International Journal of Computer and Communication Technology

Volume 1 | Issue 2 Article 9

April 2010

Functional Analysis of Artificial Neural Network for Dataset Classification

Rojalina Priyadarshini
Deptt. Of IT C.V.Raman College Of Engg, Bhubaneswar, r.priya1707@yahoo.co.in

Nillamadhab Dash
Deptt. Of IT C.V.Raman College Of Engg, Bhubaneswar, nilamadhab04@gmail.com

Tripti Swarnkar
Deptt. Of CA, SOA University Bhubaneswar, tripti_sarap@yahoo.com

Rachita Misra
Deptt. Of IT C.V.Raman College Of Engg, Bhubaneswar, rachita03@yahoo.com

Recommended Citation

Priyadarshini, Rojalina; Dash, Nillamadhab; Swarnkar, Tripti; and Misra, Rachita (2010) "Functional Analysis of Artificial Neural Network for Dataset Classification," *International Journal of Computer and Communication Technology*: Vol. 1 : Iss. 2 , Article 9.
DOI: 10.47893/IJCCT.2010.1036
Available at: https://www.interscience.in/ijcct/vol1/iss2/9

Functional Analysis of Artificial Neural Network for Dataset Classification

Rojalina Priyadarshini[1]
Deptt. Of IT
C.V.Raman College Of Engg, Bhubaneswar
r.priya1707@yahoo.co.in

Nillamadhab Dash[2]
Deptt. Of IT
C.V.Raman College Of Engg, Bhubaneswar
nilamadhab04@ gmail.com

Tripti Swarnkar, [3]
Deptt. Of CA,
SOA University
Bhubaneswar.
tripti_sarap@yahoo.com

Rachita Misra, [4]
Deptt. Of IT
C.V.Raman College Of Engg, Bhubaneswar
rachita03@yahoo.com

Abstract- **Classification is one of the most active research and application areas of artificial neural networks (ANN). One of the difficulties in using ANN is to find the most suitable combination of training, learning and transfer function for classification of data sets with increasing number of features and classified sets. In this paper we have studied the effect of different combinations of functions while using artificial neural network as a classifier and analyzed the suitability of these functions for different kinds of datasets. The appropriateness of the proposed work has been determined on the basis of mean square error, rate of convergence, and accuracy of the classified dataset. Our inferences are based on the simulation results over the datasets used.** .

Keywords- **Artificial Neural Network, Back propagation Algorithm, Classification, Transfer function**

I.INTRODUCTION

Artificial neural networks (ANN) are used in many applications to solve some real world problems. These have attracted the attention of many researchers in different areas, such as neuroscience, mathematics, physics, electrical and computer engineering, and psychology [2]. In commercial purposes, ANNs can be applied to predict the profit, market movements, and price levels based on the market's historical dataset. In medical applications, doctors can evaluate the situation of many patients depending on the historical dataset of other patients who had the same illness. In industry, engineers can apply ANNs to solve many engineering problems such as classifications, prediction, pattern recognition, and non-linear problems [3] where the issues are very difficult or might be impossible to solve through normal mathematical processes. ANNs were used to predict the rain attenuation on an Earth-space path, to predict the water quality index (WQI), and to signal predictions in a nuclear power plant. They were also used in face recognition. In medical applications, ANNs were utilized in detecting different diseases like Thalassaemia classification [8], Breast Cells Classification [9] and Diabetic Retinopathy Classification [6], in computational biology these are used in protein structure prediction, DNA motif prediction etc. Generally, such systems consist of a large number of simple neuron processing units performing computation by a dense mesh of nodes and connections.

II. LEARNING STRATEGY

ANNs gather their knowledge by detecting the patterns and relationships in data and learn (or are trained) through experience. An ANN is created from hundreds of single units, artificial neurons or processing elements (PE), connected with coefficients (weights), which constitute their architecture in a layered form. The power of neural computations comes from connecting neurons in a network. Each processing element has weighted inputs, transfer function and one output. The working mode of a neural network is determined by the transfer functions of its neurons, by the learning rule, and by the architecture itself. The weighted sum of the inputs forms the activation of the neuron. The activation signal is passed through transfer function to produce a single output of the neuron. Transfer function introduces non-linearity to the network. During training, the weights are optimized until the error in predictions is minimized and the network reaches the prespecified level of accuracy or the convergence is reached. Once the network is trained and tested, it can take new unknown input to predict the output.

Learning and generalization are perhaps the most important topics in neural network research [5]. Learning is the ability to approximate the behavior adaptively from the training data while generalization is the ability to predict the training data. Generalization is a more desirable and critical feature because the most common use of a classifier is to make good prediction on new or unknown objects. A number of practical network design issues related to learning and generalization include network size, sample size, model selection, and feature selection, transfer function to be chosen, learning rate parameter…etc. In our study we have taken Multidimensional data sets which can represent many real world problems mainly taken from medical background. The classification and clustering of these data sets are meaningful. To test our approach, we have taken different data class sets from repository of data sets. We have divided the data sets into two parts, i.e. training set and testing set which is not used in the training process, and is used to test and then we have simulated our results with these datasets. Almost $2/3^{rd}$ of the total dataset has taken as training set and $1/3^{rd}$ of the rest has taken as test set. This is done through the analysis of the accuracy achieved through testing against these data sets. Then we have simulated our network with the same data.

A. Dataset Preparation

Normalization of input data is used for ranging the values to fall within an acceptable scope, and range [1]. The features obtained from many specimens are actually considered as raw data. Some amount of pre-processing is always carried out on the input and output data in order to make it suitable for the network. These procedures help obtain faster and efficient training. If the neurons have nonlinear transfer functions (whose output range is from -1 to 1 or 0 to 1), the data is typically normalized for efficiency [5]. Each feature in each dataset is normalized using column normalization in MATLAB. The normalized data are used as the inputs to the network.

III. PROPOSED WORK

In our work we have taken different datasets by gradually increasing the no of feature attributes. We have designed the neural network to train these datasets by taking different training functions such as TRAINGD, TRAINGDM, TRAINLM, TRAINRP (Train Resilient Back propagation) and TRAINSCG with different learning function such as LEARNGDM, LEARNGD and with transfer functions TANSIG, PURELIN and LOGSIG. The neural network was trained by back propagation algorithm. Gradient descent method (GDM) was used to minimize the mean squared error between network output and the actual error rate [2]. For different datasets the no of input and output nodes in the network architecture are going on changing with the following learning and training functions. The training error continues to decrease as the number of epoch's increases. Repeated experiments were performed to get the neural network converged. Weights were initialized to random values and networks were run until at least one of the following termination conditions was satisfied

1. Maximum Epoch
2. Minimum Gradient
3. Performance Goal

For testing, the input data was presented to the ANN without weight adjustment. The output of the ANN was compared with the existing class attribute of the datasets. Then we have simulated the network with some of the inputs. We are measuring the efficiency of the network by taking the following parameters:

1. Rate of convergence
2. No of epochs taken to converge the network.
3. The calculated Mean Square Error (MSE).

Training and learning functions are mathematical procedures used to automatically adjust the network's weights and biases. The training function dictates a global algorithm that affects all the weights and biases of a given network.

Neural Network Toolbox supports a variety of training algorithms. We have tested some training functions including TRAINLM(Train Levenberg-Marquardt) TRAINLM is a network training function that updates weight and bias values to Levenberg-Marquardt optimization.,TRAINSCG (Train scaled conjugate gradient).TRAINSCG is a network training function that updates weight and bias values according to the scaled conjugate gradient method, TRAINGD(Train Gradient Descent). TRAINGD is a network training function that updates weight and bias values according to gradient descent method and TRAINGDM (Train Gradient Descent Momentum). It trains the network with Gradient descent with momentum back propagation method.

The learning function can be applied to individual weights and biases within a network. Learning algorithms are used to adapt networks. We have used LEARNGD and LEARNGDM. LEARNGD is the gradient descent weight and bias learning function and LEARNGDM is the gradient descent with momentum weight and bias learning function.

IV. RESULTS AND DISCUSSION

Data set	No of Instances (size)	No of predicting attributes (No. of features)	No of class attribute	Attribute types
1	150	4	3	Real
2	120	6	4	Real, integer
3	4177	32	2	Categorical, Integer , Real

It was found that the datasets with less number of predicted attributes are quickly converged with training function TRAINLM, where as the datasets having large number of predicted attributes are converging with training function TRAINSCG with transfer function LOGSIG and learning function LEARNGDM. But it was observed that the TRAINSCG function works on datasets having less as well as large predictive attributes Transfer functions calculate a layer's output from its net input. We have tested different Transfer functions like Log sigmoid (LOGSIG), Hyperbolic tangent sigmoid transfer function(TANSIG) and Linear transfer function(PURELIN) and from the experiments we have found out that LOGSIG transfer function makes the network converged where as with the other two network is not reaching its convergence.

Data set	Training Function	Learning function	Transfer function	No of epochs	MSE
1	TRAINLM, TRAINSCG	LEARNGDM	LOGSIG	55, 977	**0.0066667,** **0.0133415**
2	TRAINLM, TRAINSCG	LEARNGDM	LOGSIG	34, 68	**0.19375,** **0.0000004935**
3	TRAINSCG	LEARNGDM	LOGSIG	25115	**0.0000003329**

Table-2 summarizes the information about the convergence of the neural network by taking account different parameters on different datasets

The graphs below showing the results of the convergence of the network by taking the above said parameters with dataset 1, 2, 3 as inputs. In X-axis the graphs are showing the number of epochs to reach the convergence and in Y-axis these are showing the MSE(Mean Square Error)

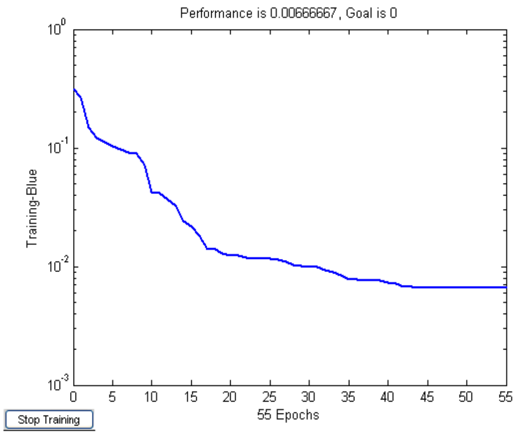

Fig-1 [Convergence of neural network with dataset-1 with training function TRAINLM and learning and Transfer function LEARNGDM and LOGSIG respectively. The network is converged with 56 epochs and with MSE as .00666667.]

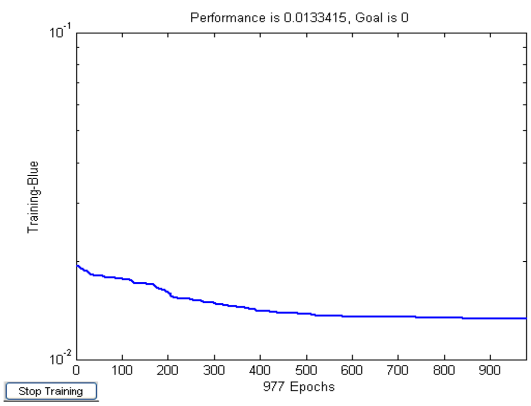

Fig-2 [Convergence of neural network with dataset-1 with training function TRAINSCG and learning and Transfer function LEARNGDM and LOGSIG respectively. The network is converged with 977 epochs and with MSE as .0133415.]

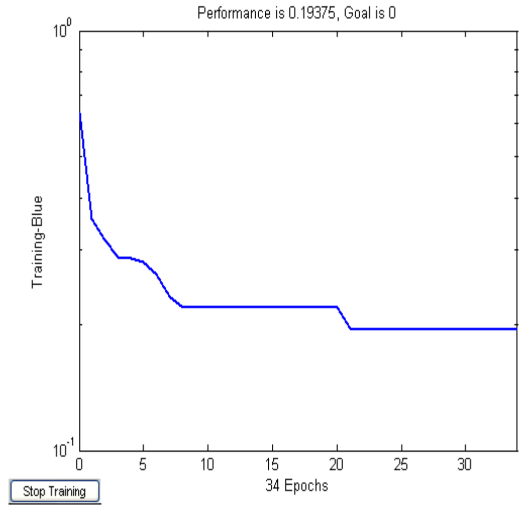

Fig-3 [Convergence of neural network with dataset-2 with training function TRAINLM and learning and Transfer function LEARNGDM and LOGSIG respectively. The network is converged with 34 epochs and with MSE as 0.19375.]

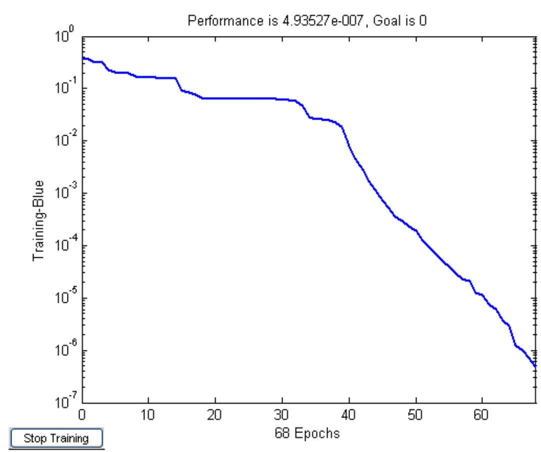

Fig-4 [Convergence of neural network with dataset-2 with training function TRAINSCG and learning and Transfer function LEARNGDM and LOGSIG respectively. The network is converged with 68 epochs and with MSE as 4.93526 e-007.]

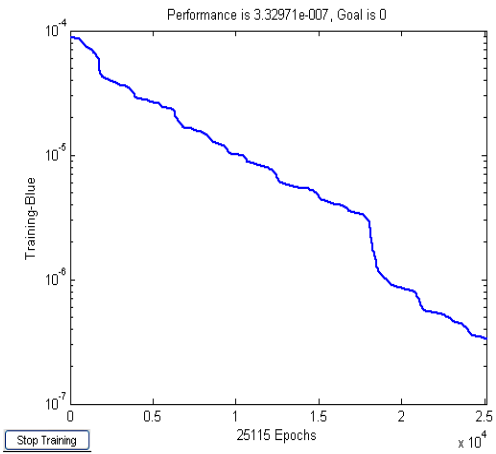

Fig-5 [Convergence of neural network with dataset-3 with training function TRAINSCG and learning and Transfer function LEARNGDM and LOGSIG respectively. The network is converged with 25115 epochs and with MSE as 3.32971 e-007.]

Data Set	Actual output	Simulated Output
1	0	0.0000
	0	0.0008
	0	0.0000
	1	1.0000
	0	0.0000
	1	1.0000
	1	1.0000
	0	0.0000
	1	1.0000
	0	0.0000
2	0	0.0000
	0	0.0000
	0	0.0000
	1	1.0000
	1	1.0000
	0	0.0000
	1	1.0000
	1	1.0000
3	0	0
	0	0.8629
	1	1
	1	1
	1	0.0064
	1	0.6824
	1	1

Table-3 represents the actual outputs vs. the outputs generated by our neural networks

V. CONCLUSION

The back propagation neural network can be used as a highly effective tool for dataset classification with appropriate combination of training, learning and transfer functions. t. In the proposed work we found that the combination of TRAINLM, LEARNGDM and LOGSIG works better for comparatively smaller datasets whereas the combination TRAINSCG LEARNGDM and LOGSIG is effective for larger datasets. However the efficiency of the neural network may be enhanced by using hybrid learning ,training and transfer functions which is the future scope of the paper.

REFERENCES

[1] Fabrizio De Nittis, Giampietro Tecchiolli, and Alessandro Zorat, “Consumer Loan Classification Using Artificial Neural Networks”, *ICSC EIS conference, held in Tenerife* ,(1998)

[2] Janusz Starzyk ,Jing Pang,” Evolvable Binary Artificial Neural Network for Data Classification”, *Evolutionary Computation* pp 5576-5783(2000)

[3] Guoqiang Peter Zhang, “Neural Networks for Classification: A Survey”, *IEEE Transaction on systems management and Cybernatics- Applications and Reviews Vol-30, No-4,*pp 451-462(2000).

[4]] M.C. Sezgin, Z. Dokur, T. Ölmez, M. Korürek,”Classification of Respiratory Sound by Using An Neural Network”, *Proceedings – 23rd Annual Conference – IEEE/EMBS*,(2001)

[5] Sean Van Osselaer, “Iris Data Analysis Using Back Propagation Neural Networks”, Journal of Manufacturing Systemsvolume 13, issue -4, Page 262(2003)

[6] R. Sivakumar,” Neural Network Based Diabetic Retinopathy Classification Using

Phase spectral Periodicity components”, *ICGST-BIME Journal, Volume 7, Issue 1*, (2007)

[7] Khaled Mohamed Almhdi, Paolo Valigi, Vidas Gulbinas, Rainer Westphal and Rainer Reuter “Classification with Artificial Neural Networks and Suport Vector Machines: Application To Oil Fluorescence “*EARSeL eProceedings 6* pp 115-129(2007)

[8]Waranyu Wongseree,Nachol Chaiaiyaratana a, Kanjana Vichittumaros , Pranee Winichagoon, Suthat Fucharoen, ”Thalassaemia classification by neural networks and genetic programming ” *Science Direct Information Sciences 177* pp 771–786 (2007)

[9] Nuryanti Mohd. Salleh, Harsa Amylia Mat Sakim and Nor Hayati Othman,” Neural Networks to Evaluate Morphological Features for Breast Cells Classification”,

IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.9, pp 51-58 (2008)