
ABSTRACT

Large number of people has being affected due to heart disease and it has become a major problem. One of such threat is Cardiovascular Disease (CVD). If it is not treated and detected at an early and proper stage may lead to death. Especially in medical sector there is no such adequate research focus on effective analysis tool to discover and trends in data. Large amount of complex clinical data of patients and other hospital resources are being generated by health care industry nowadays. In male as well as female, it is the first leading cause of death according to world health organisation (WHO). Various heart diseases are caused due to reduction in supply of blood and oxygen. Therefore main objective of our paper is to predict more accurately the percentage of possibility of cardiovascular disease in minimum number of attributes (like blood pressure, cholesterol, type of chest pain, blood sugar, etc). Many data mining techniques are used to analyse this rich collection of data from different perspectives and deriving useful information. This project intends to design and develop, diagnosis and prediction system for heart disease based on soft computing technique that is ANFIS.

KEYWORDS: Soft Computing, Fuzzy logic, Neural Network, ANFIS, Data Mining.

INTRODUCTION

As heart is an important organ of our body, our life depends on its efficient working. Many parts of our body like liver, kidney, brain, etc also effects due to improper working of heart. If normal working of heart is changed due to some disease, whole body gets disturbed. Due to heart disease normal functionality of heart is affected. There are various factors which increases the risk of heart disease. Some of them are high blood pressure, cholesterol, family history of heart disease, obesity, hypertension, smoking, etc. In today's modern world, cardiovascular diseases are the highest flying diseases and in every year 12 million deaths occur over the world due to heart problem. In India casualties are also caused due to cardiovascular diseases and its diagnosis is very difficult process. Normally, these diseases can be analysed using intuition of the medical specialist and it would be highly beneficial if the techniques used for analysis shall be improved with the medical information system. At reduced cost, if a decision support or computer based information system is developed then it will be helpful for accurate diagnosis.

Soft Computing is one multidisciplinary system as the fusion of the fields of Fuzzy logic, Neuro-computing, Evolutionary and Probabilistic Computing. It is the combination of methodologies designed to model and enable solutions to real world problems, which are not modelled or two difficult to model mathematically.

LITERATURE REVIEW

Work done in heart disease prediction using data mining techniques and fuzzy logic are discussed below: Mai Shouman, Tim Turner and Rob Stocker[1], proposed various single and hybrid data mining techniques in heart disease prediction. Using single data mining technique for heart disease has been thoroughly investigated showing the considerable levels of accuracy. Recent investigation shows that hybridizing more than one technique, will obtain enhanced result in diagnosis. Here author applies various data mining techniques like naive bayes, decision tree on various heart disease datasets and measures the accuracy of it [8]. After applying hybrid data mining techniques on different heart disease datasets shows the different accuracies [11] and [16]. After comparing both

techniques in diagnosis on heart disease datasets, hybrid datasets showing the better accuracy than single data mining techniques.

Danashree S. Madhekar, Mayur P. Bote, Shruti D. Deshmukh[2], presents a classifier approaches for heart disease prediction and shows how Naïve Bayes classification can be used for this purpose. The system proposed have categorized medical data into five distinct categories namely no, low, average, high and very high. The analysis is done on Cleveland datasets [18] and applying the classification techniques to predict group membership for data instances [15].

Vijay Choursaiya and Saurabh Pal [3] and [9], has been done a research work on several data mining techniques used to detect heart disease. Basically thirteen attributes were involved in prediction, but this were reduced to eleven and three classifier like naive bayes, decision tree and bagging algorithm are used to predict the diagnosis of heart disease patient with same obtained before the reduction of all parameters.

Nikita, Madan lal Yadav[4] have proposed the system which uses Fuzzy rule based approach for prediction of patient disease. This methodology is functional to drive predictive analysis on single as well as on large dataset [13]. As it is the intelligent soft computing approach, it can represent the probabilistic relation based on patient symptom analysis. As work is rule based, the easy estimation of the interrelated variables can be identified to understand the approach followed by the fuzzy analysis.

The paper referred for this is by Prachi Jambhulkar, Vaidehi Baporikar [7]. This research article work presents a real time WSN system for prediction and monitoring of any upcoming cardiovascular diseases. The system has a capability of monitoring of any upcoming of any upcoming cardiovascular diseases [6]. The system has a capability of monitoring multiple patients at a time and delivers remote diagnosis and prescription to the patients it also provides fast and effective warning to doctors, relatives and hospitals. From this paper we get an idea of using wireless sensor network, we can enhance and expand the model with combination of WSN system and data mining techniques for getting accuracy and more real time data sets in prediction of various cardiac diseases.

Aqueel Ahmed Shaikh Abdul Hannan [5], presents the research paper to find out the various cardiovascular disease through data mining, genetic algorithm, support vector machine (SVM), rough set theory, association rules and neural networks. After study, we examined that from the above said techniques decision tree and SVM are the most effective for cardiac disease amongst all [12]. Hence, we observed that, data mining could helpful in the identification of high or low risk cardiac diseases. The three classification function techniques in data mining like naïve bayes, decision tree and classification via clustering [10] are compared for predicting cardiac disease with reduced number of attributes has been presented by Shamsher.

Following table shows comparison of various techniques used for prediction of heart disease:

Serial No.	The Algorithm Used	Accuracy (%)
1.	Decision Tree	76
2.	Association Rule	55
3.	K-NN	58
4.	Artificial Neural Network	85
5.	Naive Bayes	69

Table 1. Performance study of different approaches used in diagnosis of heart disease

PROPOSED WORK

Soft Computing aim is to exploit tolerance for imprecision, uncertainty, approximate reasoning, and partial truth in order to achieve close resemblance with human like decision making.

Lotfi A. Zadeh 1992, has defined soft computing as: “ It is an emerging approach to computing which parallel the remarkable ability of the human mind to reason and learn in an environment of uncertainty and imprecision”.

Soft Computing is not a mixture or combination; rather it is a partnership in which each of the partners contributes a distinct methodology for addressing problems in its domain. In principal the constituent methodologies in soft computing are complementary rather than competitive.

Block Diagram of our proposed system is as follows:

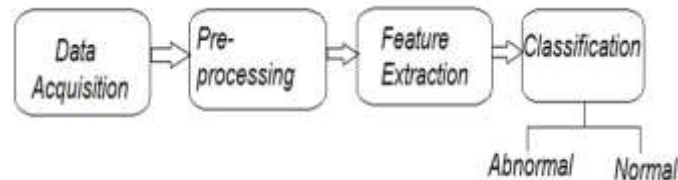


Fig 1: Block Diagram for proposed system

Datasets of heart disease patients can be collected from various Universities like UCI, Cleveland , etc for our intelligent system. The attributes like age, sex, chest pain, resting blood pressure, cholesterol in mg/dl, blood sugar, maximum heart rate, etc.

Data pre-processing is done to extract relevant data and then those data should be converted into the format necessary for the prediction of risk. Cleaning and filtering of datasets is done sometimes to remove duplicate records, normalize the values, accounting for missing data and removing irrelevant data items.

Then features are extracted for classification purpose. Now finally in last stage we will get the result on doing classification as normal or abnormal condition of the patient.

The work projected a system that uses method called ANFIS (Adaptive Neuro-Fuzzy Inference System) for heart disease diagnosis. Adaptive Neuro-Fuzzy Inference System projected by Jang in 1993, is a grouping of two machine learning approaches: Neural Network (NN) and Fuzzy Inference System (FIS).

Fuzzy Computing

In the real world there exists much fuzzy knowledge, that is, knowledge which is vague, imprecise, uncertain, ambiguous, inexact, or probabilistic in nature. Human use such information because the human thinking and reasoning frequently involve fuzzy information, possibly originating from inherently inexact human concepts and matching of similar rather than identical experiences. We want, the computing systems should not only give human like answers but also describe their reality levels. These levels need to be calculated using imprecision and the uncertainty of facts and rules that were applied.

Let us consider a simple example of Heap Paradox. This example represents a situation where vagueness and uncertainty are inevitable. If we remove one grain from a heap of grains, we will still a heap. However, if we keep on removing one-by-one grain from a heap of grains, there will be a time when we do not have a heap anymore. The question is, at what time does the heap turn into a countable collection of grains that do not form a heap? There is no one correct answer to this question. Like this questions are being solved by using this fuzzy computing by defining rules (if then or else) in it.

Neural Computing

Neural Computers mimic certain processing capabilities of the human brain. Neural Computing is an information processing paradigm, inspired by biological system, composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems.

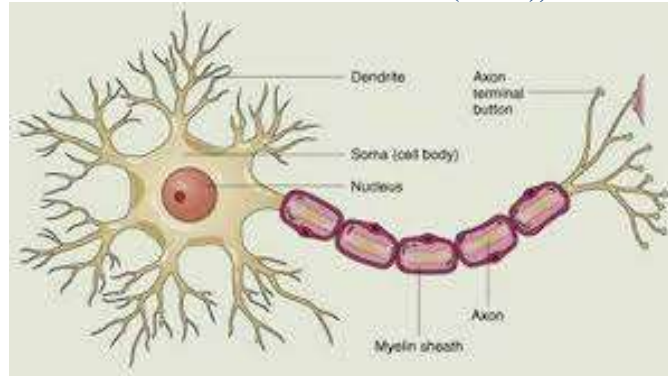


Fig 2: Structure of neuron in human brain

A neural net is an artificial representation of the human brain that tries to stimulate its learning process. The term “artificial” means that neural nets are implemented in computer programs that are able to handle the large number of necessary calculations during the learning process. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. In ANN also learning in biological systems involves adjustments to the synaptic connections that exist between the neurons.

Basic flow diagram of computations in ANFIS is as follows:

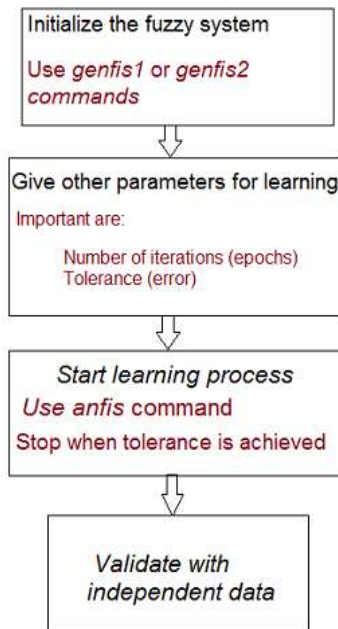


Fig 3: Flow diagram of computations in ANFIS

STANDARD VALUES OF BASIC ATTRIBUTES

1. Age: in years
2. Sex (value : Male or Female)
3. Chest Pain(value 1:typical type 1 angina, value 2: typical type angina, value 3: non-angina pain, value 4: asymptomatic)
4. Fasting Blood Sugar(value1: >120 mg/dl; value0: <120 mg/dl)

5. Cholestrol (mg/dl)
6. Blood Pressure

CONCLUSIONS

After studying various data mining and soft computing techniques, we can conclude that ANFIS is better prediction method for heart disease. So there is scope in designing a prediction system whose accuracy is more than present system using less number of attributes compared to others.

In this proposed system we will be using ANFIS as a soft computing technique for diagnosis. We are aiming to designing an intelligent system using soft computing method. As soft computing technique proves to be better in giving human like decisions.

REFERENCES

- [1] Mai Shouman, Tim Turner, and Rob Stocker “Using Data Mining Technique in Heart Disease Diagnosis and Treatment”, 2012 Japan-Egypt Conference on Electronics, Communications and Computers 978-1-4673-0484-9/12@ 2012 IEEE.pp.173-177.
- [2] Dhanashree S. Medhekar, Mayur P. Bote, Shruti D. Deshmukh “Heart Disease Prediction System using Naïve Bayes”, International Journal of Enhanced Research In Science Technology & Engineering, ISSN NO: 2319-7423 VOL. 2 ISSUE 3, MARCH- 2013.p.p. 1-5
- [3] Vikas Chaurasia and Saurabh Pal “Data Mining Approach to Detect Heart Diseases”, International Journal of Advanced Computer Science and Information Technology (IJACSIT) Vol. 2, No. 4, 2013, Page: 56-66, ISSN: 2296-1739.
- [4] Nikita, Madam lal Yadav “A Fuzzification Approach for Prediction of Heart Disease”, International Journal of Engineering Trends and Technology (IJETT)- Volume 4 Issue 5 May2013. ISSN: 2231-5381.
- [5] Aqueel Ahmed, Shaikh Abdul Hannnan “Data Mining Techniques to Find Out Heart Diseases: An overview”, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-1, Issue-4, September 2012,pp.18-23.
- [6] Deepali Chandna “Diagnosis of Heart Disease Using data Mining Algorithm”,International Journal of Computer Science and Information Technologies (IJCSIT), Vol. 5(2), 2014 pp. 1678-1680.
- [7] Prachi Jambhulkar, Vaidehi Baporikar “Review on Prediction of Heart Disease Using Data Mining Technique with Wireless Sensor Network”, International Journal of Computer Science and Applications, Vol. 8, No. 1, Jan-Mar 2015, ISSN: 0974-1011
- [8] Nidhi Bhatala, Kiran Jyothi, “ An Analysis of Heart Disease Prediction using Different Data Mining Techniques”, International Journal of Engineering Research Technology (IJERT) ISSN:2278-0181 Vol.1, Issue-8, October-2012.
- [9] Vikas Chaurasia and Saurabh Pal “Early Prediction Heart Disease Using Data Mining Techniques”, Caribbean Journal of Science and Technology, 2013, ISSN 0799-3757, Vol.1, pp. 208-217.
- [10] Shankar B. Patil, “Intelligence and Effective Heart Disease Prediction System Using Data Mining And Artificial Neural Network”, European Journal of Scientific Research ISSN: 0975-3397 Vol. 3 No. 6, June 2011.
- [11] M. A. Jabbar, “Knowledge Discovery from Mining Association Rule for Heart Disease Prediction”, Journal of Theoretical and Applied Information Technology ISSN: 1992-8643, E- ISSN: 1817-3195, 2005.
- [12] T. Srinivasan. “Knowledge Discovery in Clinical Databases with Neural Network Evidence Combination”.
- [13] T. Revathi, S. Jeevitha, “Comparitive Study of Heart Disease Prediction System Using Data Mining Techniques”, International Journal of Science and Research (IJSR) ISSN: 2319-7064.
- [14] Abhishek Taneja, “Heart Disease Prediction System Using Data Mining Techniques”, Oriental Journal of Computer Science and Technology, ISSN: 0974-6471.
- [15] Dr. Hari Ganesh S., Gajenthiram M, “Comparitive Study of Data Mining Approaches for Prediction of Heart Disease”, IOSR Journal of Engineering ISSN (e): 2250-3021, ISSN: 0974-6471.
- [16] K. Sundhakar, Dr. M. Manimekalai, “Study of Heart Disease Prediction using Data Mining ”, Internatinal Journal of Advanced Research on Computer Science and Software Engineering, ISSN: 2277 128X.
- [17] www.myreaders.info/html/soft_computing.html

- [18] Ms. Shinde S. B., Prof. Amrit Priadarshi, “Decision Support System on Prediction Using Data Mining Techniques”, iPGCON-2015.
- [19] Jyoti Soni, Ujma Ansari, Dipesh Sharma, Sunita Soni, “Prediction Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction”, International Journal of Computer Application (IJCA) ISSN: 0975-8887.
- [20] N. Sathiya Rani, K. Vimala, Dr. V. Kalaivani, “Health care Monitoring for the Detection using Soft Computing Techniques”, International Journal in Foundation of Computer Science and Technology (IJFCST).