Heart Disease Prediction Using Scaled Conjugate Gradient Back Propagation of Artificial Neural Network

BANIBRATA PAUL (paul.banibrata1@gmail.com)
Birla Institute of Technology https://orcid.org/0000-0003-2663-2260
Bhaskar Karn
Birla Institute of Technology

Research Article

Keywords: Artificial Neural Network, Cleveland Datasets, Cleveland Hungarian Statlog heart dataset, Heart Disease, Scaled Conjugate Gradient Algorithm

Posted Date: March 31st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1490110/v1

Heart Disease Prediction Using Scaled Conjugate Gradient Back Propagation of Artificial Neural Network

Banibrata Paul[*1], Dr. Bhaskar Karn[2]
1 Research Scholar, Department of Computer Sc & Engineering, BIT, Mesra, Ranchi, Jharkhand, India
2 Associate Professor, Department of Computer Science & Engineering, BIT, Mesra, Ranchi, Jharkhand, India.

Abstract-
Heart disease is a deadly disease in human life. The mortality rate from any disease is the highest in the world. Therefore, before reaching the final stage of this heart disease, all precautionary measures must be taken. For this reason without the help of any kind of traditional methods, if we can scientifically diagnose heart disease at an early stage through various decision support systems, then surely death rate of this disease will decrease in the whole world. Many researchers investigate the diagnosis of heart disease by creating various intelligent medical decision support systems. Artificial neural network concepts represent the highest predictive accuracy over medical data compared to other decision support systems.
In this paper we propose a better prediction method for the existence of heart disease through the scaled conjugate gradient back propagation of artificial neural networks using K-fold cross validation. For cardiac datasets, the University of California Irvine (UCI) Machine Learning Repository and IEEE data port have been used. For Cleveland processed heart dataset, the proposed system uses 13 input attributes and provides minimum 63.3803% & maximum 100% accurate results similarly for Cleveland Hungarian Statlog heart dataset the proposed system uses 11 input attributes and provides minimum 88.4754% & maximum 100% accurate results by estimating the presence and absence of heart disease during testing.

Index Terms- Artificial Neural Network, Cleveland Datasets, Cleveland Hungarian Statlog heart dataset, Heart Disease, Scaled Conjugate Gradient Algorithm.

I. Introduction

The heart is an important part of the human body. It is used to distribute blood to other parts of the body through the blood vessels of the circulatory system. Any disorder that affects the heart is called heart disease and it affects other parts such as the brain, lungs, liver, kidneys, etc. which have a profound effect on human survival. Several factors such as anxiety, mental depression, smoking, lack of physical exercise, blood pressure, cholesterol, obesity increase the risk of heart disease. The World Health Organization (WHO) estimates that by 2030, nearly 23.6 million people will die due to heart disease. As the problem of the disease is not identified in the earlier stages, it causes large number of deaths. If the disease is predicted at an early stage, we can prevent the death of lots of patient. The prediction depends on the symptoms, physical check-up, symptoms of the patient's body and the clinical signs of heart disease such as the presence of many functional and pathological factors. Sometimes these functional and pathological factors delay and complicate the prediction of heart disease which can have negative perceptions and unpredictable effects. Overcoming these false assumptions requires the development of an early prediction and medical diagnosis expert system that offers high accuracy with low operating costs. We therefore propose a new method by applying the scaled conjugate gradient back propagation algorithm with artificial neural networks by using Cleveland Processed database [20] and Cleveland Hungarian Statlog heart database [21], to determine the absence or presence and levels of heart disease. The Cleveland Processed Database contains 76 features and 303 instances, in this dataset all published experiments use a subset of 14 of them. According to the Cleveland processed dataset we have taken all 14 attributes, and not used feature selection method, and out of 14

*Corresponding author
E-mail address: paul.banibrata1@gmail.com

attributes, this system uses 13 input attributes and 1 Output or 'Goal' field.

On the other hand, Cleveland Hungarian Statlog dataset consists of 1190 instances with 11 attributes as input and 1 attribute as output. In this method we have taken all 12 attributes. The details of Cleveland processed dataset and Cleveland Hungarian Statlog heart dataset are explained in the subsection "dataset" under the "Experimentation" section.

II. Literature review

Works are being carried out with different varieties of methodology and reached various classification accuracies. Pronab Ghosh et al. [2021][1] invented an effective prediction of cardiovascular disease using machine learning algorithms with Relief and LASSO feature selection techniques and achieved 99.05% prediction accuracy. M.Kavitha et al. [2021] [2] predicted heart disease using hybrid machine learning techniques such as random-forest using decision tree and obtained 88.7% accuracy. Jian Ping Li et al. [2020] [3] proposed a model for heart disease diagnosis using machine learning classification techniques and achieved 92.37% accuracy. Syed Arslan Ali et al.[2020] [4]established an optimally configured and improved deep belief network (OCI-DBN) approach to predict heart disease by using Ruzzo-Tompa and stacked genetic algorithm and obtained 94.61 % prediction accuracy. Jikuo Wang et al. [2020] [5] invented a staking based model for non invasive detection of heart disease and obtained 95.43% accuracy. Mohammad Ayoub Khan [2020] [6] proposed an IoT framework for evaluation of heart disease more accurately using a modified deep convolutional neural network (MDCNN) and achieved an accuracy of 98.2%. Jayavani Vankara et al. [2020] [7] invented a predictive analysis by ensemble learning and classification for heart disease detection and obtained 93% accuracy. Senthil kumar Mohan et al. [2019] [8] predicted a model for heart disease using hybrid machine learning techniques and obtained 88.7% accuracy label. Liaqat Ali et al. [2019] [9] proposed a hybrid model named χ^2 statistical model and optimally configured deep neural network (DNN) and achieved the prediction accuracy of 93.33%. C.Beulah Christalin Latha et al. (2019) [10] established a new model for improving the accuracy of heart disease predictions based on ensemble classification techniques and achieved 85% accuracy. R.Ruhin Kouser et al. (2018) [11] proposed a model for a cardiovascular prediction system using artificial neural networks, radial base functions, and case- based reasoning, and obtained 97% to 98% accuracy. Ashok Kumar Dwivedi [2018] [12] established method for performance evaluation by using different machine learning techniques for proper diagnosis of heart disease and achieved 85% accuracy. Ji Zhang et al. [2017] [13] introduced a fast Fourier transformation-coupled machine learning ensemble model for heart disease prediction. The experimental results showed 91% to 94% prediction accuracy. N.Leema et al.(2016)[14] designed a computer-aided diagnostic system (CAD) for the prediction of heart disease using differential evolution with worldwide information and back propagation algorithms and obtained 86% accuracy. K.Rajeswari et al.(2012)[15] designed a model for feature selection in the detection of ischemic heart disease using the feed forward neural network and achieved 89% accuracy during training data and 82% accuracy during test data, respectively. Anchana Khemphila et al.(2011)[16] has achieved 89% accuracy in their model for classifying heart disease and training datasets using neural networks and feature selection. Roya Asadi et al. (2009)[17] worked on the supervised multilayer feed forward neural network model to accelerate the classification problem and got 94% prediction accuracy. Hongmei Yan et al.(2006)[18] found percentage of accuracy in the interval [88.6, 93.2] with 91% mean accuracy in multilayer perceptron-based medical decision support system for the diagnosis of heart disease. The above mentioned related works are explained in a following table.

SL NO	Method	Accuracy%	Reference
1.	Scaled Conjugate Gradient Back Propagation of artificial neural networks using Cleveland	Minimum Accuracy 63.3803 % Maximum Accuracy 100%	In this study.

	processed dataset		
2.	Scaled Conjugate Gradient Back Propagation of artificial neural networks using Cleveland Hungarian Statlog heart dataset	Minimum Accuracy 88.4754 % Maximum Accuracy 100%	In this study.
3.	Machine learning algorithms with Relief and LASSO feature selection techniques	99.05%	Pronab Ghosh et al. [2021]
4.	Hybrid machine learning techniques (random-forest + decision tree)	88.7%	M.Kavitha et al. [2021]
5.	Machine Learning Classification techniques	92.37%	Jian Ping Li et al. [2020]
6.	Optimally Configured and Improved Deep belief network (OCI-DBN) approach by using Ruzzo-Tompa and stacked genetic algorithm.	94.61 %	Syed Arslan Ali et al.[2020]
7.	Staking Based Model	95.43%	Jikuo Wang et al. [2020]
8.	Modified Deep Convolutional Neural Network (MDCNN)	98.2%.	Mohammad Ayoub Khan [2020]
9.	Ensemble Learning and Classification techniques	93%	Jayavani Vankara et al. [2020]
10.	Hybrid Machine Learning techniques	88.7%	Senthil kumar Mohan et al. [2019]
11.	Hybrid model named $\chi 2$ statistical model and optimally configured Deep Neural Network (DNN)	93.33%	Liaqat Ali et al. [2019]
12.	Ensemble Classification techniques	85%	C.Beulah Christalin Latha et al. (2019)
13.	Artificial Neural Networks, Radial Base Functions, and Case- Based Reasoning	97% to 98%	R.Ruhin Kouser et al. (2018)
14.	Machine Learning Techniques	85%	Ashok Kumar Dwivedi [2018]
15.	Fast Fourier transformation-Coupled Machine Learning Ensemble model.	91% to 94%	Ji Zhang et al. [2017]
16.	Differential Evolution with Worldwide Information and Back Propagation algorithms.	86%	N.Leema et al.(2016)
17.	Feature Selection using Feed Forward Neural Network	Training data :- 89% Test data :- 82%	K.Rajeswari et al.(2012)
18.	Neural networks + Feature Selection	89%	Anchana Khemphila et al.(2011)
19.	Supervised Multilayer Feed Forward Neural Network Model	94%	Roya Asadi et al. (2009)
20.	Multilayer Perceptron-Based Medical Decision Support System	88.6% to 93.2 %	Hongmei Yan et al.(2006)

Table 1: Related works through literature survey

III. Methodology

Proposed Model and its Design

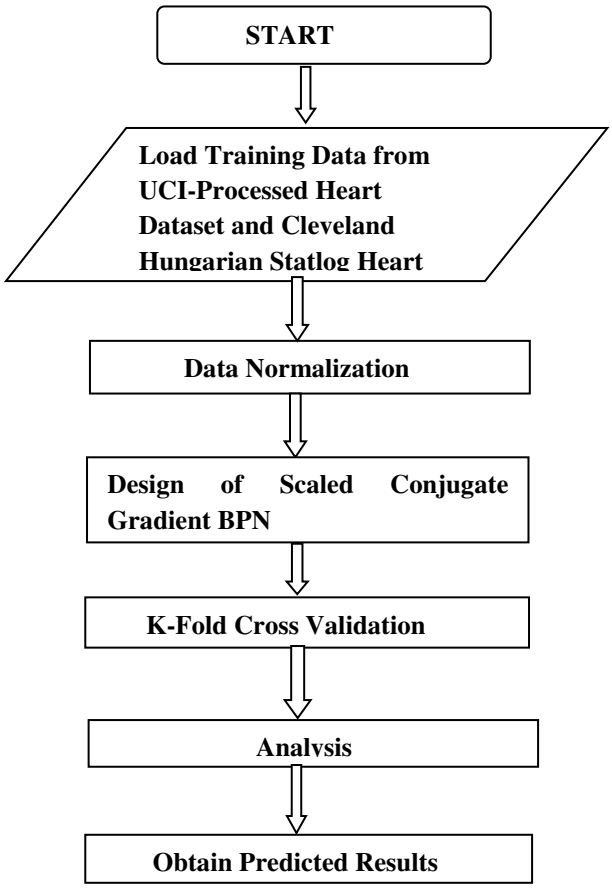

Figure 1: Block Diagram of Proposed Model

The overall methods of the proposed model are shown in Figure 1. The original datasets are taken from the Cleveland Processed Database [20] and Cleveland Hungarian Statlog heart database [21] respectively. After processing, clinical datasets are normalized using the following mathematical formula.

$$\text{Normalized}(X) = \frac{\text{Original value in the given set}}{\text{Maximum value in the given range}}$$

Where, the normalized value lies in the interval [0,1]. For example, patients' blood pressure is between 50 and 200 and if a patient's blood pressure is 145 then it can be normalized as $\frac{145}{200}$ = 0.725. After normalizing, loading it and importing into the MATLAB in table format which consists of 14 columns and 303 rows for Cleveland processed dataset, & 12 columns and 1190 rows for Cleveland Hungarian Statlog heart dataset respectively then convert the table into an array and apply the scaled conjugate gradient back propagation algorithm. In this technique, sorting the data and storing it in the 'X' variable, it sorts the rows of the matrix in ascending order based on the elements in the first column. When the first column contains repeating elements, sort the rows according to the values of the next column and repeat this process to achieve equal values. Now, take input from the user as a percentage of training. The prompt takes variable input and stores it in variable percentages. Next, perform cross-validation partitions for data training and testing. "cvpartition" constructs an object c of the "cvpartition" class defining a random non-stratified partition for k=5 fold cross-validation on n_1=303 and n_2=1190 observations. Here "cvpartition" randomly selects p numbers of observations (When $0 < p < 1$, the default value of p is 1/10) for the test set. The partition divides the observation into k disjoint subsamples (or folds), randomly selected but almost identical in size. Now, the data needs to be prepared for training data and testing data for neural network training. Enter the learning rate, number of input neurons, number of hidden neurons and number of output neurons and stored at 'Prompt'. Where 'Prompt' is the place where user type commands, formulas, and functions or perform tasks using MATLAB. Then get ready for network training. Thereafter, test our data on the trained network. Finally calculating the errors, performance, percentage errors, MSE, accuracy, percentage of accuracy, time elapse and plotting confusion matrix.

Pseudo Code of the proposed model is explained below.

Proposed Algorithm

Step1: Start

Step2.1: Load Training Data from UCI-Processed Heart Dataset and Cleveland Hungarian Statlog Heart Dataset respectively

Step2.2: Import data into the MATLAB in table format which consists of 14 columns and 303 rows for Cleveland processed dataset & 12 columns and 1190 rows for Cleveland Hungarian Statlog heart dataset respectively.

Step2.3: Convert table into the Array

Step3: Apply Scaled Conjugate Gradient Back Propagation Algorithm

Step4: Training Dataset

Step5: Calculation of error and accuracy

Step6: Testing Dataset

Step7: Calculation of error, accuracy, time elapses and plots Confusion Matrix

Step8: Stop

Here Cleveland heart dataset is taken from UCI Machine Learning Repository and Cleveland Hungarian Statlog heart dataset is taken from IEEE data port respectively.

IV Experimentation

i. Data Set

The cardiac dataset used for the experiment is taken from the UCI Machine Learning Repository and IEEE data port.The presence or absence and levels of heart disease are determined using the Cleveland processed heart database and Cleveland Hungarian Statlog heart database respectively. Cleveland processed heart data set contains 76 attributes and 303 numbers of instances, but all published experiments use a subset of 14 of them. This system uses 13 (thirteen) most significant attributes as input and one attribute as output or "target" field. Input attributes are age, gender, chest pain, blood pressure, cholesterol, blood sugar, ECG, maximum heart rate, exercise induced angina, old peak, slope (The slope of the peak exercise ST segment), CA (number of major vessels (0-3) colored by fluoroscopy), thallium scan (3 = normal, 6= fixed defect, 7= reversible defect). The output field is named as "Diagnosis of Heart Disease". It has 5 integer values (0-4). The value of the integer 0 means "No Presence of Heart Disease(healthy)", similarly the value of the integer '1' means 'sick 1', '2' means 'sick 2', '3' means 'sick 3', and '4' means 'sick 4'respectively.

On the other hand, Cleveland Hungarian Statlog dataset consists of 1190 instances with 11 input attributes and 1 target output. Input attributes are age, gender, chest pain, resting blood pressure, cholesterol, fasting blood sugar, resting ECG, maximum heart rate, exercise angina, old peak, ST-slope. The target output has two levels (0 and 1). The value of the integer 0 represent 'Normal' and 1 represent 'Heart Disease' respectively.

Two databases are explained in Table 1 and Table2 respectively.

S.No	Attribute	Description	Range
1	Age	Patient's Age	29-77
2	Gender	0=female; 1=male	0-1
3	Chest Pain	Value 1:Typical Angina Value 2: Atypical Angina Value 3: Non-Anginal Pain Value 4: Asymptotic	1-4
4	Blood Pressure	blood pressure(in mm Hg)	94-200
5	Cholesterol	Serum cholesterol in mg/dl	126-564
6	Blood Sugar	(blood sugar .120mg/dl) (0=False; 1=True)	0-1
7	ECG	Electrocardiography Results Value 0: Normal Value 1: Having ST-T wave abnormality (T wave inversions and/or ST Elevation or depression of>0.05mV) Value 2: It shows probable or definite left	0-2
8	Maximum Heart Rate	Max. Heart Rate achieved	71-202
9	Exercise Induced Angina	Exercise Induced Angina(0=no;1=yes)	0-1
10	Old Peak	ST depression induced by exercise relative to rest	0-6.2
11	Slope	The slope of the peak exercise ST segment Value 1: Up Sloping Value 2: Flat Value 3:Down Sloping	1-3
12	CA	Number of Major Vessels (0-3) Colored by Fluoroscopy	0-3
13	T. Scan	Normal, Fixed Defect, Reversible Defect	3-7
14	Output (Levels of Heart Disease)	0 : No Presence of Heart Disease 1 : Sick1 2 : Sick2 3 : Sick3 4 : Sick4	0-4

Table 2: Cleveland Processed Heart Data Set

S.No	Attribute	Description	Data type
1	Age	Patient's Age in years	Numeric
2	Gender	0=female; 1=male	Binary
3	Chest Pain Type	Value 1:Typical Angina Value 2: Atypical Angina Value 3: Non-Anginal Pain Value 4: Asymptotic	Nominal
4	Resting Blood Pressure	blood pressure(in mm Hg)	Numeric
5	Cholesterol	Serum cholesterol in mg/dL	Numeric

6	Fasting Blood Sugar	(blood sugar>120 mg/dl) (1=true; 0 = false)	Binary
7	Resting ECG	Electrocardiography Results Value 0: Normal Value 1: Having ST-T wave abnormality (T wave inversions and/or ST Elevation or depression of>0.05mV) Value 2: It shows probable or definite left	Nominal
8	Maximum Heart Rate	Max. Heart Rate achieved	Numeric
9	Exercise Angina	Exercise Induced Angina(0=no;1=yes)	Binary
10	Old Peak	ST depression induced by exercise relative to rest	Numeric
11	ST-Slope	The slope of the peak exercise ST segment Value 1: Up Sloping Value 2: Flat Value 3:Down Sloping	Nominal
12	Target	0 = Normal ,1 = heart disease	Binary

Table 3: Cleveland Hungarian Statlog Heart Data Set

ii. Data Normalization

The clinical datasets used in this work are normalized using the following mathematical formula

$$\text{Normalized(X)} = \frac{\text{Original value in the given set}}{\text{Maximum value in the given range}}$$

The normalized value 'X' lies in the interval [0,1] Numerical variables such as 'age' is normalized on to the interval [0, 1]. For example, the age of patients range from 29 to 77 years, and thus normalized value of 56 years old patient age is $\frac{56}{77} = 0.7272727$.

iii. Scaled Conjugate Gradient method

This algorithm was introduced by Martin F. Moller in the 1991. It is a feed forward neural network based on supervised learning algorithm and doesn't contain any of the user dependent parameters. This algorithm keep away from time-consuming line search per learning iteration. It leads to better performance than standard back propagation algorithm, conjugate gradient algorithm with line search and 'Broyden Fletcher Goldfarb Shanno' (BFGS) algorithm. It requires a larger number of iterations but less computational complexity. The SCG algorithm [19] is described below.

1. Choose weight vector w_1 and scalars $\sigma > 0$, $\lambda_1 > 0$ and $\overline{\lambda_1} = 0$.
   Set $p_1 = r_1 = -E'(w_1)$, $k = 1$ and
   Success= true
2. If success = true then calculate second order information:
   $$\sigma_k = \frac{\sigma}{|p_k|},$$
   $$s_k = \frac{E'(w_k + \sigma_k p_k) - E'(w_k)}{\sigma_k}$$
   $$\delta_k = p^T{}_k\, s_k$$
3. Scale s_k :
   $$s_k = s_k + (\lambda_k - \overline{\lambda_k})p_k,$$
   $$\delta_k = \delta_k + (\lambda_k - \overline{\lambda_k})|p_k|^2$$
4. If $\delta_k \leq 0$, then make positive definite Hessian Matrix
   $$s_k = s_k + (\lambda_k - 2\frac{\delta_k}{|p_k|^2})p_k,$$
   $$\overline{\lambda_k} = 2(\lambda_k - \frac{\delta_k}{|p_k|^2}),$$
   $$\delta_k = -\delta_k + \lambda_k|p_k|^2,\ \lambda_k = \overline{\lambda_k},$$
5. Calculate step size :
   $$\mu_k = p^T{}_k r_k,\quad \alpha_k = \frac{\mu_k}{\delta_k}$$
6. Calculate the comparison parameter
   $$\Delta_k = \frac{2\delta_k[E(w_k) - E(w_k + \alpha_k p_k)]}{\mu_k^2}$$
7. If $\Delta_k \geq 0$ then a successful reduction in error can be made :
   $$w_{k+1} = w_k + \alpha_k p_k,$$
   $$r_{k+1} = -E'(w_{k+1}),$$
   $\overline{\lambda_k} = 0$, success =true,
   7a. If k mod N=0, then restart algorithm: $p_{k+1} = r_{k+1}$
   Else create new conjugate direction:
   $$\beta_k = \frac{|r_{k+1}|^2 - r_{k+1} r_k}{\mu_k}$$
   $$p_{k+1} = r_{k+1} + \beta_k p_k$$
   7b. If, $\Delta_k \geq 0.75$ then reduce the scale parameter: $\lambda_k = \frac{1}{2}\lambda_k$
   Else, reduction in error is not possible: $\overline{\lambda_k} = \lambda_k$, success = false

8. If, $\Delta_k < 0.25$ then increase the scale parameter: $\lambda_k = 4\lambda_k$
9. If the Steepest Descent direction $r_k \neq 0$ then set k=k+1 and go to 2,
   Else, terminate and return w_{k+1} as the desired minimum.

Where σ ($\leq 10^{-4}$) is kept small and it implies that it is not critical for SCG's performance.

There is one call of E(w) and two calls of E'(w), for each iteration. In this algorithm the calculation complexity per iteration is of $O(3N^2)$ and it can be reduced to $O(2N^2)$.

All symbols in a SCG algorithm are explained in a table.

Sl No	Symbol	Meaning
1	w_k	Weight of Kth points
2	s_k	Kth scale.
3	r_k	Kth Steepest Descent direction
4	Δ_k	Kth comparison parameter
5	λ_k	Kth comparison parameter
6	α_k	Kth step size
7	p_k	Kth search direction
8	σ_k	Kth scalar
9	β_k	Kth new conjugate direction
10	E (w_k)	Error function of the kth point w_k in Euclidian space R^N

Table4: symbols and their meaning

iv. K- Fold Cross Validation

Here initial data sets are randomly partitioned into 'k' number of mutually exclusive subsets or folds.

"D1, D2, D3 ... DK" are almost every equal size. Training and testing are performed 'K' times.

The 1st iteration D_1, is reserved as the test set and the remaining D_2, D_3, D_4…DK are served as the training sets.

The 2nd iteration D_2, is served as the test set and the remaining D_1, D_3, D_4...D_K are served as the training sets.

Similarly the ith iteration D_i is served as the test set and the remaining D_1, D_2 .$D_{(i-1)}$, $D_{(i+1)}$, D_K are served as the training sets. Each sample is used the same number of times for training and once for testing. For classification problem

$$\text{Accuracy} = \frac{\text{overall numbers of correct classification from K iteration}}{\text{Total numbers of tuples in the initial data.}}$$

The classification, accuracy of the proposed method is evaluated with k=5 fold cross validation of the samples. The results obtained from each fold is averaged and used for comparative analysis.

v. Classification Accuracy

Classification accuracy is used to predict the performance results of the proposed method in the case of accurate diagnostic results. The accuracy of the classification is measured using the confusion matrix tool. The confusion matrix is a tool. It analyzes how well the classifier can recognize tuples of different classes. The processing outcome of the confusion matrix is shown in figure 2

Confusion Matrix

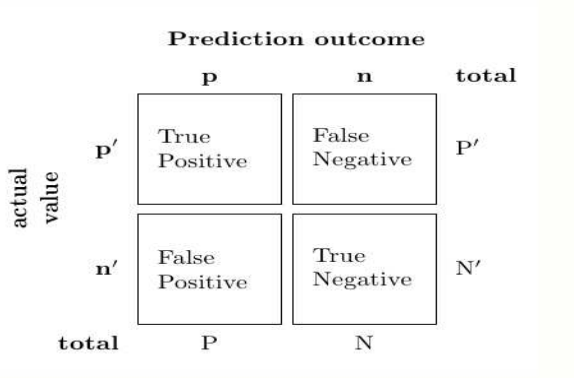

Figure 2: Confusion Matrix

There are four additional terms **"True Positive", "True Negative", "False Positive" and "False Negative",** which are explained below.

True Positive (TP): It is an outcome where the model correctly predicts the positive class.

True Negative (TN): It is an outcome where the model correctly predicts the negative class.

False Positive (FP): It is an outcome where the model incorrectly predicts the positive class.

False Negative (FN): It is an outcome where the model incorrectly predicts the negative class.

$$\text{Accuracy} = \frac{\text{True Positive + True Negative}}{\textit{True Positive+True Negative+False Positive+False Negative}}$$

V Experimental Result & Observation

All experimental results are obtained by using MATLAB coding & implementation.

1. All experimental results for the Cleveland processed heart dataset are explained below

1.1. Input percentage of training data: 70
Enter learning Rate: 0.75
Number of Input Neurons: 13
Number of Hidden Neurons: 1
Number of Output Neurons: 5

Performance =0.1667
Percent Errors = 0.3662
Accuracy =0.6338
Percentage of Accuracy = 63.3803
MSE = 0.0897
Time Elapsed = 54.2264 sec

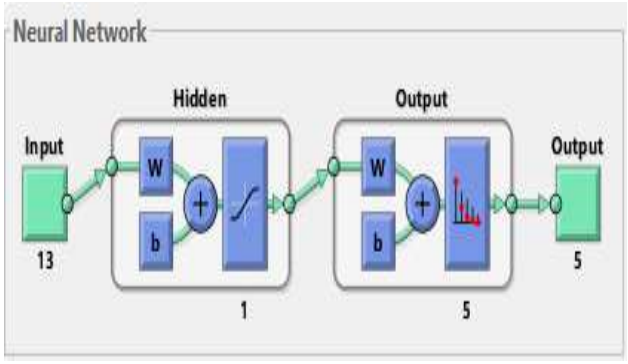

Figure 3: Three layered neural network model

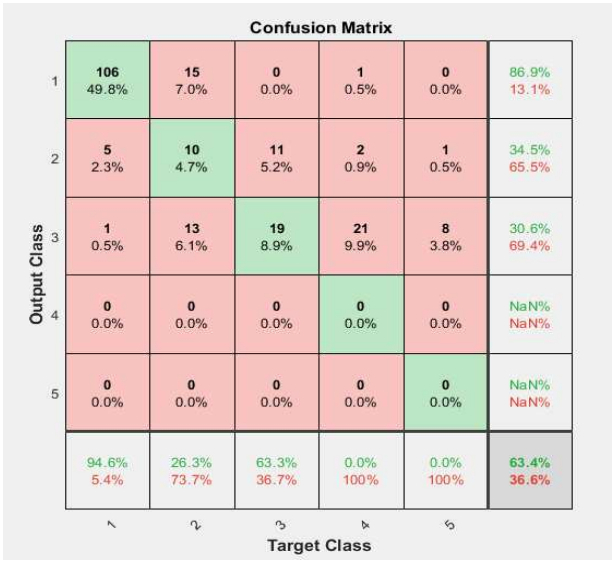

Figure 4 : confusion matrix

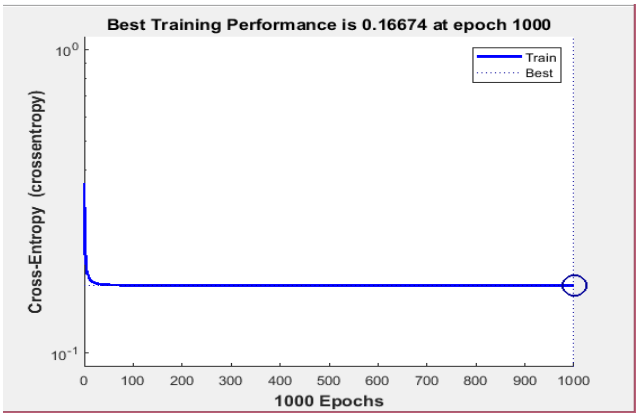

Figure5 : Neural network training performance

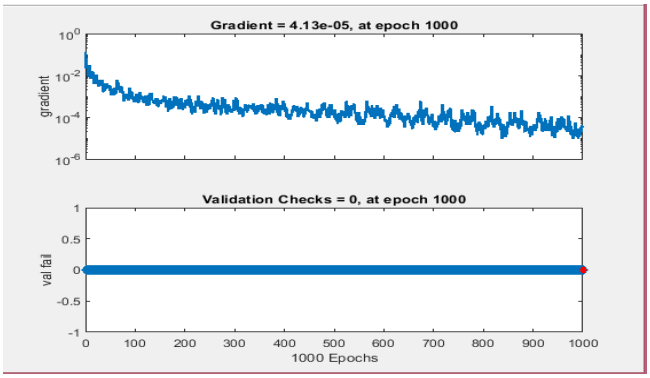

Figure6 : Neural network training state

1.2. Input percentage of training data: 70
Enter learning Rate: 0.75
Number of Input Neurons: 13
Number of Hidden Neurons: 5
Number of Output Neurons: 5

Performance =0.0484
Percent Errors = 0.1127
Accuracy =0.8873
Percentage of Accuracy = 88.7324
MSE = 0.0275
Time Elapsed = 22.0929 Sec

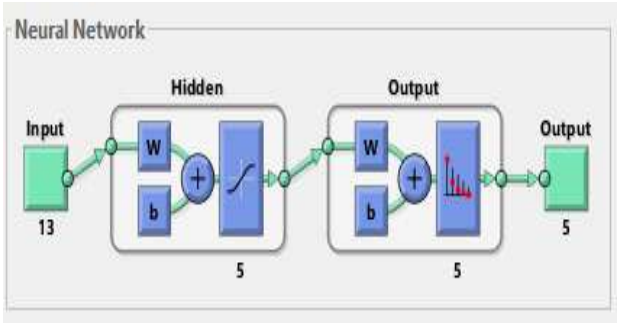

Figure 7: Three layered neural network model

Figure 8 : confusion matrix

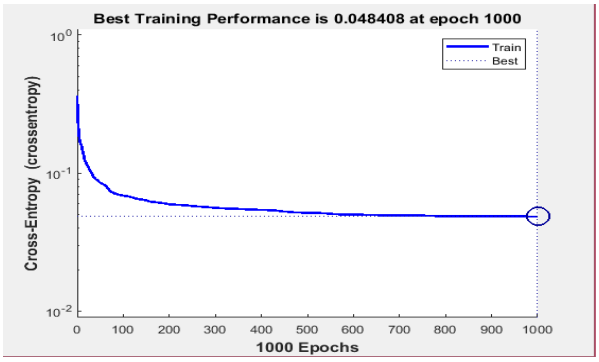

Figure 9 : Neural network training performance

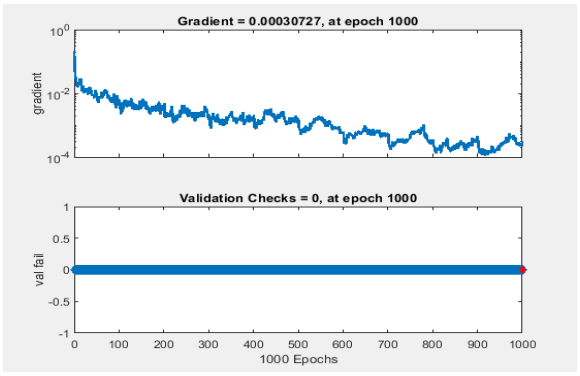

Figure10 : Neural network training state

1.3. Input percentage of training data: 70
Enter learning Rate: 0.75
Number of Input Neurons: 13
Number of Hidden Neurons: 10
Number of Output Neurons: 5

Performance = 1.6627e-06
Percent Errors = 0
Accuracy =1
Percentage of Accuracy = 100
MSE = 1.7407e-10
Time Elapsed = 14.4303 Sec

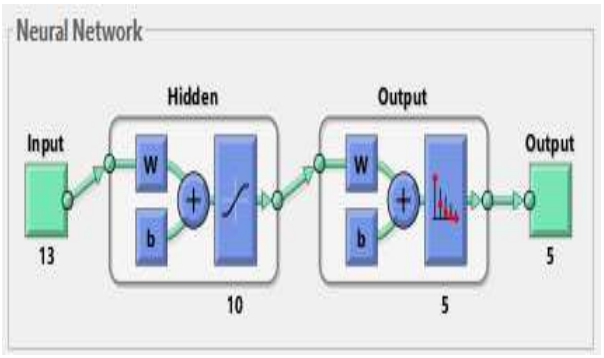

Figure 11: Three layered neural network model

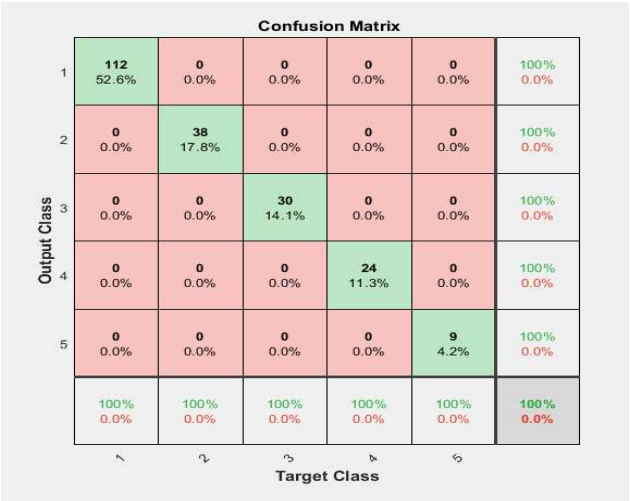

Figure 12 : confusion matrix

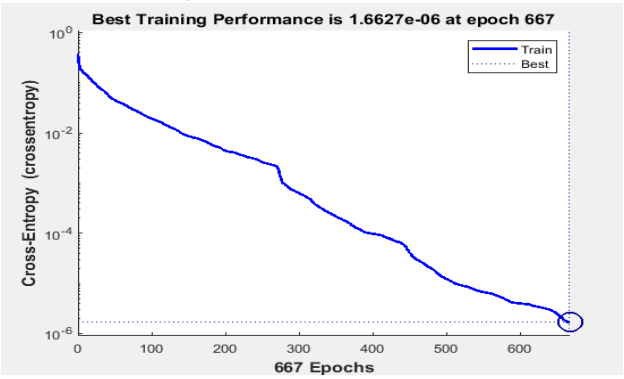

Figure 13 : Neural network training performance

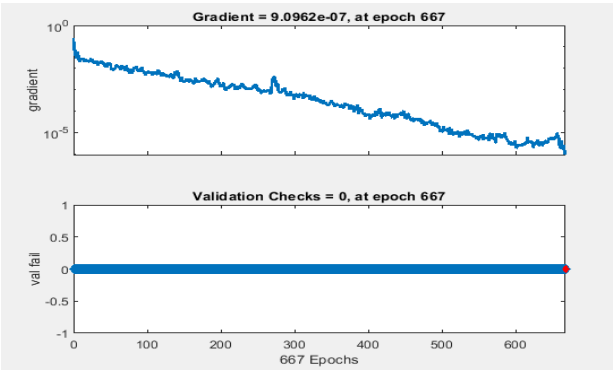

Figure14 : Neural network training state

1.4. Input percentage of training data: 70
Enter learning Rate: 0.75
Number of Input Neurons: 13
Number of Hidden Neurons: 13
Number of Output Neurons: 5

Performance = 1.4358e-06
Percent Errors = 0
Accuracy =1
Percentage of Accuracy = 100
MSE = 1.1449e-10
Time Elapsed = 18.7217Sec

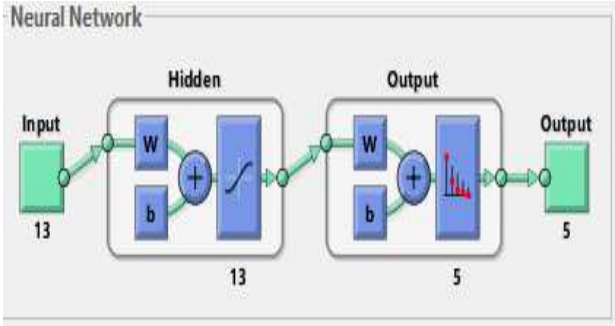

Figure 15: Three layered neural network model

Figure 16 : confusion matrix

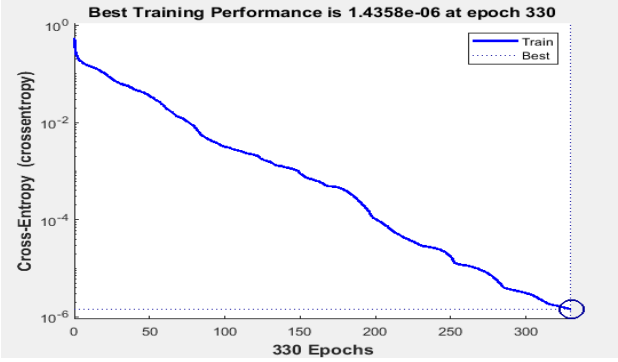

Figure 17 : Neural network training performance

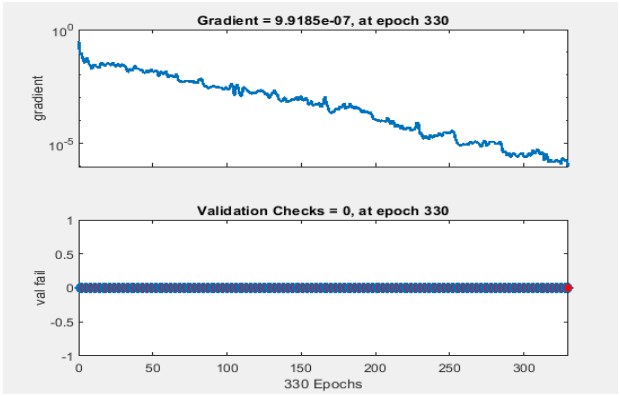

Figure18 : Neural network training state

2. All experimental results for the Cleveland Hungarian Statlog heart dataset are explained below.

2.1. Input percentage of training data: 70
Enter learning Rate: 0.75
Number of Input Neurons: 11
Number of Hidden Neurons: 1
Number of Output Neurons: 2

Performance = 0.1809
Percent Errors = 0.1152
Accuracy = 0.8848
Percentage of Accuracy = 88.4754
MSE = 0.1036
Time Elapsed = 23.5838 sec

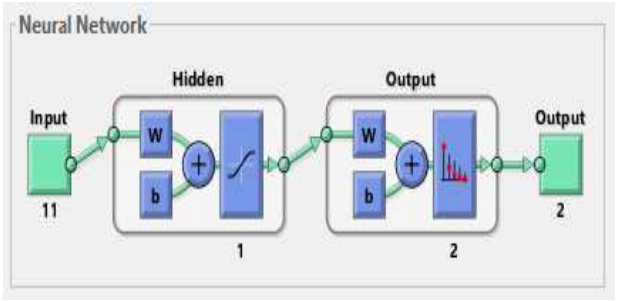

Figure 19: Three layered neural network model

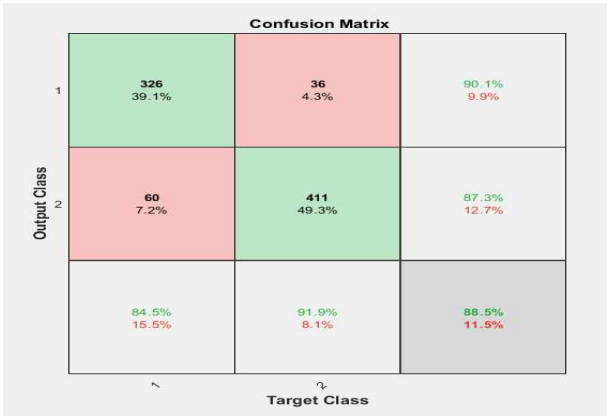

Figure 20 : confusion matrix

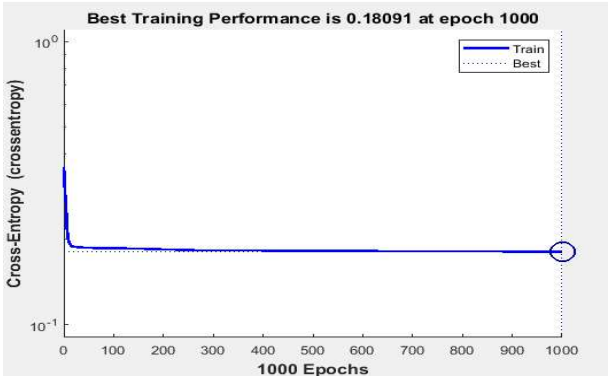

Figure 21 : Neural network training performance

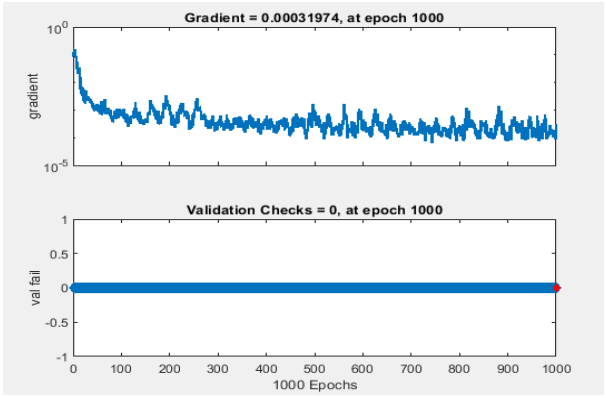

Figure22 : Neural network training state

2.2. Input percentage of training data: 70
Enter learning Rate: 0.75
Number of Input Neurons: 11
Number of Hidden Neurons: 5
Number of Output Neurons: 2

Performance = 0.0925
Percent Errors = 0.0624
Accuracy = 0.9376
Percentage of Accuracy = 93.7575
MSE = 0.0505
Time Elapsed = 17.7577sec

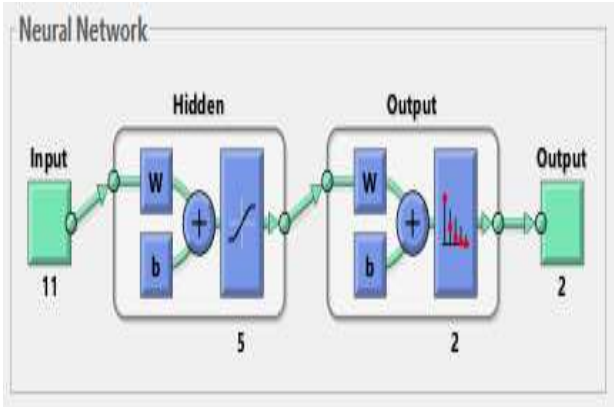

Figure 23: Three layered neural network model

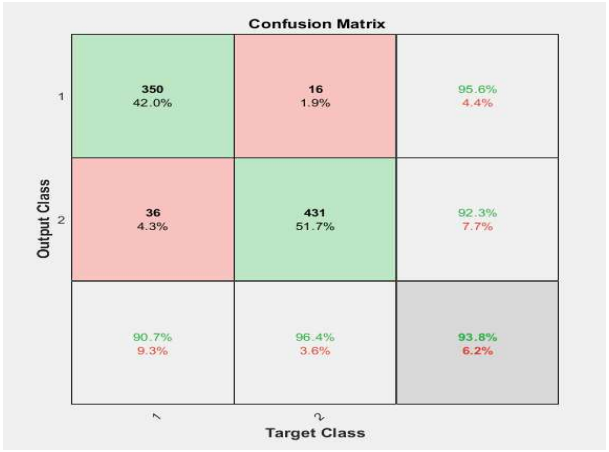

Figure 24 : confusion matrix

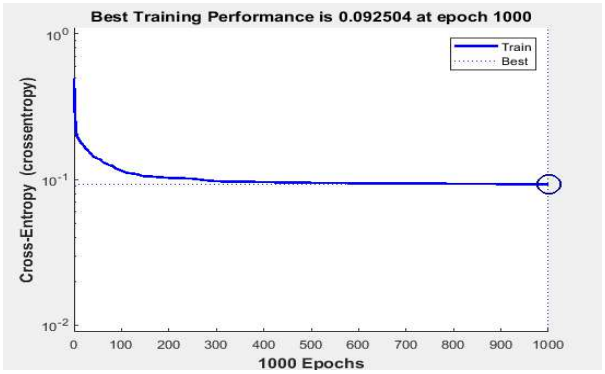

Figure 25 : Neural network training performance

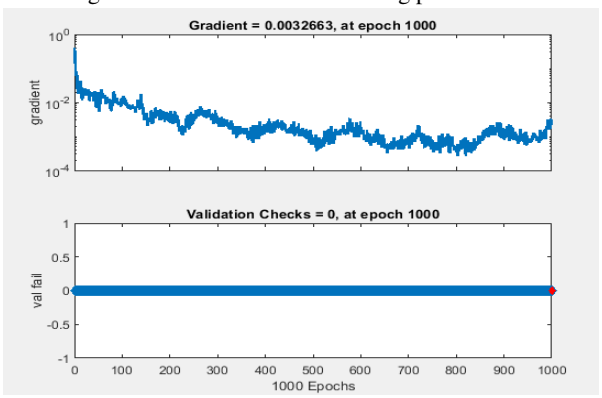

Figure 26 : Neural network training state

2.3.Input percentage of training data: 70
Enter learning Rate: 0.75
Number of Input Neurons: 11
Number of Hidden Neurons: 10
Number of Output Neurons: 2

Performance = 0.0271
Percent Errors = 0.0204
Accuracy = 0.9796
Percentage of Accuracy = 97.9592
MSE = 0.0169
Time Elapsed = 25.3934 sec

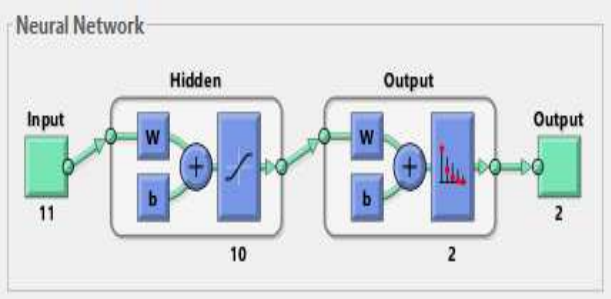

Figure 27: Three layered neural network model

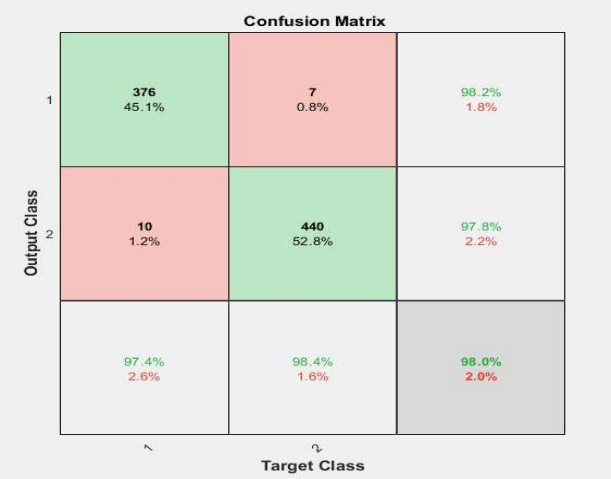

Figure 28 : confusion matrix

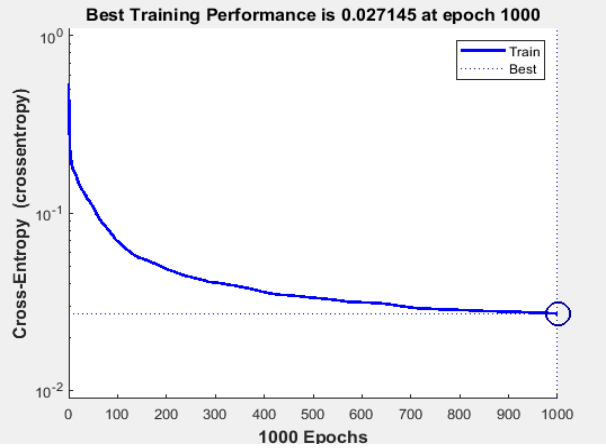

Figure 29 : Neural network training performance

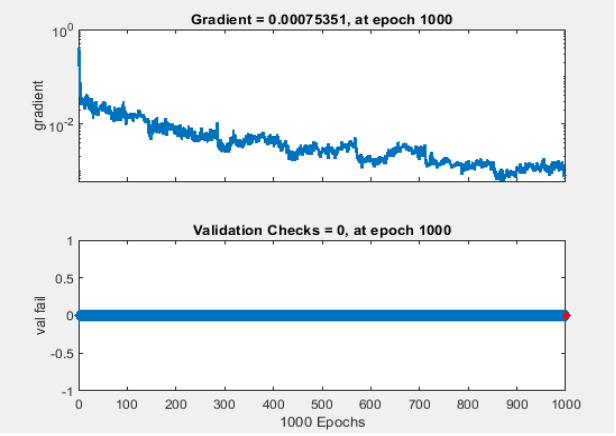

Figure 30 : Neural network training state

2.4.Input percentage of training data: 70
Enter learning Rate: 0.75
Number of Input Neurons: 11
Number of Hidden Neurons: 13
Number of Output Neurons: 2

Performance = 4.3645e-06
Percent Errors = 0
Accuracy = 1
Percentage of Accuracy = 100
MSE = 1.6701e-09
Time Elapsed = 24.2767sec

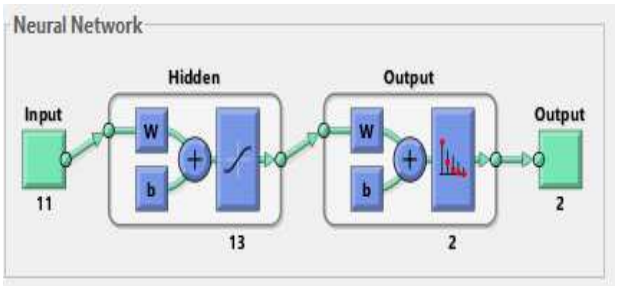

Figure 31: Three layered neural network model

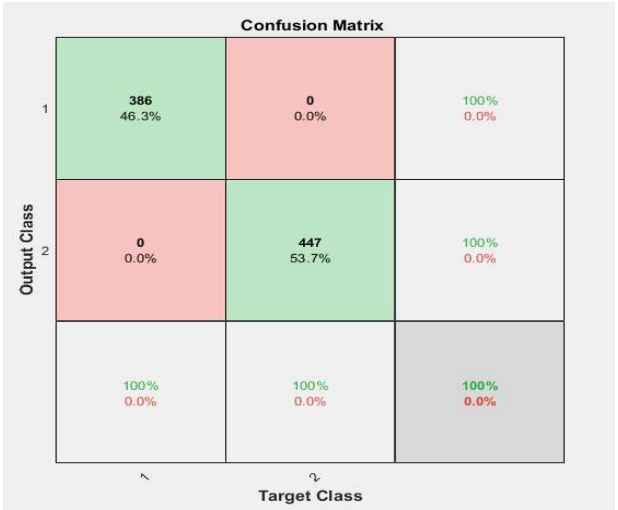

Figure 32 : confusion matrix

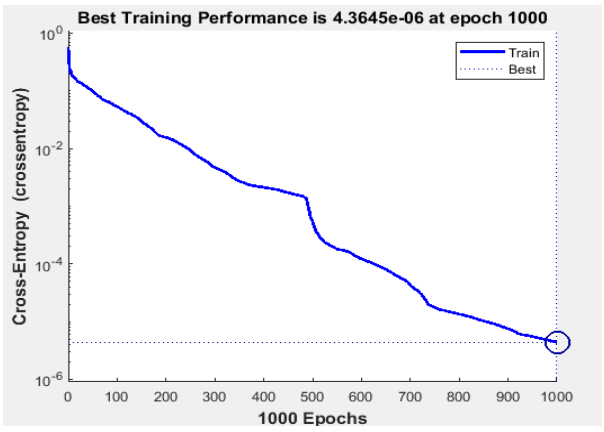

Figure 33 : Neural network training performance

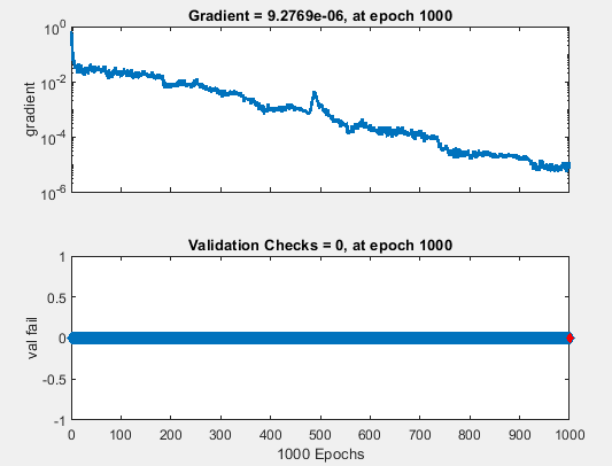

Figure 34 : Neural network training state

3. Experimental results of Cleveland Processed dataset

From above experimental result it is shown that, when input percentage of training data=70, learning rate =0.75, number of input neurons =13, number of output neurons=5, then for different numbers of hidden neurons, percentage of accuracy, time elapsed and M.S.E (Mean Squared Error) are changed accordingly.

It is explained in table 5.

Number of Hidden Neurons	Percentage(%) of Accuracy	Time Elapse(Sec)	M.S.E
1	63.3803	54.2264	0.0897
5	88.7324	22.0929	0.0275
10	100	14.4303	1.7407e-10
13	100	18.7217	1.1449e-10

Table 5: Different outputs of Cleveland Processed dataset

4. Experimental results of Cleveland Hungarian Statlog dataset

From above experimental result it is shown that, when input percentage of training data=70, learning rate =0.75, number of input neurons =11, number of output neurons=2, then for different numbers of hidden neurons, percentage of accuracy, time elapsed and M.S.E (Mean Squared Error) are changed accordingly.

It is described in table 6.

Number of Hidden Neurons	Percentage(%) of Accuracy	Time Elapse(Sec)	M.S.E
1	88.4754	23.5838	0.1036
5	93.7575	17.7577	0.0505
10	97.9592	25.3934	0.0169
13	100	24.2767	1.6701e-09

Table 6: Different outputs of Cleveland Hungarian Statlog dataset

5. Performance analysis between two datasets based on the experimental results

The following graphical representations are presented based on table 5 and table 6

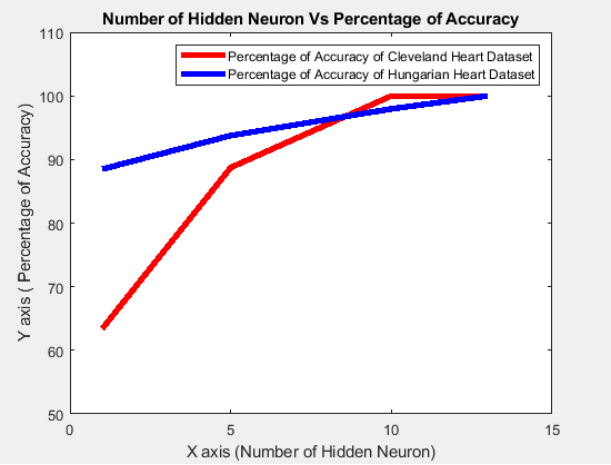

Figure35: Number of Hidden Neuron Vs Percentage of Accuracy

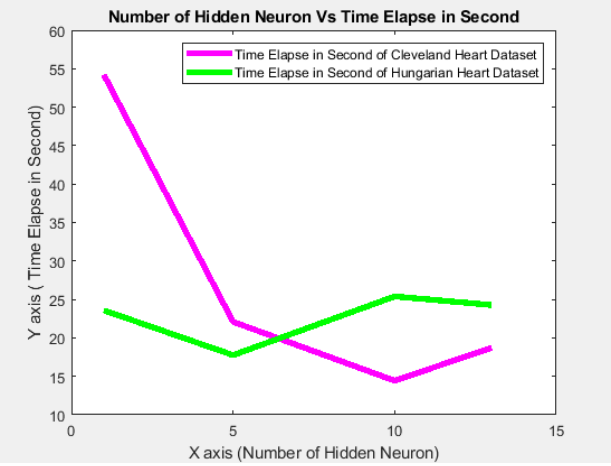

Figure36: Number of Hidden Neuron Vs Time Elapse in Second

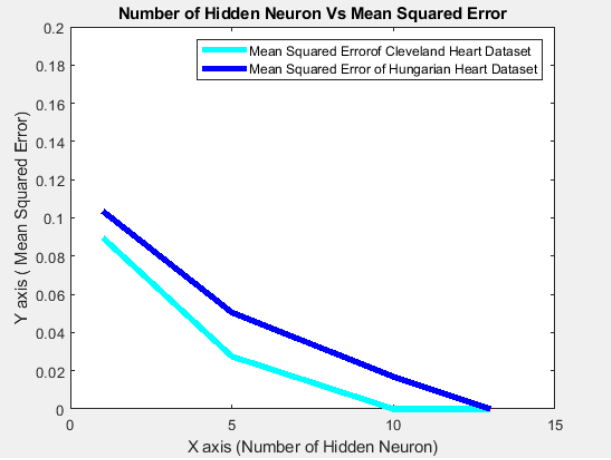

Figure37: Number of Hidden Neuron Vs Mean Squared Error

Observation

During our experimentation, we found that if the numbers of neurons in hidden layer are 10 or more for the Cleveland heart dataset and similarly numbers of neurons in hidden layer are 13 and more for Cleveland Hungarian Statlog dataset, then the accuracy of the result is 100%.

VI. Conclusion

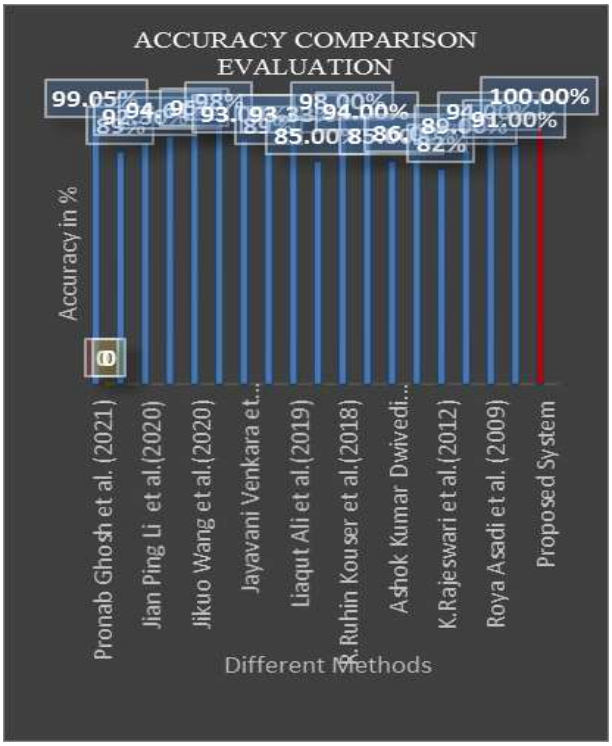

Figure: 38: Accuracy Comparison Evaluation

The proposed heart disease prediction system with accurate diagnosis has been developed using the scaled conjugate gradient back propagation neural network. Since this algorithm does not contain any user-dependent parameters whose values are crucial to the success of this method and uses the step size scaling process, this algorithm avoids time-consuming line searches per learning iteration, which makes the algorithm faster than other adaptive learning algorithms. This algorithm is repeated until the minimum error rate is observed. From figs. 37(histogram) it is clear that the proposed method has the highest maximum accuracy rate compared to various other methods. Experimental results proves that the percentage of prediction accuracy (63.3803% to 100%) for Cleveland processed dataset and (88.4754% to 100%) for Cleveland Hungarian Statlog heart dataset varies for taking of different hidden neurons. Thus the experimental results give good and encouraging results to predict heart disease with the best possible improved accuracy.

Compliance with Ethical Standards

The authors have used UCI Machine Learning Repository and which is available in public domain, and downloaded this database free of charge and utilized this dataset for only experiment purpose for ongoing Ph.D. work.

No data has been obtained directly from human.

No conflict of interest among the authors neither with any organization.

Conflict of interest statement

We have no financial or personal relationship with other individuals or organizations, which adversely affects our work.

Acknowledgement

Authors would like to thank Birla Institute of Technology, Mesra, Ranchi, Jharkhand, India for supporting Ph.D. work.

REFERENCES

[1] Ghosh, Pronab, Sami Azam, Mirjam Jonkman, Asif Karim, FM Javed

Mehedi Shamrat, Eva Ignatious, Shahana Shultana, Abhijith Reddy Beeravolu, and Friso De Boer. “Efficient prediction of cardiovascular disease using machine learning algorithms with relief and LASSO feature selection techniques.” IEEE Access 9 (2021): 19304-19326.

[2] Kavitha, M., G. Gnaneswar, R. Dinesh, Y. Rohith Sai, and R. Sai Suraj. “Heart disease prediction using hybrid machine learning model.” In 2021 6th International Conference on Inventive Computation Technologies (ICICT), pp. 1329-1333. IEEE, 2021.

[3] Li, Jian Ping, Amin Ul Haq, Salah Ud Din, Jalaluddin Khan, Asif Khan, and Abdus Saboor. “Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare." *IEEE Access* 8 (2020): 107562-107582.

[4] Ali, Syed Arslan, Basit Raza, Ahmad Kamran Malik, Ahmad Raza Shahid, Muhammad Faheem, Hani Alquhayz, and Yogan Jaya Kumar. “An Optimally Configured and Improved Deep Belief Network (OCI-DBN) Approach for Heart Disease Prediction Based on Ruzzo– Tompa and Stacked Genetic Algorithm.” IEEE Access 8 (2020): 65947-65958.

[5] Wang, Jikuo, Changchun Liu, Liping Li, Wang Li, Lianke Yao, Han Li, and Huan Zhang. “A Stacking-Based Model for Non-Invasive Detection of Coronary Heart Disease.” IEEE Access 8 (2020): 37124-37133.

[6] Khan, Mohammad Ayoub.“An IoT Framework for Heart Disease Prediction Based on MDCNN Classifier.” IEEE Access 8 (2020): 34717-34727.

[7] Vankara, Jayavani, and G. Lavanya Devi.“PAELC: Predictive Analysis by Ensemble Learning and Classification heart disease detection using beat sound.” International Journal of Speech Technology 23, no. 1 (2020): 31-43.

[8] Mohan, Senthilkumar, Chandrasegar Thirumalai, and Gautam Srivastava. “Effective heart disease prediction using hybrid machine learning techniques.” IEEE Access 7 (2019): 81542-81554.

[9] Ali, Liaqat, Atiqur Rahman, Aurangzeb Khan, Mingyi Zhou, Ashir Javeed, and Javed Ali Khan. “An Automated Diagnostic System for Heart Disease Prediction Based on χ^2 Statistical Model and Optimally Configured Deep Neural Network.” IEEE Access 7 (2019): 34938- 34945.

[10] Latha, C. Beulah Christalin, and S. Carolin Jeeva. “Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques.” Informatics in Medicine Unlocked 16 (2019): 100203.

[11] Kouser, R. Ruhin, T. Manikandan, and V. Vinoth Kumar. “Heart disease prediction system using artificial neural network, radial basis function and case based reasoning.” Journal of Computational and Theoretical Nanoscience 15, no. 9-10 (2018): 2810-2817.

[12] Dwivedi, Ashok Kumar. “Performance evaluation of different machine learning techniques for prediction of heart disease.”Neural Computing and Applications 29, no. 10 (2018): 685-693.

[13] Zhang, Ji, Raid Luaibi Lafta, Xiaohui Tao, Yan Li, Fulong Chen, Yonglong Luo, and Xiaodong Zhu. “Coupling a fast fourier transformation with a machine learning ensemble model to support recommendations for heart

disease patients in a tele health environment." IEEE Access 5 (2017): 10674-10685.

[14] Leema, N., H. Khanna Nehemiah, and Arputharaj Kannan. "Neural network classifier optimization using differential evolution with global information and back propagation algorithm for clinical datasets." Applied Soft Computing 49 (2016): 834-844.

[15] Rajeswari, K., V. Vaithiyanathan, and T.R. Neelakantan. "Feature selection in ischemic heart disease identification using feed forward neural networks." Procedia Engineering 41 (2012): 1818-1823.

[16] Khemphila, Anchana, and Veera Boonjing. "Heart disease classification using neural network and feature selection." In 2011 21st International Conference on Systems Engineering, pp. 406-409. IEEE, 2011.

[17] Asadi, Roya, Norwati Mustapha, Nasir Sulaiman, and Nematollaah Shiri. "New supervised multilayer feed forward neural network model to accelerate classification with high accuracy." European Journal of Scientific Research 33, no. 1 (2009): 163-178.

[18] Yan, Hongmei, Yingtao Jiang, Jun Zheng, Chenglin Peng, and Qinghui Li. "A multilayer perceptron-based medical decision support system for heart disease diagnosis." Expert Systems with Applications 30, no. 2 (2006): 272-281.

[19] Møller, Martin Fodslette. "A scaled conjugate gradient algorithm for fast supervised learning." Neural networks 6, no. 4 (1993): 525-533.

[20] David W.Aha. Cleveland processed heart dataset, multivariate data 07/01/1988: https://archive.ics.uci.edu/ml/datasets/Heart+Disease.

[21] Manu Siddhartha, Cleveland Hungarian Statlog heart dataset, multivariate data, 11/06/2020, https://archive.ics.uci.edu/ml/machine-learning-databases/heart-disease.