Healthcare Chatbot System using Python and NLP

Vaibhav Deshwal, Sourabh Kumar

vaibhav.deshwal.cs.2019@miet.ac.in,sourabh.kumar.cs.2019@miet.ac.in
Department of Computer Science and Engineering
DR.JAGBEER SINGH

Abstract:

To sustain a cherished and happy life healthcare plays a very important role in it. However, it is very difficult to consult a doctor if you have any health problems.Chat bots is basically interact using text or voice and toget answersusing artificial intelligence [1]A chat botis a program that is programmedin such a way that it takes some sort of input from the user then it process it using decision tree or some sort of decision making algorithm to give accurate and desired output to the user differently. On the basis of some known keywords, some patterns chat bot identifies the actualproblem and replies on the basis of those the chat bot basically need tohave a idea what the user wants and need and on the basis of thatidea it gives a suitable response with efficiency. The bot connects with potential patients who visit your site, find a specialist, make anappointment, and provide you with the right care.

Keywords:
Data Set,Symptom Description,precaution,severity, Natural Language Processing(NLP), Python using pandas, numpy , Matplotlib, sklearn,SupportVector Machine, Random forest,Decision Tree.

1. INTRODUCTION

Today, health care is very necessary in our life. People today are busy with their jobsand jobs, and they are addicted to the web. They don't care about their health. Therefore,avoid going to the hospital for minor problems. So, we propose to create a health care
chat bot system using python,NLP and machine learning algorithms. identify illnessesand tell us the very subtle information about the illness before we go for checkup with a doctor. This helps patients understand their disease better and improve their health. Users can get all relevant disease information. The system application uses a
chatbot-style question-and-answer protocol to answer user queries. answers to questionsare supported by answers to user questions. Important keywords are taken from sets to
answer these sets. If a match is found, or if a critical answer is given,

or [3] if a similar answer is seen, it determines your type of illness, supports the user's symptoms, and further identifies specific illnesses. Provide details to doctors. By sacrificing this application system, health problems can be reduced. This system is designed to reduce the asking price and the user's time. This is because users cannot goto a doctor or consultant in real time once they need it. For building the healthcare chatbot we use machine learning algorithm, python, NLP and Artificial intelligence to provide 24/7 medical support, suggestion and advices regarding the disease and symptoms of the disease.

2. LITERATURE REVIEW

The following research paper is based on the machine learning algorithm, python sklearn based chat bot and natural language processing for the user and patients for getting precaution and advice for the disease they enter symptoms for. Chat botis basically a program that interact with the user using NLP for text and speech response for the user. Eliza is the first chat bot created by Joseph Weizenbaum. [3] Eliza effect People's tendency to associate associations with terms from previous experiences. Eliza's work is based on: knowledge expression pattern recognition Keyword substitution for known phrases. Some of the chat bots are also based on medical purposes like K-Bots. Given the impact of asthma on overall health care costs, the need for a shift to preventive asthma treatment is widely recognized. The requirements for continuous monitoring of patient compliance with treatment plan, assessment of external factors, and management of asthma can be challenging for traditional clinical settings and tax clinicians.

Dataset
In this paper, [7] Natural Language Processing (NLP) based dataset is used which
consists of many datasets in the form of rows and column and comma separated values.For this project we used two types of datasets one is Data that consists of datasets,
Testing datasets, Training datasets and the other type of dataset that we used is master data that consists of datasets of symptom description, symptom precaution and symptomseverity. In this table one column has the name of Disease like fever, Diabetes, Drug,
Allergy, Jaundice etc. and for this column we have another like column which is relatedto this particular disease.

3. PROPOSED METHODOLOGY

Steps for the proposed methodology

Input data is processed using different ML algorithms based on a large number of datasets of symptoms description, symptom precaution and symptom severity and train and test on a large number of dataset for training and testing for both the known and unknown data.

1.   Input from the user
User enters his name and symptoms as input

2.   Processing the input data using different machine learning algorithms and Natural language processing

3.   Matching of the symptoms to correctly and accurately identify the disease related to symptoms

4.   On the basis of major and minor disease give advice and precaution.
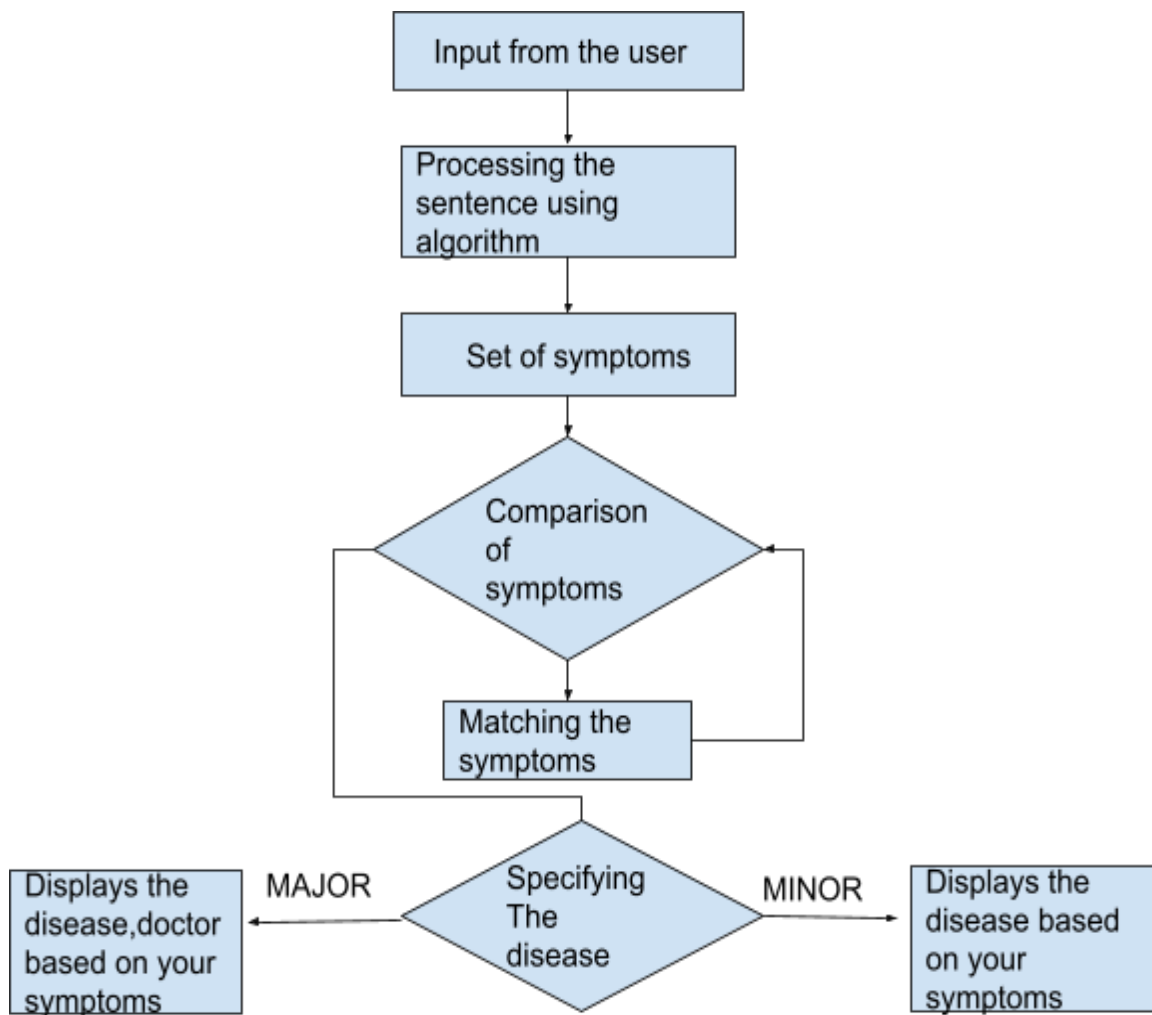
The health chat bot system must be written in Python and run Google's conversational platform, [4] Google Dialogue Flow, GUI hyperlinks, and the easily accessibleCommunity API. The machine should offer the possibility of parallel operation and the layout of the machine so as not to introduce scalability issues regarding the number of connected floor computers, medicines or presentations. The stop machine should also allow for seamless recovery without data loss after a personnel tool failure. You need a robust audit trail that records all machine movements. The interface is really worth mentioning, this machine could match what's available. With this in mind, implementations should use the most adaptable and portable technology possible. Thisis important because the machine is a permanent machine. If a machine is shut down, the client should not notice anything or indicate that the machine is recovering quickly. Machines must be reliable enough to operate more or less continuously and cause crashes and glitches. Or it should make troubleshooting bugs easier.

4. Experimental Results

   After training our model, the model is capable of displaying the Disease, Analgesics, Diet, Doctor details near us based on our symptoms. But we have no proof to examine that our health care chatbot has an increasing number of accuracy than others because wedon't have any parameters to measure that our system has more accuracy than others.

   4.1 Confusion Matrix

It is used to show the performance of a classification model.

HEALTH CHATBOT

ALGORITHMS
RF- Random Forest
DT-Decision Tree
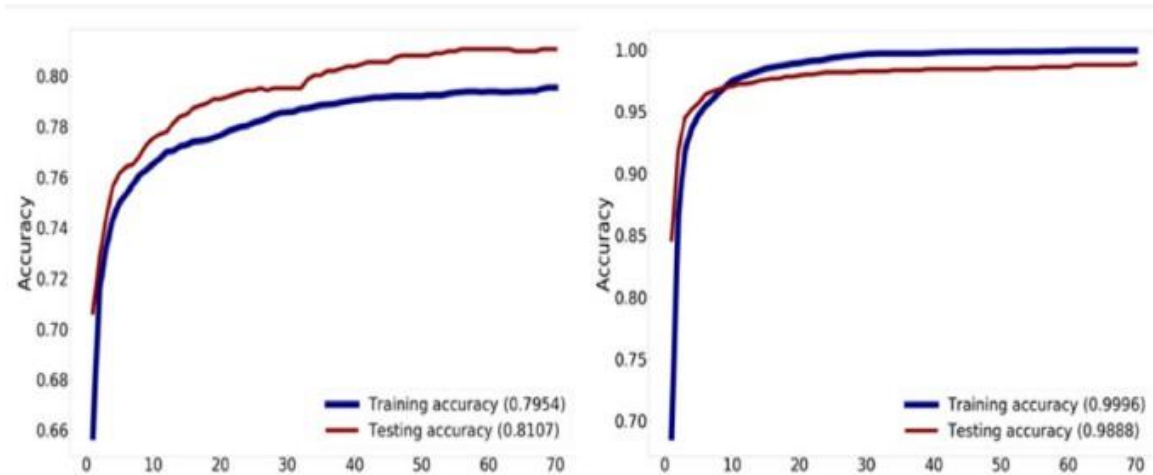NB-Naïve Bayes
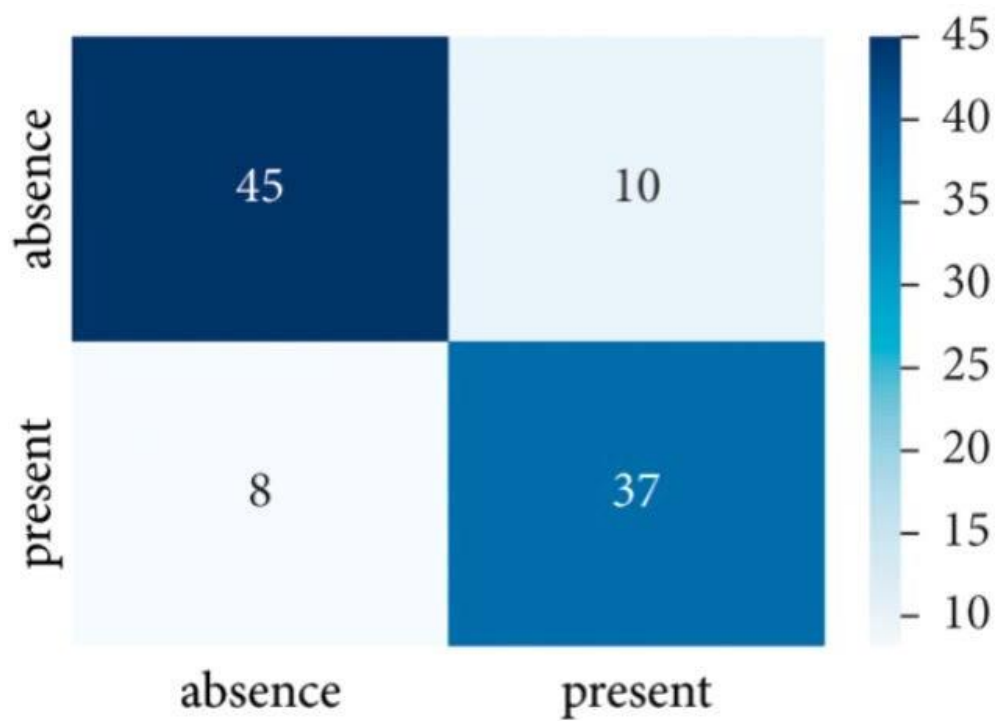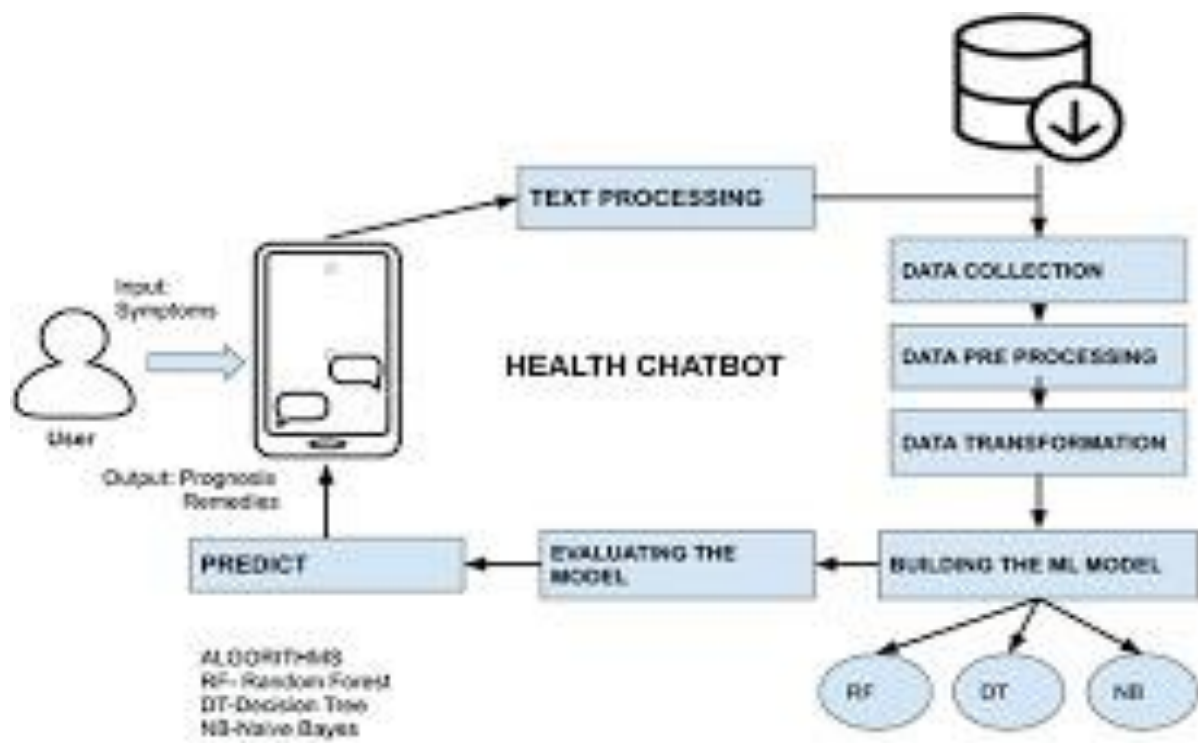
4.1   Classification of report
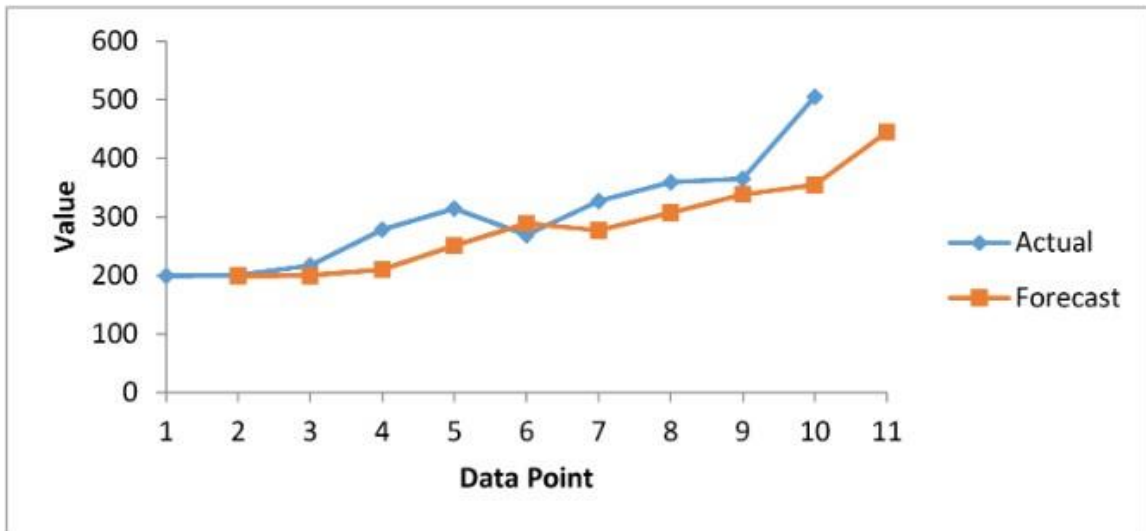
4.2   Loss in Validation and Training
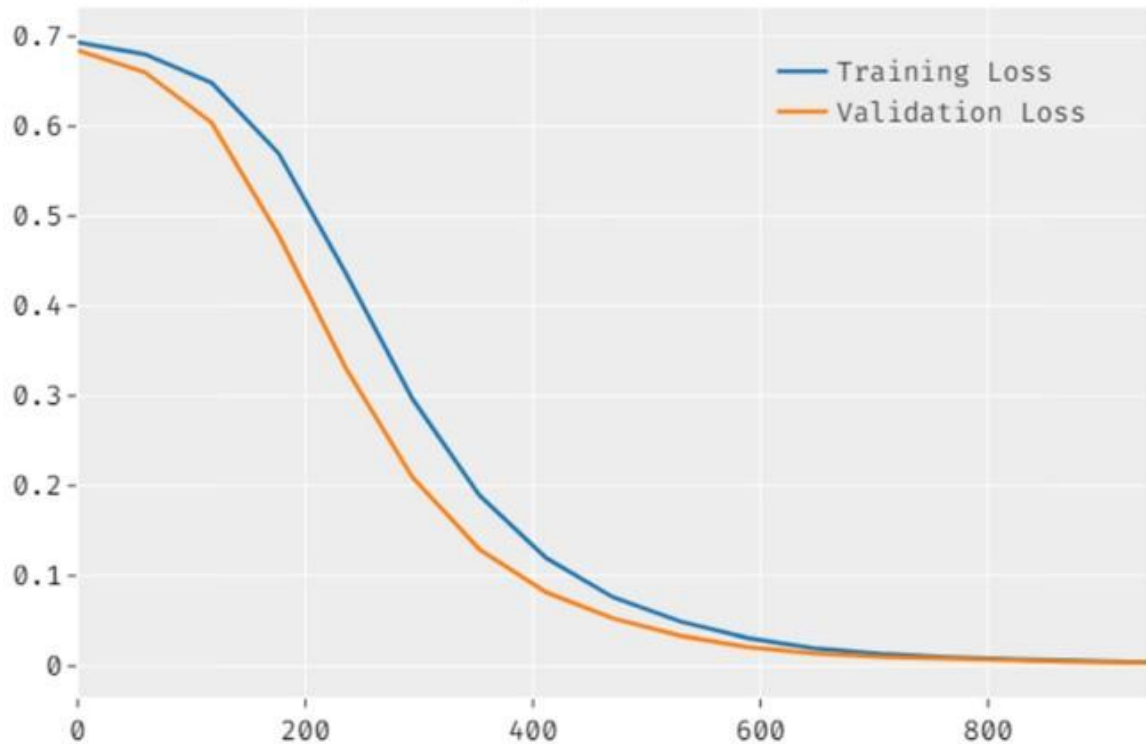
This figure shows the loss in validation and loss in training of our model, where theorange line depicts loss in validation and blue line depicts loss in training.

4.3 Accuracy in Training and Validation

The above figure contains the graph depiction between the training and testing data accuracy

4.4 Output

Below are the screenshots of our output when we use this program and tell itwhat symptoms I get. and that day I was indisposed because of some fever so I consulted my program to check the accuracy into it and then firstly it asked me about my name, and said hello to me. Now, it will ask my symptoms experience like fever and then it asks me what type of fever I have like high fever or mild fever in the form of 0-1 on that day I think I have mild fever it says okay then asks me from how many days so I type 2 days of mild fever. Now it will ask me some questions about its parameters. On the basis of these questions, it will generate my accurate prediction graph of what I should do in that situation so on that basis I have these screenshots.

In the below output screen it shows the accuracy of three different machine learning algorithms that is Decisiontree, Random forest and Support Vector Machine. From three values of each algorithm we calculate the mean value then we compare all the three algorithm accuracy in the actual digits and then plot the bar graph for eachof the algorithm accuracy for the comparisons that which algorithm shows the best accuracy. Here from the below comparison Support Vector Machine shows the best accuracy among all the machine learning algorithms.

Output 1

```
[0.96863469 0.974122   0.97227357]
Decision tree accuracy score
0.9716767500392195
[0.97601476 0.97966728 0.97966728]
random forest tree accuracy score
0.978449775256084
for svm:
1.0
```


Graph depicting test accuracy of different machine learning algorithm

```
--------------------------------HealthCare ChatBot--------------------------------
Your Name?                           ->Ravi
Hello,  Ravi

Enter the symptom you are experiencing              ->Fever
Enter valid symptom.

Enter the symptom you are experiencing              ->fever
searches related to input:
0 ) high_fever
1 ) mild_fever
Select the one you meant (0 - 1):  0
Okay. From how many days ? : 2
Are you experiencing any
muscle_weakness ? : no
stiff_neck ? : no
swelling_joints ? : no
movement_stiffness ? : no
painful_walking ? : no
It might not be that bad but you should take precautions.
You may have  Arthritis or  Impetigo
Arthritis is the swelling and tenderness of one or more of your joints. The main symptoms of arthritis are joint pain and stiffness, which typically worsen with age. The most common
Impetigo (im-puh-TIE-go) is a common and highly contagious skin infection that mainly affects infants and children. Impetigo usually appears as red sores on the face, especially arou
Take following measures :
1 ) exercise
2 ) use hot and cold therapy
3 ) try acupuncture
4 ) massage
--------------------------------------------------------------------------------
/usr/local/lib/python3.7/dist-packages/sklearn/base.py:451: UserWarning: X does not have valid feature names, but DecisionTreeClassifier was fitted with feature names
  "X does not have valid feature names, but"
```

359

In the below output screen it takes another person name as an entry and take their symptom then ask formore symptoms to accurately verify the disease situated to it and give some precaution.

Output 2

```
[0.97601476 0.97966728 0.97042514]
Decision tree accuracy score
0.975369060529792
[0.96494465 0.97966728 0.97781885]
random forest tree accuracy score
0.9741435954100761
for svm:
1.0
```


Graph depicting test accuracy of different machine learning algorithm
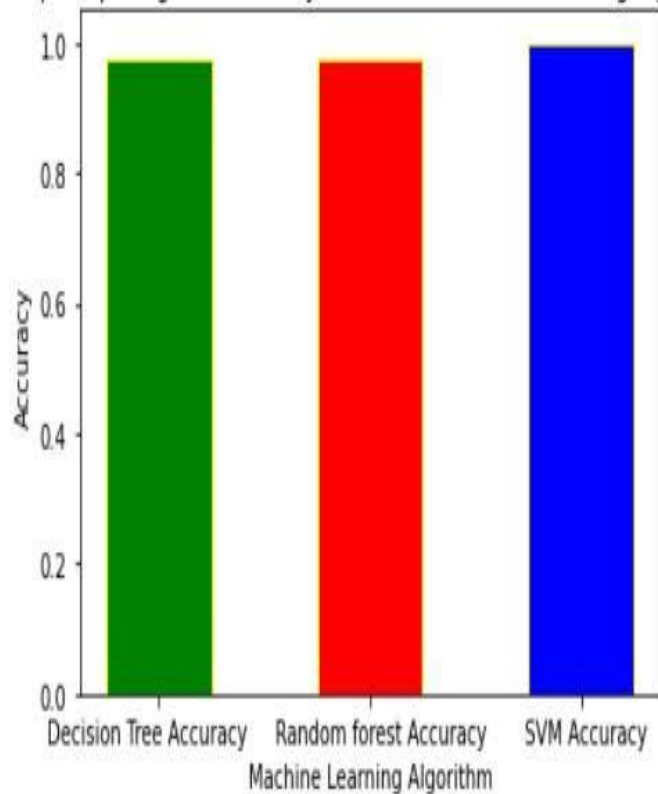
```
-------------------------------HealthCare ChatBot-------------------------------

Your Name?                            ->Vikas
Hello,  Vikas


Enter the symptom you are experiencing              ->blackhead
searches related to input:
0 ) blackheads
Okay. From how many days ? : 3
Are you experiencing any
back_pain ? : o
provide proper answers i.e. (yes/no) : no
weakness_in_limbs ? : no
neck_pain ? : o
provide proper answers i.e. (yes/no) : no
dizziness ? : no
loss_of_balance ? : no
It might not be that bad but you should take precautions.
You may have  Cervical spondylosis or  Impetigo
Cervical spondylosis is a general term for age-related wear and tear affecting the spinal disks in your neck. As the disks dehydrate and shrink, signs of osteoarthritis develop, incluc
Impetigo (im-puh-TIE-go) is a common and highly contagious skin infection that mainly affects infants and children. Impetigo usually appears as red sores on the face, especially arounc
Take following measures :
1 ) use heating pad or cold pack
2 ) exercise
3 ) take otc pain reliver
4 ) consult doctor
-------------------------------------------------------------------------------
/usr/local/lib/python3.7/dist-packages/sklearn/base.py:451: UserWarning: X does not have valid feature names, but DecisionTreeClassifier was fitted with feature names
  "X does not have valid feature names, but"
```

361

5.CONCLUSION

Further research and interdisciplinary collaboration could advance this technology, dramatically improvingthe quality of patient care, balancing the workload of clinicians, and revolutionizing medical practice.

A sklearn-based chatbot that can indicate symptoms, ask questions, and provide details and advice. A medical chatbot that quickly answers frequently asked questions by setting up a rule-based keyword chatbot.

Our medical chatbot provides medical assistance to patients with common ailments such as fever, cold, typhoid, malaria and jaundice. Due to the demands of our country's growing population, we tend to invent systems. Such systems are available abroad but not in our country. As you know, there are fewer doctors to meet the needs of patients.

By [6] using training and testing dataset of symptom description, symptom precaution and symptom severitythis project uses three machine algorithms to correctly identify the problem or disease situated to the several symptoms mentioned by the user. After all the input data entered by the user it uses its all three machine learning algorithm that is Decision tree, Random Forest and Support Vector Machine to process the input data and then it gives the response that tells or advice the user some precaution and about the disease for which useris looking for by entering his/her symptoms.

It will definitely help in medical support to advice people about the disease and some precautions related to itthat must followed by the user to maintain a good and healthy life

6. REFERENCES

1. Jadhav, K.P.; Thorat, SA The road to design dialog agent system. performance in intelligent systems andcomputing; springer: 2020 Berlin, Germany [Google Scholar]
2. Battineni, G.; Di Canio, M.; Chintalapudi, N.; Amenta, F.; Nittari, G. Physical development Trainingapp for smartphones to keep fit sailor level. international Marit. Health 2019, 70, 180–186. [Google Scholar] [CrossRef][PubMed]
3. Jan, R. "Chitty Chatbot": Deep Learning Conversational AI. Publications of the 27th Joint InternationalConference Artificial Intelligence (IJCAI-18), Stockholm, Sweden, 13-19 July 2018 [Google Scientist]
4. Tobias Kovatsch, Text-Based Medicine Chatbot that helps patients and their health", 01 October 2017
5. Ching-Yuan Huang, Ming-Ching Yang, Ching-Yu Huang, "Smart Wireless Networking with ChatbotSupport". Interactive health system for weight Health Care and Promotion", Resources IEEE, April 2018
6. Bukricha H., Waxmuth I.: Modeling Empathy for Fictional Characters: When and How Somewhat. 10th International autonomous agent and Multi-Agent Systems - Volume 3. International Autonomous Agent Fund and Multi-Agent Systems, 2011., p. 1135–1136.
7. REAL LIFE HEALTH CARE CHATBOT USING PYTHON
8. Dr.V. Vijayakumar1 V. Deeparani2, S. Gayathiri3, G. Janani4, G. Monisha5
9. International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST),Feburary 2015
10. C.P. Shabariram, V. Srinath and C.S. Indhuja, Vidhya, "Ratatta: Chatbot Application Using Expert System", International Journal of Advanced Research in Computer Science and Software

Engineering, 2017.

11. Mrs Rashmi Dharwadkar and Neeta A. Deshpande, "A Medical ChatBot", International Journal of Computer Trends and Technology (IJCTT) –, vol. 60, no. 1, June 2018.

12. Farheen Naaz and Farheen Siddiqui, "modified n-gram based model for identifying and filtering near- duplicate documents detection", International Journal of Advanced Computational Engineering and Networking, vol. 5, no. 10, Oct. 2017.

13. S.J. Du Preez, Manoj Lall and S. Sinha, An intelligent web-based voice chat bot, vol. 38, 2009.

14. Bayu Setiaji and Ferry Wahyu, "Chatbot Using a Knowledge in Database:Human-to-Machine Conversation Modeling", Intelligent Systems Modelling and Simulation (ISMS) 2016 7th Int. Conference 72-77, 2016.