Indian Journal of Science and Technology, *Vol 10(17), DOI: 10.17485/ijst/2017/v10i17/112708, May 2017*
*ISSN (Print) : 0974-6846*
*ISSN (Online) : 0974-5645*

# An Empirical Study of Applying Artificial Neural Network for Classification of Dermatology Disease

**Sanskruti Patel*** **and Atul Patel**

Faculty of Computer Science and Applications, Charotar University of Science and Technology (CHARUSAT), Changa - 388421, Gujarat, India; sanskrutipatel.mca@charusat.ac.in, atulpatel.mca@charusat.ac.in

## Abstract

**Background/Objectives:** With the growth in complexity and volume of medical data, an extensive set of information currently available in various forms related to diseases and its symptoms. Mechanisms are necessary to extract rules and patterns from these massive set of data. Identification and extraction of hidden patterns and rules in this massive data set certainly help us to understand about diseases progression facts. **Methods:** Machine learning provides an automatic way to uncover the patterns from data set and it will be helpful to health care professionals in order to provide precision medicine to their patients. Artificial Neural Network is a popular machine learning technique used for classification tasks in medical diagnosis for diseases detection. It is an eminent field of computer science which can be applied to the health care sector quite efficiently. In this study, Multi-Layer Feed Forward Neural Network has been applied to the dermatology dataset downloaded from UCI repository site to classify the dermatology diseases. **Findings:** Artificial Neural Network with back propagation algorithm produces the optimum results for classification and prediction problems. It also possesses the ability of generalization and applicable to real world problem. **Applications:** The experiment will be extended by applying on other types of diseases datasets and an automated diagnostic and advisory system with neural network integration definitely helps in diseases prediction problem.

**Keywords:** Artificial Neural Network, Classification, Disease Diagnosis

## 1. Introduction

There is an enormous volume of data available in form of structured, semi-structured and unstructured in health care sector and human analysis of these data is highly impossible. With an immense growth in complexity and volume of medical data[1], availability of decision support systems in medical applications is highly desirable[2]. Therefore, mechanisms are necessary to extract rules and patterns from these massive set of data. Identification and extraction of hidden patterns and rules in theses massive data set certainly help us to understand about diseases progression facts. Machine learning approaches are best suited where ample amount of data are available, but very less is known about the process. Machine learning provides an automatic way to uncover the patterns from data set and it will be helpful to health care professionals in order to provide precision medicine to their patients. It is an eminent field of computer science which can be applied to the health care sector quite efficiently.

Machine Learning is an emerging field of Artificial Intelligence and as its name suggests, the objective of machine learning is to recognize patterns in data to perform useful inference using those patterns that have been learned[3]. Variety of machine learning techniques are available like Artificial Neural Network (ANN), Naive Byes Classifiers, Regression algorithms, Support Vector Machine Algorithms and many more. Applications made in conjunction with machine learning algorithms are highly automated and self-modifying as they continue to improve over time with minimal human intervention as they learn with more data[4]. Artificial Neural Network is a computational model that based on the structure of brain neurons as it learns from experience just like human brain.

** Author for correspondence*

In this study, Multi-Layer Feed Forward Neural Network has been constructed and applied to the dermatology dataset downloaded from UCI repository site[5] to classify the dermatology diseases with highest accuracy.

The usage and applications of artificial neural network is widespread in various fields including searching of a gravitational-wave signals associated with short gamma-ray bursts[6], prediction of dissolution kinetics[7], time series data production[8], vibration control for vehicle active suspension system[9], economic efficiency modeling of broiler production units[10], artificial neural network based model for crop yield[11], neural network for wind turbine[12], Improvement of Quality of Service in wireless sensor network[13] and many more.

Many authors applied neural network model for medical diagnosis as a prediction technology applied neural network for diagnosis of coronary heart disease[14] used a concept of artificial neural networks for kidney stone diagnosis[15]. The work on application of artificial neural network for lung cancer cell identification on diagnosis a urinary system diseases using neural network architecture[16,17]. Usage of neural network for Parkinson's disease prediction was presented by[18]. A review on how machine learning techniques are applicable to diagnosis a thyroid disease was stated by[19].

## 2. Background

Artificial Neural Network (ANN) is made up with the simulated signal processing units called neurons. These neurons are interconnected and able to perform complex tasks[20]. A typical ANN structure consists with three components: Input Layer, Hidden layer and Output layer[21] as per shown in Figure 1. Input layers are connected with hidden layers and it is then after connected with output layer. Here, the data passed to the input layer is propagated through each hidden layer and generated output without having loop or cycles and therefore it is called feed forward neural network.

The ANN must be trained in order to use for classification problem and it is realized with an implementation of Feed Forward Back propagation Neural Network. In this model, input data is forwarded from input layer to output layer and back propagation algorithm is integrated for training purpose. It is widely used for variety of tasks including optimization, reasoning, approximation, classification etc. The data for input with the desired output is provided at the time of training. To measure errors, the actual output is compared with the desired output. These calculated errors and then weights and input threshold of neural network are altered in a way that causes the error to be reduced[20]. As it required known output in advance for training, it is mostly considered as a supervised learning. Figure 2 represents the flow of back propagation algorithm[22].

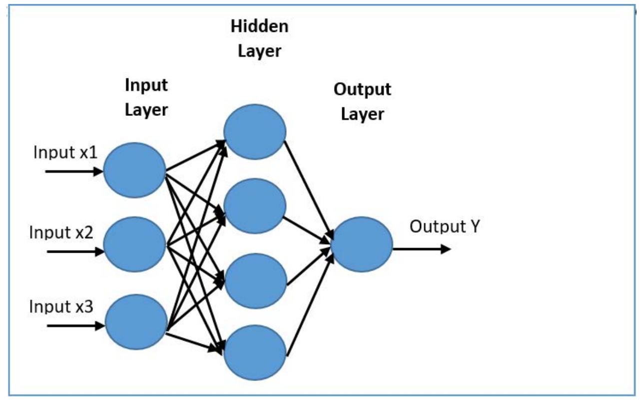



**Figure 1.** Architecture of Artificial Neural Network.

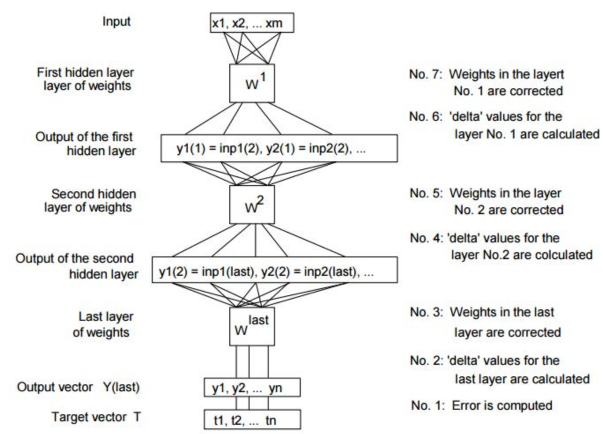



**Figure 2.** Algorithmic steps for back propagation algorithm.

## 3. Model Construction

As discussed earlier, there is a vast amount of health care data available in various repositories and for the development of accurate and efficient decision support system, meaningful information from these massive data should be extracted. Back propagation Feed Forward Neural Network is one of the proven models for classification problems as it is very good at pattern recognition problems and with enough input and output attributes (called neurons) can classify any data with arbitrary accuracy. Several data repositories are available for the researchers online. In this study, the dermatology data set from UCI is used. We have removed missing value instances from the dataset.

The aim of this study is to correctly detect the presence of erythemato squamous diseases, falls under the domain of dermatology. The dataset is having thirty four attributes as inputs and one attributes as an output variable (targeted class variable). There are total six types of diseases falls under this group and they set as an output variable. To construct, train and test the neural network, NeurophSimulator[23] is used. It is based on Java and it is an open source product which is widely used to construct neural network architectures. NeuroPhis a lightweight framework and it offers many alternatives for constructing different architectures of neural networks. It contains well designed open source library and a set of core classes that correspond to basic concepts in neural networks. It also offers good GUI to construct, train and test the model neural network[24]. The following are the steps performed during this experimental study.

## 3.1 Data Normalization and Preparation

The dataset downloaded should be first normalized before use as it is not in the form of direct use as data inputs and output for neural network. For that, all the values in the dataset should be in range of 0 to 1. The following is the equation to normalize the original values in the range of 0 to 1[23].

$$P_x = P - P_{min}/P_{max} - P_{min}$$

Here, $P_x$ is the normalized value, P is the value to be normalized, $P_{min}$ is the minimum value of P and $P_{max}$ is the maximum value of P.

We have used 70% of the data for training purpose and rest 30% for testing the neural network. To train the dataset for neural network, two types of models are available: supervised learning and unsupervised learning. In supervised learning, the training data contains inputs and its corresponding outputs. For that, a training dataset is loaded in Neuroph and the number of inputs set to 34 and number of outputs set to 6.

## 3.2 Neural Network Model Construction

For experiment, we have constructed a multilayer perceptron feed forward neural network. It maps inputs to the output nodes and may consist more than one hidden layers. Excluding the input nodes, each node is a neuron with nonlinear activation function. Multilayer Perceptron uses back propagation algorithm to train the network[23]. It is a type of supervised network and required to train in order to get a desired responses[25]. Figure 3 shows the grahical user interface prvided by Neuroph simulator for setting different parameters to construct a multilayer perceptron neural network.

The number of input neurons and number of output neurons are the same as in the dataset. We have used one hidden layer for this problem.There are generally two

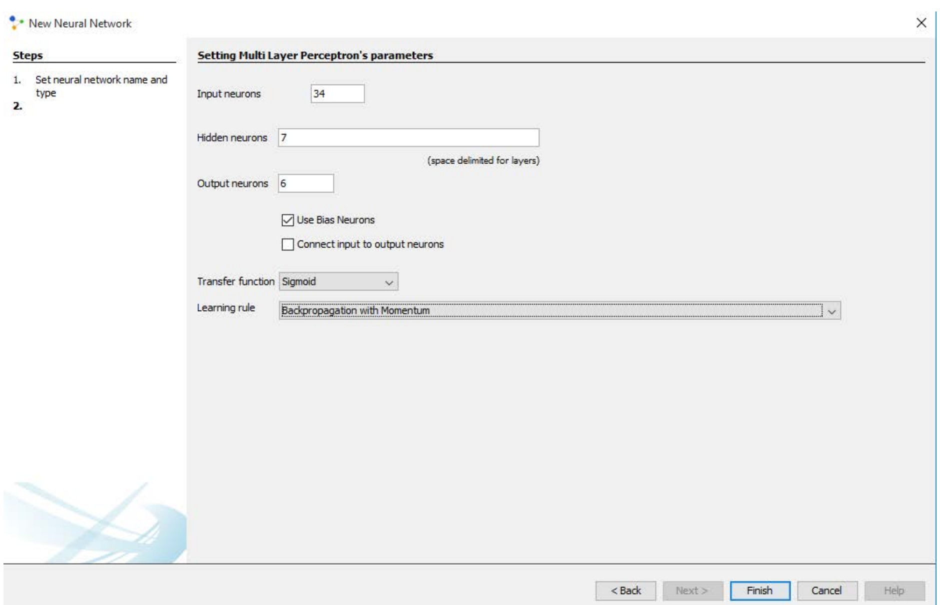
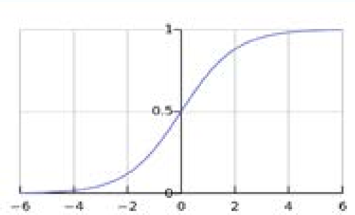



**Figure 3.** Setting parameters for multi-layer perceptron neural network.

problems occurred while adding hidden layers in neural network are overfitting and underfitting. Too many number of neurons in hidden layer generates problem of overfitting as there is a vast information processing capacity posseses by the neural network but small amout of data are avaialble. Due this reason, neural network does not provide training to all the hidden neurons and it resulted in poor performance. On the contrary, too less neurons in the hidden layers generated under fitting problem. Here, the information processing capacity is low against the large amount of data. To decide how many number of neurons should input in hidden layer, some common criteria are available[26].

- The number of hidden layer neurons is 2/3 of the size of the input layer. If this is insufficient then number of output layer neurons can be added further[27].
- The number of hidden layer neurons should be less than twice of the number of neurons in input layer[28].
- The size of the hidden layer neurons is between the input layer size and the output layer size[29].

We have followed third rule for selections of neurons in the hidden layers as per shown in Figure 3. As a transfer function, sigmoid has been selected. A sigmoid function is a mathematical function as per shown in the following Figure 4 with formula. It is having an 'S' shaped curve, known as sigmoid curve and it refers the special case of the logistic function[30]. As a learning rule, back propagation with momentum was selected as it gives better result.




$$S(t) = \frac{1}{1+e^{-t}}$$

**Figure 4.** Curvature of sigmoid function.

## 4. Training of the Neural Network

After constructing architecture for Neural Network, it is trained by setting several learning parameters. The maximum error rate stops the network training, if it is achieved. For better approximation, the smaller error rate should be selected. Learning rate sets as a control parameter for training algorithms, which controls the step size and size of weights when weights are iteratively adjusted. It also uses to control bias changes in learning of the training algorithm. Momentum is used to prevent the system from converging to a local minimum. Setting a momentum value 1.0 i.e. too high, make a system unstable. Other side, setting a momentum value 0.0 i.e. too low, stops the training of the system. With momentum m, the weight update at a given time t becomes

$$\Delta\omega_{ij}(t) = \mu_i\, \delta_i\, y_i + m\, \Delta\omega_{ij}(t-1)$$

Here, $0 < m < 1$ is a new global parameter. It is determined by trial and error[31].

In this experimental study, for training a neural network, different learning parameters were set. Each time, it generated a total network error graph to show the number of iterations and error rate. Also, after training, network was tested and it presented total mean square error to measure the accuracy of trained neural network as per shown in Figure 5 and Figure 6.

As per the standard, two types of training techniques were experimented: standard and advanced. During the standard techniques, the neural network should be first trained with the full data set and then tested with the same dataset. If accuracy is achieved, further advanced training techniques are applied. The advanced learning techniques helps to apply the neural network for generalization as per real world problem and the trained neural network should classify new data correctly that it has never seen before. For that, the dataset is divided into two parts: train dataset and test dataset. These two dataset do not have the same data for measuring the neural network accuracy of classification. Further, as per the standards, the dataset should be divided in three parts: validation, training and testing.

In this experiment, the whole dataset was divided into three sections: Validation dataset contained 10% of the data, Test dataset contained 30% of the data and Training dataset contained 70% of the data that included the 10% of Validation dataset. The instances contained by Test dataset are totally new and never appeared in other two datasets. The network was first trained by using validation dataset, then after it trained thorough training dataset. To measure the accuracy, it then after, tested with the test dataset that is newer and never before used during training.

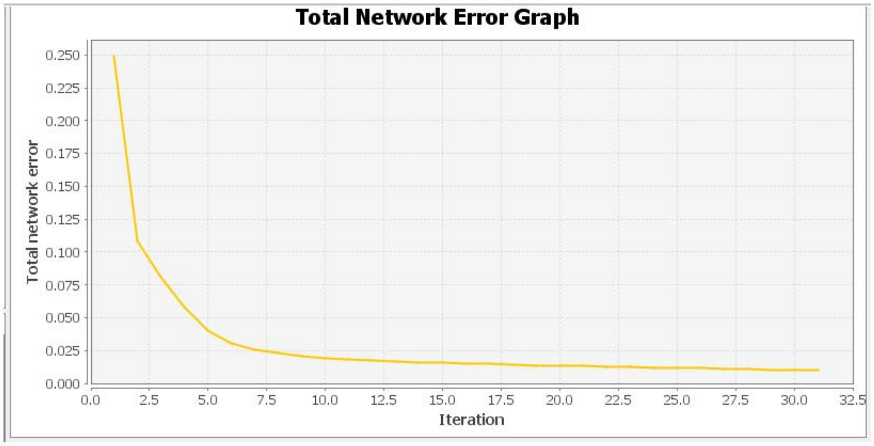



**Figure 5.** Total network error graph generated in simulator.

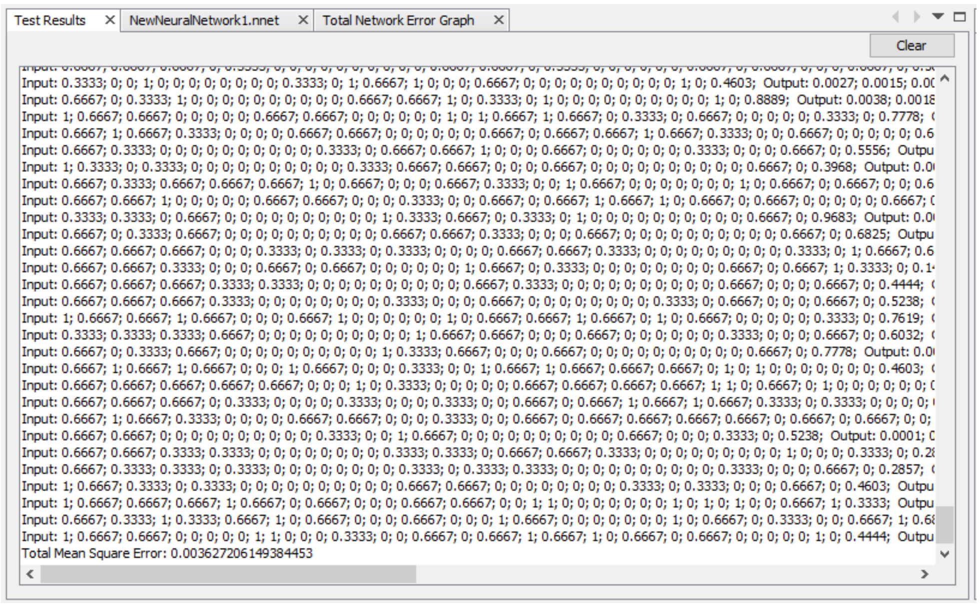



**Figure 6.** Total mean square error after testing the network.

## 4. Results and Discussion

As per the methodology, we have first trained the neural network using standard training techniques. During training, we have experimented different neural network architecture. Also, we have tried to input different numbers of neurons in the hidden layer. We have used full dataset for training purpose. To stop the network training, maximum error rate was set to 0.01. Moreover, different training rate and momentum values were tried to get the optimum result. Total mean square error was taken as the indicator and the lowest value of it leads to the optimum result. As per the outcomes presented in given Table 1. It clearly indicates that one hidden layer with more than one neurons is perfectly fits to the dataset taken for the experiment. Also, the optimum result is obtained with momentum 0.7 and learning rate 0.2.

As per the results obtained during standard training, we have experimented advanced training with the optimum architecture i.e. one hidden layer with more than one neurons. The results of the advanced training are presented in Table 2. By taking the smaller mean square error into consideration, it indicates that 7 neurons in one hidden layer with momentum 0.7 and learning rate 0.2 gets the optimum result.

**Table 1.** Results obtained after standard training techniques

| Number of Iterations | Hidden Neurons | Number of Layers | Learning Rate | Momentum | Total Mean Square Error |
|---|---|---|---|---|---|
| 37 | 7 | 1 | 0.2 | 0.7 | 0.00211 |
| 29 | 7 | 1 | 0.4 | 0.6 | 0.00423 |
| 23 | 7 | 1 | 0.4 | 0.8 | 0.00438 |
| 29 | 20 | 1 | 0.2 | 0.7 | 0.00288 |
| 27 | 20 | 1 | 0.4 | 0.6 | 0.00346 |
| 24 | 20 | 1 | 0.4 | 0.8 | 0.00355 |
| 28 | 20 12 | 2 | 0.2 | 0.7 | 0.00364 |
| 36 | 20 12 | 2 | 0.4 | 0.6 | 0.00384 |

**Table 2.** Results obtained after advanced training techniques

| Number of Iterations | Hidden neurons | Learning rate | Momentum | Total mean square error |
|---|---|---|---|---|
| 09 | 7 | 0.2 | 0.7 | 0.00216 |
| 06 | 7 | 0.4 | 0.6 | 0.02470 |
| 20 | 7 | 0.4 | 0.8 | 0.03664 |
| 12 | 20 | 0.2 | 0.7 | 0.02747 |
| 06 | 20 | 0.4 | 0.6 | 0.00296 |
| 10 | 20 | 0.4 | 0.8 | 0.02605 |

## 5. Conclusion

In this study, one of the popular machine learning techniques, artificial neural network is implemented and tested over dermatology dataset. The whole experiment was carried out using Neuroph – a neural network simulator. Neural network with back propagation algorithm produces the optimum results for classification and prediction problems. It also possesses the ability of generalization and applicable to real world problem. The experiment will be extended by applying on other types of diseases datasets and an automated diagnostic and advisory system with neural network integration definitely helps in diseases prediction problem.

## 6. References


1. Sapna S, Kumar PM. Diagnosis of disease from clinical big data using neural network. Indian Journal of Science and Technology. 2015 Sep; 8(24):1-7.
2. Chimieski1 BF, Fagundes RDR. Association and classification data mining algorithms comparison over medical datasets. J Health Inform. 2013 Apr-Jun; 5(2):44-5.
3. Clifton DA, Gibbons J, Davies J, Tarassenko L. Machine learning and software engineering in health informatics. Proceedings of 1st International Workshop on Realizing Artificial Intelligence Synergies in Software Engineering (RAISE); Zurich, Switzerland. 2012. p. 37-41.
4. Top 10 Machine Learning Algorithms. Available from: Crossref
5. UCI repository. Available from: Crossref
6. Kim K, Ian W. Harry, Kari A. Hodge, Kim YM, Lee CH, Lee HK, John J. Oh SH, Edwin JS. Application of Artificial Neural Network to Search for gravitational-wave. Classical and Quantum Gravity. 2015; 32(24):30.
7. Elcicek H, Akdogan E, Karagoz S. The use of artificial neural network for prediction of dissolution kinetics. The Scientific World Journal. 2014; 2014:9.
8. Belgrano A, Bjorn A, Malmgren, Lindahl O. Application of Artificial Neural Networks (ANN) to primary production time-series data. J Plankton Res.2001; 23(6):651-8.
9. Kalaivani R, Sudhagar K, Lakshmi P. Neural network based vibration control for vehicle active suspension system. Indian Journal of Science and Technology. 2016 Jan; 9(1):1-8.
10. Sefat MY, Borgaee AM, Beheshti B, Bakhoda H. Application of Artificial Neural Network (ANN) for modeling the economic efficiency of broiler production units. Indian Journal of Science and Technology. 2014 Nov; 7(11):1820–6.
11. Liu G, Yang X, Li M. An Artificial Neural Network Model for crop yield responding to soil parameters. Advances in Neural Networks – ISNN, Lecture Notes in Computer Science. 2005; 3498:1017-21.
12. Rajaji L, Kumar C. Neural network controller based induction generator for wind turbine applications. Indian Journal of Science and Technology. 2009 Feb; 2(2):1-5.
13. Abinaya R, Kamakshi S. Improving QOS using Artificial Neural Networks in wireless sensor networks. Indian Journal of Science and Technology. 2015 Jun; 8(12):1-5.
14. Atkov O, Gorokhova S, Sboev A, Generozov E, Muraseyeva E, Moroshkina S, Cherniy N. Coronary heart disease diagnosis by artificial neural networks including genetic polymorphisms and clinical parameters. Journal of Cardiology.2012; 59(2):190–4.

15. Kumar K, Abhishek. Artificial Neural Networks for diagnosis of kidney stone disease. I J Information Technology and Computer Science.2012; 7:20-5.
16. Zhou ZH, Jiang Y, Yang YB, Chen SF. Artificial Intelligence in Medicine.2002; 24(1):25-36.
17. Al-Shayea QK, Itedal S, Bahia H.Urinary System Diseases Diagnosis Using Artificial Neural Networks. IJCSNS International Journal of Computer Science and Network Security. 2010 Jul; 10(7):1-5.
18. Anita S, Priya AP. Early prediction of Parkinson's disease using Artificial Neural Network. Indian Journal of Science and Technology. 2016 Sep; 9(36):1-7.
19. Razia S, Rao MRN. Machine learning techniques for thyroid disease diagnosis - A review. Indian Journal of Science and Technology. 2016 Jul; 9(28):1-9.
20. Dharwal R, Kaur L. Applications of Artificial Neural Networks: A review. Indian Journal of Science and Technology. 2016 Dec; 9(47):1-8.
21. Chakravarti A, Joshi N, Panjiar H. Rainfall runoff analysis using artificial neural network. Indian Journal of Science and Technology. 2015 Jul; 8(14):1-7.
22. Zupan J. Introduction to Artificial Neural Network (ANN) methods: What they are and how to use. Acta Chimica Slovenica. 1994; 327-52.
23. Neuroph. Available from: Crossref
24. Nenkov NV, Spasova EZ. Implementation of a neural network using simulator and petri nets. IJACSA. 2016; 7(1):1-6.
25. Panchal G, Ganatra A, Kosta YP, Panchal DP. Behaviour analysis of multilayer perceptron with multiple hidden neurons and hidden layers. International Journal of Computer Theory and Engineering. 2011 Apr; 3(2):1-6.
26. Karsoliya S. Approximating Number of Hidden layer neurons in Multiple Hidden Layer BPNN Architecture. International Journal of Engineering Trends and Technology.2012; 3(6):1-4.
27. Boger Z, Guterman H. Knowledge extraction from artificial neural network models. IEEE Systems, Man and Cybernetics Conference; Orlando, FL, USA. 1997. p. 3030-5.
28. Berry MJA, Linoff G. Data Mining Techniques. NY: John Wiley and Sons; 1997. p. 1-672.
29. Blum A. Neural Networks in C++. NY: Wiley; 1992. p. 224.
30. Sigmoid function. Available from: Crossref
31. Orr G, Schraudolph N, Cummins F. Momentum and learning rate adaptation. Available from: Crossref