Prediction of Disease Level Using Multilayer Perceptron of Artificial Neural Network for Patient Monitoring

Md. Osman Goni Nayeem, Maung Ning Wan, Md. Kamrul Hasan

Abstract*— *ANN has been proved as a powerful discriminating classifier for tasks in medical diagnosis for early detection of diseases. In our research, ANN has been used for predicting three different diseases (heart disease, liver disorder, lung cancer). Feed-forward back propagation neural network algorithm with Multi-Layer Perceptron is used as a classifier to distinguish between infected or non-infected person. The results of applying the ANNs methodology to diagnosis of these disease based upon selected symptoms show abilities of the network to learn the patterns corresponding to symptoms of the person. In our proposed work, Multi-Layer Perceptron with having 2 hidden layer is used to predict medical diseases. Here in case of liver disorder prediction patients are classified into four categories: normal condition, abnormal condition (initial), abnormal condition and severe condition. This neural network model shows good performance in predicting disease with less error.

Index Terms*—*Artificial Neural Network (ANN), Multilayer Perceptron, Heart Diseases, Liver Disorder, Lung Cancer.

I. INTRODUCTION

The human brain is the main organ of the human nervous system. It is located in the head, protected by the skull. The human nervous system may be viewed as a three-stage system, as depicted in the block diagram of Fig. 1.

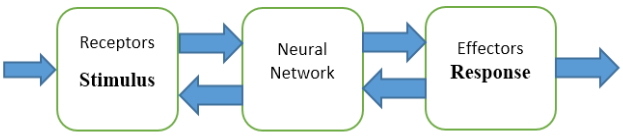


Figure 1. Block diagram representation of nervous system

Central to the system is the brain, represented by the neural (nerve) net, which continually receives information, perceives it, and makes appropriate decisions. The arrows pointing from left to right indicate the forward transmission of information bearing signals through the system. The arrows pointing from right to left signify the presence of feedback in the system. The receptors convert stimuli from human body or the external environment into electrical impulses that convey information to the neural net (brain). The effectors covert electrical impulses generated by the neural net into discernible responses as system outputs.

Revised Version Manuscript Received on July 11, 2015.

Md. Osman Goni Nayeem, Department of Electrical and Electronic Engineering (EEE), Khulna University of Engineering & Technology (KUET), Khulna-9203, Bangladesh.

Maung Ning Wan, Department of Electrical and Electronic Engineering (EEE), Khulna University of Engineering & Technology (KUET), Khulna-9203, Bangladesh.

Md. Kamrul Hasan, Department of Electrical and Electronic Engineering (EEE), Khulna University of Engineering & Technology (KUET), Khulna-9203, Bangladesh.

Artificial intelligence is not a new research field - ANNs have been in the attention of the scientists over the last 60 years. First studies on neural networks were done in 1943 by McCullough and Pitts. After a while, Rosenblatt conceived in 1959 the first learning algorithm, creating a model known as the perceptron, which was then only a solution to simple linear problems. The first non-linear processing capabilities of ANNs were reported in 1974 by Werbos, and afterwards the interest of the scientific community steadily increased, boosted in the last years by the discovery of the back propagation algorithm and by the increase in computational power, due to the exponential advances in computer technology [1]-[5]. The parallel distributed processing of the mid-1980s became popular under the name connectionism. The text by David E. Rumelhart and James McClelland [6]-[8] (1986) provided a full exposition on the use of connectionism in computers to simulate neural processes. Neural networks, as used in artificial intelligence, have traditionally been viewed as simplified models of neural processing in the brain, even though the relation between this model and brain biological architecture is debated, as it is not clear to what degree artificial neural networks mirror brain function [9]. Neural networks were gradually overtaken in popularity in machine learning by support vector machines and other, much simpler methods such as linear classifiers. Between 2009 and 2012, the recurrent neural networks and deep feed-forward neural networks developed in the research group of Jürgen Schmidhuber at the Swiss AI Lab IDSIA have won eight international competitions in pattern recognition and machine learning[10]-[13]. There are some basic advantages of Artificial Neural Network (ANN). ANN is popular strategies for supervised machine learning and classification. It has vast advantages over conventional methods of classification and prediction. These are

1. ANN is nonlinear model that is easy to use and understand compared to statistical methods.
2. ANN is non-parametric model while most of statistical methods are parametric model that need higher background of statistic.
3. ANN with Back propagation (BP) learning algorithm is widely used in solving various classification and forecasting problems.
4. Another advantages of using Artificial Neural Networks (ANN) is that it can handle large amount of data sets.
5. It has ability to detect all possible interactions between predictor variables.
6. ANN can perform massively parallel distributed operation at the same time just like human brain.

Published By:
Blue Eyes Intelligence Engineering & Sciences Publication

7. It is a non-parametric method, thus eliminates the error in parameter estimation.

There are several studies which have applied neural networks in the diagnosis of different diseases. An artificial neural network trained on admission data can accurately predict the mortality risk for most preterm infants. However, the significant number of prediction failures renders it unsuitable or individual treatment decisions. In a study [6], the artificial neural network performed significantly better than a logistic regression model (area under the receiver operator curve 0.95 vs 0.92). Survival was associated with high morbidity if the predicted mortality risk was greater than .50.There were no preterm infants with a predicted mortality risk of greater than 0.80. The mortality risks of two non-survivors with birth weights >2000g and severe congenital disease had largely been underestimated. In another study [7], an effective arrhythmia classification algorithm used for the heart rate variability (HRV) signals. The proposed method is based on the Generalized Discriminant Analysis (GDA) feature reduction technique and the Multilayer Perceptron (MLP) neural network classifier. At first, nine linear and nonlinear features are extracted from the HRV signals and then these features are reduced to only three by GDA. Finally, the MLP neural network is used to classify the HRV signals. The proposed arrhythmia classification method is applied to input HRV signals, obtained from the MIT-BIH databases. Here, four types of the most life threatening cardiac arrhythmias including left bundle branch block, fist degree heart block, Supraventricular tachyarrhythmia and ventricular triggering can be discriminated by MLP and reduced features with the accuracy of 95%. The study [8] of a functional model of ANN is proposed to aid existing diagnosis methods. This work investigated the use of Artificial Neural Networks (ANN) in predicting the Thrombo-embolic stroke disease. The Back propagation algorithm was used to train the ANN architecture and the same has been tested for the various categories of stroke disease. This research work demonstrates that the ANN based prediction of stroke disease improves the diagnosis accuracy with higher consistency. This ANN exhibits good performance in the prediction of stroke disease in general and when the ANN was trained and tested after optimizing the input parameters, the overall predictive accuracy obtained was 89%. Feed-forward neural networks are widely and successfully used models for classification, forecasting and problem solving. A typical feed-forward back propagation neural network model is proposed to diagnosis Heart Diseases, Liver Disorder, Lung Cancer. Proposed network consists of three layers:

- Input layer,
- Two hidden layer and
- Output layer.

In case of Heart Disease prediction, two hidden with 44 hidden layer neurons is created and trained. The input and target samples are automatically divided into training, validation and test sets. The training set is used to teach the network. Training continues as long as the network continues improving on the validation set. The test set provides a completely independent measure of network accuracy. The information moves in only one direction, forward, from the input nodes, through the hidden nodes and to the output nodes as shown in Fig. 2. But error calculation is performed in backward. That's why network model is termed as feed-forward back propagation neural network model. Here, use of two hidden layer helps in improving the network accuracy. Network model for other two diseases are nearly same only in the difference in the number of neurons in input and hidden layer. The network model is shown below in Fig. 2.

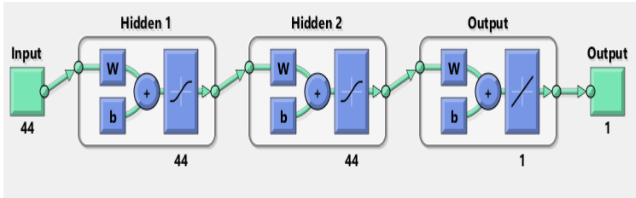


Figure 2. Network model with different layers

The hidden neurons are able to learn the pattern in data during the training phase and mapping the relationship between input and output pairs. Each neuron in the hidden layer uses a transfer function to process data it receives from input layer and then transfers the processed information to the output neurons for further processing using a transfer function in each neuron. The output of the hidden layer can be represented by equation (1), (2) and (3).

$$\mathrm{V_k} = \sum_{j=0}^{m} W_{kj} X_j \tag{1}$$

Error calculation:

$$\mathrm{E_0(k)} = Y_{d,0}(k) - Y_0(k) \tag{2}$$

Final hidden layer output:

$$\mathrm{Y_K} = \Phi(V_K) \tag{3}$$

Here, W_{k1}……..W_{km} are the value of synaptic weight.

W_{k0} is the value of bias.

E_0 (k) is the value of error.

$Y_{d,0}$ (k) is the desired value of output.

Y_0 is the actual value of output.

V_k is the output value before the process of activation function.

Ω(.) is the activation function.

Y_k is the final output for hidden layer.

Multilayer Perceptron

A multilayer perceptron (MLP) is a feed forward artificial neural network model that maps sets of input data onto a set of appropriate outputs. A MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. Except for the input nodes, each node is a neuron (or processing element) with a nonlinear activation function. MLP utilizes a supervised learning technique called back propagation for training the network. [1], [2]. MLP is a modification of the standard linear perceptron and can distinguish data that are not linearly separable.

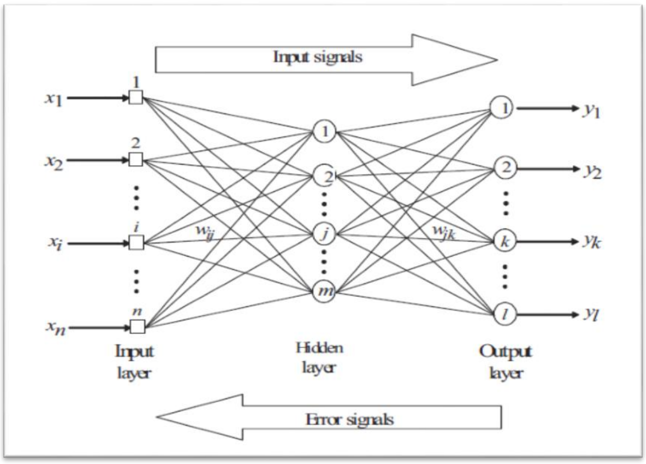


Figure 3. A multilayer perceptron neural network model.

II. METHODOLOGY AND SYSTEM ARCHITECTURE

In this paper, a typical feed-forward back propagation neural network is proposed to diagnosis diseases. It consists of three layers namely the input layer, a hidden layer, and the output layer. The input and target samples are automatically divided into training, validation and test sets. The training set is used to teach the network. Training continues as long as the network continues improving on the validation set. The test set provides a completely independent measure of network accuracy.

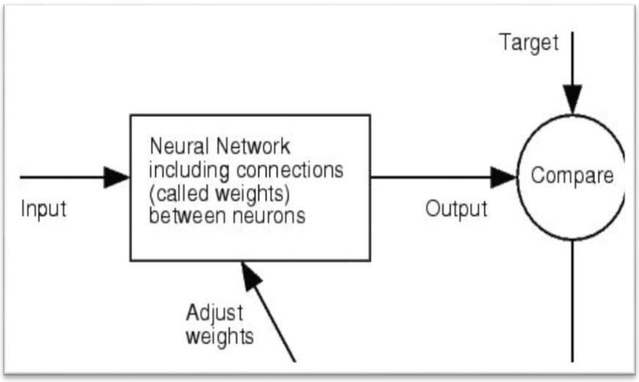


Figure 4. A simple block illustrating the whole function of the network.

As shown in **Fig. 4,** input data are fed to the network consisting of synaptic connections between neurons and including bias parameter with the network. The output obtain from the network is compare with the target output. If error exist between actual and target output then weights of the network are recalculated by moving backwards to obtain good performance by the network. By this process, propose network model is capable to learn. This refers as training the network. Training continues as long as network continues improving on the validation set. After training the network, testing is performed with a new dataset where the network is not provided with a target output. The complete block diagram presentation of training and testing a network are shown below

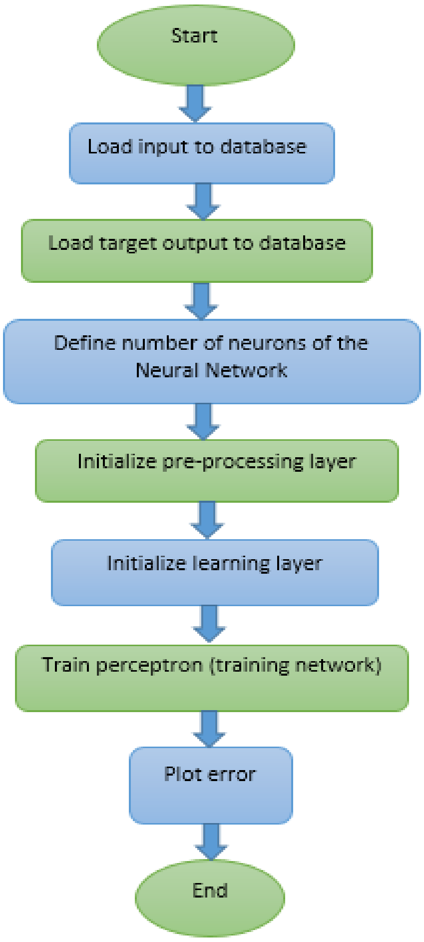


Figure 5. Block diagram for training of proposed network

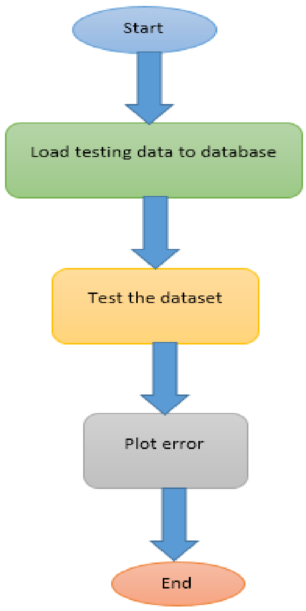


Figure 6. Block diagram for testing of proposed network

Gradient Descent Algorithm

Gradient descent is a first-order optimization algorithm as shown in Fig. 7. To find a local minimum of a function using gradient descent, one takes steps proportional to the negative of the gradient (or of the approximate gradient) of the function at the current point. If instead one takes steps proportional to the positive of the gradient, one approaches a local maximum of that function; the procedure is then known as gradient ascent. Gradient descent is also known as steepest descent, or the method of steepest descent. When known as the latter, gradient descent should not be confused with the method of steepest descent for approximating integrals.

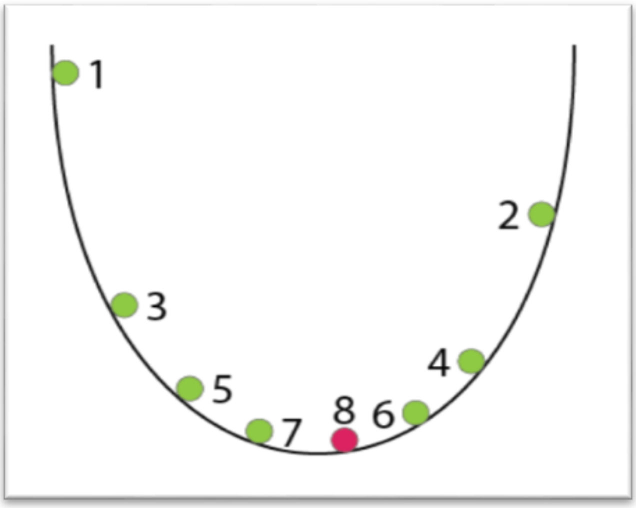


Figure 7. Error minimization using Gradient Descent Algorithm.

Levenberg Marquardt Algorithm

In mathematics and computing, the Levenberg Marquardt algorithm (LMA) also known as the damped least-squares (DLS) method, is used to solve non-linear least squares problems. These minimization problems arise especially in least squares curve fitting. The LMA interpolates between the Gauss Newton Algorithm (GNA) and the method of Gradient Descent. The LMA is more robust than the GNA, which means that in many cases it finds a solution even if it starts very far off the final minimum. For well-behaved functions and reasonable starting parameters, the LMA tends to be a bit slower than the GNA. LMA can also be viewed as Gauss–Newton using a trust region approach. The LMA is a very popular curve-fitting algorithm used in many software applications for solving generic curve-fitting problems. However, as for many fitting algorithms, the LMA finds only a local minimum, which is not necessarily the global minimum.

III. Results and Discussion

Artificial neural networks may probably be the single most successful technology in the last two decades which has been widely used in a large variety of applications. In our research, we deal with four types of dataset that is available in UCI Machine Learning Repository. The propose network shows an accuracy above 80% for each kind of dataset. During training, the following training window opens. This window displays training progress and allows you to interrupt training at any point by clicking Stop Training.

1. Heart Diseases prediction

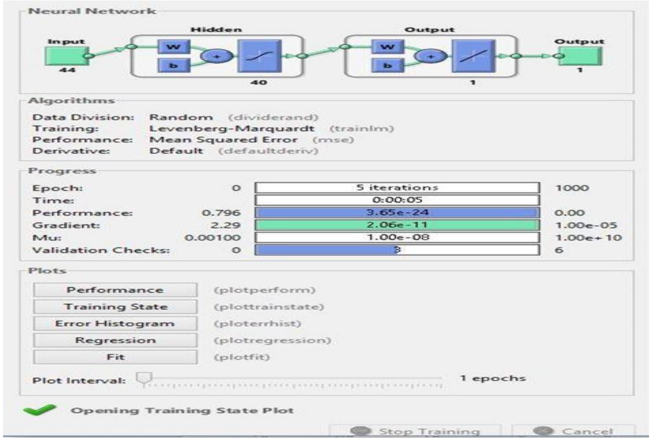


Figure 8. Neural Network train with Levenberg-Marquardt Algorithm

This training stopped when the validation error increased, which occurred at iteration 5. If you click Performance in the training window, a plot of the training errors, validation errors, and test errors appears, as shown Fig. 8. In this example, the result is reasonable because of the following considerations

- Final mean-square error is small.
- Test set error and the validation set error have similar characteristics.
- No significant over-fitting has occurred by iteration 2

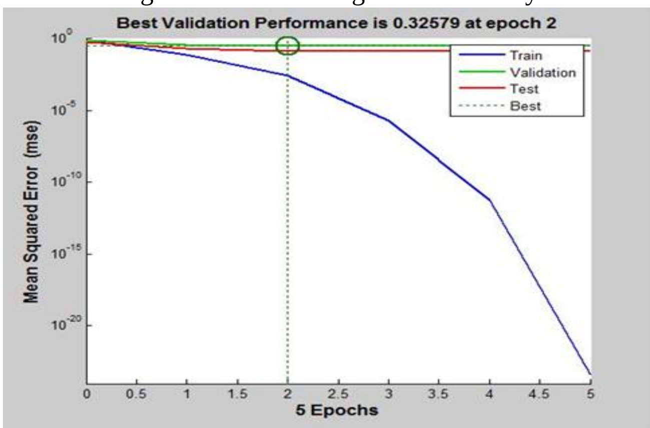


Figure 9. Neural Network performance analysis with epoch variation

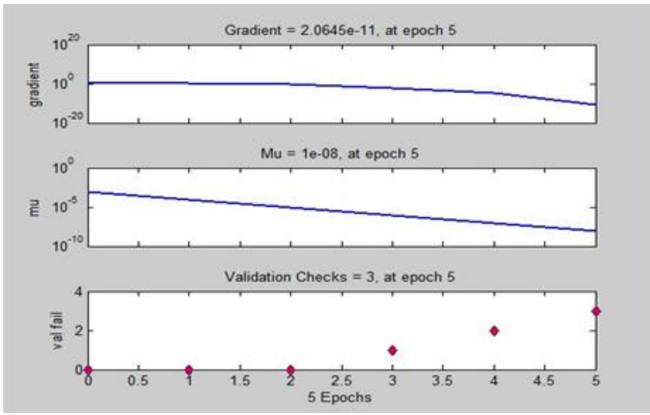


Figure 10. Neural Network Training state representation

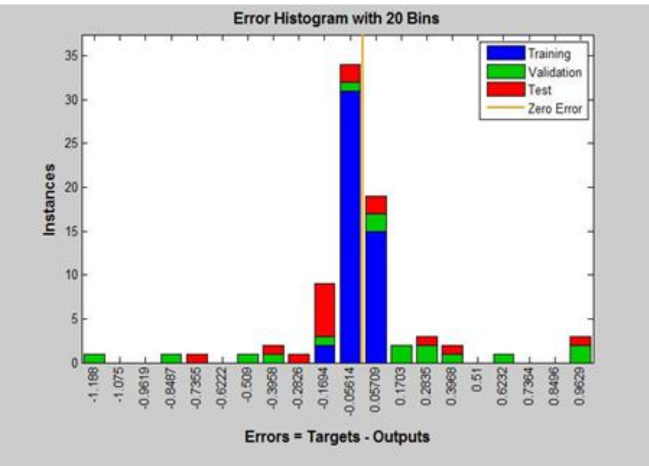


Figure 11. Error histogram presentation of Neural Network

The blue bars in Fig. 11 represent training data, the green bars represent validation data, and the red bars represent testing data. The histogram can give you an indication of outliers, which are data points where the fit is significantly worse than the majority of data. In this case, you can see that while most errors fall between -0.3968 and 0.3968, there is a validation points with errors of -1.188 and 0.6232. These outliers are also visible on the testing regression plot. It is a good idea to check the outliers to determine if the data is bad, or if those data points are different than the rest of the data set. If the outliers are valid data points, but are unlike the rest of the data, then the network is extrapolating for these points. You should collect more data that looks like the outlier points, and retrain the network. In regression plot, here you see, the output tracks the targets very well for training, testing, and validation, and overall R-value is about 0.84. This means that network has accuracy of 84% for heart disease dataset.

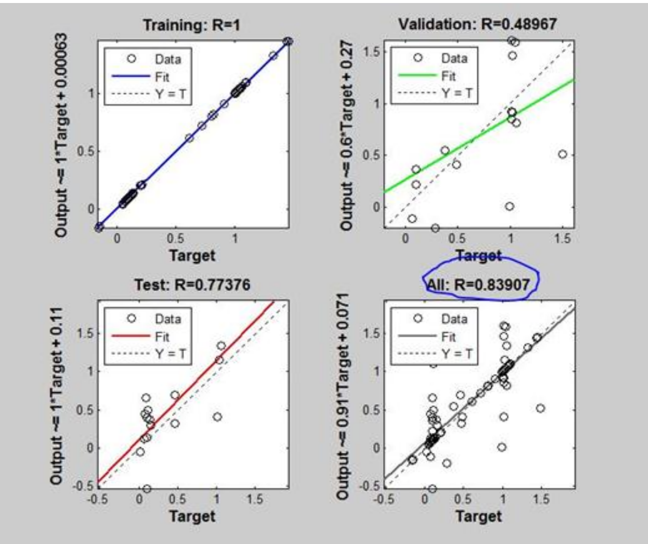


Figure 12. Neural Network Regression analysis

2. Liver Disorder prediction

In case of Liver Disorder prediction the training processes are same as the Heart disease prediction.

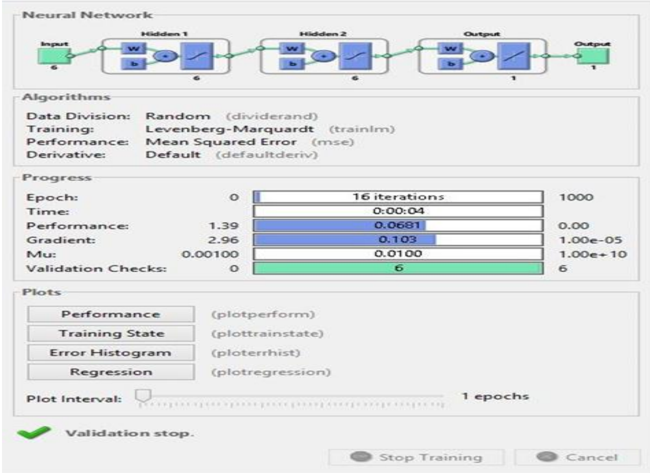


Figure 13. Neural Network Training for Liver Disorder prediction

But here patients are classified into 4 category

- Normal state
- Abnormal state (initial)
- Abnormal state
- Severe state

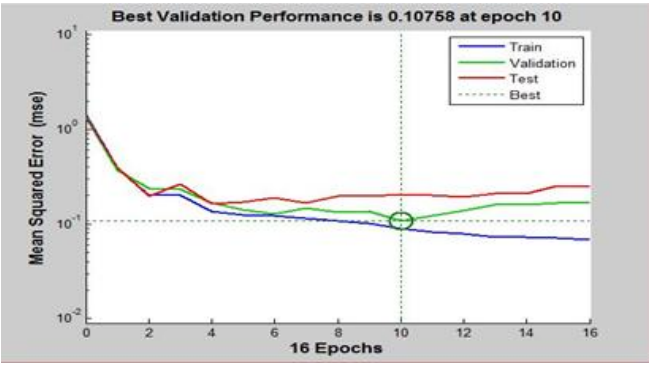


Figure 14. Neural Network performance analysis with epoch variation

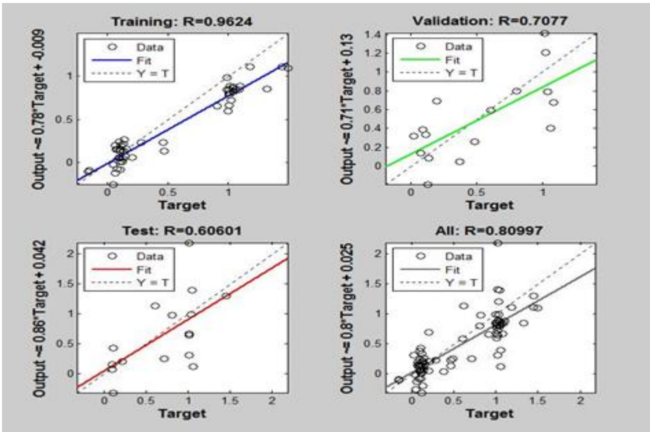


Figure 15. Neural Network Regression analysis

In regression plot of liver disorder prediction, it is seen that the overall value of R is 0.80997. Which indicates that network has accuracy of 81%.

3. Lung Cancer prediction

In case of Lung Cancer prediction the training processes are same as the Heart disease prediction. But here patients are classified into 3 category

- Normal condition.
- Abnormal condition.
- Severe condition.

Neural Network Performance plot with epoch

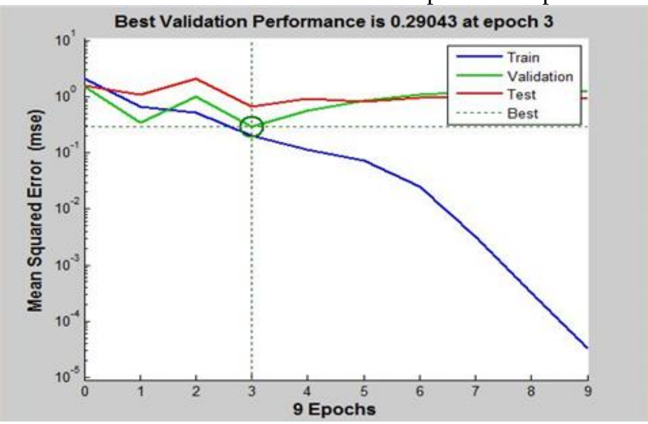


Figure 16. Error histogram presentation of Neural Network

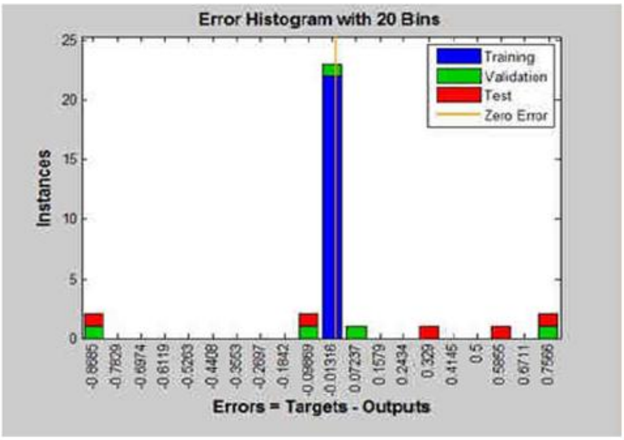


Figure 17. Neural Network Regression plot for Liver Disorder prediction

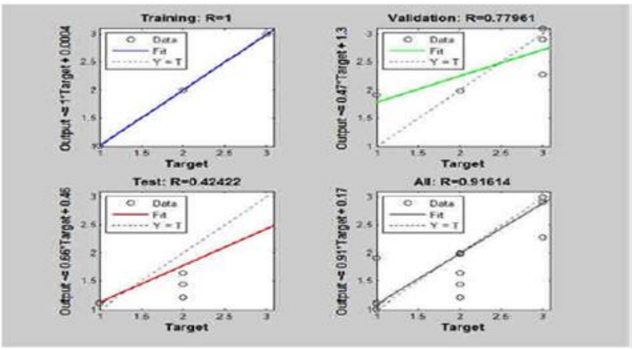


Figure 18. Neural Network Regression plot for Liver Disorder prediction

In regression plot of Lung Cancer prediction, it is seen that the overall value of R is 0.91614. Which indicates that network has accuracy of 91.6%.

4. Performance Evaluation for each Case

Neural network toolbox from MATLAB 2012 is used to evaluate the performance of the proposed networks. **Heart Disease** is the first disease to be diagnosed in which two-layer feed-forward network with 44 inputs and 40 sigmoid hidden neurons and linear output neurons was created. Such net can fit multi-dimensional mapping problems arbitrarily well, given consistent data and enough neurons in its hidden layer as shown in Fig. 8. Levenberg-Marquardt back propagation algorithm was used where training automatically stops when generalization stops improving, as indicated by an increase in the mean square error (MSE) of the validation samples. The results of applying the artificial neural networks methodology to distinguish between healthy and unhealthy person based upon selected symptoms showed very good abilities of the network to learn the patterns corresponding to symptoms of the person. The network was simulated in the testing set (i.e. cases the network has not seen before). The results were very good; the network was able to classify 84% of the cases in the testing set. Fig. 9 shows the training state values. Best validation performance is 0.32579 at epoch 2 as shown in Fig. 9. The mean squared error (MSE) is the average squared difference between outputs and targets. Lower values are better while zero means no error. In case of heart disease the percent correctly classified in the simulation sample by the feed-forward back propagation network is 84 percent. The MSE is equal to 3.65e-24 and the regression is equal to 0.83907. Liver Disorder prediction is the second disease to be diagnosed. A multilayer feed-forward network with 6 features inputs and 6 sigmoid hidden neurons and linear output neurons was created. Such net can fit multi-dimensional mapping problems arbitrarily well, given consistent data and enough neurons in its hidden layer as shown in Fig .12. Levenberg-Marquardt back propagation algorithm was used to train the network. The results of applying the artificial neural networks methodology to distinguish between normal and abnormal person showed very good abilities of the network to learn the patterns. The network was simulated in the testing set. The results were very good; the network was able to classify 81% of the cases in the testing set. Fig. 13 shows the training state values. Best validation performance is 0.10758 at epoch 10 as shown in Fig. 14. In case of Liver Disorder prediction, the percent correctly classified in the simulation sample by the feed-forward back propagation network is 81 percent. The MSE is equal to 0.0681 and the regression is equal to 0.80997. Lung Cancer prediction is the third disease to be diagnosed. A multilayer feed-forward network with 56 features inputs and 6 sigmoid hidden neurons and linear output neurons was created. Such net can fit multi-dimensional mapping problems arbitrarily well, given consistent data and enough neurons in its hidden layer as shown in Fig. 15. Levenberg-Marquardt back propagation algorithm was used with train the network. The results of applying the artificial neural networks methodology to distinguish between normal and abnormal person showed very good abilities of the network to learn the patterns. The network was simulated in the testing set. The results were very good; the network was able to classify 92% of the cases in the testing set. Fig. 16 shows the training state values. Best validation performance is 0.29043 at epoch 3 as shown in Fig. 17. In case of Lung Cancer prediction, the percent correctly classified in the simulation sample by the feed-forward back propagation network is 81 percent. The MSE is equal to 3.29e-05 and the regression is equal to 0.91614. Finally, it is seen that the proposed network predict, all of this 3 diseases with an accuracy above 82%. Which refers to a good classification result. This kind of disease prediction by using neural network is very helpful for expert doctor to cross check his result with neural network predicted result.

IV. CONCLUSION

In our research, our expectation is to utilize artificial neural network in disease diagnosis. As per wide range of applicability of ANN and their ability to learn complex and nonlinear relationships including noisy or less precise information, neural networks are well suited to solve problems in medical field. The feed-forward Back propagation neural network with supervised learning is proposed to diagnose the disease. Multi-Layer Perceptron neural network architecture is used to train the neural network for classification. Our proposed network shows an accuracy of 84% for heart disease prediction where the existing work shows accuracy of 82%. This is because of the use of more than one hidden layer, more neurons in hidden layer and use of more training data. Besides, for liver disorder and lung cancer prediction networks shows an accuracy of 82% and 91% respectively.

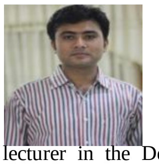
REFERENCES

1. J. Dayhoff and J. DeLeo, “Artificial neural networks opening the black box,” Cancer Supplement, vol. 91, No. 8, pp. 16151635, 2001.
2. Dimitrios H. Mantzaris, George C. Anastassopoulos and Dimitrios K. Lymberopoulos,“Medical Disease Prediction Using Artificial Neural Networks” vol.91, NO.82,pp, 16253765 , July 5, 2008, IEEE.
3. V. Piuri and F. Scotti, “Morphological classification of blood leucocytes by microscope images,” in 2004 IEEE Int. Conf. Computational Intelligence for Measurement Systems and Applications, pp. 103108.
4. D. S. Huang and S. D. Ma, “Linear and nonlinear feedforward neural network classifiers: A comprehensive understanding,” J. Intelligent Syst., vol. 9, pp. 1–38, 1999.
5. Irfan Y. Khan, P.H. Zope, S.R. Suralka, “Importance of Artificial Neural Network in Medical Diagnosis disease like acute nephritis disease and heart disease” Volume 2, Issue 2, March 2013, International Journal of Engineering Science and Innovative Technology (IJESIT).
6. Kadhim Al-Shayea, “Artificial neural network in medical diagnosis”, International journal of computer science Vol.8, Issue 2, March 2011.Qeethara
7. Qeethara Kadhim Al-Shayea and Itedal S.H. Bahia, “Urinary system Dieseases Diagnosis Using Artificial neural networks”, IJCSNS, Vol.10, No.7, July 2010.
8. S. Moein, S.A. Monadjemi, and P. Moallem, “A Novel Fuzzy Neural Based Medical Diagnosis system”, International Journal of biological and Life Science 4:3, 2008.
9. Abdel- Badeeh M. Salem, “Case Base Reasoning Technology for Medical Diagnosis”, World Academic of Science, Engineering and Technology 7, 2007.
10. Ashish Dehariya, Ilyas Khan,Vijay K. Chaudhari and Saurabh Karsoliya, “ A Novel flow Reasoning of Medical Diagnostic system using Artificial Feed Forward Neural Networks”, IJCSE, Vol.3, No-3, March 2011.
11. Zhi- Hua Zhou, Yoan Jiang, Yu- Bin Yang & Shi- Fu Chen, “ Lung Cancer Cell Identification Based On Artificial Neural Network ensembles”, Artificial Intelligence in Medicine Vol24, No-1, 2002.
12. Xin Yao Senior Member, IEEE and Yong Liu, “A new evolutionary system for evolving Artificial Neural Networks”, IEEE Transaction on Neural Network Vol.8, No-1, May 1997 Transactions on information Technology in Biomedicine Vol 11, No.3, May- 2012.
13. Baker J A, Kornguth PJ, Lo JY, Williford ME, Floyd CE Jr(1995) : “Breast cancer: prediction with artificial neural network based on BI-RADS standardized lexicon”. Radiology, 1995; 196(3): 817-22
14. P. S. Heckerling, G. Canaris, S. D. Flach, T. G. Tape, R. S. Wigton and B. S. Gerber, Predictors of urinary tract infection based on artificial neural networks and genetic algorithms, International Journal of Medical Informatics; April, 2007, Vol. 76 Issue 4, pp. 289-296.
15. S. A. Monadjemi and P. Moallem, Automatic Diagnosis of Particular Diseases Using a Fuzzy-Neural Approach, International Review on Computers & Software, Jul., 2008, Vol. 3 Issue 4, pp. 406-411
16. R. Suganya, and S. Rajaram, "Content Based Image Retrieval of Ultrasound Liver Diseases Based on Hybrid Approach”, American Journal of Applied Sciences 9 (6), 2009, pp. 938-945.
17. S. Babaei, A. Geranmayeh, "Heart sound reproduction based on neural network classification of cardiac valve disorders using wavelet transforms of PCG signals", Computers in Biology and Medicine (39), 2009, pp. 8–15.
18. Zhi-Hua Zhou, Member IEEE and Yuan Jiang, “Medical Diagnosis with C4.5 Rule Preceded By Artificial Neural Network Ensemble”, IEEE Transactions on information Technology in Biomedicine Vol-7, No-1, March 2003.
19. D. Gil, M. Johnsson, J. M. Garicia Chemizo, A. S. Paya and D. R. Fernandez, Application of Artificial Neural Networks in the Diagnosis of Urological Dysfunctions, Expert Systems with Applications, April, 2009, Vol. 36 Issue 3, pp. 5754-5760.
20. W. G. Baxt, “Application of artificial neural networks to clinical medicine,” Lancet, vol. 346, no. 113, pp. 58, 1995.
21. J. S. Chiu, Y. C. Li, F. C. Yu, and Y. F. Wang, “Applying an artificial neural network to predict osteoporosis in the elderly,” Studies in Health Technology and Informatics, vol. 124, pp. 609614, 2006.
22. G. Zhang and V. Berardi, “An investigation of neural networks in thyroid function diagnosis,” Health Care Management Science, vol.1, pp. 29–37, 1998.

AUTHOR PROFILE

Mr. Md. Osman Goni Nayeem received his B.Sc. in Electrical and Electronic Engineering from Khulna University of Engineering & Technology (KUET) in 2012 and currently he is pursuing his M.Sc. from here. He is serving as a lecturer at Dept. of Electrical and Electronic Engineering in KUET, Khulna, Bangladesh from September, 2012. His current research interests include Biomedical Engineering, Artificial Intelligence and Reversible Watermarking.

Maung Ning Wan received his B.Sc. in Electrical and Electronic Engineering from Khulna University of Engineering & Technology (KUET) in 2015. His current research interests include Biomedical Engineering and Artificial Intelligence.



Mr. Md. Kamrul Hasan received his B.Sc. in Electrical and Electronic Engineering (EEE) from Khulna University of Engineering & Technology (KUET). He is currently a student of M. Sc. in Electrical & Electronic Engineering at the same University. After Graduation, Mr. M. K. Hasan joined at Daffodil International University (DIU) as a lecturer in the Dept. of Electrical & Electronic Engineering (EEE) from January, 2015 to April 2015. Currently, M. K. Hasan Working at Khulna University of Engineering & Technology (KUET) as a Lecturer in the Dept. of Electrical & Electronic Engineering in the month of April, 2015. His research interests are Biomedical Materials and Devices, Biomedical Signal Processing, MRI and f-MRI, EEG Brain Mapping. He has published several IEEE sponsored International Conference Paper and International Journal.