International Journal of Soft Computing and Engineering (IJSCE)
ISSN: 2231-2307, Volume-3, Issue-2, May 2013

Performance Analysis of Different Feed Forward Networks in Non-Linear Classification

Poornashankar

Abstract:-Artificial Neural Networks (ANN) are recognized extensively as a powerful tool for most of the research applications including classification of heterogeneous data using function approximators. Identifying better neural classifier architecturefor a given input data depends on many factors, including the complexity of theproblem, the training set, the number of weightsand biases in the network and the error goal. Feedforward networks frequently exercise classification techniques for complex non-linear data. This paper presents a comparative study of different type of Feedforward neural networks such as Simple Feedforward networks, Pattern recognition networks and Cascade forward networks in classifying the global carbon emissions data. In this study the percapita carbon emissions of several countries are classified into low, medium and high category. Levenberg-Marquardt learning algorithm is used to train these networks as it is the fastest and first choice supervised learning algorithm with less training errors. Hyperbolic tangent activation function is used in this study because of their massive interconnectivity and enhanced processing performance. Experimental results show that simple Feedforward network outperformed in less number of epochs with higher classification accuracy.

Keywords: Green House Gases (GHG), Feed Forward network, Pattern Recognition Network, Cascade forward network

I. INTRODUCTION

Observations are now evident that the change in climate leads to rise in global average sea level, global average air and ocean temperatures. Widespread melting of glaciers will create risk of flash flood and occurrence of extremely disastrous events.Rain-fed farms provideup to 80% of food in developing countries. Climate change induce rain patterns threaten food security.

Sea levels rose by about 17 centimeters during 20^{th}century putting coastal dwellers in jeopardy.Most of the observed increase in global averaged temperature is very likely due to increase in Green House Gases (GHG) concentrations [1].The energy balance of the climate system is affected by the major climate process drivers like aerosols, emissions and concentrations in GHG.

The latest analysis of observations from Global Atmosphere Watch (GAW) network shows that the globally averaged mixing ratios of carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O) reached new heights with CO_2 at 389 ppm in 2010 [2]. It is essential to identify and mitigate the factors that affect global warming metric values and future carbon emissions.

Manuscript received on May, 2013.

Poornashankar, MCA department, Indira College of Engineering & Management, Pune, India.

Artificial Neural Networks are promising alternative to various conventional classification methods due to its self-adaptive data driven methods. The universal approximator in neural networks can approximate any function with arbitrary accuracy [3].

Since classification procedure seeks a functional relationship between the group membership and the attributes of the objects, accurate identification of this underlying function is doubtlessly important. for the Artificial Neural Network undergoes a process of training in which previously recorded inputs are presented to the network and tuning is done to produce the desired target outputs. Different neural classifiers are available for dynamic and diverse data classificationin a systematic manner. The choice of neural architecture for classification problems depends upon the characteristics of the given input data. If more neurons are selected, they require more computations which may have a tendency to over-fit the data. This study is conducted to compare and analyze the performance attributes of Feedforward neural types and identify a better network that classifies a huge volume of global carbon emissions.

Multilayer perceptron networks with sigmoid activation functions are used in generating a model for methane emissions of solid waste landfills at Turkey [4]. Unsupervised learning network such as self-organizing map is also suitable to solve complex climate classification problem which empirically based and often mix the mutual impact between climate, soil and vegetation [5].The ensembleneural network can be easily developed to performmulti-class classification problems without increasingthe calculation complexity in forecasting the values of the temperature, windspeed and humidity for the four seasons. Multilayer perceptrons and regression networks are used in the ensemble model in the research work done by Imran Maqsoodet. al [6].

Feedforward network models are broadly used in hydrological problems and climate change projections such as classifying the areal extent of snow cover in western United States by James J. Simpson and Timothy J. McIntire [7].In a study done by I.S Isa et.al [8], Levernberg-Marquardt Training algorithm with hyperbolic tangent function is proved as efficient classifier model for weather classification. Energy consumption predictor for greenhouses from a MLP neural network was developed by Mario Trejo Perea et al. [9] using real data obtained from a greenhouse located at the Queretaro State University, Mexico. The results showed that the selected ANN model gave a better estimation of energy consumption with a 95% significant level. But very few research applications are available on classification of climate data and carbon emissions through Neural Networks.

This paper is organized as follows. Section II describes the different types of feed forward networks and its architectural details that are used in this research. Section III presents the experiments conducted using various classifiers and their results. Section IV describes the discussions and conclusion.

II. METHODOLOGY

The layered Feedforward networks with nonlinear transfer functions can be used in function approximation and patter recognition. Hence an attempt is made to compare the classification efficiency of different types of Feedforward network.

A. Simple feedforward network

Feedforward neural network has an interconnection of processing elements with layered architecture operating in parallel. The network function is determined largely by the connections between elements. The training is imparted to neural network to perform a particular function by adjusting the values of the connections (weights) between elements

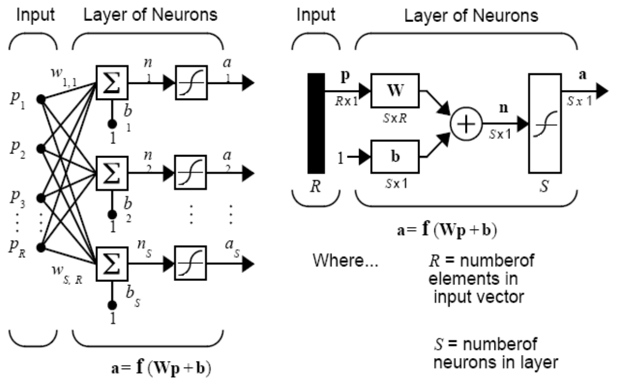

Figure1: A Simple Feedforward Network

The weights (w) and biases (b) together constitute the adaptive parameters in the network for the given input vector p. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. The linear output layer lets the network to produce values outside the range from –1 to +1.The goal of the network is to train and achieve a balance between the ability to respond correctly to the input patterns and provide good response.

B. Pattern recognition network

In addition to functional approximation Feedforward networks are also good at recognizing patterns. The task performed by a network trained to respondwhen an input vector close to a learned vector is presented. The network“recognizes” the input as one of the original target vectors. Pattern recognition networks can be trained to classify inputs according to target classes. Such network classifies target vectors having N elements.

The target data for pattern recognition networks should consist of vectors of all zero values except for a 1 in element i, where i is the class they are to represent.By default this network uses hyperbolic tangent functions in hidden and output layers.When an input vector of the appropriate category is applied to thenetwork, the correspondingneuron should produce a 1, and the otherneurons should output a 0.

C. Cascade forward network

Cascade-forward networks are similar to Feedforward networks, but include a weight connection from the input to each layer and from each layer to the successive layers. As with feed-forward networks, a two-or more layer cascade-network can learn any finite input-output relationship arbitrarily for the given hidden neurons.The additional connections might improve the speed at which the network learns the desired relationship.

The first layer has weights coming from the input. Each subsequent layer hasweights coming from the input and all previous layers. All layers have biases.The last layer is the network output.Each layer’s weights and biases are initialized in the network.Adaption is done with hyperbolic tangent functions which update weights with the specifiedlearning function. Performance is measured according to the specified performance function.

D. Levenberg-marquardt (lm) training

The network training was imparted with Levenberg-Marquardt(LM) non-linear optimization algorithm. It is reputably the fastest back propagation algorithm with a combination of steepest descent and the Gauss Newton method [8]. When the current solution is far from the correct one, the algorithm behaves like a steepest descent method: slow, but guaranteed to converge. When the current solution is close to the correct solution it becomes Gauss-Newton method. Thus it continuously switches its approach and can make very rapid progress. At each iteration in the learning process, the weight vector w will be updated as follows:

$$W_{k+1} = W_k + d_k \qquad (1)$$

$$d_k = -[J^T J + \mu I]^{-1} J^T \varsigma \qquad (2)$$

where, d_k is search direction, μ is damping parameter of k-th iteration, ς is a vector of network errors and J is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights. When the scalar μ is zero, this is just Newton’s method, using the approximate Hessian matrix. When μ is large, this becomes gradient descent with a small step size. Newton’s method is faster and more accurate near an error minimum, so the aim is to shift towards Newton’s method as quickly as possible. Thus, μ is decreased after each successful step and is increased only when a tentative step would increase the performance function.

E. Activation functions

A transfer function provides a means of further processing the output of a neuron after the initial processing has taken place. They are the non-linear functions that transform the weighted sum of the inputs to an output value and do the final mapping of the activations of the output neurons into the network outputs [10].

The hyperbolic tangent function is applied to each neuron in the layer. This will squash the range of each neuron in the layer to between -1 and 1. Such nonlinear elements provide a network with the ability to make soft decisions. The Tanh function will typically be used as hidden and output layers in MLP topologies. If used in the output layer, it is important to verify that the desired signal is normalized to between -1 and 1. f(x) = tanh(x) where tanh(x) calculates its output according to:

$$\tanh(x) = 2/((1+e^{(-2x)}) \qquad (3)$$

But the outputs from a single cycle of operation of a neural network may not be the final outputs; since more number of iteration is required till the network attains convergence

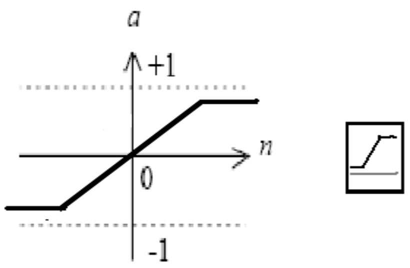

Figure 2: Hyperbolic tangent transfer function.

III. EXPERIMENTS AND RESULTS

Carbon emissions from fossil fuel and other industrial processes were calculated by Carbon Dioxide Information Analysis Center, Environmental Sciences Division, Oak Ridge National Laboratory, Tennessee, and United States. Percapita carbon emissions of 183 countries are obtained from 1981 to 2007. All emission estimates are expressed in metric tons of carbon.

Carbon emission data of 183 countries for the period of 27 years are classified into a set of target categories as low, medium and high groups depending on the percapita emissions. The target is quantized such that the countries with standard emission rate less than 3 metric tons are classified into low emitting group and standard emission rate between 3 and 5 metric tons are classified as medium emitting group and above 5 metric tons are grouped as high emitting countries.

Before training it is often useful to normalize the inputs and the targets so that they always fall within a specified range. Two preprocessing strategies are used in this experiment. One is for pattern normalization to process the inputs by normalizing the minimum and maximum values of each row to a specified range and the other strategy is to remove rows with constant values from input and target data as they cause numerical problems for some algorithms.

Various combinations of ANN structures were investigated to get the optimum output. Finally athreelayer Feedforward network with hyperbolic activation function and Levenberg-Marquardt training algorithm is constructed with three hidden layers and three output layers. 27 processing elements were used in the input. Same architecture is used for pattern recognition network and cascade forward network.

Results

The performance measures and outcome of all three experiments are depicted below

Table 1: Performance measures of three network models

Classifier Models	No. of Epochs	Training	Testing	Validation
Simple Feedforward Network (SFN)	20	0.0056	0.0191	0.0031
Pattern Recognition Network (PRN)	50	0.0329	0.0373	0.0265
Cascade Forward Network (CFN)	25	0.0057	0.0225	0.012

The number of epochs required to train the network was very less for simple Feedforward network when compared to other types. The classification accuracy is high and error levels are low the simple Feedforward network. The following graph depicts the performance measures of all three experiments

From the above graph it is evident that simple Feedforward network and cascade forward network are performing better with respect to training, testing and validation respectively, as they are using function approximation techniques.

In Feedforward networks training set is used to define an error function in terms of the discrepancy between the predictions of the network, for given inputs and the desired values of the outputs given by the training set. The learning process then involves adjusting the values of the parameters to minimize the value of the error function. The Mean Squared Error is computed as the squared difference between desired and actual output, summed over all outputs and summed over all patterns in the training set. Percentage error indicates the fraction of samples which are misclassified. Simple Feedforward network and Cascade networks exhibit very low error levels.

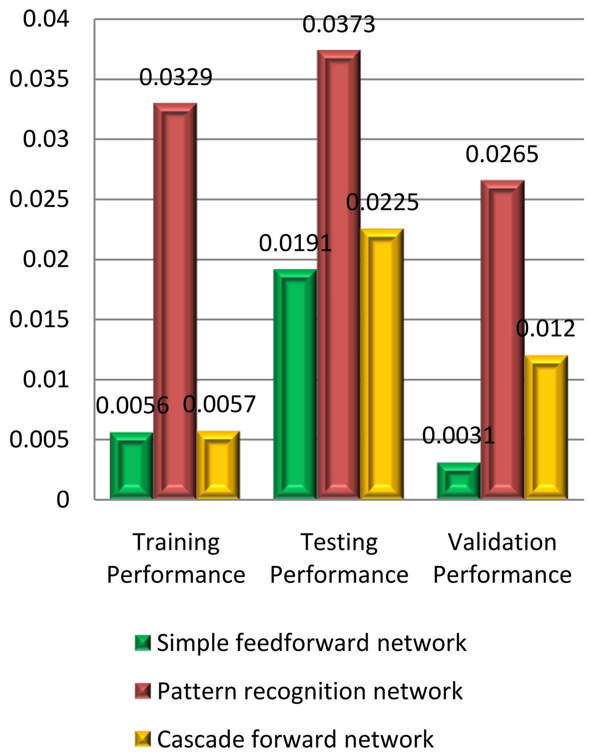

Figure 3: Performance measures of classifier models

Table 2: Error measures of three network models

Classifier Models	Mean Squared Error	Percent Error
Simple Feedforward Network(SFN)	0.0073	0.5464
Pattern Recognition Network(PRN)	0.0264	0.8784

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Cascade Forward Network(CFN)	0.0119	0.6245

Following graph depicts the error measures of all three classifiers

Table 3: Classification results of Countries using three network models

Classifier Models	Low Emission		Medium Emission		High Emission	
	Classified	Misclassified	Classified	Misclassified	Classified	Misclassified
SFN	99(100%)	0	21(95.5%)	1(4.5%)	62(100%)	0
PRN	97(99.0%)	1(1.0%)	20(87.0%)	3(13.0%)	62(100%)	0
CFN	99(100%)	0	19(95.0%)	1(5.0%)	62(96.9%)	2(3.1%)

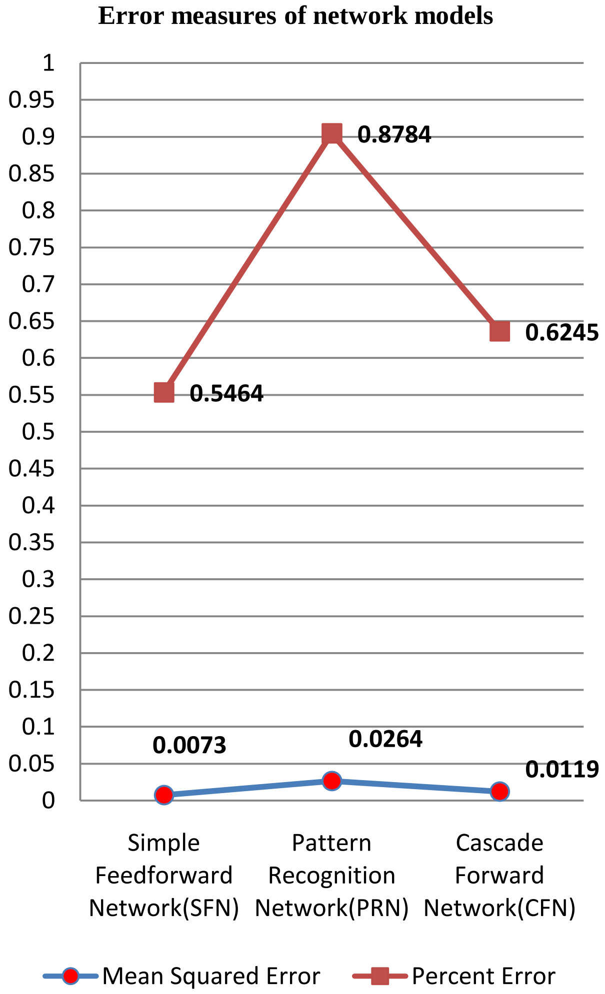

Figure 4: Error measures of classifier models

The performance of all networks is evaluated with respect to its classification rate, which is calculated as follows.

$$\text{Classification rate} = \frac{CorrectClassification}{ActualClass} * 100 \qquad (3)$$

The results obtained from the confusion matrix and the classification of countries according to standard emissions isdepicted below.

The following graph illustrates the accuracy of classification of all three types of Feedforward classification networks. The simple Feedforward network surpassed with 99.5% accuracy. The summary of results is depicted in the following table.

Table 4: Comparison results of three classifier's accuracy

Classifier Network Model	Classified	Misclassified
SimpleFeedforward Network(SFN)	99.5%	0.5%

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Pattern Recognition Network(PRN)	97.8%	2.2%
Cascade Forward Network(CFN)	98.4%	1.6%

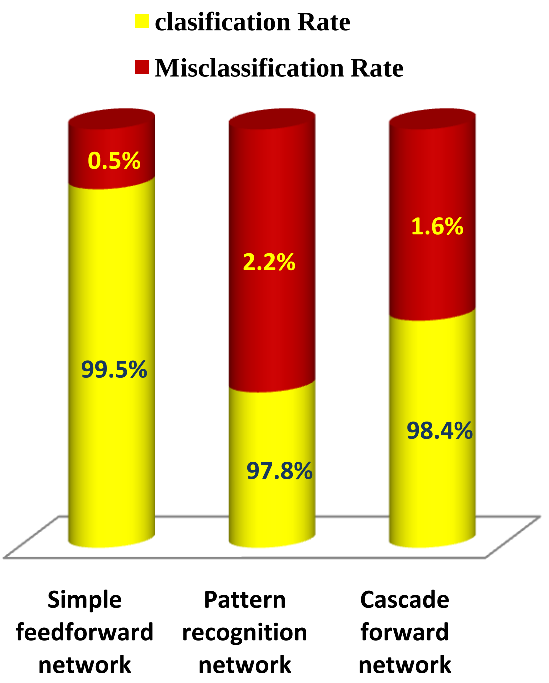

Figure 5: Classification Results

IV. DISCUSSION

The main objective of this study is to compare the performance of function approximation and pattern recognition networks. Simple Feedforward network has the highest accuracy and swift convergence. The time taken to adjust the weights and reduce the errors is also very minimal when compare to other network classifiers. The number of epochs required by the network was only 20. The cascade forward network took 25 epochs to train and adjust the weights. As the interconnected layers have to be assigned biases it increased the number of iterations. The performance measures and error measures are optimum in both the networks.

The performance of Levernberg-Marquardt algorithm is relatively poor on Pattern recognition networks due to which the time taken for network convergence has increased. They are more suitable for image recognition and identification problems as they infer an underlying regularity in input patterns which can subsequently be used to solve new instances of the problem. Function approximator network models are more suitable for classification of real-time complex numeric data.Nevertheless the classification accuracy levels of all three network models are above 97% which reveals the suitability of Feedforward networks in classification problems.

There are several algorithm characteristics that we can deduce from theabove experiments. In general, on function approximationproblems, for networks that contain up to a few hundred weights, theLevenberg-Marquardt algorithm will have the fastest convergence. Thisadvantage is especially noticeable if very accurate training is required.The performance of the various algorithms can be affected by the accuracyrequired of the approximation. Hyperbolic tangent functions improve the performance of the network to solve classification problems with less number of hidden nodes.

V. CONCLUSION

In this paper, Performance analysis of Function approximation networks such as Feedforward and Cascade forward networks and Pattern recognition networks were done to solve classification problem. All classifiers achieved a reasonable performance with minimum accuracy rate of 97%. The best result was obtained from simple Feedforward network using Levernberg-Marquardt learning algorithm with hyperbolic tangent activation function. The network was stabilized early in 20 epochs attaining 95.5% classification accuracy.Only one country was misclassified as medium category. The error measures and training time was relatively low in simple Feedforward network. Due to interconnected biases in all layers Cascade forward network got delayed. The results illustrates that 99 countries are in low emitting category, 21 countries are in medium emitting category and 62 countries are in high emitting category. Integrated International level increased adaptation of mitigation policies in all sectors and improving the efficiency of our economy and moving to renewable energy will reduce future vulnerability globally.

REFERENCES

1. Thomas Stocker , University of Bern, Switzerland and Gian-Kasper Plattner, The Physical Science Basis of Climate Change:- *IPCC Working Group I: University of Bern presented in Copenhagen Climate summit, 9 December 2009.*
2. World Meteorological Organization (WMO), *Atmospheric Environment Research Division, Geneva, Vol. No.7, November 2011.*
3. Guoqiang Peter Zhang, Neural Networks for Classifcaiton: A survey, *IEEE Transactions on systems,man and cybernetics-Applications and Reviews, Vol.30,No.4, Nov 2000.*
4. H KurtulusOzcan, Osman N Ucan, UlkuSahin, Mehmet Borat , Artificial neural network modeling of methane emissions at Istanbul Kemerburgaz-Odayeri landfill site, *Journal of Scientific & Industrial Research, Vol. 65, February 2006, pp. 128-134.*
5. Kropp, J. A neural solution: a data driven assessment of global climate and vegetation classes,*Neural Information Processing, 1999. Proceedings. ICONIP '99. 6th International Conference ,Digital Object Identifier: 10.1109/ICONIP.1999.844000*
6. Imran Maqsood, Muhammad Riaz Khan, Ajith Abraham, An ensemble of neural networks for weather forecasting, *Neural Computing & Applications, pp 112-122, DOI 10.1007/s00521-004-0413-4*
7. James J. Simpson and Timothy J. McIntire, "A Recurrent Neural Network Classifier for Improved Retrievals of Areal Extent of Snow Cover," *IEEE transactions on Geosciences And Remote Sensing, Vol. 39, No. 10, October 2001*
8. I.S. Isa, S. Omar, Z. Saad, M.K. Osman, Performance Comparison of Different Multilayer Perceptron Network Activation Functions in Automated Weather Classifcation, *2010 Fourth Asia International Conference on Mathematical/Analytical* Modeling and Computer Simulation, Digital Object Identifier: 10.1109/AMS.2010.27

9. Mario Trejo-Perea1, Gilberto Herrera-Ruiz, Jose Rios-Moreno, Rodrigo Castañeda Miranda And Edgar Rivasaraiza, "Greenhouse Energy Consumption prediction using neural network models," *International journal of agriculture & biology, Vol. 11, No. 1, 2009.*
10. Soria D, Garibaldi J.M, Biganzoli E, Ellis I.O, A Comparison of Three Different Methods for classification of Breast Cancer Data, Machine Learning and Applications, 2008,*ICMLA '08, Seventh International Conference, Digital Object Identifier: 10.1109/ICMLA.2008.97*

Dr. Poornashankarreceived MCA degree from University of Madras, Chennai and M.Phil degree from BharatiVidyapeeth Deemed University, Pune, and Ph.D from S.R.T.M University, Nanded, India. She has more than 15 years of experience in handling post graduate courses in Computer Science & Applications. Her research interests are in the area of Artificial Intelligence, Computer Architecture, Data mining, Artificial Neural Networks and Cloud Computing. She is presently serving as a Professor in MCA department of Indira College of Engineering and Management, Pune, India.