International Journal of Engineering Science Invention (IJESI)
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org ||Volume 7 Issue 7 Ver III || July 2018 || PP 07-11

Heart Disease Detection: A Neural Networks Application

Chaitanya Roygaga[1], Suraj Punjabi[2], Swapnil Sampat[3], Ms. Tanuja Sarode[4]
[1,2,3]*(B.E., Semester 8, Computer Engineering, ThadomalShahani Engineering College/Mumbai University, India)*
[4]*(Head of Department, Department of Computer Engineering, ThadomalShahani Engineering College/Mumbai University, India)*
Corresponding Author:ChaitanyaRoygaga

Abstract :*The heart is the most vital organ in the human body. Any heart related disease can be very harmful and can often lead to death if not treated properly in the early stages. Therefore if heart diseases can be diagnosed as soon as possible, the doctor could immediately begin with the treatment. Also, that would mean less suffering time for the patient. To predict a heart disease or a type of heart disease the doctors needs to run a couple of tests before reaching a conclusion. These tests might cost a lot of money and might require advanced technology. Every patient might not be able to afford the tests. Also the technology used is available mainly in urban areas. Artificial Neural Networks provide an economical and accurate approach to diagnose heart diseases. Using Neural Networks we have developed a system that could predict the absence or presence of heart disease. This paper deals with the analysis of parameters of Error-Back Propagation algorithm that would provide the best accuracy for diagnosing heart disease in a patient. The system uses 13 attributes of a patient as input. The system diagnoses heart disease with an accuracy of 88.5%.*

Keywords –*ANNs, Classification Accuracy, Error Back Propagation algorithm, Feed-Forward, Heart Disease Diagnosis*

Date of Submission: 27-06-2018 Date of acceptance: 12-07-2018

I. Introduction

The human heart is roughly the size of a fist, and is used to pump blood to almost every single part of the body. Oxygen that is carried by the blood is needed by every cell in the human body to survive. Compared to any physical muscle in the body, the heart does the maximum physical work. These facts indicate that the heart is veritably the most important part of the human body. Therefore good care needs to be taken of the heart in order to live a good healthy life. Heart diseases can occur due to various reasons and it is necessary to diagnose them as soon as one comes to know about the initial symptoms. However the symptoms do not directly indicate the presence of a heart disease. A woman feels nausea, indigestion pain and shoulder ache while suffering from a heart disease which could be a symptom to any other disease and might confuse the doctor. Thus the proposed system proves to be an economical way to detect heart diseases at early stages.

II. Existing Work

Oleg Yu. Atkov, Svetlana G. Gorokhova [1] developed an artificial neural networks-based (ANNs) diagnostic model for coronary heart disease (CHD). They provided the system with data of 487 patients (327 having CHD). They obtained the best accuracy with multi-layer perceptron network with two hidden layers.

Dr. K. Usha Rani [2] analysed a heart disease dataset using neural network approach, for data classification. She used parallel approach in the training phase to increase the efficiency of the classification process. The classification task of the medical data was done using the backpropagation algorithm.

Jae Kwon Kim, Sanggil Kang [3] devised a neural network based prediction of CHD risk using feature correlation analysis, using two stages. They used data of 4146 individuals in Korea for their network. NN-FCA provided better CHD risk prediction than FRS.

NouraAjam [4] implemented classification using artificial neural network, to diagnose heart disease.
Sandra Cako, Angelina Njegus, Vladimir Matic [5] used kNN and LDA algorithms in MLP to predict heart disease with higher accuracy.

Miss. Chaitrali S. Dangare, Dr.Mrs.Sulabha S. Apte [6] developed a Heart Disease Prediction System (HDPS) using data mining and artificial neural network techniques.

III. System Flow

The training and testing data has been collected from the UCI Machine Learning Repository [7].The data has 297 tuples that are divided equally into training and testing. The data is normalized and fed to the system [8]. The system uses the training data to learn.Learning involves updating the weights of the network after every iteration, based on the error(1) produced after the feed-forward phase [8]:

- For every iteration, the weights are updated so that in the next iteration we get minimum error.
- The algorithm stops when E (Error) of iteration is less than the limit of E_{max}assigned at the start of iteration 1, else the iterations continue.
- If the error criteria is satisfied and the algorithm stops, the final weights from the training will be obtained, which can now be used for diagnosing the Test data from the data set.

$$\mathrm{Error} = \mathrm{DesiredOutput} - \mathrm{ActualOutput}(1)$$

The mean square error function is chosen as the stopping criteria for the algorithm during training. It is the mean of the square of the differences between the desired output and the actual output(2).

$$\mathrm{E_{MSE}} = \frac{1}{\mathrm{M}}\sum(\mathrm{Desiredoutput} - \mathrm{Actualoutput})^2 \; (2)$$

Where, E_{MSE} = Mean square error
M = Number of training inputs

After the training phase is completed, the system is given the testing data:

- Each pair of the test data is tested using the weights obtained in training, and minimal error is expected from each test.
- The accuracy of the algorithm is then calculated from the mean error obtained after testing all of the testing data pairs.

This phase helps us to identify the number of data tuples that were correctly classified by the system.

IV. Implementation

The Error-Back Propagation algorithm was run on Octave. Initially, there were 297 data samples, from which outliers were removed. After this, 278 data samples were left, which were equally divided into training and testing data pairs. Each data pair has 13 attributes and 1 desired output, the output being the result of diagnosis according to the given data.
The implementation was conducted in 3 stages:

Stage 1:

This stage was to analyse which of the two activation functions – Unipolar and Bipolar – provided better prediction accuracy. A random number of hidden layer neurons was selected between 2 and 26. The network was then trained using Unipolar (Sigmoid) activation function, and then using Bipolar activation function. The network was then tested, and the Sigmoid activation function provided better accuracy. This activation function was then selected for further implementation.

Stage 2:

This stage was to analyse which would be the best number of hidden layer neurons for the network, so that the network would give the best prediction accuracy. A range of 2 to 26 hidden neurons was analysed. It was concluded that the number of hidden layer neurons that provided the best network accuracy for diagnosis was 13.

Stage 3:

Now, momentum was used to improve the accuracy of the network. The momentum which combined well with the already chosen network parameters was selected. Also, analysis on the number of training epochs was done at which the best accuracy is obtained.
Table 1 shows the final parameters obtained from our analysis.

Table 1 – Summary of the final network parameters

Learning Rate	0.05
Momentum	0.75
Activation Function	Sigmoid
Error Function	Mean square error
Number of hidden layer neurons	13
Number of training epochs	6

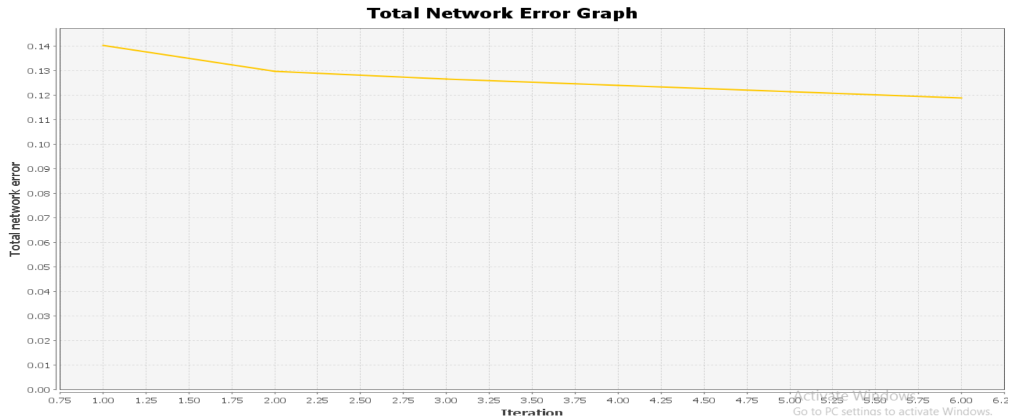

Fig.1 - Network Error Graph for network with 13 Hidden Neurons (Without Momentum)

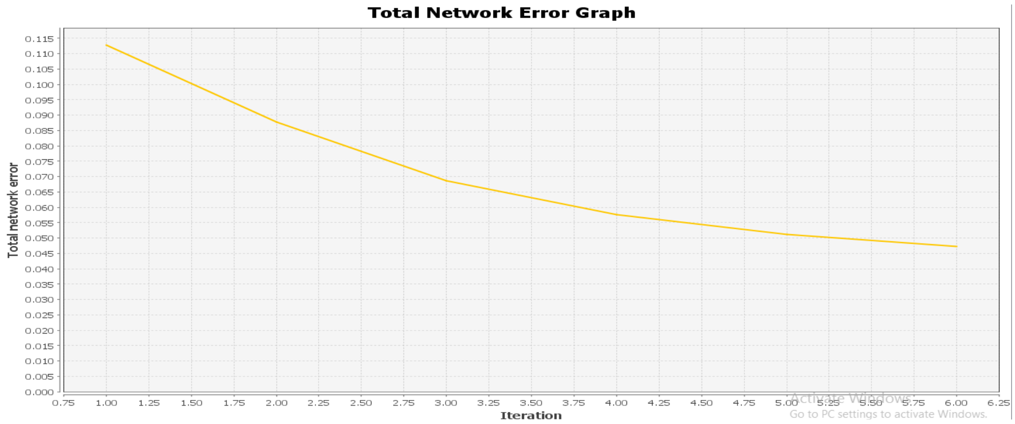

Fig. 2 - Network Error Graph for network with 13 Hidden Neurons (With Momentum)

Fig. 1 plots the rate of decreasing error, when momentum is not considered. In our further analysis, after considering momentum, the error has decreased at a faster rate, as shown in Fig. 2. Both the graphs have been plotted against the number of iterations.

Table 2 shows the complete comparison of statistical measures considered for analysis of the best number of hidden layer neurons for the network.

Table 2 – Comparison between different number of hidden layer neurons (After Stage 3)

Number of hidden layer neurons	**True Positive (TP)**	**False Negative (FN)**	**True Negative (TN)**	**False Positive (FP)**	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**
2	57	6	64	12	87.05	90.47	84.21	82.61	91.43
3	56	7	63	13	85.61	88.88	82.89	81.16	90
4	55	8	63	13	84.89	87.3	82.89	80.88	88.73
5	55	8	63	13	84.89	87.3	82.89	80.88	88.73
6	52	11	66	10	84.89	82.54	86.84	83.87	85.71
7	58	5	61	15	85.61	92.06	80.26	79.45	92.42
8	57	6	62	14	85.61	90.47	81.57	80.28	91.18
9	57	6	61	15	84.89	90.47	80.26	79.17	91.04
10	57	6	61	15	84.89	90.47	80.26	79.17	91.04
11	58	5	62	14	86.33	92.06	81.57	80.56	92.54
12	55	8	62	14	84.17	87.3	81.57	79.71	88.57
13	**55**	**8**	**68**	**8**	**88.49**	**87.3**	**89.47**	**87.3**	**89.47**
14	56	7	62	14	84.89	88.88	81.57	80	89.86
15	55	8	62	14	84.17	87.3	81.57	79.71	88.57
16	54	9	63	13	84.17	85.71	82.89	80.6	87.5
17	54	9	62	14	83.45	85.71	81.57	79.41	87.32
18	54	9	64	12	84.89	85.71	84.21	81.82	87.67
19	54	9	67	9	87.05	85.71	88.16	85.71	88.16
20	54	9	64	12	84.89	85.71	84.21	81.82	87.67
21	54	9	66	10	86.33	85.71	86.84	84.38	88
22	57	6	64	12	87.05	90.48	84.21	82.61	91.43
23	54	9	63	13	84.17	85.71	82.89	80.6	87.5

24	54	9	63	13	84.17	85.71	82.89	80.6	87.5
25	54	9	64	12	84.89	85.71	84.21	81.82	87.67
26	55	8	65	11	86.33	87.3	85.53	83.33	89.04

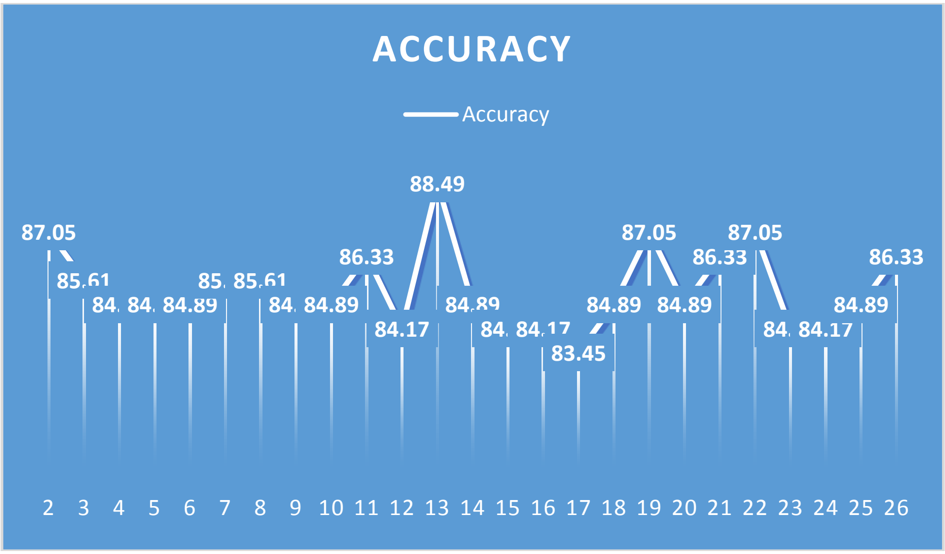

Fig. 3 – Accuracy of diagnosis for different number of hidden layer neurons.

V. Conclusion

The diagnostic analysis is conducted with the help of the following factors: TP, TN, FP and FN (Table 2) [8]. These factors help in evaluating the system based on statistical measures such as Accuracy, Sensitivity, Specificity, Positive Prediction Value and Negative Prediction Value.

From the above mentioned analysis, it has been concluded that the neural network design with 13 hidden layer neurons gives the best accuracy in correctly diagnosing presence or absence of heart disease.

Accuracy is considered as the primary measure to decide which network provides the best diagnosis of heart disease. The other measures can be considered as primary measures if the diagnosis demands the best out of these measures, i.e., if there is a need for more specificity from the system, then the network with more specificity is considered. Overall, the network with 13 hidden neurons provides the best Accuracy, Specificity and Positive Prediction Value.

References

[1]. Oleg Yu. Atkov, Svetlana G. Gorokhova, Alexandr G. Sboev, Eduard V. Generozov, Elena V. Muraseyeva, Svetlana Y. Moroshkina, Nadezhda N. Cherniy, "Coronary heart disease diagnosis by artificial neural networks including genetic polymorphisms and clinical parameters", Journal of Cardiology (2012) 59, 190-194, Journal homepage: www.elsevier.com/locate/jjcc, doi: 10.1016/j.jjcc.2011.11.005.

[2]. Dr. K. Usha Rani, "Analysis of Heart Diseases Dataset using Neural Network Approach", International Journal of Data Mining and Knowledge Management Process (IJDKP) Vol. 1, No. 5, September 2011, doi: 10.5121/ijdkp.2011.1501.

[3]. Jae Kwon Kim, Sanggil Kang, "Neural Network-based Coronary Heart Disease Risk Prediction using Feature Correlation Analysis", Hindawi, Journal of Healthcare Engineering, Volume 2017, Article ID 2780501, 13 pages, https://doi.org/10.1155/2017/2780501.

[4]. NouraAjam, "Heart Disease Diagnoses using Artificial Neural Network", Network and Complex Systems, ISSN 2224-610X (Paper) ISSN 2225-0603 (Online), Vol. 5, No. 4, 2015, www.iiste.org.

[5]. Sandra Cako, Angelina Njegus, Vladimir Matic, "Effective Diagnosis of Heart Disease Presence using Artificial Neural Networks", Sinteza 2017, International Scientific Conference on Information Technology and Data Related Research, doi: 10.15308/Sinteza-2017-3-8.

[6]. Miss Chaitrali S. Dangare, Dr.Mrs.Sulabha S. Apte, "A Data Mining Approach for Prediction of Heart Disease using Neural Networks", International Journal of Computer Engineering and Technology (IJCET), ISSN 0976-6367 (Print), ISSN 0976-6375 (Online), Vol. 3, Issue 3, Oct-Dec 2012, pp. 30-40, www.iaeme.com/ijcet.asp.

[7]. Dataset source: http://archive.ics.uci.edu/ml/datasets/heart+Disease

[8]. ChaitanyaRoygaga, Suraj Punjabi, SwapnilSampat, Ms.TanujaSarode, "Diagnosis of Heart Disease Using Neural Networks", i-manager's Journal on Information Technology, June-August 2018, www.imanagerpublications.com

[9]. SamehGhwanmeh, Adel Mohammad, Ali Al-Ibrahim, "Innovative Artificial Neural Networks-Based Decision Support System for Heart Diseases Diagnosis," Journal of Intelligent Learning Systems and Applications, 2013,5,176-183 http://dx.doi.org/10.4236/jilsa.2013.53019 Published Online August 2013 (http://www.scirp.org/journal/jilsa).

[10]. Ebenezer ObaloluwaOlaniyi and OyebadeKayodeOyedotun, "Heart Diseases Diagnosis Using Neural Networks Arbitration," International JorunalIntelligent Systems and Applications, 2015, 12, 75-82 in MECS (http://www.mecs-press.org/) DOI:10.5815/ijisa.2015.12.08.

[11]. N. Guru, A. Dahiya and N. Rajpal, "Decision Support System for Heart Disease Using Neural Network," Delhi Business Review, vol. 8, No 1, pp. 1 – 6, Jan – June (2007).

[12]. M. Kumari, S. Godara, "Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction," International Journal of Computer Science and Technology, vol. 2, Issue 2, June (2011).

[13]. N. Al-Milli, "Backpropagation neural network for prediction of heart disease," Journal of theoretical and applied information Technology, vol. 56, pp.131-135, Oct 10, 2013.

[14]. A. 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.

ChaitanyaRoygaga "Heart Disease Detection: A Neural Networks Application"International Journal of Engineering Science Invention (IJESI), vol. 07, no. 07, 2018, pp07-11