General Article

Impelling Heart Attack Prediction System using Data Mining and Artificial Neural Network

J.E.Nalavade[A], M.L.Gavali[A], N.D.Gohil[A] and S.C.Jamale[A*]

[A]Department of Computer Engineering, Sinhgad Institute of Technology, Pune-410401, India

Abstract

Diagnosis of diseases is an important and intricate job in medicine. The identification of heart disease from diverse features is a no of layered problem that is not free from the wrong assumptions and is frequently accompanied by impulsive effects. Thus to exploit knowledge and experience number of specialists and clinical screening data of patients inserted in databases to assist the diagnosis procedure is regarded as a valuable option. This system work is the extension of our previous system with intelligent and effective heart attack prediction system by using neural network. A professional methodology for the extraction of easiest patterns from the heart disease warehouses for heart attack prediction has presented. Data warehouse is preprocessed in sequence to make it easy for the mining process. Processing gets finished, then heart disease warehouse is clustered with aid of K-means clustering algorithm, which will extract data, appropriate to heart attack from warehouse. The frequent patterns applicable to heart disease are mined with aid of the algorithm from data extracted. The patterns important to heart attack prediction are selected on basis of the significant weight. The neural network is well trained with selected significant patterns for effective heart attack prediction system. We have implemented the Multilayer Neural Network with Back-propagation training algorithm. Results obtained have illustrated that designed prediction system is capable of predicting the heart attack more effectively.

Keywords: Hidden layer, Back Propagation, Data Mining, Artificial Neural Network (ANN), K-means clustering.

1. Introduction

As previous, the prediction of heart attack was manual process for doctors and it was very time consuming for new doctors to identify the level of heart attack that is was low or high. Always practitioners new doctors had to meet with the experienced doctors for their proposed and also it was going very more costly.

Following as result of the above illustrated issues, some group of experienced doctors thought to developed their considerations and charts related heart attack prediction system. They have extracted all the scene indication and consideration regarding to the heart attack (age, smoking, tobacco etc) and accordingly they have suggested the total amount of risk to particular combination of indication. This approach of doctors was very helpful for the new practitioner's doctors to identify the amount of risk level of patient related heart beats. Regarding this work some software team presented the system of prediction level of the heart attack using different approaches wise using naïvebaiyes, decision trees. But we present system of intelligent and effective heart attack prediction system using artificial neural network and data mining. In this heart attack Prediction system we are using K-means clustering algorithm and

Artificial Neural Network (ANN). Presented system provides 100% accuracy in that prediction system as compare to the decision trees and naïvebaiyes system. System should provide correct and accurate results. Medical diagnosis is important but complicated task that should be performed accurately. Why we are implementing the reasons behind that All Doctors Do Not Possess Expertise, Make Intelligent Clinical Decisions, Reduces Financial Pressures, Influence Revenue and Operating Efficiency. Medical diagnosis is an important but complicated task that should be performed accurately and efficiently and its automation would be very useful and advantageous. Unfortunately, all doctors do not possess expertise in every sub specialty and moreover they are in many places a scarce resource. However, appropriate computer-based information and/or decision support systems can aid in enhancing medical care and in achieving clinical tests at a reduced cost.

The main motivation of this research is to process data in order to get useful information that can enable healthcare practitioners to make intelligent clinical decisions.

Such analysis has become increasingly essential as financial pressures have heightened the need for healthcare organizations to make decisions based on the analysis of clinical and financial data. Insights gained data mining can

*Corresponding author: **S.C.Jamale**

influence cost, and operating efficiency while maintaining a high level of care. The World Health Organization has estimated every year due to the Heart diseases. Half deaths in the Uk and other developed countries occur due to cardio vascular. It is reason of deaths in developing countries. It is regarded as the primary reason deaths in adults. The Heart disease is the diverse diseases that affect the heart. Heart disease was the major cause of disease in the different countries in also India. It kills one person every 34 seconds in the United States.

2. Related Work

We are implementing The Heart attack prediction system using artificial neural network and in that we are using k-means and artificial neural network which provides up to 100% accuracy. In the past, the prediction of heart attack was the manual task for the doctors to predict the risk level of the attack but nowadays it is very difficult for the new doctors to predict the risk level of patient because they do not have certain amount of experience like expertise hence they need to communicate to the expertise always for their feedback or their suggestions. For this the naïve bays was the existing system in the prediction of the heart attack risk level. but the drawback of the naïve bays system is that it cannot predict the risk level accurately. hence we are implementing the intelligent and effective heart attack prediction system using neural network and in this concept we are using the k-means algorithm and artificial neural network in that we are providing the various and k-means will determine the effective set of values as a input to the Ann algorithm and the output of the k-means algorithm will be given to the artificial neural network in this the possible combination of the inputs get proceeded and accordingly the prediction level will get determined. And we are making this system as a two sided like client and server means if patient wants to predict the risk level of their heart attack personally then patient can use the feeded database in the system but he cannot do any modification in their reports as already generated by their doctors.

Firstly we have referenced the paper design and analysis of computational model for heart disease diagnosis published in international journal paper of computer science and mobile computing 2013 this paper presents the prediction of heart attack using the clustering algorithms and also states the best accuracy among all the specified algorithms. This research paper proposed a frequent feature selection method for Heart Disease Prediction. Using medical profiles such as age, blood pressure, blood sugar. it can identify the patients getting a heart disease. The objects which is having same meaning, proposed approach improves the accuracy and reduces the computational time and the advantages of this paper is that it provides the better accuracy as compare to other traditional algorithms also it considers the low and high level of LDL that is of the bad cholesterol and gives efficient prediction but the disadvantage is that it cannot predict up to 100% accuracy.

Then we have looked the prediction system for heart disease using naïve bayes published in international journal of advanced computer and mathematical sciences in 2012 by A.S.pattekari and asmaparveen and The main objective of this paper is to develop an Intelligent System using data mining and Naive Bayes. It is implemented as web based application in this user answers the predefined questions. It retrieves hidden data from stored database and compares the user values with trained data set. It can answer complex queries for diagnosing heart disease and thus assist healthcare practitioners to make intelligent clinical decisions which traditional decision support systems unable.. the advantage of this paper is that it reduces the medical errors also enhance the patients safety and decrease unwanted practice variations but the disadvantages of this implementation is that it is not up to the accurate level means it cannot predict the accurate result and the output varies by the same input hence this implementation is not effective.

The next paper which we have referenced that is Intelligent and effective heart attack prediction system using data mining and artificial neural network in that they have illustrated the k-means and MAFIA algorithm and for the prediction they have used the back propagation algorithm and it provides the accuracy up to 100% also gives proper grouping of data set and data mining for retrieving the results hence we are also using the same concept and making it as a client server application in the system so that patient can also predict their risk level and also updation of system database will be done in the computer as application get update by the administrator. Hence we have determined the difference between the naive baiyes and artificial neural network is that naïve bayes uses the same inputs but the prediction of the output result is not that much accurate and if we provide same inputs at different time then output is also getting variases instead getting the same result by providing the same input and the artificial neural network gives the 100% accuracy by providing the various inputs and it predicts the various combination of inputs and gives the correct output.

Hence we are implementing the two sided intelligent and effective heart attack prediction system using data mining and neural network for giving the accuracy of heart attack prediction up to 100%

3. Overview Description

Efficient Prediction

As we are considering K-means clustering and artificial neural network algorithms. So level of prediction becomes more efficient to understand.

Import Data Set

Always we are importing data set from our database in this proposed system.

Authentication

For every upcoming doctor, unique user id and password is provided. Also at the time of when accessing system, verification is done on the basis of this user id and

password.

Apply clustering

Well formed data set will be generated by using K-means clustering algorithm.

Predict new records

Every time new records are made according to the upcoming new patient, that records will match with stored records.
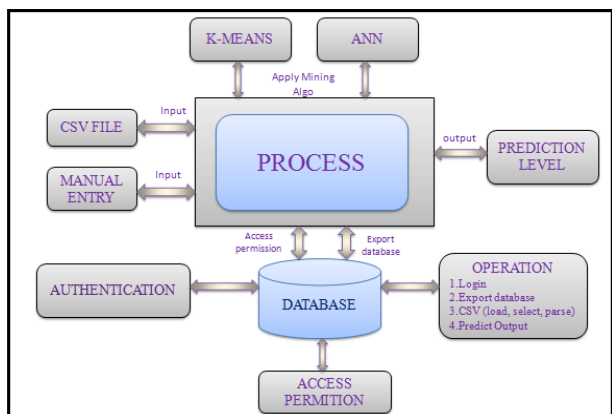
System Architecture

Fig.1 System Architecture

In this system we are implementing effective heart attack prediction system using K-means and ANN algorithm. We can give the input as in csv file or manual entry to the system. After taking input the two algorithm apply on that input that is K-means and ANN. From the database authentication is checked out. Database access permission is also granted from database. In this system our back end is serialization. After accessing database operation is performed and effective heart attack level is produced. System architecture is shown in fig 1.

K-means Algorithm

In k-means problem, we having given a number of set N of points in Kp, and integer $p \geq 1$, and we want to find Kp points ("centre") so minimize the summation of the square of the Euclidean distance each point in N to its closest center. This is well-known defined, and most popular clustering problem that has received allot of attention in the clustering algorithms community.

Kp = # of clusters per cluster Interval data Initialize means by taking Kp samples randomly.
Complexity is given as O(n * Kp * I * d)
n = it denotes number of points,
Kp= it denotes number of clusters,
I = it denotes number of iterations
d = it denotes number of attributes
Easily parallelized and Use efficient spatial data structures for some situations.

A very good clustering with smaller Kp (clusters) can have a lower SSE than a poor clustering with higher Kp (clusters). K-means has problems when clusters are of Differing Sizes , Densities, Non-globular shapes. The flowchart of the K-means algorithm is given in fig 2.
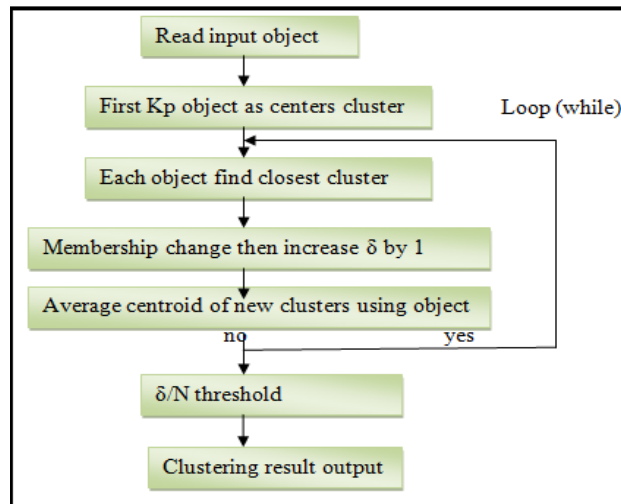

Fig. 2 K-means Flowchart

Artificial Neural Network (ANN)

ANN called **neural network** is a computational model that is inspired by the structure and functional aspects of biological neural networks. Network term 'Artificial Neural Network' considering to the interconnections between the neurons on the different layers of each system. ANN is defined by three types of parameters:
1. First pattern is to interconnection pattern between different layers of neurons.
2. The second pattern is to learning process for updating the weights of the interconnections.
3. The third pattern is to activation function that converts a neuron's weighted input to its output activation.
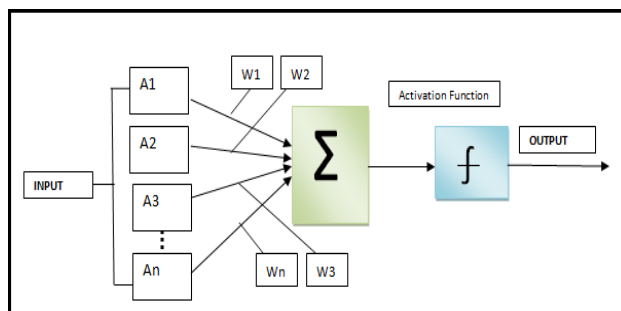
MLPNN

Fig.3 Multilayer Perceptron Neural Network

It is the Multilayer Perceptron Neural Network in that Perceptron is one of the pioneer neural network system models. Perceptron is a neuron by taking a weighted addition of inputs and passing the output 1 when if the addition is higher than some adjustable threshold value. Otherwise it passes 0. Nothing spiking described in the

biology and also see the neuron throwing section above fig.)it is also called an activation function.

Inputs **(A1,A2,A3,A4,A5..An)** and the connection weights **(w1,w2,w3..Wn)** in Figure are typically real values and that both positive (+) and negative (-) values. The feature of some **(Ai)** tends to cause the perceptron to throws and the weight **(Wi)** will be positive; if the feature **(Ai)** inhibits the perceptron, the weight **(Wi)** will be negative.

New Network show Like

New Artificial neural Network that solve the XOR problem and look like to a single layer network. The new network shown in fig 4.
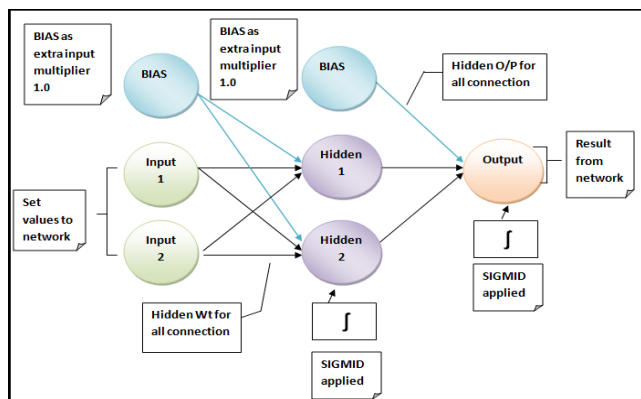

Fig.4 Artificial Neural Network with hidden layer

In this one input layer, one output layer, and in the middle one hidden layer. Inputs and weights that we are capable to work out with the activation for a node. Effectively achieved for the hidden layer and it has directly connected to the links real input layer.

$$A = \sum_{n=1}^{N+1} Wi * Ii$$

Where Wi is the weight(i) value and Ii is the input(i) value.

5. Implementation Results

The implementation result analysis in identifying significant patterns for heart attack prediction system are presented in this part. We are implemented this system approach in Java. Existing work, the heart attack dataset we have used for our experiments was obtained from. Data set is pre-worked successfully by removing duplicate results, records and also supplying missing parameter values to the system. Purified heart disease system data set, implementation result from preworked and then clustered using K-means clustering algorithm with K value as 2. The significant patterns are discribed with the aid of the significance weightage which is greater than the pre-defined threshold values. The parameter values related to each attribute in the significant patterns are as follows : 1. Higher heart rate is greater than 100 beats/ minute. 2. blood pressure range is greater than 140/90 mm Hg.

3.cholestoral range is greater than 240 mg/dl. 4. The age range in youth, middle age, senior. 5. abnormal ECG and unstable angina etc.

In our proposed system we added some more parameters significant to heart attack with their weightage and the priority levels are by consulting expertise doctors and the medical experts. Heart attack prediction system implemented using the MLPNN with Back-propagation as training algorithm . our heart attack prediction system designed to help the identify different risk levels of heart attack like normal, low or high and also giving the prescription details with related to the predicted result. Combinations of the heart attack parameters values for normal, low and risk level along with their values and weightage are detailed shown in below Table.1 shows the parameters of heart attack prediction with corresponding values and their weightage.

Table 1: Parameters with related values and their weightage

Parameter	Weightage	
Male and Female	Age < 30	0.1
	>30 to <50	0.3
	Age>50 and Age <70	0.7
	Age>70	0.8
Exercise	Never	0.6
	Regular	0.1
	High If age < 30	0.1
	High If age > 50	0.6
Overweight	Yes	0.8
	No	0.1
High salt diet	Yes	0.9
	No	0.1
Smoking	Never	0.1
	Past	0.3
	Current	0.6
Hereditary	Yes	0.7
	No	0.1
Blood Pressure	Normal (130/89)	0.1
	Low (< 119/79)	0.8
	High (>200/160)	0.9
Heart Rate	Low (< 60bpm)	0.9
	Normal (60 to 100)	0.1
	High (>100bpm)	0.9
Sedentary Lifestyle/inactivity	Yes	0.7
	No	0.1
Alcohol Intake	Never	0.1
	Intake Past	0.3
	Current	0.6

The level of heart prediction is predicted using the parameter result. The eg of heart attack levels shows in Table 2.

Future Scope

In this paper, we have enhanced the method of effective heart attack prediction system using data mining and the ANN which provide the more accurate result. It is desktop application where running on stand alone in desktop system. In future work, We intend to apply on high

Table 2: Level of Prediction system

Risk level - Normal	Risk level – Low or High	Risk level - High
Male And age 30	Male And Age > 30 age < 50	Female And Age >70
Smoking = Never	Smoking= Current	Smoking = Never
Overweight = No	Overweight = Yes	Overweight = Yes
Alcohol = Never	Alcohol= Current	Alcohol = Past
High saturated fat diet = No	High saturated fat diet = No	High saturated fat diet = Yes
High salt diet =No	High salt diet = Yes	High salt diet = Yes
Exercise= Normal	Exercise = High	Exercise = Never
Bad Cholesterol =Low	Bad Cholesterol = High	Bad Cholesterol = High
Blood Sugar = Normal	Blood Sugar = High	Blood Sugar = High
Blood Pressure = Normal	Blood Pressure = Low	Blood Pressure = High
Heart Rate = Normal	Heart Rate = Low	Heart Rate = High

performance client server or parallel processing computer and optical neural network as a classifier model.. Using data mining techniques to help healthcare professionals in diagnosing and providing suitable treatments for heart disease and Continuous data can also be used instead of just categorical data. The client server application installed that request and receives information over the network so it would easy to access this application to all users.

Conclusion

We proposed system to predict the risk level of heart attack for the patients. We are using Data Mining and Artificial neural network, by using these techniques the prediction level will increase and more accuracy is provided. This system is very helpful for the new practitioner's doctors to identify the amount of risk level of patient related heart beats.

References

B. Shantakumar (2009), Intelligent and Effective Heart Attack Prediction System Using Data Mining and Artificial Neural Network, *European journal of scientific research*,vol.3.1 no.4, pp 642-656

G. Subbalakshmi, K.Ramesh M.ChinnaRao (Apr-May 2011), Decision Support in Heart Disease Prediction System using Naive Bayes, *indian journal of computer science and engineering*

S. Adam, P. Asma (2012), Prediction System For Heart Disease using Naïve Bayes, *International journal of advanced computer and mathematical sciences*, vol 3,pp 290-294.

S.Chaitrali, S.Apte (June 2012), Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques, *journal of computer science and application,* vol 47-no10, pp 0975-888

N.Bhatla,K.Jyoti (October – 2012) An Analysis of Heart Disease Prediction using Different Data Mining Techniques, ISSN: 2278-0181 Vol. 1 Issue 8,

Authors

Jagannath Nalavade- Asst.Professor, Department of Computer Engineering , Sinhgad Institute of Technology, Pune-410401, India
Gavali Monica- Department of Computer Engineering, Sinhgad Institute of Technology, Pune-410401, India
Gohil Neha- Department of Computer Engineering, **Jamale Sujata-** Department of Computer Engineering, Sinhgad Institute of Technology, Pune-410401, India