relevant information and displays it on the standalone app

5]Supervised Machine Learning Chatbots for Perinatal Mental Healthcare by Ruyi Wang

Perinatal mental health (PMH) problems that are a form of mood disorders which occur during Pregnancy and up to two years post partum, thus impacts pregnant women, newborns and families. Such problems may start appearing at any period of maternal women. The nature of PMH diagnosis is primarily behavioral, with less emphasis on subjective observation and symptom self-reports as well as the PMH testing using behavioral scale tests. is an effective technology. With assistance of HRI, it can continuously assess the perinatal women's mental health status and simultaneously gather user health information.The application of human-robot interaction in mental health services has attracted widespread attention. Compared with conventional approaches, robot involvement in mental health interventions can lead to potential barriers for subjects in seeking care for mental health and collects more data and details of patients on patient's mental status to refer the level of mental health of the user; it assists clinicians to diagnose and recommend correct precautionary measures more efficiently and effectively. Effectively in this article, the author suggests development of a perinatal women's mental health chatbot. In this article, authors apportion the 31 characteristics of 223 samples using supervised machine learning in order to train a model that is able to reveal the anxiety, depression as well as hypomania index of perinatal women. On the other hand, psychological tests are written calibrated used in the assessment of the populations hence provided with treatment recommendation to enhance on the user's mental health.

CHAPTER 3

SYSTEM

DESIGN

3.1 GENERAL

In this section, we would like to show how the general outline of how all the components end up working when organized and arranged together. It is further represented in the form of a flow chart below.
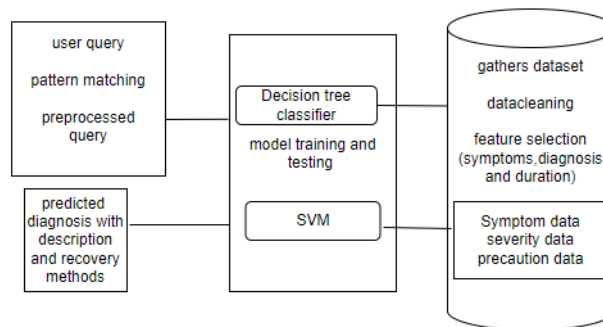
3.2 SYSTEM ARCHITECTURE DIAGRAM

Fig 3.2.1: System Architecture

DATASET

The dataset used for our system gathered from kaggle. It consists of a training set which includes symptom data and corresponding diagnoses to train performance. Each dataset features columns for various symptoms and a diagnosis column. The system utilizes the system description dataset which links symptoms to detailed descriptions and the symptom severity dataset which assigns severity scores to symptoms. The symptom precaution dataset lists precautionary measures for each diagnosis, providing actionable advice to users. These datasets collectively enable comprehensive symptom analysis, disease prediction, and user guidance.

PREPROCESSING OF DATASET:

Before the actual hooping and elimination process, the data are preprocessed to remove missing values or an incompatible format of the data.Next, the LabelEncoder from the sklearn library is utilized to encode the disease labels into numerical format that is easily understandable by the model.It will be further subdivided into training set with which the model is built and the testing set which is used for validation.

MODEL SELECTION:

The suitable machine learning models will be chosen such as Decision Tree Classifiers for primary predictions and Support Vector Machines (SVM) for secondary validations.

TRAINING THE MODEL:

Tune the selected features with the chosen algorithm; the base algorithms to use include decision tree and SVM. Final step in tuning is to reach improved performance of the model. Calculations are made to determine and compare the degree to which the result is accurate. Automate input typification by applying the feed-forward method of regex-based pattern matching.

DEPLOYMENT:

Build a chatbot application with the integration of the trained models to achieve natural human interaction and real-time conversation with the users

3.3 DEVELOPMENTAL ENVIRONMENT

3.2.1 HARDWARE REQUIREMENTS

The hardware requirements may serve as the basis for a contract for the system's implementation. It should therefore be a complete and consistent specification of the entire system. It is generally used by engineers as the starting point for the system design.

PROCESSOR	Intel Core i5
RAM	8 GB RAM
GPU	NVIDIA GeForce GTX 1650
MONITOR	15" COLOR
HARD DISK	512 GB
PROCESSOR SPEED	MINIMUM 1.1 GHz

3.2.2 SOFTWARE REQUIREMENTS

1. Jupyter notebook
2. Visual studio code

CHAPTER 4

PROJECT DESCRIPTION

4.1 MODULE DESCRIPTION

4.4.1 SYMPTOM COLLECTION AND INPUT PROCESSING

The chatbot begins by engaging with users to collect their symptoms and the duration of these symptoms. This interaction is facilitated through a user-friendly interface that employs natural language processing (NLP) techniques to understand and interpret user input. The chatbot utilizes regex-based pattern matching to enhance the accuracy of the input, ensuring that the symptoms are correctly identified and recorded. This module is crucial for gathering precise data, which forms the foundation for subsequent analysis and diagnosis.

4.1.2   Symptom Severity Evaluation

Once the symptoms are collected, the system evaluates their severity. This evaluation is based on predefined criteria sourced from a structured dataset. The severity assessment is essential for determining the urgency of the condition and plays a pivotal role in the chatbot's decision-making process. By accurately assessing symptom severity, the chatbot can prioritize cases that may require immediate medical attention.

4.1.3 Disease Prediction Models: Decision Tree Classifier and SVM

The core of the chatbot's diagnostic capability lies in its use of two machine learning models: the Decision Tree Classifier and the Support Vector Machine (SVM). These models are trained on extensive medical datasets to predict the most probable disease based on the reported symptoms and their duration. The Decision Tree Classifier offers an intuitive and interpretable model, while the SVM provides high accuracy, particularly in cases with clear margin separation. The combination of these models ensures robust and reliable

disease prediction.

4.1.4 Precautionary Measures and Recommendations

Based on the predicted disease, the chatbot suggests appropriate precautionary measures. These   recommendations are sourced from a structured dataset containing medically verified preventive guidelines. This module aims to provide users with actionable steps they can take to manage their symptoms and prevent the condition from worsening. By offering these measures, the chatbot not only helps in diagnosis but also in initial disease management.

CHAPTER-5   IMPLEMENTATION AND RESULTS

5.1 IMPLEMENTATION

The process of diagnosing a disease through a chatbot with the help of one's reported symptoms entails some several processes. Firstly, it was important to decide on some important parameters such as the general type of medical assistance or if the chatbot would be designed for certain category of symptoms.Then, a set of parameters such as symptoms for the disease, the diagnosis, possible severity levels, detailed descriptions, and precautionary measures are obtained. This data is preprocessed by some data transformation activities like formatting of the data, standardization of the symptoms as the features and the diagnosis as the target variable. In the first prediction, the Decision Tree Classifiers would be applied while in the second prediction layer, the Support Vector Machines (SVM) would be obtained.The training data is developed when the preprocessed data is divided into feature and labels to train through the network accurately the symptoms and diagnoses. This symptom data is gathered through chatting with user and exclusion of the user's slang through the usage of the regex-based pattern matching to increase the input accuracy. It then employs the trained Decision Tree model to make future diseases(predictions) and then checks the authenticity of these predictions from the SVM model. Of importance is that the chatbot's decision making process is informed by the severity of reported symptoms and the

prevalence time of the ailment where the user is advised on whether or not to seek medical attendance or take preventive measures.

5.2 OUTPUT SCREENSHOTS

The user will the get the input from the user to predict the health status.It will be done by having some default questions and those questions will get the status of the patient's health condition.

```
Your Name?                              ->sylvia
Hello,  sylvia

Enter the symptom you are experiencing                   ->fever
searches related to input:
0 ) high_fever
1 ) mild_fever
Select the one you meant (0 - 1):  1
Okay. From how many days ? : 2
Are you experiencing any
joint_pain ? : yes
vomiting ? : yes
yellowish_skin ? : no
dark_urine ? : yes
nausea ? : no
loss_of_appetite ? : no
abdominal_pain ? : yes
diarrhoea ? : no
mild_fever ? : yes
yellowing_of_eyes ? : yes
muscle_pain ? : yes
C:\Users\Sylvi\anaconda\Lib\site-packages\sklearn\base.py:464: UserWarning: X does not have valid feature names, but DecisionTr
eeClassifier was fitted with feature names
  warnings.warn(
It might not be that bad but you should take precautions.
You may have  hepatitis A
Hepatitis A is a highly contagious liver infection caused by the hepatitis A virus. The virus is one of several types of hepati
tis viruses that cause inflammation and affect your liver's ability to function.
Take following measures :
1 ) Consult nearest hospital
2 ) wash hands through
3 ) avoid fatty spicy food
4 ) medication
```

```
Arthritis is the swelling and tenderness of one or more of your joints. The main symptoms of arthritis are joint pain and stiff
ness, which typically worsen with age. The most common types of arthritis are osteoarthritis and rheumatoid arthritis.
Peptic ulcer disease (PUD) is a break in the inner lining of the stomach, the first part of the small intestine, or sometimes t
he lower esophagus. An ulcer in the stomach is called a gastric ulcer, while one in the first part of the intestines is a duode
nal ulcer.
Take following measures :
1 ) exercise
2 ) use hot and cold therapy
3 ) try acupuncture
4 ) massage
```

```
Enter the symptom you are experiencing                    ->vomiting
searches related to input:
0 ) vomiting
Okay. From how many days ? : 2
Are you experiencing any
muscle_weakness ? : yes
stiff_neck ? : no
swelling_joints ? : no
movement_stiffness ? : no
painful_walking ? : no
```

```
PS C:\Users\Sylvi\User\healthcare-chatbot-master\healthcare> python chat_bot.py
Decision Tree accuracy: 97.00%
SVM accuracy: 98.00%
Decision Tree cross-validation accuracy: 96.74%
SVM cross-validation accuracy: 96.45%
```

CHAPTER 6

CONCLUSION AND FUTURE ENHANCEMENT

6.1 CONCLUSION

Needless to say, with the recent advancements in chatbot and its accessibilities, health carebot made way for diagnosing the disease by entering the symptoms reported by the user in a more convenient and comfortable way. The dynamic bonding between the user and improves upon the quality of input fed to the algorithm. With the help of Decision Tree Classifier and Support Vector Machine models, the chatbot can diagnose diseases correctly and suggest solutions for those problems to the users. This interaction proves that the chatbot is capable of having a conversation with users about their ailments, correctly interpreting the symptoms provided, and advising the appropriate safety measures thus highlighting the role of chatbots in increasing the availability and effectiveness of healthcare. The good thing about a chatbot is that it requires regular enhancements to know more on the diseases and hence perform proper diagnosis of the disease based on the symptoms presented. If correctly integrated and deployed to healthcare contexts in accordance with the law and ethics, then the chatbot function can become a reliable companion for patients.

6.2 FUTURE ENHANCEMENTS

The chatbots are the better way to interact faster and giving personalized service. Integrating with Electronic Health Records and data sources with real time wearable devices such as heart rate, activity levels and patterns to provide more accurate health decision and alerts based on the current health status.Another way is to incorporate multiple language for easy accessible for all people hereby increasing usability and diagnosing diseases, which enhances the positive impact and usefulness. Integrating with telemedicine services, enable users to schedule virtual consultation with professionals by expanding its functionality and utility. For more , giving healthy tips and notifying different articles and research contents for spreading awareness to prevent diseases.

REFERENCES

1]Wang, Ruyi, Jiankun Wang, Yuan Liao, and Jinyu Wang. "Supervised machine learning chatbots for perinatal mental healthcare." In *2020 International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI)*, pp. 378-383. IEEE, 2020.

2]Shedthi B, Shabari, Vidyasagar Shetty, Rajalaxmi Chadaga, Rashmi Bhat, B. Preethi, and Priyanka Kini K. "Implementation of Chatbot that Predicts an Illness Dynamically using Machine Learning Techniques." International Journal of Engineering 37, no. 2 (2024): 312-322.

3]Shree, Ritu, Ajay Rastogi, and C. Kalaiarasan. "Machine Learning-Driven Cutting-Edge Approach for Designing a Healthcare Chatbot." International Journal of Intelligent Systems and Applications in Engineering 11, no. 8s (2023): 198-205.

4]Goel, Rhythm, Ratnesh Puri Goswami, Somesh Totlani, Parv Arora, Rahul Bansal, and Dinesh Vij. "Machine learning based healthcare chatbot." In 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), pp. 188-192. IEEE, 2022.

5]Solanki, Ram Kumar, Anand Singh Rajawat, Amit R. Gadekar, and Manoj Eknath Patil. "Building a conversational chatbot using machine learning: Towards a more intelligent healthcare application." In *Handbook of research on instructional technologies in health education and allied disciplines*, pp. 285-309. IGI Global, 2023.

6]Swain, Suprabha, Sneha Naik, Apoorva Mhalsekar, Harshada Gaonkar, Deepmala Kale, and Shailendra Aswale. "Healthcare chatbot system: A survey." In *2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)*, pp. 75-80. IEEE, 2022.

7]Naim, Iram, Aanya Raj Singh, Anjali Sen, Anurag Sharma, and Devesh Mishra. "Healthcare CHATBOT for diabetic patients using classification." In *Soft Computing: Theories and Applications: Proceedings of SoCTA 2021*, pp. 427-437. Singapore: Springer Nature Singapore, 2022.

8]Reshma, R. "An improved chatbot for medical assistance using machine learning." In *2022 International Conference on Inventive Computation Technologies (ICICT)*, pp. 70-75. IEEE, 2022.

9]Badlani, Sagar, Tanvi Aditya, Meet Dave, and Sheetal Chaudhari. "Multilingual healthcare chatbot using machine learning." In *2021 2nd International Conference for Emerging Technology (INCET)*, pp. 1-6. IEEE, 2021.

10]Ayanouz, Soufyane, Boudhir Anouar Abdelhakim, and Mohammed Benhmed. "A smart chatbot architecture based NLP and machine learning for health care assistance." In *Proceedings of the 3rd international conference on networking, information systems & security*, pp. 1-6. 2020.References