ANALYSIS OF HEART DISEASE PREDICTION SYSTEM USING ARTIFICIAL NEURAL NETWORK

[1]Akruti Dave, [2]Prof. Gayatri Pandi
[1] Master of Engineering Student, [2]Head of Department
[1] Department of Computer Engineering
[1]L. J Institute of Engineering &Technology (Gujarat Technological University), Ahmedabad, India

Abstract :According to World Health Organization (WHO) , 17.7 Million people die every year due to various Cardiovascular diseases. The prediction of heart disease system can become a boon for the population to predict the health complexity at early stage .Various methods and techniques are undergoing the research phase for developing the most accurate heart disease prediction system. Machine Learning based Artificial Neural Network (ANN) is the most promising one in this area due to its computing functions , neuron functionality and machine learning rules. Various layers in ANN are input, hidden & output layer ,which make the system nearly most accurate for the prediction approach. Existing systems are working well in predicting the accurate result, but still more attributes of data and complexity of health parameters make the root layer for the innovation of new approaches. This study focuses on the analysis of various neural network approaches for the prediction of heart disease. The dataset for experiment used here is Cleveland Heart Disease dataset available on UCI machine learning repository.The proposed model uses a combined approach neural network in which one multilayer perceptron neural network(MLP) is used for prediction and another MLP neural network is used for error reduction with backpropagation approach.Also,the input parameters are the combination of clinical attributes as well as additional attributes.

***IndexTerms* – Artificial Neural Network, Heart Disease Prediction, Combined Approach**

I. INTRODUCTION

Heart is the most important component of the cardiovascular system. It weighs around 250-350 grams i.e. about the size of a fist. It beats around 2.5 billion times during as lifespan of 66-68 years. Heart pump 51% of the blood to different parts of the body providing human body renewed materials. Any failure or defect in the heart may result into sudden death. This is one of the reasons why researcher has focus more in designing intelligent system that can be used to diagnose heart diseases with high accuracy, to avoid misdiagnosis. Millions of people die every year because of heart disease and large population of people suffers from heart disease. Prediction of heart disease early plays a crucial role for the treatment. If heart disease could be predicted before, lots of patient deaths would be prevented and also a more accurate and efficient treatment way could be provided. A need to develop such a medical diagnosis system arises day by day. The important key points of such medical diagnosis systems are reducing cost and obtaining more accurate rate efficiently. Developing a medical diagnosis system based on machine learning for prediction of heart disease provides more accurate diagnosis than traditional way and reduces cost of treatment. In this paradigm Artificial Neural Network (ANN) stands as a top most widely used approach, producing accurate prediction results.

1.1 PREDICTION USING ARTIFICIAL NEURAL NETWORK

Prediction of heart disease is called supervised learning problem. As the output variables are in category type, the prediction of heart disease is “classification type of supervised learning”. An artificial neural network (ANN) learning algorithm, usually called "neural network" (NN), is a learning algorithm that is vaguely inspired by biological neural networks.Computations are structured in terms of an interconnected group of artificial neurons, processing information using a connectionist approach to computation.

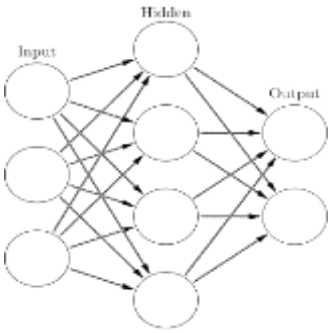

Figure 1: Artificial Neural Network

ANN is an interconnected group of nodes. Here, each circular node represents an artificial neuron.The arrow represents a connection from the output of one artificial neuron to the input of another. Artificial neural networks are relatively crude electronic networks of neurons based on the neural structure of the brain. They process records one at a time, and learn by comparing their prediction of the record (largely arbitrary) with the known actual record. The errors from the initial prediction of the first record is fed back to the network and used to modify the network's algorithm for the second iteration. These steps are repeated multiple times. The basic architecture is as shown in figure 2.

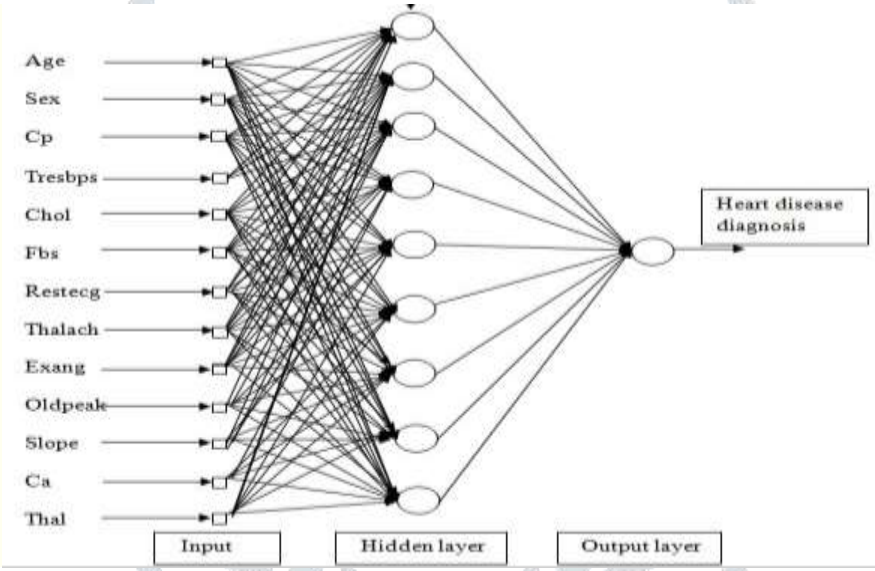

Figure 2: Heart Disease Prediction using Artificial Neural Network[10]

II. LITERATURE REVIEW

2.1 DATASET

For the heart disease prediction system, the dataset in all the existing research methodologies discussed below is cleveland hospital heart disease dataset available on UCI machine repository. The features taken into considerations are as below.The dataset contains 303 instances and 76 attributes, but only 14 of them are referred by all published studies.

Table 1: Attributes of the UCI Cleveland dataset

1	age	Age in years
2	sex	Sex
3	cp	Chest pain type
4	trestbps	Resting blood pressure on admission to the hospital in mm Hg
5	chol	Serum cholesterol in mg/dl
6	fbs	Fasting blood sugar is greater than 120 mg/dl or not
7	restecg	Resting electrocardiographic results
8	thalach	Maximum heart rate achieved
9	exang	Exercise induced angina
10	oldpeak	ST depression induced by exercise relative to rest

11	slope	The slope of the peak exercise ST segment
12	ca	Number of major vessels (0-3) colored by fluoroscopy
13	thal	The heart status
14	num	Diagnosis of heart disease

2.2 METHODOLOGY

Artificial neural network approaches are widely used for developing the most accurate predictive model for Heart Disease.MATLAB tool is used in all the published researches for implementing the ANN. In 2014 Muhammad Fathurachman proposed the Heart Disease Diagnosis using Extreme Learning Based Neural Networks approach. Extreme Learning Machine is artificial neural network with Single Hidden Layer Feedforward (SLFNs). The method of Extreme Learning machine created to solve the limitations of artificial neural networks specially in learning process. Different training and testing dataset which had been divided using K-Fold Cross validation further produced the average prediction Accuracy 83% based on 5 experiments[1].

D. K. Ravish proposed the artificial neural network approach based on Data acquisition stage, Processing stage, Network training stage, Disease predicting stage and finally Data transmission stage. ECG data of the diseased patients from the Physio net ECG database was used along with the clinical features.The ECG features used for this study were 1) QRS duration 2) R-R interval 3) P-R interval 4) Q-T interval 5) R-wave amplitude 6) Pwave duration 7) T wave duration. The clinical features were 1) Mean arterial pressure (in mm/hg) 2) Fasting blood sugar (in mg/dL) 3) Heart rate (in BPM) 4) Cholesterol levels (in mg/dL) 5) Age factor (in years) 6) Smoking/Drinking/Tobacco factors 7) Diabetes factor. Furthermore they used backpropagation neural network approach to predict the heart disease accurately in the form of columnised 0 & 1 combination-000,001,010,100-normal,110,011,101-abnormal,111-predicted[2].

Jayshril S. Sonawane proposed the Prediction of Heart Disease Using Multilayer Perceptron Neural Network along with back propagation neural network. In this paper a useful and accurate technique for the classification and retrieval of image by using self organizing map (SOM) is proposed and developed. In this technique the image texture is classified in two main phases, in first phase the color features are extracted and classification is done based on color features using self-organizing map. In second phase the images in each class of first phase are again classified using self-organizing map based on texture features extracted using GLCM-gray level co-occurrence matrix. The SOM is trained with different topology sizes and no. of iterations to improve the performance of the system. The experiments were performed on Wang's database comprising of total 1000 images including 10 categories. The experimental results show that the proposed method gives increased accuracy and improved retrieval performance for each category of the image database The prediction system in this paper gives higher accuracy of 93.39% for 5 neurons in hidden layer with running time of 3.86s[3].

In 2015,Ebenezer O. Olaniyi proposed the diagnosis prediction based on multilayer neural network trained with backpropagation and simulated on feedforward neural network. The number of neurons used at the hidden layer of the research work was obtained by experimenting which number of neurons will best represent the patterns. Also the neurons in the hidden layer were varied. The learning rate which is the learning power of the network and the momentum rate which determines the speed of the network were also varied to allow the network to learn and prevent the network from settling at the local minima. The neurons in the hidden layer were experimented with variation of 2 neurons to 6 neurons to represent the pattern accurately for a better performance. In this research work, recognition rate of 85% was obtained. Multilayer neural network trained with backpropagation was discovered to have the best result suitable for diagnosis of heart disease without misdiagnosis[4].

In 2016,Majid and Omid proposed the feed forward neural network approach along with particle swarm optimization for improving the Heart Disease Diagnosis prediction. The methodology followed the Pre-processing unit,PSO- Particle Swarm Optimization,Feature ranking,FFNN- Feed Forward Neural Network.In the pre-processing step two main activities were performed: noise-elimination and determining training and testing data samples.The Particle Swarm Optimization is the algorithm to optimize particles used for feature selection.After cost normalization; the features are put in categories of- Free of Charge, Least Cost, Average Cost and Costly. Accuracy of the FFBP Classifiers was combined by PSO algorithm and the most optimized set of particles were found in terms of time, accessibility, venture and precision.Finally, the selected best set through PSO algorithm, has the following features: Age, Sex, Chol, FBS, RestECG, Exang, Oldpeak, Slope and the accuracy of FFNN was 91.94% [5].

Jairam P. Kelwade proposed the Prediction of heart abnormalities using Particle swarm optimization in Radial basis function neural network.The Prediction procedure is structured into four levels. First procedure involves the preparation of time series data of RR intervals involves the cardiac arrhythmia database acquisition, reprocessing of the ECG records, recognition of QRS complexes and detection of R peaks. The Data segmentation is done in a second stage. Third stage contains application of linear and nonlinear techniques. Final stage contains presentation of features to RBF neural network for training and testing purpose to obtain an accurate prediction.Results demonstrate that a combination of PSO and RBFN is more effective for prediction with an overall accuracy of 96.3% compared to RBFN[6].

In 2017,Tulay and Ozkan proposed prediction of heart disease using multilayer perceptron back propagation neural network.To optimize network and to get better performance pruning which defines a set of techniques for trimming size of network by nodes was used. Hidden layer size of the network was changed from 3 neurons up to 12 neurons.To improve performance, dimensionality reduction with Principal Component Analysis (PCA) was done by reducing number of neurons of the input layer from 13 neurons to 8 neurons.The proposed system gives 95% accuracy rate which means a very good rate according to related studies on this field.[7].

Jackielyn proposed Cardiovascular Health Pre-Diagnosis Based on a BP Profile Using Artificial Neural Network.He used Exercise Stress Test(EST)-treadmill,BP –Profile+ Body Mass Index BMI,Confusion matrix evaluation.The profile was made on the basis of Systolic blood pressure (SBP) - value of the arterial blood where the heart contracts,Diastolic blood pressure (DBP) - measured when the heart relaxes.Increase of > 20mmHg in the SBP during exercise or >10mmHg in DBP- measured as a hypertensive response,decrease of >20mmHg in the SBP during exercise or >10mmHg in DBP –measured as a hypotensive response .To produce a proper profile, the average of all the parameters, except gender, for the first two EST runs was obtained.It exhibited an overall accuracy of 95.2% with accuracies of detecting normal, hypotensive and hypertensive responses of 94.0%, 99.0% and 92.5% respectively [8].

Kann and Ahmet proposed Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks.The RFNN used in this study had 13 inputs, 7 hidden neurons and 1 output neuron. The weights and biases of the RFNN were coded as 64 bits long genes. GA used here is with mutation probability 0.05, multipoint crossover with probability 0.25 and size of population 100 .The testing set (45 instances) results showed that the GA based trained RFNN approach had a sensitivity of 100%, a specificity of 95.24%, a precision of 96%, and Fscore of 0.9796 and accuracy rate is 97.78%.The approach has an excellent rate of 100% for the patients without heart disease that were found to have no heart disease in the testing set (20 instances). The overall (297 instances) results had a sensitivity of 97.74%, a specificity of 95.734%, a precision of 94.89%, and F-score of 0.9626 and accuracy rate is 96.63%[9].

2.3 PERFORMANCE EVALUATION

The performance of the proposed system was computed by different metrics like accuracy, precision and recall. Accuracy is computed dividing number of predictions which are correct by number of all predictions. The obtained result is multiplied by 100 to get value as percentage.

Accuracy=(TP+TN)/(TP+TN+FP+FN)

where TP, TN, FP and FN demonstrate in order of the number of True Positives, True Negatives, False Positives and False Negatives.

- TP demonstrates the number of instances which are sick and diagnosed accurately.
- FP demonstrates the number of instances which are healthy and diagnosed wrongly as they are sick.
- FN demonstrates the number of instances which are sick but the instances are diagnosed wrongly.
- TN contains a number of instances which are healthy and the instances are diagnosed accurately

Table 2: Comparative Analysis

PUBLICATION & YEAR	TITLE	METHOD	RESULT ACCURACY %
ELSEVIERE 2017	Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks(RFNN)	RFNN	96.63
IEEE 2017	Cardiovascular Health Pre-Diagnosis Based on a BP Profile Using Artificial Neural Network(ANN)	BP Profile + ANN	95.2
IEEE 2017	Prediction of Heart Disease Using Neural Network	ANN+BPNN	Min 82, Max 91
IEEE 2016	Prediction of heart abnormalities using Particle swarm optimization in Radial basis function neural network	PSO+RBFNN	96.3
IEEE 2016	Improving the Heart Disease Diagnosis by Evolutionary Algorithm of PSO and Feed Forward Neural Network(FFNN)	PSO+FFNN	Min 76 Max 91.9
IEEE 2015	Neural Network Diagnosis of Heart Disease	Multilayer BPNN	85
IEEE 2014	Heart Disease Diagnosis using Extreme Learning Based Neural Networks	EXTREME LEARNING ANN	79
IEEE 2014	Prediction of Heart Disease Using Multilayer Perceptron Neural Network	MULTILAYER PERCEPTRON ANN	92
IEEE 2014	Heart Function Monitoring, Prediction and Prevention of Heart Attacks: Using ANN	Normalization +ANN	Highly precise 1,0 Combination

III. PROPOSED METHOD

Uptil now,all the research methodology has been focused on clinical parameters ,but we can use additional attributes in consideration too.The proposed method is focused on 4 stages. Stage 1 will be the data profiling and attribute selection in which clinical parameters-as shown in UCI parameters as well as risk factors –medical history,smoke,hypertension,hypotensive etc will be taken into consideration. Stage 2 will be the optimization phase to reduce the noise elimination in dataset..Stage 3 will be the neural network implementation with a combined approach where recurrent fuzzy neural network or radial basis function neural network can be used for prediction and backpropagation for error reduction.Stage 4 will show the final predicted output showing presence or absence of the disease.

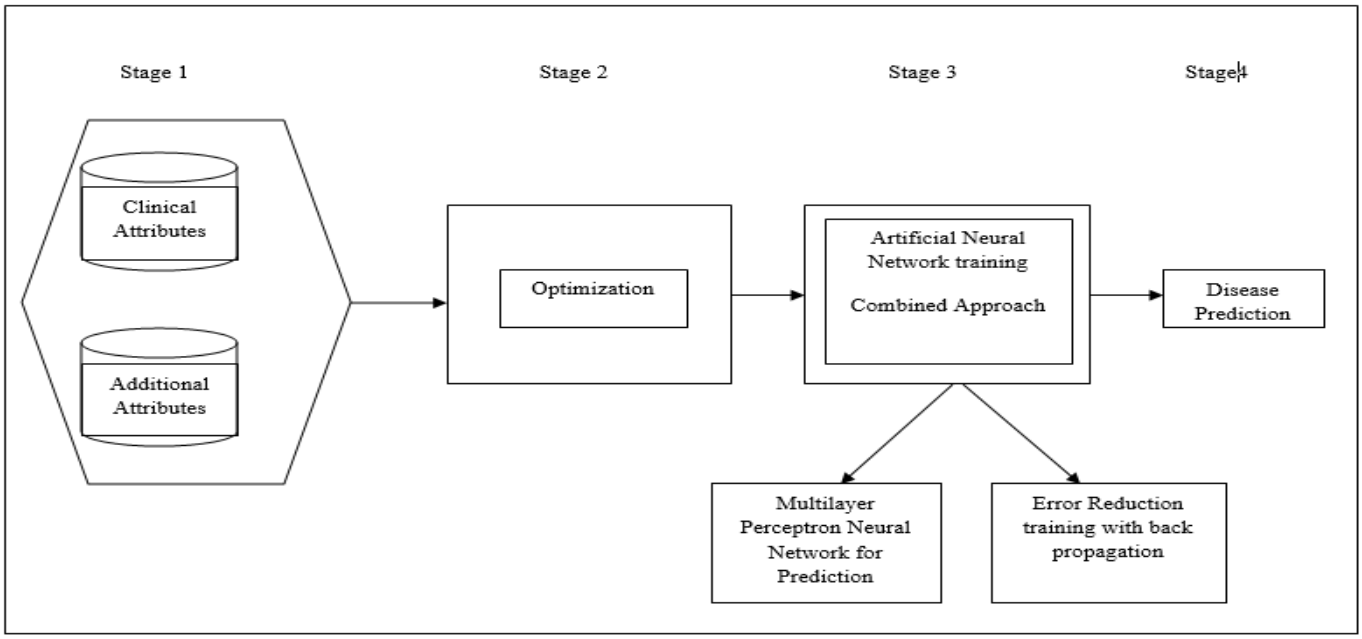

Figure 3: Block Diagram of Proposed System

3.1 STAGE 1:Data profiling and attribute selection

Existing published researches are dealing with clinical attributes only which are mentioned in Table 1.But if we add some more attributes the predictive output will be better than before.These additional attributes will be smoke, hypertension-hypotension, hereditary disease presence, workout habbits ,stress and heart condition at birth(heart valve dysfunction,hole in valve,abnormalities,etc).

3.2 STAGE 2:Optimization

Various optimization techniques can be useful for the noise elimination and reducing the missing dataset impact with greatly fitting the dataset to produce accurate prediction.Feature selection can be done with widely used approach Particle Swarm Optimization which gives fitness value of each attributes.

3.3 STAGE 3: Artificial Neural Network training Combined Approach

Tool for implementing and training the neural network is MATLAB.The prediction alone is not sufficient for the better accuracy. The model must have low error rate in its training and testing set. The error rate is measured in mean squared error (MSE).

3.4 STAGE 4: Disease Prediction

The accuracy and error rate will be calculated. The output will show whether the disease is present or not.

IV. CONCLUSION AND FUTURE WORK

The proposed system will definitely help in improving the prediction of heart disease system by improving its accuracy as well as the prediction capability with reduced cost. The existing systems are focused on clinical attributes only whereas the proposed system is working with additional risk factor attributes which is quite advantageous for improving the prediction. The

future research may focus on hybrid approaches as well as ensemble neural network approach to improve the prediction capability with better accuracy.

REFERENCES

[1] Muhammad Fathurachman et al 2014. Heart Disease Diagnosis using Extreme Learning Based Neural Networks.IEEE:23-27

[2] D. K. Ravishet al 2014. Heart Function Monitoring, Prediction and Prevention of Heart Attacks: Using Artificial Neural
Networks.IEEE:1-6

[3] Jayshril S. Sonawane 2014.Prediction of Heart Disease Using Multilayer Perceptron Neural Network.IEEE

[4] Ebenezer O. Olaniyi et al 2015**.**Neural Network Diagnosis of Heart Disease.IEEE:21-24

[5] Majid and Omid 2016:Improving the Heart Disease Diagnosis by Evolutionary Algorithm of PSO and Feed Forward Neural Network
IEEE:48-53

[6] Jairam P. Kelwadeand Suresh Salankar 2016.Prediction of heart abnormalities using Particle swarm optimization in Radial basis
Function neural network.IEEE:793-797

[7] Tulay and Ozkan 2017.Prediction of Heart Disease Using Neural Network.IEEE:719-723

[8] Jackielyn G. Domingo et al.2017.Cardiovascular Health Pre-Diagnosis Based on a BP Profile Using Artificial Neural Network.IEEE.

[9] Kaan Uyar and Ahmed 2017.Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks.
Procedia Computer Science 120 (2017) 588–593

[10]Kanita and Rasit 2014. Diagnosis of heart disease using a committee machine neural network. Proceedings of the 9th International
Conference on Applied Informatics Eger, Hungary 9(1): 351–360

[11]AHA, 2017. American Heart Association website. http://www.heart.org (11.05.2017)

[12]Aliev, R.A., Aliev R.R., Guirimov, B.G. ,Uyar K., 2007. Recurrent Fuzzy Neural Network Based System for Battery Charging. Lecture Notes in Computer Science 4492-II. 307-316.

[13]Aliev R.A., Aliev R.R., Gurimov B.G., Uyar K., 2008. Dynamic data mining technique for rules extraction in a process of battery charging, Applied Soft Computing 8 (3).1252-1258

[14]Alsalamah, M., Amin, S., Halloran, J., 2014.Diagnosis of heart disease by using a radial basis function network classification technique on patients' medical records.2014 IEEE MTT-S International Microwave Workshop Series on RF and Wireless Technologies for Biomedical and Healthcare Applications (IMWS-Bio2014), London.1-4

[15]Amiri A.M., Armano, G., 2013. Early Diagnosis of Heart Disease Using Classification And Regression Trees. The 2013 International Joint Conference on Neural Networks (IJCNN), Dallas, TX.1-4. Anooj, P.K., 2013.

[16]Implementing Decision Tree Fuzzy Rules in Clinical Decision Support System after Comparing with Fuzzy Based and Neural Network Based Systems. 2013 International Conference on IT Convergence and Security (ICITCS). 1-6.

[17]Bouktif, S., Hanna, E.M., Zaki, N., Abu Khousa, E., 2014. Ant Colony Optimization Algorithm for Interpretable Bayesian Classifiers Combination: Application to Medical Predictions. PLOS ONE 9(2).e86456, 12

[18]Çiftçi, F.B., Incekara, H., 2011. Web based medical decision support system application of Coronary Heart Disease diagnosis with Boolean functions minimization method. Expert Systems with Applications 38.14037–14043.

[19]Das, R., Turkoglu, I., Sengur, A., 2009. Effective diagnosis of heart disease through neural networks ensembles.Expert Systems with Applications 36. 7675–7680.

[20] De Falco, I., 2013. Differential Evolution for automatic rule extraction from medical databases. Applied Soft Computing 13.1265–1283. Hedeshi, N.G., Abadeh, M.S., 2014.

[21]Coronary Artery Disease Detection Using a Fuzzy-Boosting PSO Approach. Computational Intelligence and Neuroscience. Article ID 783734, 12

[22]Jabbar, M.A., Deekshatulu, B.L, Chndra, P., 2014. Alternating decision trees for early diagnosis of heart disease.Proc.of Int. Conf. on Circuits, Communication, Control and Computing (I4C 2014), MSRIT, Bangalore, India.322-328.

[23]Jabbar, M.A., Deekshatulu, B.L, Chandra P., 2015. Computational Intelligence Technique for early Diagnosis of Heart Disease. 2015 IEEE International Conference on Engineering and Technology (ICETECH), 20th March 2015, Coimbatore, TN, India.1-6.

[24] Kim, J.K., Lee, J.S., Park, D.K., Lim, Y.S., Lee, Y.H., 2014. Adaptive mining prediction model for content recommendation to coronary heart disease patients. Cluster Compute 17(3). 881-891.

[25]Miao K.H., Miao, J.H., Miao, G.J, 2016. Diagnosing Coronary Heart Disease Using Ensemble Machine Learning. (IJACSA) International Journal of Advanced Computer Science and Applications 7(10).30-39.

[26]Nahar, J., Imam, T., Tickle K.S., Chen Y.-P. P., 2013a. Computational intelligence for heart disease diagnosis: A medical knowledge driven approach. Expert Systems with Applications 40.96–104.

[27]Nahar, J., Imam, T., Tickle K.S., Chen Y.-P. P., 2013b. Association rule mining to detect factors which contribute to heart disease in males and females. Expert Systems with Applications 40.1086–1093.

[28]Nguyen, T., Khosravi, A., Creighton, D., Nahavandi, S., 2015.Medical data classification using interval type-2 fuzzy logic system and wavelets. Applied Soft Computing 30.812–822.

[29]NHLBI, 2017.National Heart, Lung, and Blood Institute website. https://www.nhlbi.nih.gov/health/health-topics/topics/hf (11.05.2017)

[30]Olaniyi, E.O., Oyedotun, O.K., Helwan, A., Adnan, K., 2015. Neural network diagnosis of heart disease. 2015 International Conference on Advances in Biomedical Engineering (ICABME), Beirut. 21-24.

[31] Paul, A.K., Shill, P.C., Rabin, M.R.I., Kundu, A., Akhand, M.A.H., 2015. Fuzzy membership function generation using DMS-PSO for the diagnosis of heart disease. 2015 18th International Conference on Computer and Information Technology (ICCIT), Dhaka. 456-461.

[32]Paul, A.K., Shill, P.C., Rabin M.R.I., Akhand, M.A.H., 2016. Genetic algorithm based fuzzy decision support system for the diagnosis of heart disease. 2016 5th International Conference on Informatics, Electronics and Vision (ICIEV), Dhaka. 145-150.

[33]Polat, K., Güneş, S., Tosun, S., 2006. Diagnosis of heart disease using artificial immune recognition system and fuzzy weighted preprocessing. Pattern Recognition 39. 2186 – 2193.

[34] Reddy, G.T., Khare, N., 2017. An Efficient System for Heart Disease Prediction using Hybrid OFBAT with Rule-Based Fuzzy Logic Model. Journal of Circuits, Systems, and Computers 26 (4). 1750061-1-21

[35]Sagir, A.M., Sathasivam, S., 2017. A Novel Adaptive Neuro Fuzzy Inference System Based Classification Model for Heart Disease Prediction. Pertanika J. Sci. & Technol. 25 (1). 43 – 56.

[36] Samuel, O.W., Asogbon , G.M., Sangaiah, A.K., Fang, P. , Li, G., 2017. An integrated decision support system based on ANN and Fuzzy_AHP for heart failure risk prediction. Expert Systems With Applications 68. 163–172.

[37] Shaoa, Y.E., Houa, C.-D., Chiuba C.-.C., 2014. Hybrid intelligent modeling schemes for heart disease classification. Applied Soft Computing 14:47–52.

[38] Uyar, K., 2006. Modeling and Simulation of NiCd Batteries Behaviour Under Fast Charging with Genetic Algorithm Based Trained Recurrent Fuzzy Neural Network. Proc. of the 7th Int. Conf. on Application of Fuzzy Systems and Soft Computing (ICAFS2006), Siegen, Germany:247-255