IJCTA, 9(22), 2016, pp. 223-231

Prediction of Diseases Using Neural Network

Bharti Yadav* Shilpi Sharma Ashwani Kumar Yadav*** Vaishali******

Abstract : It has been observed that a better health is playing a vital role in human happiness. There are various developments done in the field of medical sciences, neural networks are quite important for the prediction of diseases. ANNs are used to process the data in the same way that people do, learn by illustrations. As individuals have interest in their wellbeing advancements in medical research has been a standout amongst the most dynamic research areas. This is made out of an extensive number of much neurons working for the consideration of a specific issues. In this paper, we approached the machine learning techniques i.e Feed forward back propagation neural network for the prediction of diseases. For the proposed work a datasets of the 164 people were used as testing and training by taking 11 number of input attributes. The datasets are collected from UCI machine learning respiratory. The data has been processed with the help of MATLAB, simulation results demonstrates that execution of NN methodology is much imperative and easily understandable.

Keyword : Back propagation, Diseases, Techniques, Artificial Neural Network

1. INTRODUCTION

Over the past decades, a consistent development identified to diseases prediction has been performed. Researchers connected distinctive techniques [1].For eg- screening in early stages, in mind to discover type of disease before they cause manifestation. Additionally they made new system for the early detection of diseases such as cancer, Heart Attack, Skin disease or Diabetes [2]. However to get the precious prediction of disease result is a most challenging assignment. In this Machine learning, we require complete training data to construct the model [3]. Particularly in the medical research area, the complete training data portraying shows the relation between input and output. Machine learning algorithms need memory and computation time for the learning procedure [4]. This is a branch of artificial Intelligence that utilises a variety of probabilistic an optimisation procedure that allows system to learn from past illustrations [5]. This ability is an especially appropriate to medical field, particularly those that rely on complex proteomic and genomic estimations. This machine learning is used in prediction of disease. From the globally accepted studies, we can say that machine learning technique can be used to achieve the higher accuracy of disease prediction. According to the paper [6] one of the best methods is an artificial neural network, this is a highly effective tool used to predict the heart diseases. The best feature of the neural network is adaptively. K et al.,[7] states that multilayer perceptron model gives best results in comparison with others models of the artificial neural network. This paper concentrate on the advancement of NN system in light of both multi-layer perceptron's and addition feed forward and the use of this model is effortlessly detection of disease [8]. Particularly this system has been created to help physicians, since identification if disease in early stages is very essential for cure of any disease. The principle goal of this project is to create the frame work for prediction of

* Department Of ECE Chandigarh Group Of Colleges Chandigarh, India Email -bhartiyadav60@gmail.com

** Department of ECE Chandigarh Group Of Colleges Chandigarh, India Email -shilpisharma.ece@cgc.edu.in

*** Department of ECE, ASET Amity University Rajasthan Jaipur, Rajasthan, India Email - ashwaniy2@gmail.com

**** Department of CSE, ASET Amity University Rajasthan Jaipur, Rajasthan, India Email – Vaishaliyadav26@gmail.com

different diseases-heart attack, liver disorder, and diabetes which are the most common problems nowadays. There is lots of work done in the artificial neural network for different diseases but the system made by them is for prediction of single diseases e.g cancer or skin diseases or heart attack or diabetes. But this paper presented a system which is used for prediction of three diseases. So by this system, we detect diseases only by putting dataset. This will help the doctor to diagnosis the diseases in a very rapid manner [9].

1.1. Structure of NN

The neural system fuses three specific layers with feed forward setup. The input layer of this system is an arrangement of information these are totally connected with the hidden layers [10]-[11]. Further the hidden layers totally connected the next layer (i.e. output layer). The last layer (Output layer) sends the response to the framework [12]. Architecture and Structure decide for disease diagnosis Of ANN is based feed forward network. In this input layer send their signal to hidden layer and hidden sends to Output layer now the input symptoms received by network in the input layer and the outputs sent by the neuron to the output layer. In this, supervised learning method and back propagation algorithm is implemented. In this input is provide as an example and output is computed by the network, once this is done, error will be. Our intension is to minimise the error till the time frame work get familiarized with training.

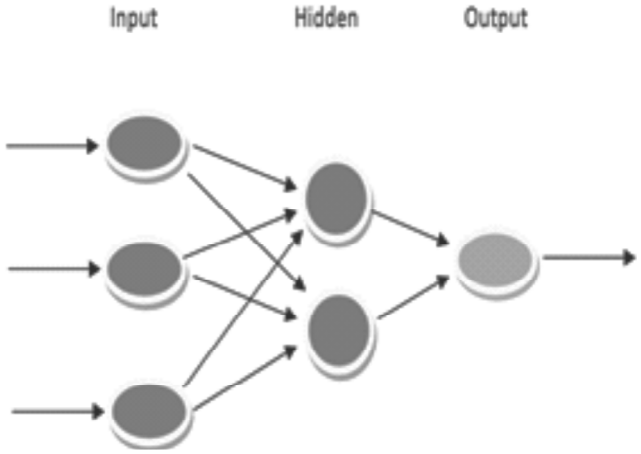

Fig. 1. ANN architecture for disease diagnosis.

1.2. Training the model

A system has composed for a specific application. To begin this methodology, beginning weights picked subjectively. By then, training & learning, begins. The ANN is trained by exhibiting it with an arrangement of existing data (in light of the subsequent history of patients) where the outcome is known. Multilayer system utilizes an assortment of learning strategies; the most renowned is back–propagation (BP) algorithms. It is a standout amongst the most ideal approaches to manage with machine learning algorithm data which streams from the course of the input layer towards the output layer.

Two approaches are -Supervised and Unsupervised in this [13], supervised learning categorize into many types :

1. Back propagation network
2. Perceptron network
3. Radial basis function network
4. Time delay neural network
5. Wavelet neural network
6. Adaptive linear neuron

1.3. NN Features

There are various features of Neural Networks:

- The NNs can outline examples to their related output designs in this manner showing mapping capabilities.

- NNs are trained with known example of an issue subsequently recognizing new objects beforehand untrained.
- The NNs have the capacity to sum up in this way anticipating new results from past patterns.
- The NNs process data in parallels and in a distributed manner [14].

2. LITERATURE REVIEW

Various researchers has been presented various techniques and algorithms to diagnose various diseases, some of them are discussed here for references. L.G et al., [15]-[16] uses the feed forward design for training to the neural system to show the results of skin diseases. This paper includes different types of skin diseases- leprosy, benign, skin tumor, impetigo, acne, scabies and diaper candidacies these all can be treated by this system. The system shows an accuracy of 90%. Muhammad et al., [17] introduced a method of diabetes prediction using the different type of supervised learning techniques of the neural network. In this results are evaluated using 250 diabetes dataset and the age varies from 25 to 78 years. They use 27 input variables for analysis of diabetes and different training algorithms used are- BFGS Quasi-Newton algorithm, Levenberg Marquardt, and Bayesian Regularization. The accuracy obtained from these different methods is 63%.

P.K Anooj [18] presented a weighted fuzzy rule for detection of heart diseases using k-fold cross validation. The dataset got from UCI Respiratory. This consists of 14 attributes of input and its output value is vary from 0 to 4 where 0 means no presence of diseases and from 1 to 4 it shows the existence of diseases. In this three different datasets are used for the analysis of heart disease and finally got the accuracy of 57.85%

Seema et al., [19] adopted a technique adaptive resonance neural network (ARNN) which is a type of unsupervised learning to detects cancer. This dataset contains a total of 699 cases out of which 600 cases used to train the network. This database contains 9 attributes and it is categorized into two outputs benign or malignant. This technique shows the accuracy of 75%.

Table 1. Summary of previous work discussed above.

Title	*Authors*	*Methods*	*Accuracy*	*Year*
Diagnosing Skin diseases using an ANN	L.G Kabari, F.S Bakpo	Feed Forward Network	90%	2009
Prediction of diabetes using ANN	Muhammad Akmal Sapon Khadijah Ismail, Suehaglyn Zainudin	Levenberg Marquardt, Newton Algorithm, Bayesian Regularization	63%	2011
Clinical decision support system- Risk level prediction of Heart diseases	P.K Anooj	K- fold cross validation	57%	2012
Cancer detection using adaptive Neural Network	Seema Singh, Sunita Saini, Mandeep Singh	Adaptive Resonance Neural Network	75%	2012

This table is the summary of previous paper, in this different methods are used to diagnose the different diseases but best result is shown by feed forward method which is 90% for the diagnose of skin diseases [16] that why in this paper same method which is feed forward network is implemented to get the best results.

3. SYSTEM IMPLEMENTATION

Correct analysis of any diseases depends on upon the various and normally disjointed information.

3.1. Diagnosis of Disease using ANN System

The diagnosis of disease is done by collecting data. If the real-time input data is matched with the preloaded data which we used to train the system. If the match is more than a threshold then the doctor will assume and suggest that we are having that particular disease. After then the doctor recommends going for a laboratory test to the patient in order to sure about the result. By the test it confirmed that which particular disease we have and what is the stage of that [20].

This system will work like this : First of all, we listed all the attributes required to train the system, then we collected the real-time data of patients and then we uploaded that to the system so as to train it. Once the system is got trained then in next data input, the result will be formulated on the basis of how much the data is matched with the preloaded data. On the basis of a match , this is presumed that the person may have Heart attack , Liver Disorder or Diabetes , and if the match is less than a certain limit , then in, case the test is treated as negative and the person is informed about his Health Status.

At this stage, NN design must be decided. This is typically based on by testing frameworks with a different number of hidden layers and input layers. After then frame work is set up to be trained. The illustrations are subsequently passed into training, validation and testing sets. Training continues till as long as the framework continuing upgrading the validation sets. The test sets given as absolutely autonomous of measure of stem accuracy. A multilayer feed forward network is eleven input attributes and three outputs.

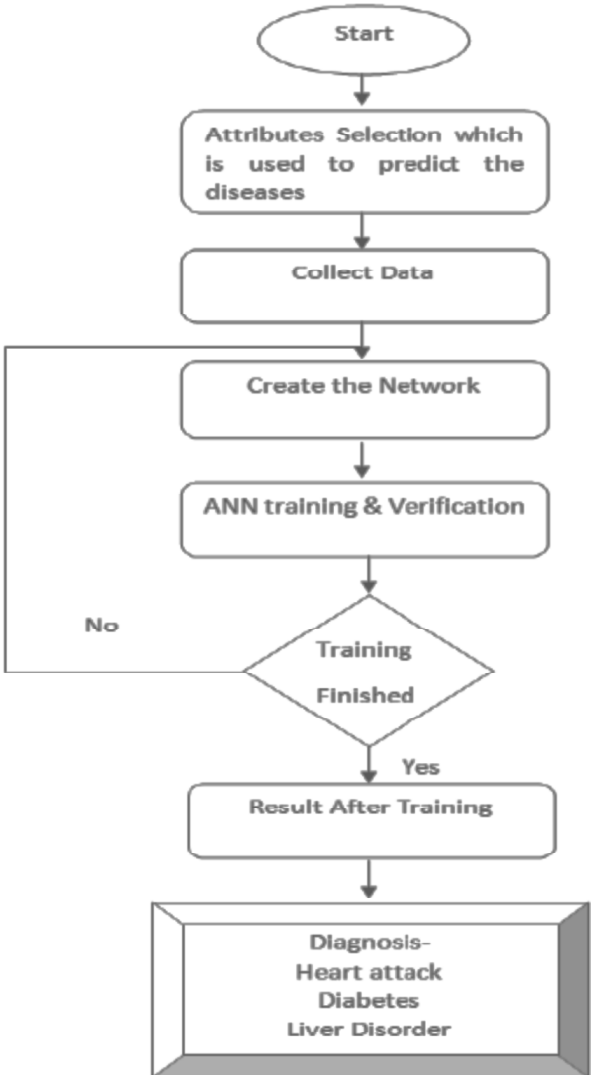

Fig. 2. Flow chart of proposed work.

The quantity of inputs depends on the final sets of components for every patient .Now, the number of units must be chosen accordingly for which the training is quick or fast and system gives the best outputs. The initial step is to initialise the weights of neural networks. At the point these designed weights are passed to the genetic calculations for Optimization. After the weights are upgraded, the feed forward back propagation algorithm has been used for training and learning process [20].

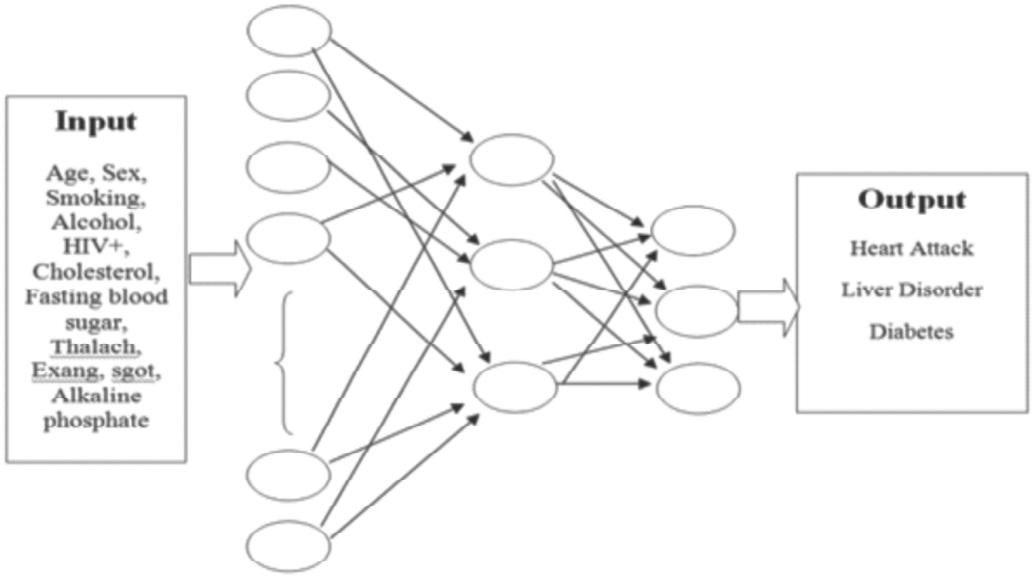

Fig. 3. Network Structure.

4. RESULTS

The input attributes used for the diagnosis of the disease contains the following component :

Age, Sex, Smoking Habit, Alcohol consumption, HIV status ,cholesterol, Fasting Blood Sugar, Thalach, Exang, SGOT and Alkphos[21].

Table 2. Attributes and description of input data

Attributes	*Description*	*Value of Attributes*
Age	Age in Years	
Sex	Sex (M/F)	1-Yes , 0-No
Smoke	Smoking	1-Yes 0-No
Alcohol	Drinking	1-Yes0-No
HIV Status	Positive or Negative	1-Yes 0-No
Chol	Cholesterol	125-570 mg/dl
Fbs	Fasting Blood sugar	[94-200]
Thal	Thalach(Max heart rate achived)	[71-202]
Exang	Exercise induced angine	1-Yes 0-No
SGOT	Serum Glutamic OxaloceticTransamine	[18-324]
Alkphos	Alkaline Phosphate	[62-180]

Diabetes, Heart Disease and Liver Disorder dataset of UCI Machine Learning [22] among the 164 instances of data, 114 instances are used for tanning and 25 were used for testing and 25 for the validation purpose. Here data are classified into 3 Categories- heart attack, Liver disorder & diabetes.

Table 3. Description of output data

Output	*Description*
Class 1	Heart Attack
Class 2	Liver Disorder
Class 3	**Diabetes**

Performance calculation NN toolbox from MATLAB2012 is used for the proposed network performance. This proposed work shows an accuracy of 70% for the dataset. During training of NN tool following windows are opened for training. These windows display training step for the diseases diagnosis.

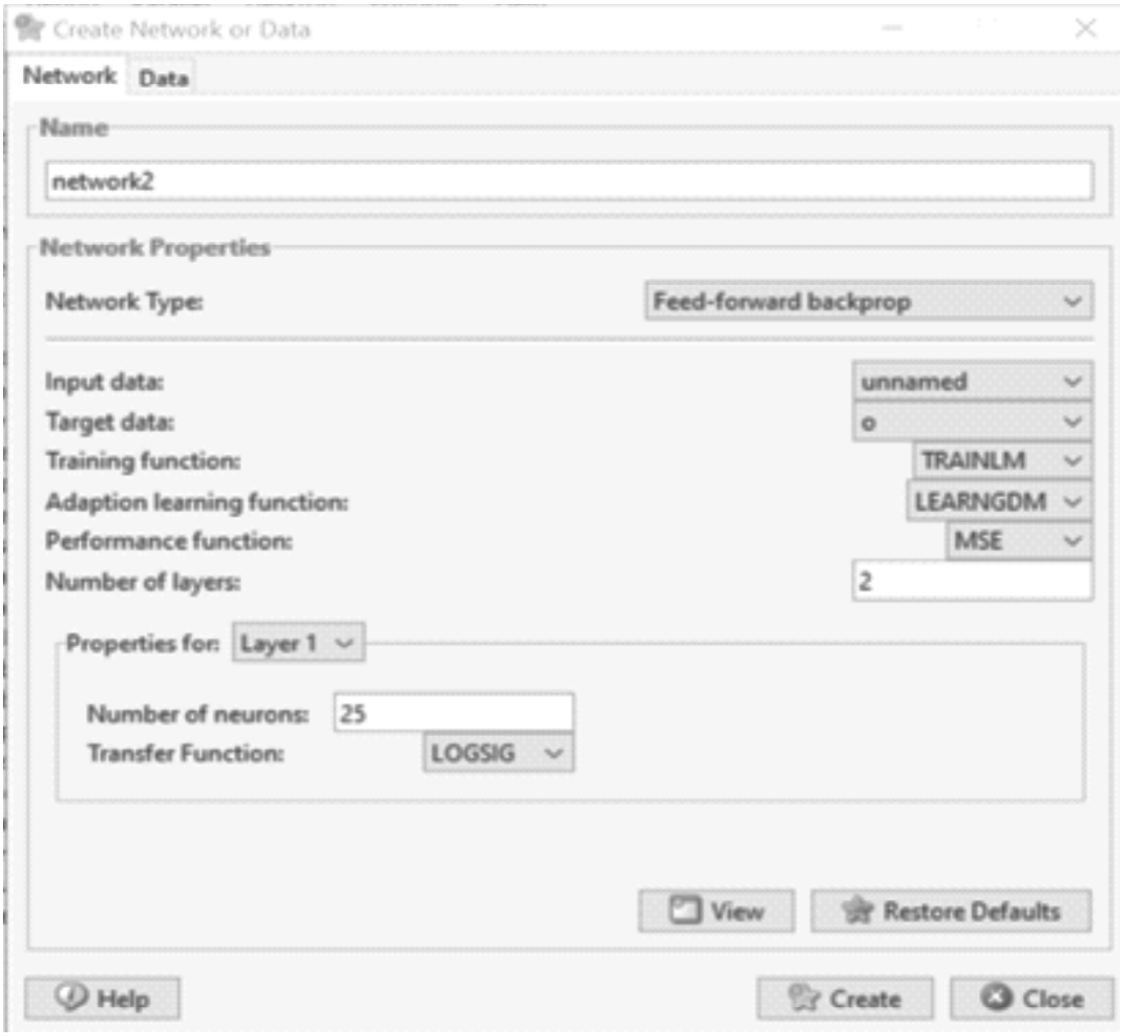

Fig. 4. NN Tool Input Values.

We used TRAINLM is the most reliable and fastest method in the NN tool box. "Trainlm" is a system training function used to update the weight and bias value as according to feed forward back propagation optimisation. We used the "Tainlm" because it the fastest back propagation algorithm in the tool box. This is profoundly recommended as a best choice in supervised algorithm. The learning stops a predefined least error subsequent to changing system weights and adjusting them to an ideal amount at which arrangement is precise. This best method for supervised learning although it requires less space from the rest of the algorithm. MSE used for Network's performance based on mean squared errors

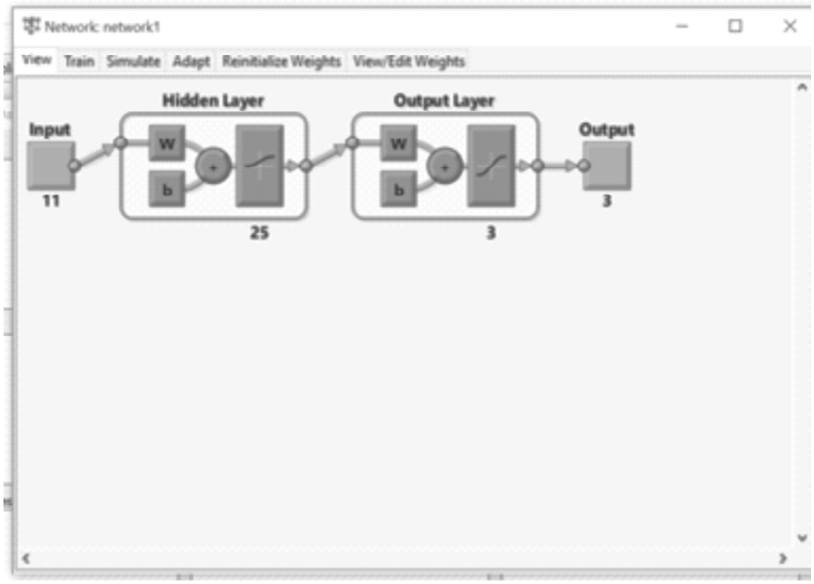

Fig. 5. Neural Network Model with Different Layers.

This is an NN model which consists of 11 input attributes in the input layer and two intermediate layers are used where we take 25 neurons and it gives three outputs.

Table 4. MSE and R values.

	Samples	***MSE***	***R***
Training	114	3.31597e-1	8.65235e-1
Validation	25	1.93237e-1	3.05772e-1
Testing	25	1.8799e-1	2.18168e-1

In this table mean square error and regression values of training of 114 samples, validation of 25 samples and testing of 25 samples are given.

MSE : Mean Squared Error referred as the average(mean) squared difference amongst the output and target values. Thus, lower values are considered much better (consideredmore valuable) to higher values. So, a zero value means no error.

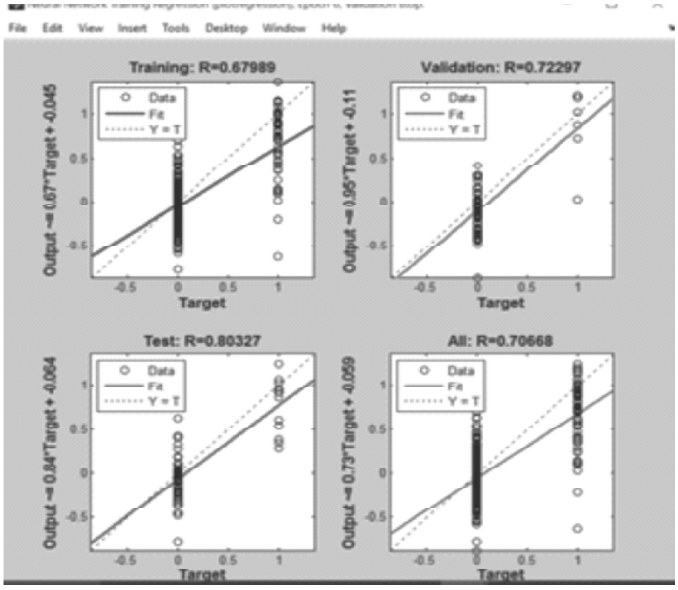

Fig. 6. Neural network regression values.

R : Regression R Values determine the correlation among output and target values. R value of 1 refers to a close relationship and 0 refers to a random relationship.

This figure represents training, validations and testing graphs of regression analysis. In this, Training shows an accuracy of 67%, Validation shows an accuracy of 72%, Testing shows an accuracy of 80% and overall result for this implementation is 70%.

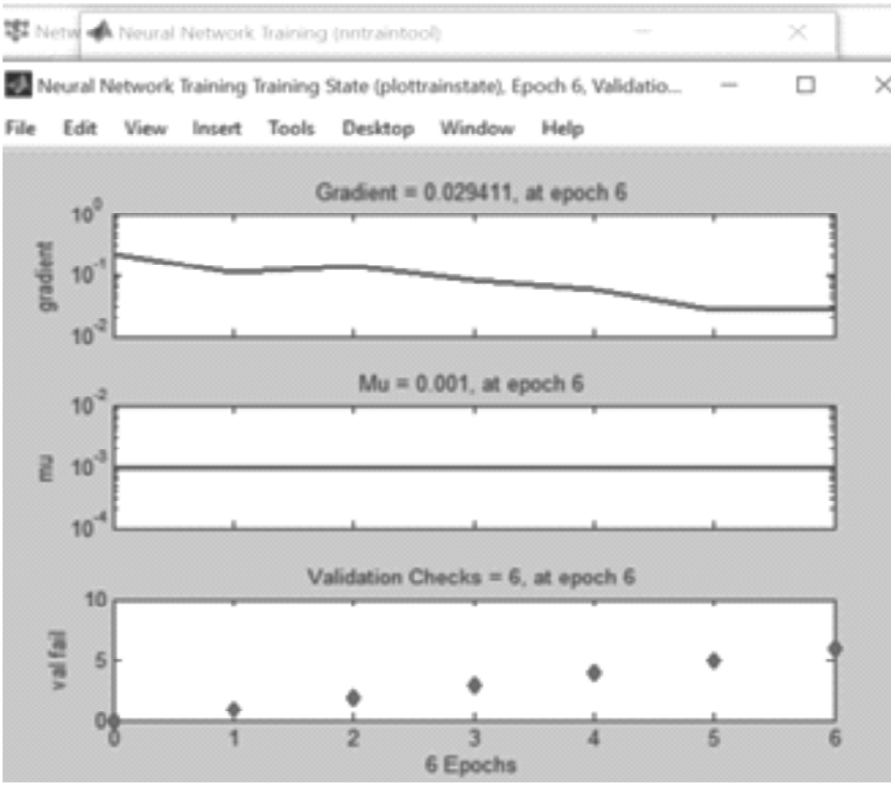

Fig. 7. Neural Network Training state representation.

Finally, it is seen that proposed network predict 70% accuracy. This prediction of diseases using NN is very effective for the specialist.

5. CONCLUSION

This paper presented a system for diagnosis Heart, Liver and diabetes disease using ANN. ANN is a powerful tool to help specialist to perform diagnosis. The Development of ANN technique is based on multilayer perceptron with feed forward neural system. An accuracy of 70% achieved by the proposed network. ANN has proved that it has ability to process a very large amount of information. From the simulated results of proposed work it has been concluded that diagnosis time will be reduced and this system provides great flexibility. ANN is adaptive this makes the diagnosis more reliable and accurate.

6. REFERENCES

1. Konstantina Kourou , Themis P. Exarchos, Konstantinos P. Exarchos , Michalis V. Karamouzi and Dimitrios I. Fotiadis "Machine learning applications in cancer prognosis and prediction", Computational and Structural Biotechnology journal, Elsevier, 2014.
2. Giduthuri Sateesh Babu and Sundaram Suresh "Sequential Projection-Based Metacognitive Learning in a Radial Basis Function Network for Classification Problems", IEEE transactions on neural networks and learning systems, vol. 24, NO. 2, 2013.
3. Atkov O, Gorokhova S, Sboev A, Generozov E, Muraseyeva E, Moroshkina S and Cherniy N."Coronary heart disease diagnosis by artificial neural networks including genetic polymorphisms and clinical parameters. J Cardiol", Elsevier, vol 59: 190–194, 2012.

4. Ms. K Sowjanya, Dr. Ayush Singhal and Ms. Chaitali Choudhary "MobDBTest: A machine learning based system for predicting diabetes risk using mobile devices", IEEE International Advance Computing Conference (IACC), (2015)
5. Xiaoyi Xu,Ya Zhang, Liang Zou, Minghui Wang and Ao Li "A Gene Signature for Breast Cancer Prognosis Using Support Vector Machine", 5th International Conference on BioMedical Engineering and Informatics, 2012.
6. Dimitrios H. Mantzaris. "Medical disease prediction using Artificial Neural Networks", 2008 8th IEEE International Conference on Bio Informatics and Bio Engineering, 2008.
7. K.Balachandran, and R. Anitha. "Dimensionality reduction based on the classifier models: Performance Issues In The Prediction of Lung cancer", 2012 CSI Sixth International Conference on Software Engineering (CONSEG) on IEEE, 2012.
8. Jayshril S. Sonawane and D. R. Patil "Prediction of Heart Disease Using Multilayer Perceptron Neural Network", ICICES2014 - S.A.Engineering College, Chennai, Tamil Nadu, India, IEEE, 2014.
9. Ashwani Kumar Yadav, R. Roy, Archek Parveen Kumar, Ch. Sandesh Kumar, Shailendra Kr. Dhakad, " De-noising of Ultrasound Image using Discrete Wavelet Transform by Symlet Wavelet and Filters", in Proceedings of the 2015 IEEE, pp. 428- 432, 2015.
10. Miss Ankeeta R. Patel and Maulin M .Joshi "Heart diseases diagnosis using Neural Network", 4th ICCCNT 2013 July 4-6, 2013.
11. Sivachitra.M and Vijayachitra.S "Planning and Relaxed state EEG signal classification using Complex Valued Neural Classifier for Brain Computer Interface", IEEE, 2015.
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, 2012.
13. Shikha Agrawal, Jitendra Agrawal "Neural Network techniques for cancer prediction: a survey" 19th International Conference on knowledge based and intelligent information and engineering systems, Procedia computer science, Elsevier (769-774), 2015.
14. Kaur K. "Optical Multistage Interconnection Networks Using Neural Network Approach", Master of Engineering, Thapar University, 2009.
15. K.. Subramanian, S. Suresh, and N. Sundararajan, "A metacognitive neuro-fuzzy inference system (McFIS) for sequential classification problems," IEEE Transactions on Fuzzy Systems, vol.21, no.6, pp.1080-1095, 2013.
16. L.G Kabari, F.S Bakpo "Diagnosing skin diseases using an artificial neural network", Proceedings of the IEEE IInd International conference on adaptive science and technology, 2009.
17. Muhammad Akmal Sapon, Khadijah Ismail, Suehazlyn Zainudin "Prediction of diabetes by using artificial neural network" International Conference on circuits, system and simulation, Volume 7, 2011.
18. P.K.Anooj "Clinical decision support system: Risk level prediction of heart diseases using weighted fuzzy rules", Journal of king saud university- computer and information science, Elsevier, 2012.
19. Seema Singh, Sunita Saini, Mandeep Singh "Cancer detection using adaptive neural network", International journal of advancements in research and technology, Volume 1, 2012.
20. 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, 2003.
21. Syed Umar Amin, Kavita Agarwal and Dr. Rizwan Beg "Genetic Neural Network Based Data Mining in Prediction of Heart Disease Using Risk Factors", Proceedings of 2013 IEEE Conference on Information and Communication Technologies, 2013.
22. https://archive.ics.uci.edu/ml/datasets.html