



INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X

Impact factor: 4.295

(Volume 5, Issue 2)

Available online at: www.ijariit.com

Virtual doctor

Divyansh Tiwari

tiwari.divyansh007@gmail.com

IMS Engineering College,
Ghaziabad, Uttar Pradesh

Arpit Kumar

arpit.gupta151@gmail.com

IMS Engineering College,
Ghaziabad, Uttar Pradesh

Ayush Tripathi

ayushtripathi1024@gmail.com

IMS Engineering College,
Ghaziabad, Uttar Pradesh

ABSTRACT

The healthcare environment is still 'Information Rich' but 'Knowledge Poor'. There is a wealth of data available within the health care systems. However, there is a lack of effective analysis tools to discover hidden relationships in data. The aim of this work is to design a GUI based Interface to enter the patient symptoms and predict which disease the patient is having using various machine learning algorithms. The prediction is performed from mining the patient's symptom data or data repository. This paper has analyzed prediction systems for disease using more number of input attributes. The system uses medical terms such as fever, pain, cholesterol-like attributes to predict the likelihood of a patient getting a particular disease. Until now, over 100 attributes are used for prediction. The data mining classification techniques, namely Decision Trees, Naive Bayes, and Random Forest are analyzed on disease database. The performance of these techniques is compared, based on accuracy.

Keywords— Predictive analysis, Data mining Machine Learning

1. INTRODUCTION

Data mining is the method for finding unknown values from an enormous amount of data. As the patient's population increases the medical databases also increasing every day. The transactions and investigation of these medical data are difficult without the computer-based analysis system. The computer-based analysis system indicates the mechanized medical diagnosis system. This mechanized diagnosis system supports the medical practitioner to make a good decision in treatment and disease. Data mining is the huge platform for the doctors to handle the huge amount of patient's datasets in many ways such as making sense of complex diagnostic tests, interpreting previous results, and combining the dissimilar data together. In today's computerized world considering automatic and dynamic requirements healthcare system should be more efficient by predicting the disease and providing appropriate medications through user-friendly mobile applications. This study aims mainly for the health concerns and the ones who want to be their own Doctor. It is an interactive service for users who wants to know about what health issues they are going through as per the symptoms. It is easy to access and use for searching medicines for the diseases predicted.

2. LITERATURE SURVEY

2.1 Comparative analysis

In the paper "Disease Prediction System using data mining techniques" the author has discussed the data mining techniques like association rule mining, classification, clustering to analyze the different kinds of heart-based problems. The database used contain a collection of records, each with a single class label, a classifier performs a brief and clear definition for each class that can be used to classify successive records. The data classification is based on MAFLA algorithms which result in accuracy, the data is estimated using entropy-based cross-validations and partition techniques and the results are compared. C4.5 algorithm is used as the training algorithm to show the rank of heart attack with the decision tree. The heart disease database is clustered using the K-means clustering algorithm, which will remove the data applicable to a heart attack from the database. Some limitations are faced by the system like, the time complexity is more due to DFS traversal, C4.5- Time complexity increases while searching for insignificant branches and lastly no precautions are defined. In the paper "A study on data mining prediction techniques in the healthcare sector" [2] the fields which discussed are, Knowledge Discovery Process (KDD) is the process of changing the low-level data into high-level knowledge. Hence, KDD refers to the nontrivial removal of implicit, previously unknown and potentially useful information from data in databases. The Knowledge Discovery in Databases process comprises of a few steps leading from raw data collections to some form of new information. The iterative process consists of the following steps: Data cleaning, Data integration, Data selection, Data transformation, Data mining, Pattern evaluation, Knowledge. Healthcare data mining prediction based on data mining techniques are as follows: Neural network, Bayesian Classifiers, Decision tree, Support Vector Machine. The paper states the comparative study of different healthcare predictions, Study of data mining techniques and tools for the prediction of heart disease, various cancers, diabetes, eye disease and dermatological conditions. Data mining based prediction system reduces the human effects and

cost-effective one. Few limitations are that if attributes are not related then Decision trees prediction is less accurate and ANN is computationally intensive to train also it does not lead to a specific conclusion. The paper "Predicting Disease By Using Data Mining Based on Healthcare Information System" [4] applies the data mining process to predict hypertension from patient medical records with eight other diseases. The data was extracted from a real-world Healthcare Information System database containing medical records. Undersampling technique has been applied to generate training data sets, and data mining tool Weka has been used to generate the Naive Bayesian and J-48 classifiers created to improve the prediction performance, and rough set tools were used to reduce the ensemble based on the idea of second-order approximation. Experimental results showed a little improvement of the ensemble approach over pure Naive Bayesian and J-48 in accuracy, sensitivity and F-measure. Initially, they had a classification and then ensemble the classifiers and then the reduction of Ensemble Classifiers is used. But the decision trees generated by J-48 is sometimes lacking in the balancing so the overall improvement of using the ensemble approach is less. And for the prediction of heart disease level, no precautions are defined. The paper "An approach to devise an Interactive software solution for smart health prediction using data mining" [5] aims in developing a computerized system to check and maintain your health by knowing the symptoms. It has a symptom checker module which actually defines our body structure and gives us liability to select the affected area and check out the symptoms. Technologies implemented in this paper are: The front end is designed with the help of HTML, JavaScript and CSS. The back end is designed using MySQL which is used to design the databases. This paper also contains the information of testing like Alpha testing which is done at the server side or we can say at the developer's end, this is actual testing done with potential users or as an independent testing process at the server end. And Beta testing is done after performing alpha testing, versions of a system or software known as beta versions are given to a specific audience outside the programming team. Only the limitation of this paper is it suggests only the award-winning doctors and not the nearby doctors to the patient.

3. EXISTING SYSTEM

The existing system is health care prediction using data mining techniques. It takes some input parameters like symptoms and again provides with the related symptoms from which users need to select the symptoms from which he/she is suffering from and then finally providing the user with the predicted disease name.

4. PROPOSED SYSTEM

The proposed system will not only predict the diseases but will also suggest the appropriate medications. The MedlinePlus datasets will be used for both the symptoms checking, prediction of diseases and for providing medications. This system will predict the common-cold, cough, viral, like basic diseases including supplementary illness like acidity and weakness. It will be more efficient than the existing system. The application is user-friendly and mobile. It will use three algorithms namely Random Forest, Naïve Bayes and Decision Tree. The result will be that with maximum votes.

5. CONCLUSION

Thus in this paper, we have successfully studied the existing architecture of the system and through the proposed architecture we are designing a system which will help in predicting the disease using user-friendly application. The system should be efficient to predict the diseases and suggestion of medications using machine learning techniques like Classification, Association rule mining. It has a very wide range of scope in case of diseases. Further, the system can be extended to N number diseases existing with proper medications. It can be made available with the doctors through live chat so the end user can freely interact with the doctor in case of emergency.

6. REFERENCES

- [1] Aditya Tomar, "Disease Prediction System using data mining techniques", in International Journal of Advanced Research in Computer and Communication Engineering, ISO 3297, July 2016
- [2] Dr. B. Srinivasan, K. Pavya, "A study on data mining prediction techniques in the healthcare sector", in International Research Journal of Engineering and Technology (IRJET), March-2016.
- [3] Megha Rathi, Vikas Pareek, "An integrated hybrid data mining approach for healthcare", in IRACST - International Journal of Computer Science and Information Technology Security (IJCSITS), ISSN: 2249-9555, Vol.6, No.6, Nov-Dec 2016.
- [4] Feixiang Huang, Shengyong Wang, and Chien-Chung Chan, "Predicting Disease By Using Data Mining Based on Healthcare Information System", in IEEE 2012.
- [5] M.A. Nishara Banu, B Gomathy, "An approach to devise an Interactive software solution for smart health prediction using data mining, in International Journal of Technical Research and Applications, EISSN, Nov-Dec 2013.
- [6] Akshay Raul, Atharva Patil, Prem Raheja, Rupali Sawant, "Knowledge Discovery, Analysis And Prediction in Healthcare using Data Mining And Analytics", in IEEE, 2nd International Conference on Next Generation Computing Technologies, 2016.
- [7] M.A. Nishara Banu, B.Gomathy, "Disease Forecasting System Using Data Mining Methods", in IEEE, 2014 International Conference on Intelligent Computing Applications.
- [8] Sellappan Palaniappan, RafiahAwang, "Intelligent Heart Disease Prediction System Using Data Mining Techniques", in IEEE 2008.
- [9] Theresa Prince, R Research Scholar, J.Thomas, "Human Heart Disease Prediction System using Data Mining Techniques", in 2016 International Conference on Circuit, Power and Computing Technologies [ICCPCT], @2016 IEEE. [10] Ranganatha S.1, Pooja Raj H.R.2, Anusha C.3, Vinay S.K, "Medical Data Mining and Analysis for Heart Disease Dataset Using Classification Techniques", in IEEE 2009