

Healthcare Chatbot

Sunil Tayde, Bhavit Khandelwal, Om kakde, Karitkey Dwivedi, Vedant Khandekar, Mandar Kulkarni

Department of Electronics and Telecommunication Engineering

Abstract — *Healthcare chatbot is nothing but the normal chatbot which helps in guiding the users about their health conditions and predicting the disease from which the patient is suffering. The project comes under AI and ML. The chatbot also guides by suggesting which Doctor to consult and the address of his clinic or the hospital he is working in. The system also gives more information about the predicted disease. For creating the chatbot, “Decision Tree algorithm” is used. Also, techniques like NLTK and Cosine Similarities were also used. Confidence Level is also calculated.*

Keywords — *Decision Tree Algorithm, NLTK, Cosine Similarity and Confidence Level .*

I. INTRODUCTION

Today in the fast-moving world, every individual has loads of work. Due to which a person is bound to have small or big health issues in their life, and they should go to the doctor once a year. In this job-oriented world, they must attend to their work and while doing so they will try to skip the long doctor's queue for appointments. We have a brilliant solution to tackle this problem, just bring the doctor home with our Health Care Assistant powered by AI and Deep Learning.

So, the Health Care is Assistant is the chatbot which chats with users/patients about their health problems and tries to find out practical solution. Healthcare chatbot helps in 3 different ways. First it asks some basic information about the patient like his name and the symptoms the patient is observing. 1] Based on that the system predicts the disease and if serious then suggests the name of the specialist Doctor for consultation. 2] Also, it shares the address of the Doctor's clinic. 3]The bot also provides information about the disease which has been predicted using NLTK and Cosine Similarity techniques. So that patient can get clear idea about the illness.

“Decision Tree Algorithm” is the algorithm on which the whole project is based. After asking the name, it asks about the main symptom observed by the patient. Based on the symptom given by the patient the system takes the decision about which symptom to ask next. After asking sufficient no. of questions, it predicts the disease. Datasets helps the algorithm to give right results. Confidence level of the project is also calculated.

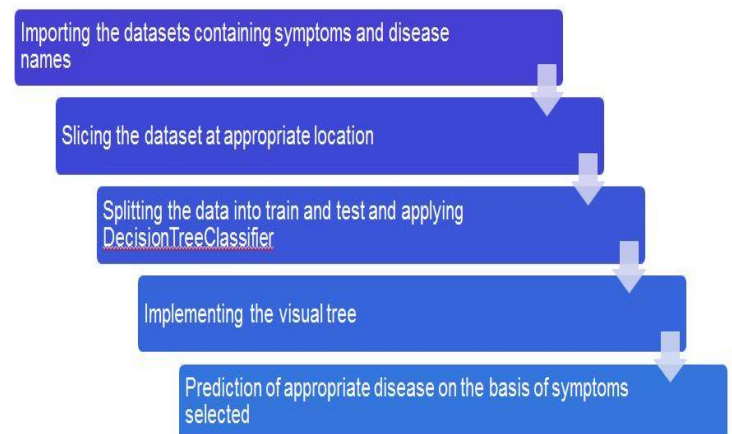
II. METHODOLOGY/EXPERIMENTAL

ALGORITHMS

Problem Statement:

1. Predicting the disease name based on various symptoms present in the patient using ML algorithms and reliable datasets

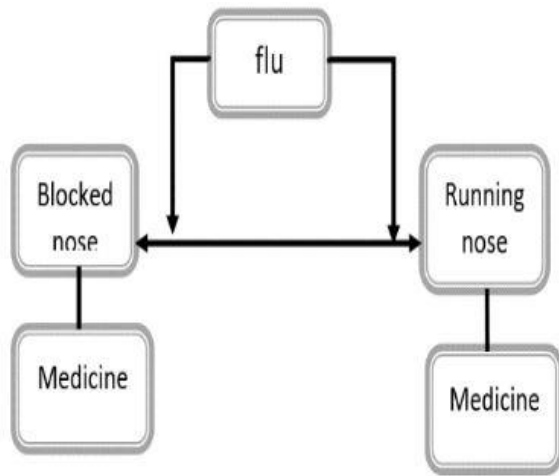
Flowchart



Decision Tree Classifier

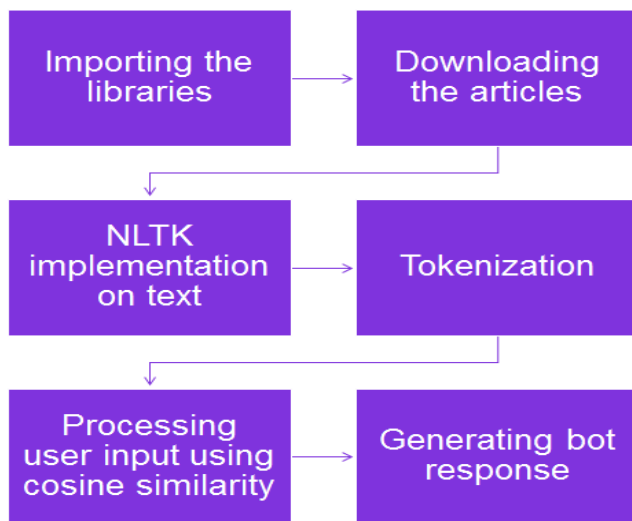
Decision Trees (DTs) are a non-parametric supervised learning method used for classification and regression. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features.

The decision algorithm used for simple disease uses a basic tree structure with minimum nodes as shown in below figure



2. To provide information of diseases or conditions when searched by the user

Flowchart



Natural Language Processing

Computers and machines are great at working with tabular data or spreadsheets. Much information that humans speak or write is unstructured. So it is not very clear for computers to interpret such. In natural language processing (NLP), the goal is to make computers understand the unstructured text and retrieve meaningful pieces of information from it. Natural language Processing (NLP) is a subfield of artificial intelligence, in which its depth involves the interactions between computers and humans.

Natural Language Understanding involves the process of extracting meanings from text inputs. In this direction, the basic design steps include

- Syntactic Parsing- determines the function of each word (part-of-speech), the way words are related to each other,

how they are grouped into phrases and how they can modify each other. The context-free 2grammar (CFG) definition and parsers implementation are the common NLP techniques used in this step.

- Semantic Parsing-The role of a semantic parser is to extract the context- independent meaning of a written sentence. The discriminative methods such as support vector machines (SVM) and statistical methods such as decision trees and classification and regression trees (CART) are used to find the most probable parse tree that fits the sentence.
- Contextual Interpretation-refines the semantic interpretation by taking advantage of information at the discourse level and, removing remaining ambiguities such as anaphors, pronouns and ellipses. Discourse entity (DE) list maintains a set of constants referring to objects that have been evoked in previous sentences and can subsequently be referred implicitly.

Cosine Similarity

Cosine similarity is a metric used to measure how similar the documents are irrespective of their size. Mathematically, it measures the cosine of the angle between two vectors projected in a multi-dimensional space. The cosine similarity is advantageous because even if the two similar documents are far apart by the Euclidean distance (due to the size of the document), chances are they may still be oriented closer together. The smaller the angle, higher the cosine similarity.

Dimensionality Reduction

The number of input variables or features for a dataset is referred to as its dimensionality. Dimensionality reduction refers to techniques that reduce the number of input variables in a dataset. Now in this part of project we want just diseases column to be copied in new Data Frame, so reduction is done in the way that just diseases column gets copied.

diseases = dimensionality_reduction. index

III. RESULTS/DISCUSSIONS

```

0.9747574647658358
for svm:
1.0
Your Name                                -->Naman

hello Naman
Enter the symptom you are experiencing      -->cough

searches related to input:
0 ) cough
Okay. From how many days ? : 3
Are you experiencing any
muscle weakness ? : yes
stiff_neck ? : no
swelling_joints ? : no
movement_stiffness ? : yes
painful_walking ? : yes
It might not be that bad but you should take precautions.
You may have Arthritis
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.
Take following measures :
1 ) exercise
2 ) use hot and cold therapy
3 ) try acupuncture
4 ) massage
confidence level is 0.6
  
```

IV. LIMITATIONS

- 1] There are times when people are ill, and they don't realize it or don't know what actually is happening in their body so these kinds of situations the best option is to consult a doctor.
- 2] The accuracy of the bot may not be accurate if the user selects a wrong option.
- 3] Self-assessment can always be dangerous, and the bot can not always be a 100% sure.
- 4] People can take wrong dosages of medicine after self-assessment which is dangerous.

V. FUTURE SCOPE

- 1] Making an android app with a more interactive format.
- 2] Improving the accuracy of the code.
- 3] Making a new dataset for doctors and clinics in Pune.
- 4] Starting a subscription-based service to avoid the likes of scammers and such.
- 5] Voice input can make the program to work even better.

VI. CONCLUSION

Artificial Intelligence is becoming a norm of every technological development. Through this project the gap between Ai and health science has been narrowed down and a new possibilities have been introduced by which AI can be used to facilitate human development. The project uses a AI system to develop an efficient chatbot system based on rational agent with tree searching algorithm and knowledge base of medical data. The chatbot asks sever questions from the user or patient and then based on the replies suggests a possible remedy. In future, the system need to increase its database and machine learning capabilities for much improved diagnostics. The use voice as input would also be an effective feature. The system needs to go to deeper level of tree from 3 to 7 and be more accurate.

REFERENCES

- 1] Michalski, R. S., Carbonell, J. G., & Mitchell, T. M (Eds.),— Machine learning: An artificial intelligence approachl Springer Science & Business Media, pp.92, ,2013.
- [2] Oliver Kharraz, Nick Ganju, Cyrus Massoumi , —Zocdoc: Find a Doctor – Doctor Reviews & RatingslDate retrieved: 2017,pp.23-25, August 2017 . URL:https://www.zocdoc.com
- [3] Babylon, —Babylon Health: Online Doctor Consultations & Advicel Retrived on 2017, pp. 34-38, 2017.URL:https://www.babylonhealth.com.
- [4] Khoury Consulting, —Doctor Pocketl Retrieved on August-2017,vol.8(2):76-79, 2017, URL https://play.google.com/store/apps/details?id=com.docpoc.doctor&hl=en
- [5] Baidu Doctor App, —Baidu's - Medical Robotl, Retrieved on July 2017.
- [6] S.M. Monzurur Rahman and Xinghuo Yu, —An unsupervised neural network approach to predictive data miningl, Int. J. Data Mining, Modelling and Management, Vol. 3, No. 1,pp 18-41, 2011.
- [7] Lev V. Utkin,lRegression analysis using the imprecise

Bayesian normal model, International Journal of Data Analysis Techniques and Strategies,Vol. 2, No.4 pp 356 – 372, 2010.

- [8] Z.A. Al-Hemyari and I.H. Hussain, —Improved shrinkage testimators for the simple linear regression modell, International Journal of Information and Decision Sciences, Vol. 4, No.1 pp 87 – 103, 2012.
- [9] E. Heierman, III, and D. Cook, —Improving home automation by discovering regularly occurring device usage patternsl, in Proc. 3rd IEEE Int. Conf. Data Mining, pp. 537–540, 2003.
- [10] M. Ruotsalainen, T. Ala-Kleemola, and A. Visa, —Gais: A method for detecting interleaved sequential patterns from imperfect data,___ in Proc. IEEE Symp. Comput. Intell. Data Mining, pp. 530–534, 2007.
- [11] Abel, David, James MacGlashan, and Michael L. Littman, —Reinforcement Learning As a Framework for Ethical Decision Makingl. Workshops at the Thirtieth AAAI Conference on Artificial Intelligence, 2016.
- [12] Cserna, Bence, et al.,lAnytime versus Real-Time Heuristic Search for On-Line Planningl. Ninth Annual Symposium on Combinatorial Search, 2016.
- [13] Hamet, Pavel, and Johanne Tremblay, —Artificial Intelligence in Medicinel. Metabolism, pp.10-15, 2017.
- [14] Forestier, Germain, et al. ,lAutomatic matching of surgeries to predict surgeons_ next actionsl, Artificial Intelligence in Medicine,pp.12-15,2017.
- [15] Jha, Saurabh, and Eric J. Topol, —Adapting to artificial intelligence: radiologists and pathologists as information specialistsl. JAMA 316.22, pp.2353-2354, 2