International Journal of Innovative Research in Engineering
Volume 3, Issue 3 (May-June 2022), PP: 167-171.
www.theijire.com
ISSN No: 2582-8746

Disease Forecast Method using Decision Tree Classifier

Jitendra Kumar[1], Palak Chanana[2], Ankita[3], Vishwajeet Behera[4]
[1] *Professor, Computer Science, Panipat Institute Of Engineering & Technology, Haryana, India.*
[2,3,4] *Student, Computer Science, Panipat Institute Of Engineering & Technology, Haryana, India.*

How to cite this paper:
Jitendra Kumar[1], Palak Chanana[2], Ankita[3], Vishwajeet Behera[4], "Disease Forecast Method using Decision Tree Classifier", IJIRE-V3I03-167-171.

Abstract: *The development of latest technologies like data science, machine learning, and artificial intelligence has created new path for medical organization and healthcare institutions, to forecast the illness early as possible and it provide better treatment with advanced facilities. Quality of Forecast of illness goes down with less data. But with data everything is possible. Even detecting any illness with help of just symptoms is now possible. Many illness are area based, so accuracy goes down there also. Human body always give some kind of signal when there is something wrong in our body, when we have those signal like small fever or headache we need to take care of those signs else it can be too late to treat them. Many times patients treat symptoms as a minor issue and later on, he/she gets to know it was a major disease that led her to something very serious. So we are offering an illness forecast method that can forecast the illness depending on symptoms so they can be treated at the early stage, which is necessary as with help of this we can help someone to save his/her life or his/her friend life too. In this way with this forecast we can get treated only for those signs and illness which patient does have. This forecast method is not a replacement for real diagnosis but this method in place help patient to get treated for what is necessary and avoid complication on time rather than getting the cure for some other disease with help of symptoms he/she can get treatment for the exact disease which he/she likely to have. The accuracy of offered method is 95.12%. The method has huge possibility in detecting possible illness quite accurately. The objective of this offered method is to give a tool to people with less technical knowledge and a new doctor starting in his/her career can treat patient effectively.*
Keywords : *Disease Forecast Method, Machine Learning, Decision Trees Classifier (DTC)*

I. INTRODUCTION

It is a method that is made by the use of machine learningalgorithm for predicting the possibility ofillness based on thesymptoms entered by the user[1].The rapid rise in data science technology has been improving our lives so far. It provides many tools that canbe useful for many purposes, one such use case is helping in getting the disease with just the help of symptoms.Machine Learning is used to develop methods that can helpus predict so many illness based on symptoms.Itcan be used to predict the exact accuracy of the disease. How likely according to your method you are to have that disease and according to that based of how likely you have one can get treated for the exact disease.

For example, based on symptoms two illness are most likely to be the same that is Alzheimer's and Parkinson's. If a person has one of these illness and gets treated for another disease, he/ she will likely have the bad treatment and the situation can get worsen, So with help of this model and based on symptoms we can likely predict how likely you are to have that particular disease and according to that, a person can get treated accordingly. So our proposed in boon even in dilemma situation too when just based on symptoms a person is in confused mind if he/she has just the normal disease or it's something serious. He/she can get results freely and then decide accordingly.It can provide the possible forecast to the expert and the chances of possible illness and treatement can be done on basis of this suggestion which will help in taking further action.

Here we are in 21st century and today everything is possible with this advanced technological advnancement, today we have enormous different ways to get data from any area of this world about any person just by sitting in one place with help of one tap if we have exact softwares to do that.Even with a health card, one can get a person's complete health history with just one click. Here today we have ways that can give data in our hand over the internet when we use those tools.

Today over google in just a minutes billion of data searches are happening where some searches give right information some confuses the person by provding extra information and in those queries lakhs of queries are related to illness. Humans have tendency to know whats happening with them, they want to know why their body is behaving like this so they do those search queries over the internet but they do not want to go to medical expert as they have fear of having something serious so they look for some easy solution over the internet.Sometimes even after knowing the disease people tend to take cheap medication and think they will get better by it and later on they suffer due to bad effects of those medications.So we are trying to provide a tool for the right treatment based on symptoms.

II. LITERATURE REVIEW

Before this study there have been many studies done to forecast the illness using different method for helping the healtcare expert. Some of the reviews of those studies done in papers using those methods and results are used by them are following:

1. Swarupa, A. N. V. K., V. Hema Sree, S. Nookambika, Y. Kiran Sai Kishore, and U. Ravi Teja. "Disease Prediction: Smart Disease Prediction System using Random Forest Algorithm." In 2021 IEEE International Conference on Intelligent Systems, Smart and Green Technologies (ICISSGT), pp. 48-51. IEEE, 2021.
2. Ali, Liaqat, Awais Niamat, Javed Ali Khan, Noorbakhsh Amiri Golilarz, Xiong Xingzhong, Adeeb Noor, Redhwan Nour, and Syed Ahmad Chan Bukhari. "An optimized stacked support vector machines based expert system for the effective prediction of heart failure." IEEE Access 7 (2019): 54007-54014.
3. Pattekari, Shadab Adam, and Asma Parveen. "Prediction system for heart disease using Naïve Bayes." International Journal of Advanced Computer and Mathematical Sciences 3, no. 3 (2012): 290-294.
4. Parajuli, Anil. "DISEASE PREDICTOR." PhD diss., Tribhuvan University, 2016.
5. Dahiwade, Dhiraj, Gajanan Patle, and Ektaa Meshram. "Designing disease prediction model using machine learning approach." In 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), pp. 1211-1215. IEEE, 2019.
6. Repaka, Anjan Nikhil, Sai Deepak Ravikanti, and Ramya G. Franklin. "Design and implementing heart disease prediction using naives Bayesian." In 2019 3rd International conference on trends in electronics and informatics (ICOEI), pp. 292-297. IEEE, 2019.
7. Gao, Jianliang, Ling Tian, Jianxin Wang, Yibo Chen, Bo Song, and Xiaohua Hu. "Similar disease prediction with heterogeneous disease information networks." IEEE Transactions on NanoBioscience 19, no. 3 (2020): 571-578.
8. Rudra A. Godse, [6] studied multiple disease forecasts using differentmachine learning algorithm involving techniques like KNN Algorithm, SVM, and Decision Tree.

III. MATERIALS AND METHODS.

The offered method consist of data of 4920 patients' records who were treated with 41 illness were chosen for further investigation. Our dependent variable consisted of 41 illness. 95 of 132 independent variables (signs) closely related to illness were chosen. We uses ML algorithm to forecast illness depending on patients' signs. In this method, we are forecasting illness depends on signs but if we provide information of other illness to the method then it can also forecast other illness.

Procedure methodology:

a. There are following steps are there to perform the computation:
b. First It will take the person's symptoms and problems they are having in their body.
c. Then It will take information that will match the given symptoms and it will collect information related to that particular disease.
d. Then It will take the data of the symptoms as input and process it by the Decision Tree Classifier Method.
e. After that Decision Tree predicts the illness that can be accurate according to given symptoms and best fit and most likely to occur to that patient.
f. Then the method will give the name of the disease along with the link to prescribe to the doctor via practo.

The flowchart of the methodology is given below:

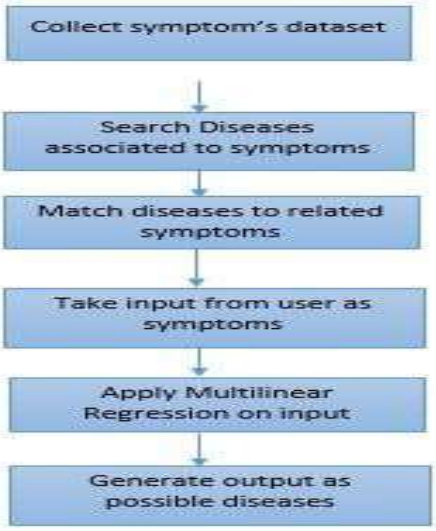

Fig1: Flowchart of the proposed method

Algorithm Used: Just by seeing the name one can do some guess about our forecast method. Here we are using Decision Tree Classifier for forecasting the illness. Decision Tree Classifier is an algorithms of ML which uses different rules to get to a decision, as we humans also make decisions, based on choices.
The thinking behind our method is to use the information features to provide option like either yes/no decision and based on our decision we have to break the datasets till the point where each points of data belongs to its class.

IV.THE ARCHITECHURE OF ILLNESS FORECAST METHOD

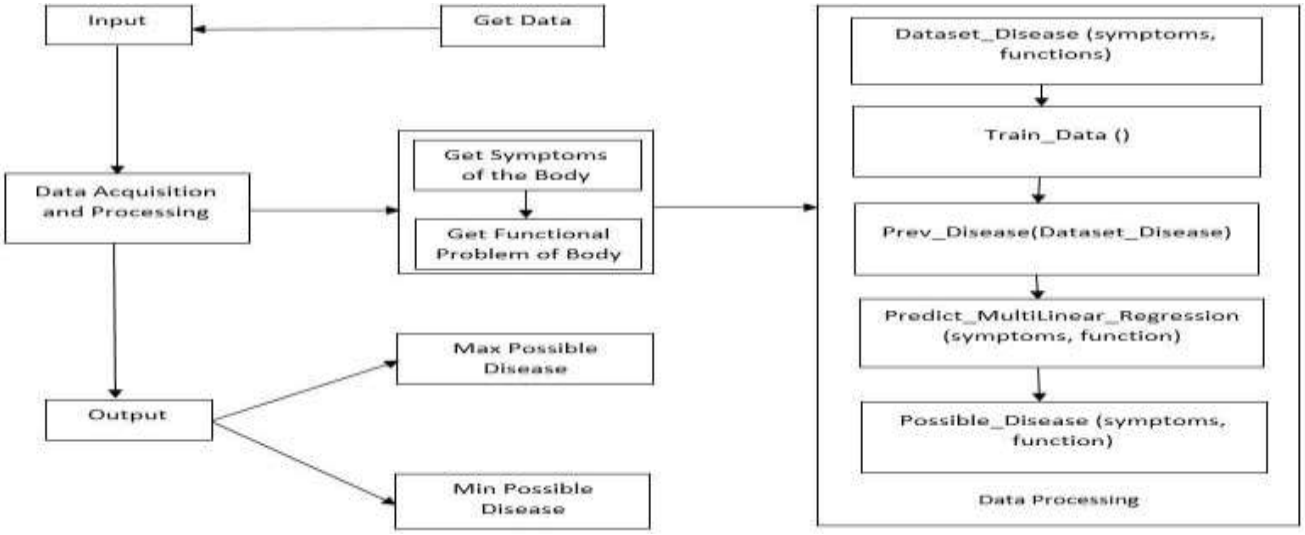

Fig2 : Illness Forecast Method

Input: We are taking input from the user of the diseaseforecast method as a symptoms list.
GetData: In this field, the user will provide data about theirsymptoms.
Data Acquisition andProcessing: In this field, the input is provided for processing. Data acquisition and processing perform two operations, the first is acquiring the data and the second is processing the data and extracting information based on that acquireddata.
Get Symptomsof theBody: In this field symptoms of the body are gathered and analyzed. So that this information can be used by the algorithm to predict the possible illness.
Get Functional Problem of Body: In this field, functional problems of the body that is associated with the symptoms are gathered. So that it is analyzed to get the possible disease.

Dataset_Disease(symptoms,functions): In this field, we have a predefined dataset of illness that involves symptoms and functions that are caused by the disease. This dataset is further used to match the data that has been obtained from the user and if matched properly then the method will suggest the possible illness.
Train_Data(): In this field training in the method is performed. Our disease forecast method is trained using the SVM (support vector machine) algorithm. Here we are using the SVM algorithm to solve a problem related to regression.
Prev_Disease(Dataset_Disease): In this field Dataset of the illness is provided as a parameter and processing is performed based on this dataset.
Predict_Decision Tress Classifier(symptoms,function): In this field, the forecast is performed using the Decision Tree Classifier algorithm. In DTC, you ask a question, and based on yes/no you further divide it till you reach your answer. Symptoms and their functions in the user's body are involved in the forecast.
Possible_Disease(symptoms,function): In this field symptoms and functions are passed as a parameter and possible illness are calculated based on these parameters.
Data Processing: This field contains the above five data processing fields and is the main part of our disease forecast method. It has all the necessary fields for processing the data.
Output: After Data Acquisition and Processing, possible illnessare generated as output.

V. RESULT

Result analysis in our proposed method is an essential part of this research paper. By the analysis of the results, we can compare how much better this proposed method is performing. In the result analysis, we will see the accuracy of different illness that are predicted using our proposed method. We have taken datasets of 100 cases for result analysis.

Disease based accuracy analysis for 100 cases:

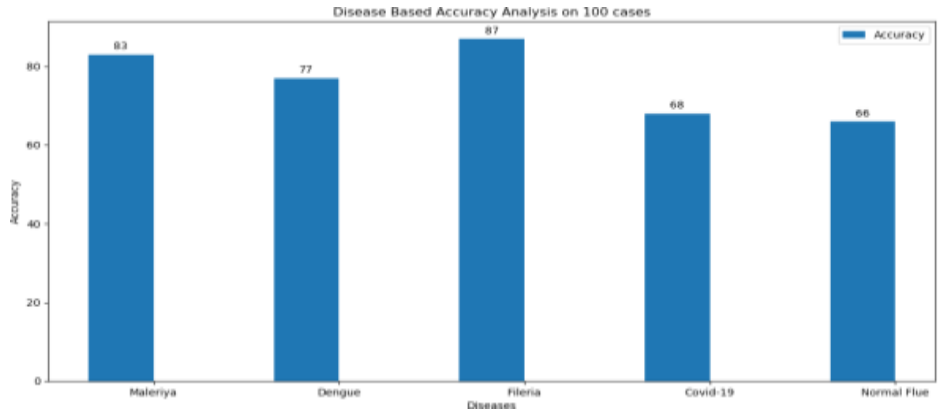

Fig 3: Disease Based Accuracy analysis on 100 cases Above diagram shows the accuracy of 5 illness that are malaria, dengue, filaria, covid-19, and normal flu.

Comparative analysis between algorithm for our disease forecast method:

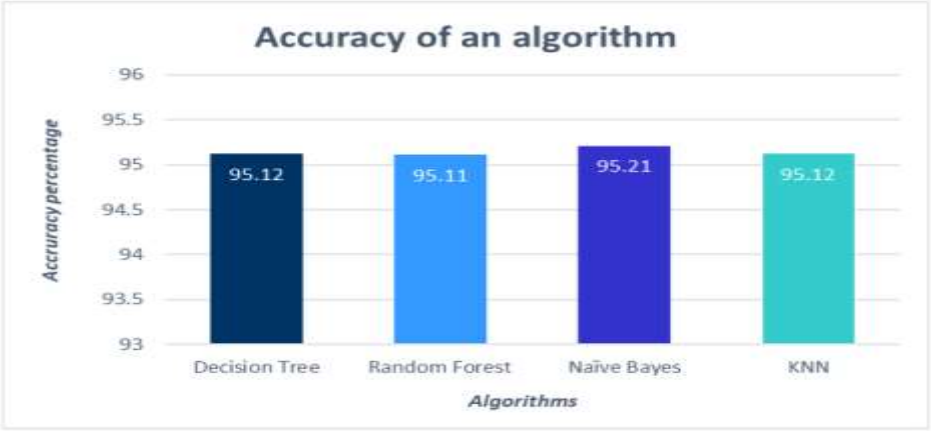

Fig 4: Comparative Analysis between Algorithms

VI.CONCLUSION

This program aims to predict the disease on the basis of the symptoms. The project is Disease Predictor was successfully implemented using grails framework and designed in such a way that the system takes symptoms from the user as input and produces output i.e. predict disease. In our research, we have used a support vector machine and a multilinear regression algorithm to predict illness and test the multiple algorithm attached to them for better result.

The purpose of this research was to provide medical diagnosis information based on symptoms to normal people, fresher doctors, medical students, and anyone who wants to know about a set of symptoms and associated illness. In this research, we have found that possible disease forecast can go up to 87% for some illness and a minimum of 68% for some illness these results are obtained using the minimum amount of data set but if we can feed the method a humongous amount of data set then this disease forecast method can give accuracy up to 95%. Obtaining a humongous amount of data sets related to illness and their symptoms is very time-consuming and it cannot be done within one or two years it requires multiple years to collect those data sets and train the method using those data searches. This method can be used by Ph.D. scholars to do further research.

With the help of a disease forecast method, it was possible todiagnose people based on symptoms. A disease forecast method provides only possible outcomes it does not guarantee that it will predict the disease correctly. But it has significantly higher accuracy for predicting possible illness. In our research, we have analyzed the accuracy of this method for 5 different illness and our accuracy can go up to 87%.

References

1. *Disease forecast from various symptoms using machine learning. By RinkalKeniya, Aman Khakharia, Vruddhi Shah*
2. *Min Chen, Yixeu Hao,"DiseaseForecast by MachineLearning Over Big Data From Healthcare Communities"IEEE.*
3. *Rahul Chaurasia, Saksham Gupta, "Forecast of Heart disease using machine learning algorithm"IEEE,ISSN: 2320-2882, Volume 6, Issue 2 April 2018*
4. *K. Arumugam, MohdNaved, Priyanka P Shinde,"MultipleDiseaseForecastusingMachine Learning Algorithm"*
5. *Rudra A. Godse. and Smita S Gunjal., "Multiple Disease Forecast Using Different Machine Learning Algorithm"*
6. *Akash C Jamgade, Prof S.D. Zade,"Disease Forecast UsingMachineLearning"IRJET,e-ISSN: 2395-0056.*
7. *Heart Disease Forecast using Logistic Regression Algorithm using Machine Learning"Reddy Prasad,Pidaparthi Anjali, S.Adil, N.Deepa (2019).*