COMPETING DATA MINING METHODS USED FOR HEART DISEASE PREDICTION

Ambresh Bhadrashetty[1] Sudhir Anakal[2] Priyanka Biradar[3]

[1,2]*Assistant Professor,* [3]*Postgraduate Student*
[1, 2, 3] *Department of Computer Applications, Visvesvaraya Technological University,*
Postgraduate Centre, Kalaburagi, India

Abstract— Data Mining is a process of extracting the data from a large set of any raw data. Data Mining is one of the well-known technology that is used in the Health Organization. Data mining techniques plays an important role for the Prediction of the disease by which the health care professionals can diagnosis the diseases easily. Heart Disease is caused when the normal functionality of a heart is affected. There are various factors that affect the functionality of the heart. In this paper, the different data mining techniques are used to predict the heart disease. Data Mining techniques like Neural Network, Decision Tree and Naives Bayes Theorem are used. The main goal of this paper is to analyze the various data mining techniques used, comparing these techniques, to achieve the better accuracy of the techniques and to identify the risk level of the heart disease.

Keywords— Heart Disease, Data Mining, Soft Computing, Decision Tree Techniques, Neural Network.

1. INTRODUCTION

Heart is an important organ of our body. Heart Disease is due to the abnormal condition or disorder of the heart. The various factors which increase the risk of heart disease are high blood pressure, cholesterol, family history of heart disease, obesity, hypertension, smoking etc. The most common symptom of heart disease is angina. Other symptoms include shortness of breath, palpitations (irregular heart beats or "flip-flop" feeling in your chest), a faster heart beat, sweating.

Heart disease is the leading disease that causes the death globally. Every year 12 million deaths occur over the world due to heart problem. According to the National Institutes of Health, there are about 425,000 deaths each year from heart disease.

Researchers have been applying the different data mining techniques to help health care professionals with improved accuracy in the diagnosis of heart disease. Neural Network, Naive Bayes, Genetic algorithm, Decision Tree, Classification via clustering and direct kernel self organizing map are some techniques used by the researchers. Here in this paper, the different data mining techniques used for the prediction are Neural Network, ID3 (Iterative Dichotomiser 3) a type of Decision Tree, Naives Bayes Theorem. Here, the Neural Network with 15 attributes has outperformed over all other data mining techniques. Decision Tree has shown good accuracy with ID3. Naive Bayes algorithm gives an average prediction with 90% accuracy.

Here in this paper, we have developed a decision support or computer based information system for the prediction of the risk level of the heart. The prediction of the risk level can be classified into three levels namely low risk, high risk which indicates lesser than 50%, greater than 50% and 0 respectively.

The risk level prediction system predicts the risk by reading the input attributes. Age, Gender, Blood Pressure, Cholesterol, Pulse rate are the input attributes. Here, classifying the age and gender of the patient using Artificial Neural Network (ANN) algorithm. It shows the number of patients as per gender and age category. Next, it checks the BP, Cholesterol, Pulse rate by using ID3 algorithm. If it is abnormal then verifying the patients previous history using current history. The patients previous history includes to check whether the patient smokes, consumes alcohol and tobacco.

2. LITERATURE REVIEW

According to the Indian Heart Association, 50% of all heart attacks in Indians occur under 50 years of age and 25% of all heart attacks in Indians occur under 40 years of age. Populations living in cities are three times more prone to heart attacks than people living in villages. About 610,000 people die of heart disease in the United States every year.

Kumaravel et.al. has proposed automatic diagnosis system for heart diseases using neural network. In this system ECG data of the patients is used to extract features and 38 input parameters are used to classify 5 major types of heart diseases with accuracy of 63.6-82.9%

Aditya Methaila et.al. In their research work concentrated on utilizing different algorithms and combinations of few target attributes for effective heart attack prediction using data mining. Decision Tree has outperformed with 99.62% accuracy by using 15 attributes.

B. Venkatalakshmi and M.V Shivsankar in year 2014 performed an analysis on heart disease diagnosis using data mining techniques Naive Bayes and Decision Tree techniques. Different sessions of experiments were conducted with the same datasets in WEKA 3.6.0 tool. Data set of 294 records with 13 attributes was used and the results revealed that the Naive Bayes outperformed the Decision tree techniques.

Chaitrali S. Dangare and Sulabha S. Apte showed that Artificial Neural Network outperforms other data mining techniques such as Decision Tree and Naive Bayes. In this research work, Heart disease prediction system was developed using 15 attributes. The research work included two extra attributes obesity and smoking for efficient diagnosis of heart disease in developing effective heart disease prediction system.

3. PROBLEM DEFINITION

Normally, the heart diseases can be analyzed using intuition of the medical specialist. The diagnosis is very difficult process and expensive. In order to overcome these difficulties, we have developed a decision support or computer based information system which predicts the risk level of the heart. The prediction of the risk level leads to the early involvement in the prevention of the heart diseases. If the predicted risk is high, necessary steps are taken to diagnosis the diseases.
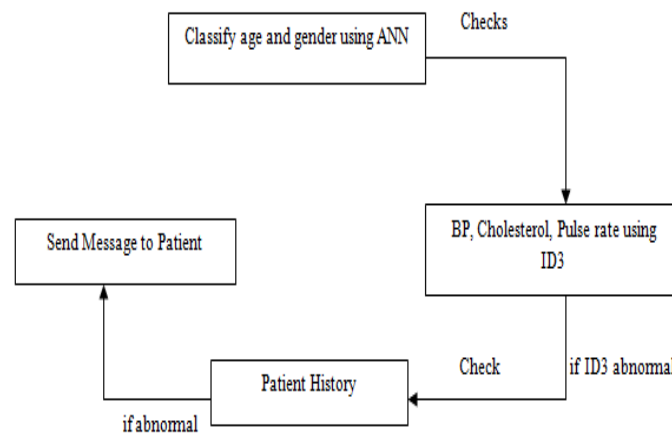
4. ARCHITECTURE

Fig. 1. System Architecture

Figure.1 represents the system Architecture of the decision support or computer based information system. Here, it classifies age and gender using ANN then checks the BP, Cholesterol, Pulse rate using ID3 algorithm. If ID3 is abnormal, then checks the patient history. If patient history is abnormal, then the message is sent to patient.
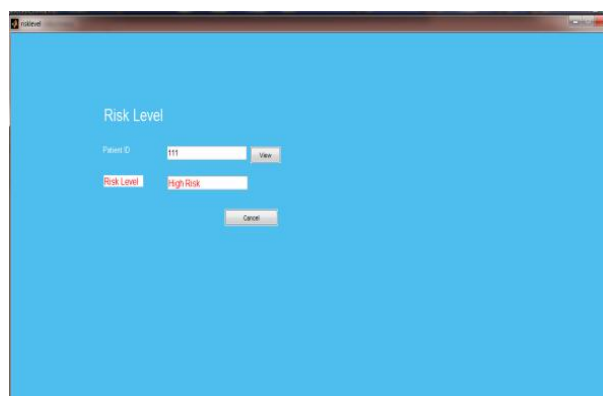
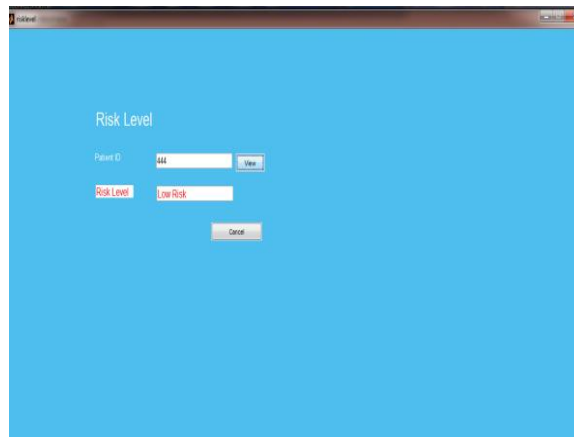## 5. RESULTS

Fig. 2. Results showing the high risk

Fig. 3. Results showing the low risk

Decision support or computer based information system predicts high risk for the patient ID 111(Fig. 2) and low risk for the patient ID 444(Fig. 3)

6. CONCLUSION

Decision support or computer based information system is more efficient so that it can predict the chances of heart attack. This system is useful for the doctors to check the risk level of the heart and further the necessary steps are taken for the accurate diagnosis of the heart diseases.

REFERENCES

[1] G. E. Sakr, I. Elhajj, and H. Huijer, "Support vector machines to define and detect agitation transition", IEEE Transactions on Affective Computing, vol. 1, pp. 98–108, December 2015.

[2] Neha Chauhan and Nisha Gautam "An Overview of heart disease prediction using data   mining techniques"

[3] Chaitrali S. Dangare, Sulabha S. Apte, ―Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques; International Journal of Computer Applications (0975 – 888) Volume 47– No.10, June 2014.

[4] Sujata Joshi and Mydhili K.Nair,"Prediction of Heart Disease Using Classification Based Data Mining Techniques", Springer India 2015, volume 2.

[5] Beant Kaur and Williamjeet Singh.," Review on Heart Disease Prediction System using Data Mining Techniques", IJRITCC, October 2016.

[6] Frank Lemke and Johann-Adolf Mueller, "Medical data analysis using self-organizing data mining technologies," Systems Analysis Modeling Simulation , Vol. 43, Issue No. 10, 2003, pp. 1399-1408

[7] B.Venkatalakshmi, M.V Shivsankar, Heart Disease Diagnosis Using Predictive Data mining; International Journal of Innovative Research in Science, Engineering and Technology Volume 3, Special Issue 3, March 2016.

[8] Beant Kaur h, Williamjeet Singh, "Review on Heart Disease Prediction System using Data Mining Techniques", International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 2 Issue: 10, pp.3003-08, October 2015.

[9] Nidhi Bhatlet and Kiran jyothi, "An Analysis of Heart disease prediction system using different data mining techniques", International Journal of Engineering Research and Technology, ISSN, volume 1,October-2016

[10] Aditya Methaila, Early Heart Disease Prediction Using Data Mining Techniques; CCSEIT, DMDB, ICBB, MoWiN, AIAP pp. 53-59, 2016.