IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 11, November 2016

Survey on Optimization of Number of Hidden Layers in Neural Networks

Dinesh P. Pitambare

Assistant Professor, Dept. of Computer Engineering, AISSMS IOIT, Savatribai Phule University, Pune, India

Abstract: Artificial neural networks attempt to mimic the basic operations of the brain. Information is pass between the neurons, and based upon the structure and synapse weights, a network behavior or output mapping is provided. Challenges in artificial neural network are finding number of hidden layer and hidden neurons in respected layer, design of a network with more hidden layers, designing optimal network, finding minimal network structure in less time, determination of processing and storage resources in implementation of large and effective network. Among these challenges, finding number of hidden layers and hidden neurons in the respected layer is the core objective of this system. Multi layer neural network permit more complex, non linear relationship of input data to output results. Hidden layer is the intermediate unit between input and output unit used to calculate weighted sums of the inputs. The output result of the network is dependent on hidden layer result. In a network with a hidden layer and an output layer, the hidden layer is computed first and then the result of the hidden layer are used to compute the output layer.

Keywords: Artificial Neural Networks; hidden neuron; hidden layer.

I. INTRODUCTION

Clustering is unsupervised and classification is supervised approach. Neural networks uses supervised approach hence it is vital to find number of neuron in hidden layer of artificial neural networks [1]. Multi layer neural network permit more complex, non linear relationship of input data to output results. Hidden layer is the intermediate unit between input and output unit. It is used to calculate weighted sums of the inputs. The output result of the network is dependent on hidden layer result. In a network with a hidden layer and an output layer, the hidden layer is computed first and then the result of the hidden layer are used to compute the output layer.

II. RELATED WORK

In [2] authors proposed a new method to determine the optimal architecture by using a pruning technique. Delta values are used to identify the irrelevant neurons in hidden layers. Small number of neurons in modified network exhibit superior generalization, which ultimately increases the speed relative to the back propagation training. Generalization power of ANN is directly proportional to the Number of Hidden Layers in a network.General Approximation Theorem on Feedforward Networks proposed standard feed-forward neural networks with lesser single hidden layer and arbitrary bounded nonlinear activation functions. Generalized sigmodial function can be used as activation functions for a single hidden layer neural network to approximate any arbitrary mappings. On activation functions at an infinity conditions boundedness and unequal limits are sufficient conditions in [3].In [4] Optimization of Neural Networks Using Variable Structure Systems proposed a new blend training algorithm along with error back propagation (EBP) and variable structure systems (VSSs).A new term based on the output of the hidden layer is added to the cost function as a consequence to make optimal use of hidden units related to weights corresponding to each unit in the hidden layer to achieve the optimization of the number of neurons in the hidden layer. Main challenge in multilayer artificial neural networks is nothing but finding number of neurons in different number of hidden layer, there is no exact solution for optimization of multilayer ANN in hidden layer in [5]. In Finding the number of hidden neurons for an multilayer perceptron neural network using coarse to fine search technique by the number of neurons in the inputs and output layers are ascertained by the number of existing inputs and required outputs respectively. This approach uses binary search method for finding the number of hidden neurons with least mean squared error (MSE). In [5] author proposed K-means-Greedy Algorithm (KGA) model that constitute the process of automatically finding the optimal number of neurons in the hidden layer. The premise is that a back propagation (BP) network which has this optimal number of neurons in its hidden layer. KGA model is effective in finding the optimal number of neurons for the hidden layer of a backpropogation network which is used to perform prediction of a time series. In [6] author proposed novel training algorithm for artificial neural networks that combines the gradient descent technique with variable structure systems approach. This combination is performed by conventional weight update rule in contiguous time and applications of sliding mode control method to the gradient based training procedure. This combination provides robustness with respect to unmolded multivariable internal dynamics of gradient descent. A key to construction of a network structural design of ANN.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 11, November 2016

Selection of optimal number of hidden layer and hidden node affects performance of neural network. For that purpose, it use Particle Swarm Optimization (PSO) algorithm, which is a straight forward and easy to implement but highly effective evolutionary algorithm [7].In [8] author proposed the correlation between the number of hidden layer and the performance of the back propagation neural networks. The selection of learning rate is also examined using the 3-layer BPNN and the same non-linear system. This paper verifies how the number of hidden layers affects the performance of a pack propagation neural network. The 3- layer BPNN is considered the best performer compared with other multi-hidden layer BPNNs. A New Adaptive Merging and Growing Algorithm for Designing Neural Networks presented a new algorithm, called adaptive merging and growing algorithm (AMGA), which combine and add together hidden neurons during the training process of ANNs. This algorithm emphasis on autonomous functioning in the design process of ANNs. AMGA can design compact ANN architectures with good generalization as of algorithms in [9].A universal neural net with guaranteed convergence to zero system error invented a learning algorithm with guaranteed convergence to zero system error. This algorithm has high potential to converge fast. A mathematical proof of the guaranteed learning of the neural networks is presented, which provides high potential for fast learning in [10].

In [11] an algorithm that automatically design compact two hidden layered artificial neural networks proposed algorithm for automatic design of two hidden layered artificial neural networks, which is also known as cascade neural network design algorithm (CNNDA), uses constructive approach to form network by addition of node one by one at a time. In order to improve the computational efficiency and convergence] CNNDA adopts an training process which temporarily freezes the input weights of a hidden node when the output of that node does not change much in the successive few iterations. In [12]effects of the sample size in artificial neural network classifier design Stated the effects of the sample size on reckon of the error rate of the artificial neural network classifiers. ANN classifier is used to fix the size of the network by stopping condition for learning. Stopping condition can be specified using number of steps (m). Usually m can be selected by ANN classifier using heuristic approach.Multilayer feed forward network is superior in artificial neural network. It uses back propagation method. A BPN network is method is efficient, reliable that yields the absolute optimal network in [13]. For optimal network training process must continuous until absolute minimum error is reached. Generally the number of neurons in the hidden layer of multilayer neural network is determined by trial and error method. Here 2 phase method is used to determine the number of neurons viz. candidate number of neurons is determined by error back propagation and optimal number of neurons is from generalization capacity. Using this 2 phase method number of neurons are determined in short time and so this method is effective in [14].In [15] author proposed, an iterative pruning algorithm for feedforward neural networks proposed a new pruning method is developed that is used to train larger network. The verification of this approach shows effectiveness of the proposed approach.In [16], Constructive algorithms for structure learning in feedforward neural networks for regression problems, in this survey paper, verification of constructive algorithm for structure learning in feedforword neural networks for regression problem is specified. This algorithm concentrates on search strategy.VSS is used to optimize parameter updating of neural networks, to control the dynamic model of the training process and to provide robustness and conventional way for learning process. For minimizing the cost function EBP is used. Proposed work is beneficiary as of improved robustness, guaranteed convergence and lower sensitivity to initial weights of the neural network. Four different procedures are used to fix network structure automatically viz. pruning, constructive, regularization & cross validations in [17].In [18] author proposed Optimization of one and two hidden layer neural network architectures for handwritten digit recognition that uses back propagation for training up to 1000 iterations.In [19] author proposed the Marquardt algorithm for nonlinear least squares and is incorporated into the back propagation algorithm for training feedforward neural networks. This work concerned with the speed of convergence of neural network. The Marquardt algorithm is effective than other techniques when the network contains no more than a few hundred weights. In First and Second-Order Methods for Learning: between Steepest Descent and Newton's Method [20], reviews of first- and second-order optimization methods for learning in feedforward neural networks are given in this paper. In [21] author uses incremental training procedure. In these paper different problems in neural network training are recognized viz. finding number of neurons, convergence and its speed. Training procedure will lead to the solution is called as convergence.

III.PROPOSED WORK

Artificial neural network is nothing but computational models that are inspired by an animal's central nervous systems, these models are used to calculate or approximate unknown functions. Computational model is depending on a large number of inputs. eg. human brain. Artificial neural network is also known as "neural network" (NN). Artificial neuron is also called as node.

ANN uses three types of parameters:

1. The interconnection pattern between the different layers of neurons.
2. The learning process for updating the weights of the interconnections.
3. The activation function that converts a neuron's weighted input to its output activation.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 11, November 2016

Multi layer neural network permit more complex, non linear relationship of input data to output results. Multilayer neural network is made up of an input layer, an intermediate or hidden layer and an output layer. The input layer is fixed length vector containing user defined numbers. Hidden layer performs intermediate calculations that trained the network. This layer is used to calculate weighted sums of the inputs. Output layer represents the output containing fixed length vector of the numbers. The output result of the network is dependent on hidden layer result. In a network with a hidden layer and an output layer, the hidden layer is computed first and then the result of the hidden layer are used to compute the output layer. More complex systems will have more layers of neurons with some having increased layers of input neurons and output neurons

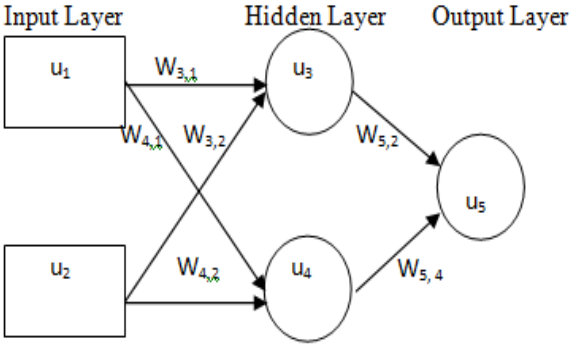

Figure 1. Multilayer network

Here in figure 1, input layer nodes are u_1 and u_2, u_3 and u_4 are hidden layer nodes and u_5 is output layer node. In a neural network, each neuron has an activation function which specifies the output of a neuron to a given input. One of the activation function commonly used for neurons is the sigmoid function. This function looks like an S hence called as sigmoid function.

$$f(x) = \frac{1}{1+e^{-x}} \qquad (1)$$

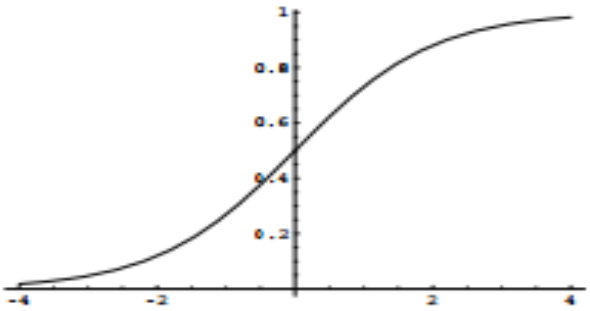

Figure 2. Sigmoid activation function.

When one uses the neural network to learn the updating of the contributions depends on the steepness slope of the activation function. For adjusting connection weights learning rate is used. It determines how quickly the backpropogation algorithm converges toward a solution.

Multi layer neural network permit more complex, non linear relationship of input data to output results. Multilayer neural network is made up of an input layer, an intermediate or hidden layer and an output layer. The input layer is fixed length vector containing user defined numbers. Hidden layer performs intermediate calculations that trained the network.

This layer is used to calculate weighted sums of the inputs. Output layer represents the output containing fixed length vector of the numbers. The output result of the network is dependent on hidden layer result. In a network with a hidden layer and an output layer, the hidden layer is obtained first and then the results of the hidden layer are used to compute the output layer. Exact number of neurons can be obtained using feedforword backpropogation algorithm. More complex systems will have more layers of neurons with input neurons and output neurons. Uniqueness of the proposed system is it will be efficiently finding the number of hidden neurons in more than four hidden layer. Neural network software is used to simulate research, develop and apply artificial neural networks. Microsoft visual studio, matlab, eclipse, netbean and jcrator are different tools used for the proposed system.

IV. APPLICATION

Applications of the proposed system are same as of the neural networks. Mainly the proposed system will be more useful in rural areas of India to identify different diseases patients quickly like cancer, thyroid and diabetes patients.

V. CONCLUSION AND FUTURE WORK

Finding the number of neurons and hidden layer in a network is vital part.To the best of my knowledge there is no exact solution for finding optimal number of hidden layer and hidden neuron which is one of future concern. The proposed system will efficiently discover the number of hidden neurons in numerous hidden layers using optimal design of feed forward backpropogation neural network. Approach may be extended to Nemours hidden layers, for improving the performance in the network changes in weights of connections, learning rate and number of hidden layers required.

REFERENCES

[1] Dinesh P. Pitambare ,Pravin M. Kamde," EfficientlyExploring Clusters using Genetic Algorithm and Fuzzy Rules", International Conference on Computer Communication and Informatics (ICCCI - 2014), Jan. 03 – 05, 2014, Coimbatore, INDIA

[2] N. M.Wagarachchi and A. S. Karunananda "Optimization of Multi-layer Artificial Neural Networks Using Delta Values of Hidden Layers" Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2013 IEEE Symposium, pp. 80 – 86, Apr 2013.

[3] Ken Chen, Shoujian Yang and Batur, C., "Effect of Multi-hidden-layer Structure on Performance of BP Neural Network: Probe", IEEE-Natural Computation (ICNC),pp. 1-5, May 2012.

[4] Seyed Alireza Mohseni and Ai Hui Tan,"Optimization of Neural Networks Using Variable Structure Systems"", Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions , vol. 42, issue: 6, pp. 1645 – 1653, May 2012.
[5] Bong, D.B.L. ; Tan, J.Y.B. ; Rigit, A.R.H.," Optimization of the back propagation hidden layer by hybrid K-means-Greedy Algorithm for time series prediction", IEEE conference, pp. 669-674,Oct. 2010.
[6] Kazuhiro Shin-ike and Maizuru, "A Two Phase Method for Determining the Number of Neurons in the Hidden Layer of a 3-LayerNeural Network", IEEE-SICE Annual Conference , pp. 238 - 242, Aug 2010.
[7] Doukim, C.A. , Dargham, J.A.and Chekima, A. ,"Finding the number of hidden neurons for an MLP neural network using coarse to fine search technique", Information Sciences Signal Processing and their Applications (ISSPA), IEEE Conference, pp. 606 – 609, May 2010.
[8] N. P. Suraweera and D. N. Ranasinghe, "A Natural Algorithmic Approach to the Structural Optimization of Neural Networks", Information and Automation for Sustainability ICIAFS 2008. 4th IEEE Conference, pp. 150 – 156, Dec. 2008.
[9] Islam, M.M. , Amin, M.F., Ahmmed, S. and Murase, K.,"A New Adaptive Merging and Growing Algorithm for Designing Neural Networks", IEEE World Congress on Computational Intelligence, pp. 2003 – 2008, June 2008.
[10] Derong Liu ,Chang, Tsu-Shuan ; Yi Zhang ,"A New Learning Algorithm for Feed forward Neural Networks", IEEE International Symposium on Intelligent Control, ISIC 2001 ,pp. 39 – 44, 2001.
[11] Islam, M.M. ,Shahjahan, and Murase, K.,"An algorithm that automatically design compact two hidden layered artificial neural networks", International Joint Conference on Neural Networks. IJCNN, IEEE-INNS-ENNS ,vol. 6, pp. 467 – 472, Jul 2000.
[12] Fee, M.O. and Kayak, O., "Hybrid training procedure for artificial neural networks leading to parametric stability and cost minimization", IEEE International Symposium on Emerging Technologies and Factory Automation ETFA ,vol. 1, pp. 71 – 77 ,1999.
[13] Cooke, M.J., Lebby, G.L., "An Optimal Design Method for Multilayer Feed forward Networks", IEEE Conference on System Theory, pp. 507 - 511, Mar 1998.
[14] Guang-Bin Hiiang and Haroon A Bari, "General Approximation Theorem on Feed forward Networks", IEEE Information, Communications and Signal Processing ICICS, pp.698 – 702, vol.2, Sep 1997.
[15] Castellano, G., Fanelli, A.M.and Pelillo, M.,"An iterative pruning algorithm for feedforward neural networks", IEEE Transactions on Neural Networks, vol. 8 , issue: 3 , pp. 519 – 531, May 1997.
[16] Tin-Yau Kwok and Dit-Yan Yeung,"Constructive algorithms for structure learning in feed forward neural networks for regression problems", IEEE Transactions on Neural Networks ,vol. 8 , pp. 630 – 645,May 1997.
[17] Uchimura, S., Hamamoto, Yoshihiko ; Tomita, S., "Effects of the sample size in artificial neural network classifier design", IEEE Transactions on Neural Networks, vol. 4 ,pp. 2126 - 2129 ,Dec 1995.
[18] Perez, C.A., Holzmann, C.A., Morelli, I. "Optimization of one and two hidden layer neural network architectures for handwritten digit recognition", IEEE Conference on Systems, Man and Cybernetics,1995. Intelligent Systems for the 21st Century. vol. 3, pp. 2795 - 2799, Oct 1995.
[19] Hagan, M.T. ; Menhaj, M.B.,"Training feed forward networks with the Marquardt algorithm", IEEE Transactions on Neural Networks ,vol. 5, issue: 6, pp. 989 – 993, 1994.
[20] Battiti, R, "First- and Second-Order Methods for Learning: Between Steepest Descent and Newton's Method", Neural Computation, IEEE Transactions, vol. 4, issue: 2, pp. 141 – 166, Mar 1992.
[21] Chang, Tsu-Shuan, Abdel-Ghaffar and K.A.S. ," First- and Second-Order Methods for Learning: Between Steepest Descent and Newton's Method" , IEEE Conference on Neural Computation ,vol. 4 , Issue: 2 ,pp.141 – 166, 1992.

BIOGRAPHY

Dinesh Praksah Pitambare is an Assistant Professor in the Computer Engineering Department, All India Shri Shivaji Memorial Society's Institute of Information Technology-01, Pune, Savatribai Phule University, Pune, India. He received Master of Engineering (ME) degree in 2014 from PU, Pune, MS, India. Her research interests are Artificial Intelligence.