Global Journal of Computer Science and Technology: E
Network, Web & Security
Volume 16 Issue 5 Version 1.0 Year 2016
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

An Extensive Investigation on Coronory Heart Disease using Various Neuro Computational Models

By D. Rajeswara Rao & Dr. JVR Murthy
KL University

Abstract- The diagnosis of heart disease at the early time is important to save the life of people as it is absolutely annoying process which requires extent knowledge and rich experience. By and large the expectation of heart infections in conventional method for inspecting reports, for example, Electrocardiogram - ECG, Magnetic Resonance Imaging- MRI, Blood Pressure- BP, Stress tests by medicinal professionals. Presently a-days a huge volume of therapeutic information is accessible in restorative industry in all maladies and these truths goes about as an incredible source in foreseeing the coronary illness by the professionals took after by appropriate ensuing treatment at an early stage can bring about noteworthy life sparing. There are numerous systems in ANN ideas which are likewise contributing themselves in yielding most elevated expectation precision over medical information. As of late, a few programming devices and different techniques have been proposed by analysts for creating powerful decision supportive systems.

More over many new tools and algorithms are continued to develop and representing the old ones day by day. This paper aims the study of such different methods by researchers with high accuracy in predicting the heart diseases and more study should go on to improve the accuracy over predictions of heart diseases using Neuro Computing.

Keywords: *artificial neural networks (ANN), heart diseases, neuro computing.*

GJCST-E Classification : *C.1.3 F.1.1*

Strictly as per the compliance and regulations of:

An Extensive Investigation on Coronory Heart Disease using Various Neuro Computational Models

D. Rajeswara Rao [α] & Dr. JVR Murthy [σ]

Abstract - The diagnosis of heart disease at the early time is important to save the life of people as it is absolutely annoying process which requires extent knowledge and rich experience. By and large the expectation of heart infections in conventional method for inspecting reports, for example, Electrocardiogram - ECG, Magnetic Resonance Imaging- MRI, Blood Pressure- BP, Stress tests by medicinal professionals. Presently a-days a huge volume of therapeutic information is accessible in restorative industry in all maladies and these truths goes about as an incredible source in foreseeing the coronary illness by the professionals took after by appropriate ensuing treatment at an early stage can bring about noteworthy life sparing. There are numerous systems in ANN ideas which are likewise contributing themselves in yielding most elevated expectation precision over medical information. As of late, a few programming devices and different techniques have been proposed by analysts for creating powerful decision supportive systems.

More over many new tools and algorithms are continued to develop and representing the old ones day by day. This paper aims the study of such different methods by researchers with high accuracy in predicting the heart diseases and more study should go on to improve the accuracy over predictions of heart diseases using Neuro Computing.

Keywords: artificial neural networks (ANN), heart diseases, neuro computing.

I. Introduction

Heart is most paramount part of the body. Life itself is dependent on the efficient working of the heart. Any realistic predicament in heart has an instantaneous impact on the survival of concerned person that it affects different components of the body. Heart disease is the disease predicated on the performance of the heart. Several factors increases risk of Heart disease like cholesterol, lack of physical exercise, high blood pressure, smoking and exorbitant corpulence. At present, most of the people are suffering from heart disease so there is a need to precise diagnosis at early stages then followed by subsequent treatment that can result the preserving of life. The incipient data relinquished by the National Heart, Lung, and Blood Institute (NHLBI) shows that especially women in the older age groups are in peril of getting heart disease than other people. Recent study withal verbalized that Heart disease can be controlled, if it is diagnosed at an early stage. But it's not facile to control and to do precise diagnosis because of many perplexed factors of heart diseases, like many clinical symptoms are linked with other human organs also, often heart diseases can exhibit sundry syndromes. In-order to scale back the analysis time and to amend the analysis precision, it has end up extra of a inductively sanctioning, to enhance the nontoxic and the puissant clinical determination aid methods to improve the analysis decision approach. Medical diagnosis is intricate process; hence the approach is to develop an accurate system.

II. Literature Survey

To gain the background knowledge this paper presents a literature survey on neuro computing Techniques for diagnosing Heart disease.

Laercio Brito Gonçalves etal (2016) shown that the Inverted Hierarchical Neuro-Fuzzy Binary house Partitioning which was centred on the Hierarchical Neuro-Fuzzy Binary area Partitioning model (HNFB) that gave an proposal of recursive partitioning to allowed a large number of inputs. The classification method of HNFB-1 has been evaluated with exclusive benchmark databases akin to heart ailment datasets. For interpretable fuzzy ideas it allowed the knowledge extraction [1].

Durairaj M, Revathi V (2015) proposed newly system to obtain more accuracy using back propagation multilayer perceptron (MLP) Algorithm of neural networks than the other neural networks. It is a popular effective method of ANN training network with some optimized techniques like gradient desent where it propagates back to hidden layer. This learning rule moves the network down the steapest slope in error space. The method computes the depth of the loss function in the input data with respect to all the weights in the network. As back propagation algorithm necessitates the activation function as it is applied to multilayer feed forward networks which needs differentiable activation functions. The dataset used for experimentation is Information of heart disease dataset taken from UCI machine learning repository called

Author α: Department of Computer Science and Engineering, KL University, Andhra Pradesh. e-mail: rajeshduvvada@kluniversity.in
Author σ: Professor, Dept of computer science & engineering, JNTU-Kakinada, Andhra Pradesh.

Cleveland dataset with 14 attributes, 303 Instances and after cleaning of data they have taken 297 Instances out of 303 Instances. MLP back propagation is trained with the inputs that are adjusted purelin function automatically for increasing the output accuracy. The main aim is to minimize the average sum of errors. The feed forward back propagation algorithm secures highest accuracy of 96.30%. so that the experimental results showed that this algorithm can be effective to predict the heart disease with more accuracy [2].

Noura Ajam (2015) stated that artificial neural networks (ANN) shows the significant results in heart disease diagnosis. The architecture of neural network is formed by number of processing units (Neurons) and connections between them. A subgroup of processing elements is called layer. The number of neurons and the layers depends upon the complexity of the system studied. ANN is widely used in medical diagnosis and health care applications because of their predictive powerful classifier for tasks, fault tolerance, generalization and learning from environment. ANN is unsupervised learning type provided only with inputs, but no known targets. It is self organized. The Dataset used is Cleveland dataset which consists of 14 attributes and 303 instances. ANN is trained using back propagation learning algorithm on heart disease data. Input and target samples are divided as 60% training set, 20% validation set, 20% test set. The activation function of tangent sigmoid for hidden layers and linear transfer function for output layer is used. Mean square error "MSE" is calculated which is equal to 0.1071 and the classification accuracy for heart disease is 88% [3].

S. Florence, N.G. Bhuvaneswari Amma, G.Annapoorani, K.Malathi (2014) proposed the system which uses the Neural Network and the Decision Tree (ID3) for the prediction of heart Attacks. The dataset used is acath Heart attack dataset provided by the UCI machine Learning Repository which has attributes are considered to diagnose the heart attacks. CART , ID3, C4.5 decision tree algorithms used Gini index to measure the impurity of a partition or set of training attributes. The dataset contains 6 attributes like age, sex, cardiac duration, signal, possibility of attack. The final one is the class label. Depends upon the attribute values present in the dataset the corresponding class label that is the prediction is happening at the final stage. For training and testing the network where 75 percent is used for training and 25% is used for testing the system. The knowledge obtained from the classification is used to test the system. When comes to the neural network the input layer has 6 nodes, the hidden layer has 3 nodes and the output layer consists of 2 nodes. Finally it shows 2 outputs, that is the possibility of heart attacks. The prediction is done using the tool called Rapid Miner Studio. Results have been generated by using the Decision Tree as well as the Neural networks. The graphs have been plotted and these are generated in a simulator called Rapid Minor Tool. They had just drawn the results and have predicted whether is there is an attack or not using these Networks [4].

Hlaudi Daniel Masethe etal (2014) discussed data mining classification algorithms for predicting the heart attacks are J48, Naïve Bayes, REPTREE, CART and Bayes. The aim is to predict possible heart attacks using data mining techniques from the patient dataset and determines the best model which gives the correct predictions of highest percentage for diagnoses. From medical practitioners, the dataset of eleven attributes are collected for the predictions of heart disease. The attributes are labelled as Patient Id-Number, Gender of the patient, Cardiogram report, Age of the patient, Chest Pain type, Blood Pressure Levels, Heart Rate, Cholesterol levels, Smoking habitat, Alcohol consumption and levels of Blood Sugar. For the prediction, the weka data mining tool is used to analysis to discover the patterns. The algorithms have been applied on the data set and thus the results have been obtained and they had observed that the J48, REPTREE and SIMPLE CART show a prediction model of 89 cases with a risk factor positive for heart attacks. The best classification technique to be J48, REPTREE and SIMPLE CART algorithm perform similar in this data set, while Bayes Net algorithm out-performed the other techniques. Thus the algorithms results do not show any difference in the prediction when using different classification algorithms in data mining [5].

Jayshril S. Sonawane, D. R. Patil (2014) evaluated the prediction approach for coronary heart sickness making use of studying LVQNN algorithm. The neural network on this algorithm had 13 attributes of input and predicts the presence or absence of heart ailment of sufferer. The prediction method is based on ANN. Synthetic neural network is an understanding processing procedure that strategies the expertise in an identical method because the biological apprehensive system techniques. On this technique neural network considers that they've got competencies to derive that means from tricky or imprecise capabilities which possibly used to extract specified patterns and discover trends which can be extra elaborate to be seen through both humans or other pc techniques and approaches. The essential cause of utilizing LVQ is that it creates prototypes which can be effortless to interpret for specialists within the respective utility area. Learning vector Quantization is aggressive network uses the supervised learning methodology which contains two layers specifically competitive layer and linear layer. The learning vector quantization algorithm is applied on the Cleveland heart ailment database. This suggests that the prediction approach offers higher performance consequently giving us an effective method for the prediction for the heart sickness [6].

Jesmin Nahar and Tasadduq Imam etal (2013) examined the actual fact of computational smart techniques in coronary disorder discovering. Cleveland knowledge was once used to participate in connection with six comprehended classifiers. For most classifiers and higher section knowledge set the execution was once elevated by way of encouraged feature decision. They developed an efficient algorithm for rule extraction test on coronary heart ailment information for various associative rule mining algorithms such as Apriori, Predictive Apriori and Tertius for the analysis of coronary heart diseases diagnosis [7].

Sanjeev Kumar, Gursimranjeet Kaur (2013) detected the heart ailments in individual by means of utilising the Fuzzy expert system. The designed approach is based on two hospitals dataset and international lab database. Comparative analysis is finished between these two hospitals dataset and the lab database methods. Via utilising the fuzzy knowledgeable approach the diagnosis of heart disease has been carried out which consists of six inputs and a pair of outputs. The six attributes are chest anguish, Blood strain, ldl cholesterol, Blood Sugar, Max coronary heart expense and old peak. Rule base is the important phase in fuzzy inference method and first-rate of results in a fuzzy process depends upon the fuzzy principles. These principles have been applied on enter variables to verify how effective the fuzzy approach works. Using the informed method the established results for prognosis of coronary heart disorder have the foundations that if the worth is low then the danger is low if the value is excessive then the hazard is excessive. For this reason the trained procedure has been carried out and suggests that it is extra effective for analysis of heart ailment [8].

Dhanashree S. Medhekar, Mayur P. Bote, Shruti D. Deshmukh (2013) presented a classifier technique for the heart sickness prediction and likewise they've confirmed how the Naïve Bayes can be used for the classification cause. They will categorise clinical knowledge to five distinct classes namely no, low, normal, excessive, very excessive. If discovered any unknown sample the method will classify into its respective class label of that sample. The dataset used here is the Cleveland medical institution groundwork coronary heart disease set which contains 303 observations and 14 parameters. The system works in two phases: coaching phase, testing phase. In the training segment the classification is supervised, classifies knowledge situated on the training set and sophistication labels as a classifying attribute and classifies into new knowledge. In the checking out segment it involves the prediction of the unknown knowledge or the lacking values. The Naïve Bayes algorithm is used and it is situated upon the Bayesian theorem. The outcome has proven that the accuracy has been obtained through altering the number of occasions within the given dataset [9] .

Akin Ozcift and Arif Gulten (2011) developed a Random forest “RF” ensemble classifier to assess their classification of performances utilizing Parkinson’s, diabetes and coronary heart diseases data sets. Using correlation situated characteristic determination algorithm three knowledge sets were minimized after which performances of 30 machine learning algorithms were estimated for three information sets and constructed situated on RF algorithm [10].

Mai Shouman, Tim Turner, Rob Stocker (2011) have applied a wide range of techniques to different types of Decision tree seeking better performance in diagnosing the heart disease. They have proposed a model that outperforms J4.8 Decision tree and bagging algorithm in diagnosing heart disease patients. They have proposed a model that involves different discretization techniques, multiple classifiers voting technique and different Decision tree type for diagnosis of heart patients. Different combinations of discretization methods, decision tree types and voting are tested to identify which combination will provide the best performance in diagnosing heart patients. Data discretization is divided into supervised and unsupervised methods. The unsupervised methods involve equal width and equal frequency while the supervised discretization methods involve chi merge and entropy. The data partitioning involves testing with and without voting. Three Decision Tree types are tested: Information Gain, Gini Index, and Gain Ratio. Finally, reduced error pruning is applied on all the Decision Tree rules extracted from the training data. The Dataset used is the Cleveland Clinic Foundation heart disease consists of 76 raw attributes. The results show us that highest accuracy is obtained by the equal width discretization Information Gain Decision Tree with 79.1%. Different partitions of voting were applied to the data. The highest accuracy achieved by the equal frequency discretization in Gain Ratio Decision Tree is 84.1%. When compared with the existing system this model has shown the best results and has achieved highest accuracy [11].

Shashikant Ghumbre, ChetanPatil, Ashok Ghatol (2011) developed a decision help approach using RBF and SVM. RBF networks are beneficial for continuous or piecewise continuous actual-valued mapping approximations. Three parameters particularly the quantity of basis capabilities, their place and their width determine the measure of accuracy of the RBF networks. SVM is a class of common feed ahead networks like Radial-groundwork perform networks. SVM can be utilized for pattern classification and nonlinear regression. Extra precisely, support vector computer is an approximate implementation of the system of structural risk minimization. This principle is based on the actual fact the error expense of a learning computer

on scan information is bounded with the aid of the sum of the educational-error price and term that will depend on the Vapnik- Chervonenkis (VC) dimension. During experimentation, it's observed that, proper and complete data assortment process is the right route for the choice of high-quality classifier. For evaluating generalization performance with appreciate to accuracy, sensitivity, and specificity dataset is partitioned into quantity of subsets (i.E train set and test set). Overall natural performance will depend on accuracy of SVM and RBF utilizing subsets of coaching and test sets. SVM offers highest accuracy with increase in measurement of training data, while RBF gives minimal accuracy with scale down in dimension of test information [12].

Pasi Luukka and Jouni Lampinen (2010)have carried out classification manner situated on pre-processing the info with principal aspect evaluation (PCA) and then utilising differential evolution classifier to the prognosis of coronary coronary heart ailment. This system used to be utilized for predicting prognosis from clinical data units. The outcomes indicated that pre-processing the info before classification would not simplest help with the curse of increasing information dimensionality, but additionally furnish one more enhancement in classification accuracy [13].

Nazri Mohd Nawi etal (2010) have proposed a novel technique to increase the effectiveness of back propagation neural network. In Gradient Descent with Momentum and Adaptive gain proposed calculation, for every hub the addition high-quality used to be modified adaptively to alter initial search. The coronary health problem of the sufferer was predicted productively and the calculations had been firmly developed and can upgrade the computational productiveness [14].

Resul Das and Ibrahim Turkoglu et al (2009) have encouraged unique tools and quite a lot of methodologies to create powerful scientific decision supportive network. A framework used to be offered which makes utilization of Statically analysis procedure (SAS) base programming for diagnosing of the coronary ailment [15].

[16] Hongmei Yan and Jun Zheng (2008) have provided a exact coded GA established framework to decide on the elemental medicinal accessories key to the coronary heart sicknesses choice. It has been proposed to prefer the basic elements and aid the finding of 5 principle heart infections which have been hypertension, coronary health problem, rheumatic valvular coronary sickness, perpetual pulmonale and innate coronary illness.

Kemal Polat and Salih Gunes (2007) offered a hybrid method on medical diagnosis using feature decision, fuzzy weighted pre-processing and artificial Immune Recognition System. The hybrid method have two stages. The datasets of heart ailment and hepatitis disease have been reduced to 9 from 13 & 19 within the feature selection by means of C4. 5 decision tree algorithm. The heart sickness and hepatitis sickness datasets are utilized from UCI database as clinical dataset [17].

Other types of methods which are widely employed in the diagnosis of Heart disease are: Hongmei Yan etal [18] developed a multilayer perceptron based decision support system to support the diagnosis of heart disease. A.T.Sayad etal [19] employed the Multi-layer Perceptron Neural Network with Back-propagation as the training algorithm on heart disease diagnostic system. Resul Das etal [20] introduced a methodology which uses SAS base software for diagnosing of the heart disease. Sunila Godara etal [21] presented a decision support system based on MLP neural network architecture for diagnosing heart disease.

Summary

Table 1 : Summarized table shows proposed models for diagnosing the heart disease

S.NO	Author	PROPOSED MODEL	EXPECTED RESULTS *ACCURACY*
1.	Hlaudi Daniel Masethe etal [5]	Data Mining Techniques	Depends upon conditions provided.
2.	Mai Shouman etal[11]	Nine subsets voting model	79.1%-Info gain 84.1%-Gain ratio
3.	Durairaj M	MLP + Back propagation	Highest accuracy of 96.30%
4.	S.Florence etal[4]	Neural Network+ Decision Tree	Depends upon conditions provided
5.	Sanjeev Kumaretal[8]	Fuzzy Expert System	Risk Factors 0-low,1-high
6.	Dhanashree S. Medhekar etal [9]	Naïve Bayes	Based upon different instances-89.98%
7.	Shashikant Ghumbreetal[12]	RBF + SVM	Depends upon size of training data
8.	Jayshril S. Sonawanel	LVQ+ Neural Network	85.55%-highest accuracy

	etal[6]		
9.	Sumit Bhatiaetal[22]	SVM + GA	accuracy of 90.57%
10.	Vidyullatha.p, D.Rajeswara Rao[23]	Rough Set Model	good clarity and more Accuracy over the incomplete data set.

III. Conclusion

This learn applied a literature survey of comparative studies on neural networks, machine learning procedures and statistical techniques used for prediction and classification intent of the heart sickness. These evaluation facets out the knowledge of neural networks being employed for classification and prediction of heart attack. In this regard, these systems end up complementary approaches for model constructing rather of competing approaches.

References Références Referencias

1. Laercio Brito Gonçalves and Marley Maria Bernardes Rebuzzi Vellasco.Inverted Hierarchical Neuro-Fuzzy BSP System: A Novel Neuro-Fuzzy Model for Pattern Classification and Rule Extraction in Databases.Journal of IEEE Transactions On Systems, Man, And Cybernetics2016 March; Vol. 36(2)
2. Durairaj M, Revathi V, "Prediction Of Heart Disease Using Back Propagation Mlp Algorithm", International Journal Of Scientific & Technology Research, Issn 2277-8616 ,Volume 4, Issue 08, August 2015
3. Noura Ajam, "Heart Diseases Diagnoses Using Artificial Neural Network", Network And Complex Systems, ISSN 2224-610X (Paper) ISSN 2225-0603 (Online).Vol.5, No.4, 2015
4. S.Florence, N.G.Bhuvaneswari Amma, G.Annapoorani, K.Malathi, "Predicting The Risk Of Heart Attacks Using Neural Network And Decision Tree", International Journal Of Innovative Research In Computer And Communication Engineering, ISSN (Online) : 2320-9801, Vol. 2, Issue 11, November 2014.
5. Hlaudi Daniel Masethe, Mosima Anna Masethe, "Prediction of Heart Disease using Classification Algorithms", Proceedings of the World Congress on Engineering and Computer Science 2014 Vol II WCECS 2014, 22- 24 October, 2014, San Francisco, USA.
6. Jayshril S. Sonawanel, D. R. Patil, "Prediction of Heart Disease Using Learning Vector Quantization Algorithm", 978-1-4799-3064-7/14/$31.2014 IEEE.
7. Jesmin Nahar and Tasadduq Imam. Association rule mining to detect factors which contribute to heart disease in males and females.Journal of Expert Systems with Applications. **2013**Vol.40:1086–1093.
8. Sanjeev Kumar, Gursimranjeet Kaur, "Detection Of Heart Diseases Using Fuzzy Logic", International Journal Of Engineering Trends And Technology (IJETT), ISSN: 2231-5381,Volume 4 Issue 6- June 2013, Pp:2694-2699
9. Dhanashree S. Medhekar, Mayur P. Bote, Shruti D. Deshmukh, "Heart Disease Prediction System Using Naive Bayes", International Journal Of Enhanced Research In Science Technology & Engineering, Issn No: 2319-7463, Vol. 2 Issue 3, March.-2013
10. Akin Ozcift and Arif Gulten, "Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms.Journal of Computer Methods and Programs in Biomedicine2011:Vol.104:443-451.
11. Mai Shouman, Tim Turner, Rob Stocker, "Using Decision Tree for Diagnosing Heart Disease Patients", Proceedings of the 9-th Australasian Data Mining Conference (AusDM'11), Ballarat, Australia, Conferences in Research and Practice in Information Technology (CRPIT), Vol. 121.
12. Shashikant Ghumbre, Chetan Patil, And Ashok Ghatol, "Heart Disease Diagnosis Using Support Vector Machine", International Conference On Computer Science And Information Technology (ICCSIT'2011) Pattaya Dec. 2011
13. Pasi Luukka and Jouni Lampinen A Classification Method Based on Principal Component Analysis and Differential Evolution Algorithm Applied for Prediction Diagnosis from Clinical EMR Heart Data Sets.Journal of Computer Intelligence in Optimization Adaption, Learning and Optimization. **2010**;Volume 7: 263-283.
14. Nazri Mohd Nawi and Rozaida Ghazali.The Development of Improved Back-Propagation Neural Networks Algorithm for Predicting Patients with Heart DiseaseIn proceedings of the first international conference ICICA.2010;Vol.6377:317-324.
15. Resul Das and Ibrahim Turkoglu, "Effective diagnosis of heart disease through neural networks ensembles", Journal of expert system with applications.**2009**:Vol.36, :7675–768.0
16. Hongmei Yan and Jun Zheng, "Selecting critical clinical features for heart diseases diagnosis with a real-coded genetic algorithmJournal of Applied Soft Computing 2008 :Vol.8:1105-1111.

17. Kemal Polat and Salih Gu nes.A hybrid approach to medical decision support systems: Combining feature selection, fuzzy weighted pre-processing and AIRS. Journal of Computer Methods and Programs in Biomedicine.2007;Vol.88:164-174
18. Hongmei Yan, Yingtao Jiang, Jun Zheng, Chenglin Peng, Qinghui Li," A multilayer perceptron-based medical decision support system for heart disease diagnosis", Expert Systems with Applications 30 (2006) 272–281.
19. T. Sayad, P. P. Halkarnikar," Diagnosis of Heart Disease Using Neural Network Approach", International Journal of Advances in Science Engineering and Technology, ISSN: 2321- 9009, Volume- 2, Issue-3, July-2014,pp:88-92.
20. Resul Das, Ibrahim Turkoglu , Abdulkadir Sengur," Effective diagnosis of heart disease through neural networks ensembles", Expert Systems with Applications 36 (2009) ,ISSN:7675– 7680,pp:7676-7680.
21. Sunila Godara, Nirmal," Yanwei, X et.al [7] developed data mining algorithms for predicting survival of CHD patients. Intelligent and Effective Decision Support System Using Multilayer Perceptron", International Journal of Engineering Research and Applications (IJERA), ISSN: 2248-9622, Vol. 1, Issue 3, pp.513-518.
22. Sumit Bhatia, Praveen Prakash, and G.N. Pillai, "SVM Based Decision Support System for Heart Disease Classification with Integer-Coded Genetic Algorithm to Select Critical Features", Proceedings of the World Congress on Engineering and Computer Science 2008 WCECS 2008, October 22 - 24, 2008, San Francisco, USA.
23. Vidyullatha Pellakuri, D.RRao, "Progressive Decision Making In The Department Of Cardiology By Optimized Rough Set Model", Int J Pharm Bio Sci 2016 April; 7(2): (B) 658 – 665.