Classification of Global Carbon Emissions using Artificial Neural Networks

Poornashankar
College of Computer Science, Latur
Affiliated to S.R.T.M. University, Nanded
Maharashtra, INDIA.

Vrushsen P. Pawar
Associate Professor, S.R.T.M. University,
Nanded, Maharashtra, INDIA

ABSTRACT
Artificial Neural Networks (ANN) are good at recognizing patterns and proven themselves as proficient classifiers for addressing problems that are non-linear in nature which belong to the real world phenomena. The greatest environmental challenge on the earth is to mitigate Global Warming. Carbon dioxide is the most anthropogenic Green House Gas in the atmosphere which is growing rapidly since three decades and decreasing the global energy. This research paper applies classification techniques for global carbon emissions using ANN by grouping the countries based on the quantum of carbon emissions.

The global percapita carbon emissions of 183 countries are classified using Generalized Feedforward Networks (GFF) based on the emission rate into three categories namely - low, medium and high. The low carbon emitting countries sharing complex boundaries are further categorized using Support Vector Machines (SVM) with Radial Basis Function (RBF) kernel. It is found that the GFF training performance was exemplary with the classification rate of 0.9950 with testing error rate of 0.0191. SVM classifiers mapped the non-linear input feature space into high dimensional space by constructing an optimal hyper-plane with the classification rate of 0.9796. Various performance measures of experiments and accuracy of classification in grouping countries based on the emission rate are discussed.

Key words
Artificial Neural Networks (ANN), Generalized Feedforward (GFF), Green House Gases (GHG), Multilayer Perceptron (MLP), Support Vector Machines (SVM).

1. INTRODUCTION
Industrial Revolution has cascading impact on the environment due to the release of heavy quantity of pollutants leading to global warming, climatic changes and health hazards. Technological and Engineering advancements clubbed with higher growth rate in the developing countries has brought heavy reliance on fossil fuels which is leading to environmental degradation . Fossil fuels are extensively used as a prime energy source by power sectors due to which atmospheric accumulation of carbon dioxide has increased [1]. The largest points of sources of carbon dioxide emissions are from the fossil fuels, coal mining and oil & gas based thermal plants. Research Initiatives are taken up world-wide to reduce the Carbon dioxide Emissions and consequently mitigate the global climate changes. Most popular research applications implements Artificial Neural Networks (ANN) extensively to perform accurate and simple data classification in heterogeneous databases. It facilitates dynamic non-linear state space through function approximations in constructing a classification model with optimum conversion rate [2]. ANN offer efficient modeling to solve complex problems in which there may be hundreds of predictor variables that may have many interactions. As this data relates to carbon emissions of 80% of the countries, ANN classification techniques are selected to classify them. The objective of this study is to categorize the data relating to global carbon emissions depending on the quantum of per capita emissions and classify them using ANN. The training performance and testing error rates during classification are also discussed.

1.1. Classification Using ANN
In classification problem a set of $P_j = \{x_{j1}, x_{j2}, \ldots x_{jn}\}$ unlabelled patterns are given where j is the pattern number and that has to be assigned to one of k possible classes from a pre-identified set $C=\{c_1, c_2, \ldots c_k\}$ of classes. The input set may have any number of patterns having no control over the size [3]. In ANN classification of an input pattern in relation to stored patterns is attempted, and if unsuccessful, a new stored classification is generated. It classifies input vectors on the basis of a set of stored or reference vectors. ANNs are widely used in data classification [4], text classification [5] pattern classification and image classification [2][6]. Two phase ANNs are used to recognize three commonly faced environmental noise signals: highway, subway and airport. Even though the recognition of these noises is difficult due to their similar acoustical features, by using ANN 89% accuracy rates are achieved by two-phase environment noise classification [3].

1.2. Carbon Emission Trends
Carbon dioxide concentration in atmosphere from anthropogenic sources has increased considerably worldwide recently. Carbon dioxide emissions are those stemming from the burning of fossil fuels and the manufacture of cement [1]. They include carbon dioxide produced during consumption of solid, liquid, and gas fuels and gas flaring According to the latest statistical reports from CDIAC (Carbon dioxide Information Analysis Centre), China is leading in emissions having 19 lakh tons per year and USA is in the second position having 15 lakh tons. India is holding the third position in emissions having 4.7 lakh tons. Russia and Japan are in the fourth and fifth position in emissions having 4.3 lakh tons and 3.5 lakh tons respectively.

1.3. Climatological Research Using ANN

ANN models are broadly used in hydrological problems and climate change projections. James J. Simpson and Timothy J. McIntire used Feedforward Networks and Recurrent networks for classifying the areal extent of snow cover in western United States [7]. Satellite images of clouds are classified by Bin Tian, Mukhtiar A. Shaikh, using Probabilistic Neural Networks (PNN) and Kohonen Self Organized Maps(SOM) [8] . 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 selection procedure of appropriate ANN model for this study. Section III portrays experiment process like architecture of Neural Network and training procedures. Section IV depicts the experimental results. Section V discusses the performance measures and accuracy levels of classification. Finally, conclusions are drawn in section VI.

2. ANN MODEL SELECTION

Before selecting the Generalized Feedforward model popular classification techniques are studied like Generalized Regression Neural Networks (GRNN), Linear Vector Quantization (LVQ) and Radial Basis Function Networks (RBF). In theory each one of these methods has disadvantages. The main drawback of GRNN is that, they suffer seriously from the curse of dimensionality like kernel methods. It can cause bad performance when the data is non-spherical distribution and especially contains noises or outliers.LVQ suffers from the prototype under-utilization problem, (i.e.) only the winner is updated for each input and because of adopting Euclidean distance measure for linear data. Radial Basis Function (RBF) networks represent any non-linear functions as a sum of weighted radial basis functions and received significant attention most commonly with a Gaussian form that can be used to approximate functions with a smaller dataset [10]. As the input data set used is high dimensional and non-linear in nature a swift and efficient model like Multilayer Feedforward Networks are used in this experiment. It updates the weights and reduces the error value for non-linear optimization based on incremental gradient descent of mean-square error [11].

2.1. Support Vector Machines

The foundations of SVM have been developed by Vapnik and are gaining popularity in the field of machine learning due to many attractive features and promising empirical performance [12]. SVM method does not suffer the limitations of data dimensionality and limited samples. SVM uses a kernel function to transform the input data into high-dimensional space using RBF network that places a Gaussian at each data sample in which it searches for a separating hyperplane [13]. Thus the feature space has large margin classifiers for training. This decouples the capacity of the classifier from the input space and at the same time provides an ideal combination for classification. RBF Networks represents any non-linear functions as a sum of weighted radial basis functions. RBF Networks have weights of unity between input and hidden layers. In other words, it is not necessary to learn the weights.

Specifically, RBFN employs the Gaussian distribution as the radial basis function such as

$$a_j(x) = \exp(-\frac{|x-\mu_j|}{2\sigma_j^2}) \qquad (1)$$

where

a_j : output of the j-th node at the hidden layer
x : input vectors.
μ_j : center vector of the j-th node at the hidden layer
σ_j : width of the j-th cluster

Classical techniques utilizing RBF employ some method of determining a subset of centers. An attractive feature of the SVM is that this selection is implicit, with each support vectors contributing one local Gaussian functions, and centered at that data point [14]. More than 50% of the low carbon emitting countries shared complex boundaries. To normalize the low emission data and to separate lowest and low data inputs which fall close to the data boundaries in the hyperplane SVM technique is adopted in this study. MLP kernel function is used to map the training data into kernel space.

3. EXPERIMENT

3.1. Data Set & Source

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 for the period 1981 to 2007 of 183 countries are obtained. All emission estimates are expressed in metric tons of carbon.

3.2. Research Technique

The research is carried out with technical computing tool Matlab 7.11 with Neural Toolbox. 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, 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 group. As the low emitting countries contribute more than half of the input data and they share complex margins, they are further grouped using Support Vector Machines (SVM). The training performance, error rates and classification rates are discussed in Section IV.

3.3. Network Architecture & Training

A two layer Feedforward network with sigmoid activation function is constructed with three hidden layers and three output layers. The percapita emissions of 183 countries for 27 years are given as input to the network.

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.

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. 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. In each iteration in the learning process, the weight vector *w* is updated as follows:

$$W_{k+1} = w_k + d_k \qquad (2)$$

$$d_k = -[J^T J + \mu I]^{-1} J^T \varsigma \qquad (3)$$

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. Activation function has important role in training stage.

70% of the input data set is presented to the network during training, so that the network can be adjusted according to its error. 15% data are used to measure the network generalization and to halt the network when generalization stops improving. Remaining data performs testing of an independent measure of network performance during and after training. The training stops when a classifier gives a higher accuracy value with minimum training and testing errors.

4. EXPERIMENTAL RESULTS

4.1. Performance

The LM training algorithm outperformed in this experiment by classifying the high dimensional data in 20 epochs with the average training time of 0.2 seconds. The performance measures and outcome of the network are depicted below.

Table 1. GFF Classifier Performance Measures

Number of Epochs	20
Training Performance	0.0056
Testing Performance	0.0191
Validation Performance	0.0031
Classification Rate	0.9950
Mean Squared Error	0.0073
Percent Error	0.5464

The error measures like Mean Squared Error (MSE) and Percent Error (PE) are recorded. MSE is the mean of the squared error between the desired output and the actual output of the neural network.

The MSE is computed as follows.

$$MSE = \frac{\sum_{j=0}^{P}\sum_{i=0}^{N}(d_{ij} - y_{ij})^2}{N \quad P} \qquad (4)$$

where P = number of output processing elements
N = number of exemplars in the training data set
y_{ij} = estimated network emissions output for exemplar i at processing element j
d_{ij} = actual output for emissions exemplar i at processing element j

In this experiment the obtained MSE value is 0.0073 which was attained at 14th epoch. Percent error is calculated as error percent between the desired output and the actual output obtained from the network. Percent Error indicates the fraction of samples which are misclassified.

$$\%Error = \frac{100}{N \quad P}\sum_{j=0}^{P}\sum_{i=0}^{N}\frac{|dy_{ij} - dd_{ij}|}{dd_{ij}}$$

where P = number of output processing elements
N = number of exemplars in the training data set
dy_{ij} = denormalized network emissions output for exemplar i at processing element j
dd_{ij}= denormalized desired network emissions output for exemplar i at processing element j

The Percent Error obtained in this experiment is 0.5464

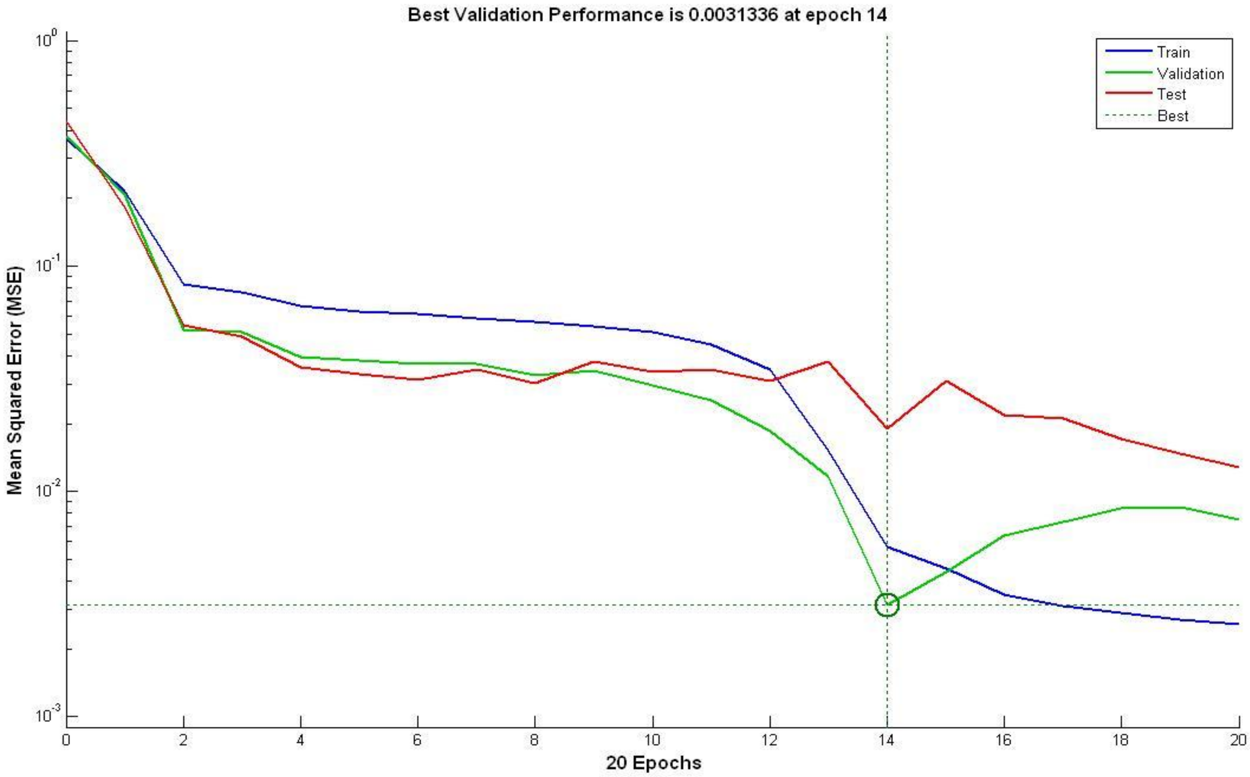

Fig. 1: Performance Graph

The Error histogram depicts the training errors, testing errors and validation errors of the GFF network.

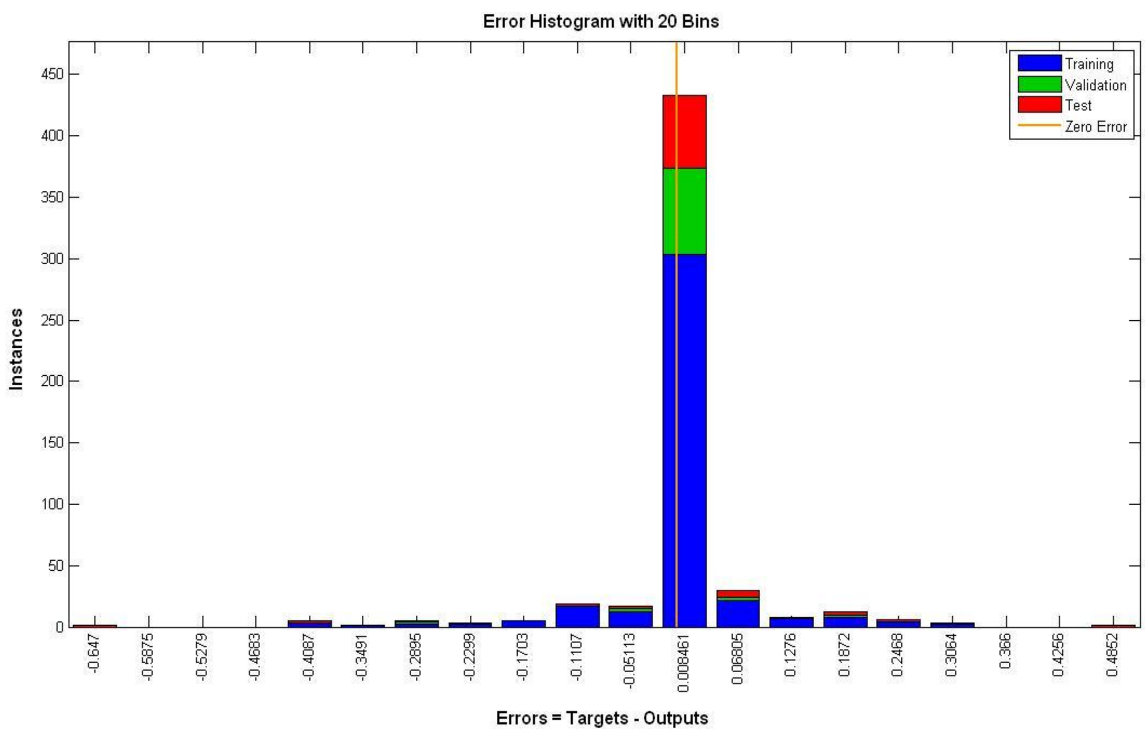

Fig 2: Error Histogram

4.2. Confusion Matrix

The confusion matrix depicts the accuracy of the classification problem. 183 countries are classified into three categories with 99.5% accuracy. The diagonal elements in the confusion matrix depict the classified groups. The matrix result illustrates that, Low emitting countries are 100 and Medium emitting countries are 21 and High emitting countries are 61. As one country in the medium group having complex margin could not be classified, the accuracy level has decreased by 0.5%.

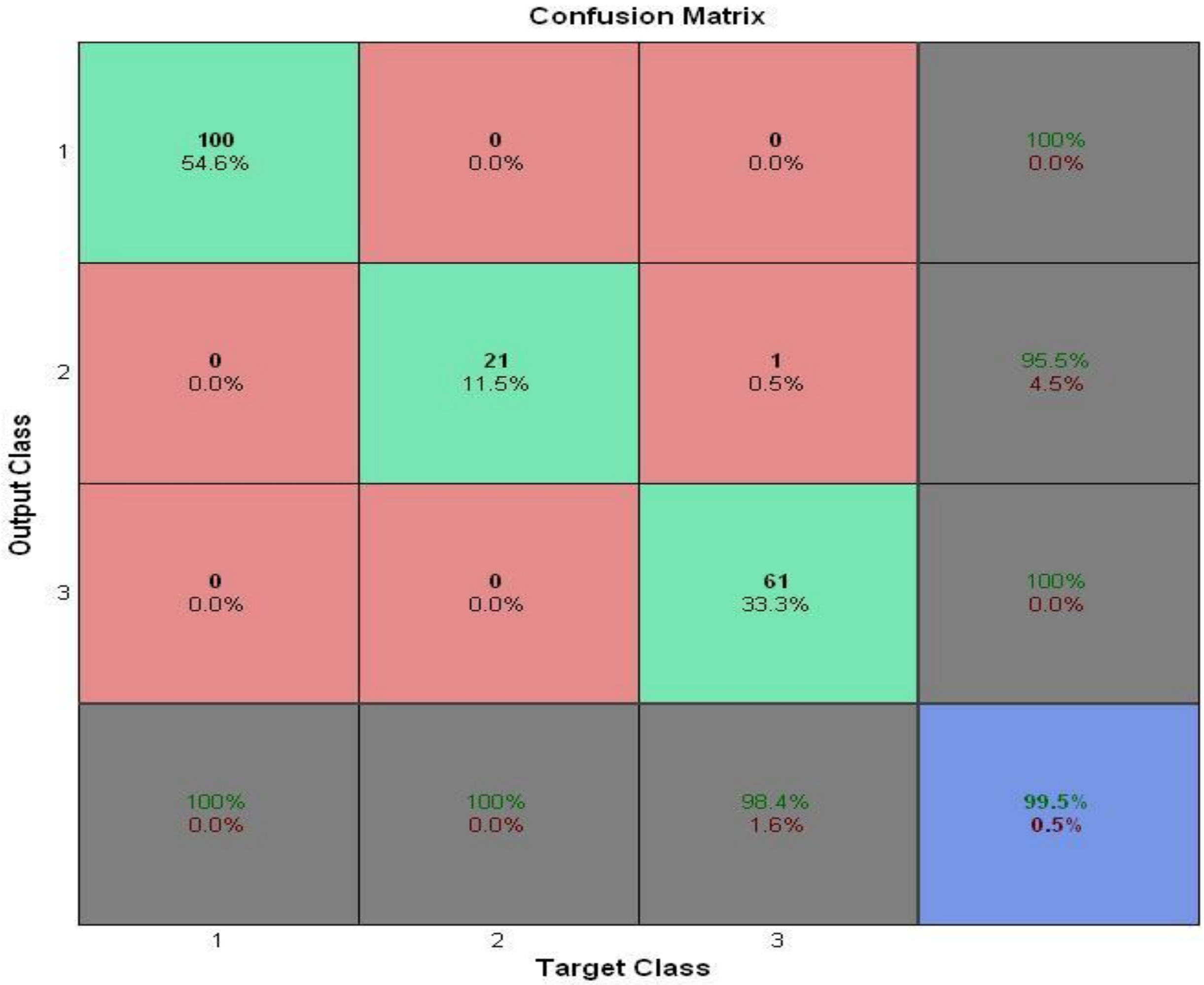

Fig3: Confusion Matrix

4.3. Receiver Operator Characteristic(ROC)

Classification accuracy is usually measured by Receiver Operator Characteristic (ROC) curve which shows the relationship between false positives and true positives. For each threshold, two values are calculated, the True Positive Ratio (the number of outputs greater or equal to the threshold, divided by the number of one targets), and the False Positive Ratio (the number of outputs less than the threshold, divided by the number of zero targets).The following figure plots the receiver operating characteristic for each output class. The more each curve cuddles the left and top edges of the plot, the better the classification.

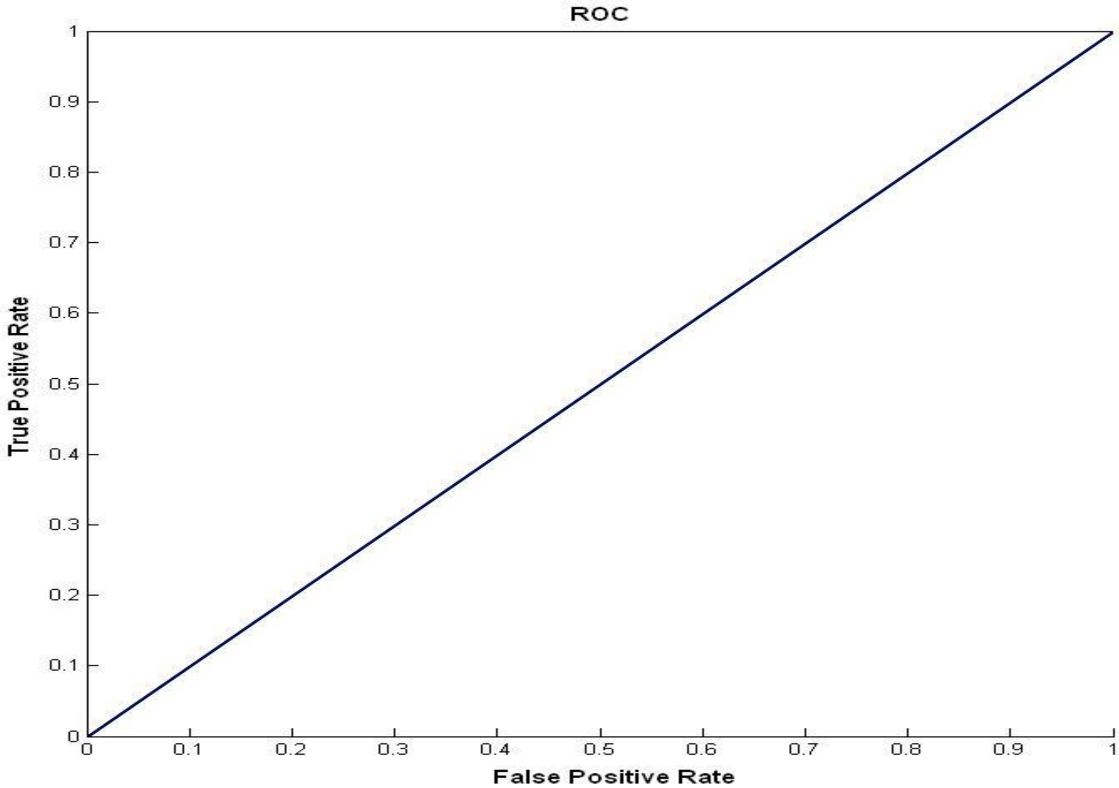

Fig 4: ROC curve

4.4. SVM Results

The low emitting countries are 100 and they share very complex boundaries. Hence SVM classifier is used to further categorize the group into very low emitting and low emitting with a target of standard emission rate less than 1.5 and less than 3 respectively. The SVM structure was constructed with 100 x 27 matrix having 100 countries carbon emission for 27 years with RBF kernel function. The outcome of the experiment separated 61 countries as very low emitting countries and remaining 39 countries in low group with the classification rate of 0.9796.

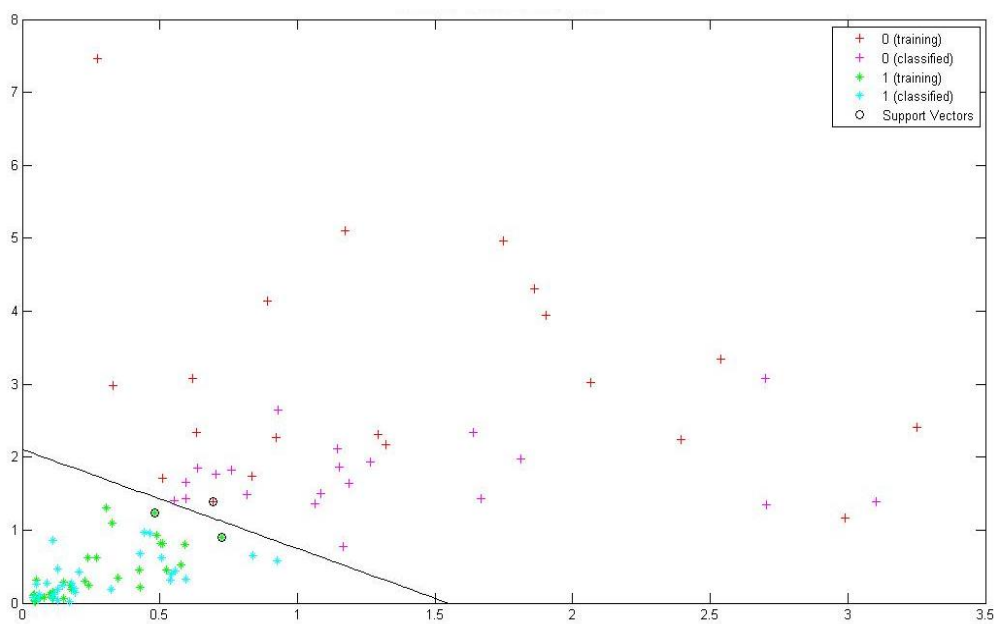

Fig 5: Low Emitting Countries Classification

5. FINDINGS

According to the results GFF has classified the high dimensional emission data with 99.5% accuracy yielding the best results with less number of epochs. The average training time was also low with least MSE value as low as 0.0073.On the other hand, when SVM was implemented for low emitting countries, the classification rate was 0.9796. However SVM are the most appropriate networks for data having closer boundaries. Based on the results obtained, the classification of countries and its percapita carbon emissions are depicted below.

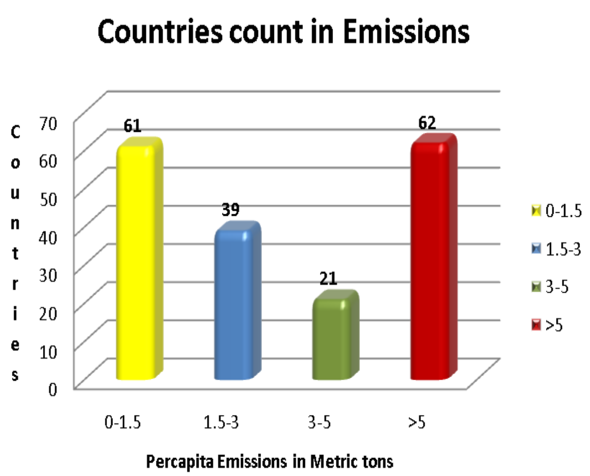

Fig 6: Range of Percapita Emissions

Around 34% of the countries form the high emitting group; 12% form the medium emitting group; 21% form the low emitting group and 33% form the very low group. Qatar is leading in per capita carbon emission with 46.5396 Metric tons and Chad is the least per capita emitting country with 0.0263 Metric tons.

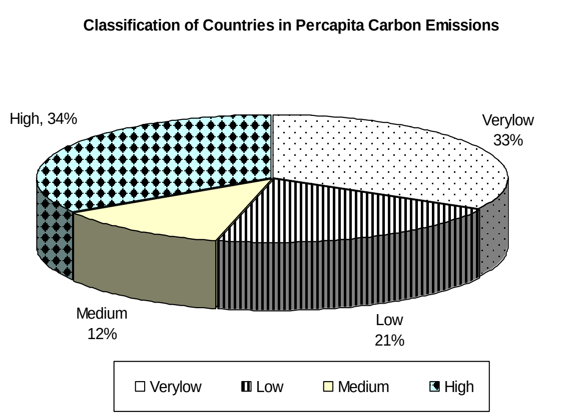

Fig 7: Classification based on Percapita Emission

This study demonstrated the efficiency and accuracy of Artificial Neural Networks in classifying high dimensional non-linear data.

6. CONCLUSION

In this study three important attributes are dealt with classifiers: accuracy, the ability to train using large datasets, and the speed at which classifiers are trained. This research has presented two classifier models using ANN. These models have been applied to classification problems. As it has been demonstrated experimentally, the Generalized Feedforward network proved the efficiency and accuracy with the highest classification rate of 99.5% and SVM with the classification rate of 97.96% further separated the low emitting countries data in a hyperplane by maximizing the margin between the classes. About 34% of the countries are emitting high carbon dioxide and 11% of the countries are releasing medium carbons. The impact of all emissions and climate change can be reduced, delayed or avoided by implementing significant mitigation policies in all sectors. The widespread diffusion of low-carbon technologies may take many decades. If early investments in these technologies are made attractive, the risks and exposure to vulnerability from extreme climate change can be managed for a radiant future. However, in the real-time applications this study recommends using GFF due to the low response time and best training performance with lesser error rate yielding good classification accuracy.

7. REFERENCES

[1] B.Kumar, “Carbon capture and storage initiatives for sustainable energy future,” Page no 3, International Conference on climate change perspectives and projections A systems Approach, December 2010

[2] K.Umamaheswaril, S. Sumathil, S.N. Sivanandaml and K.K.N. Anburajan “Efficient Finger Print Image Classification and Recognition using Neural Network Data Mining,” IEEE - ICSCN 2007, MIT Campus, Anna University, Chennai, India. Feb. 22-24, 2007. pp. 426-432.

[3] Buket D. Barkana, Inci Saricicek, “Environmental Noise Source Classification Using Neural Networks,” 978-0-7695-3984-3/10 © 2010 IEEE DOI 10.1109/ITNG.2010.118, 2010 seventh International Conference on Information Technology.

[4] Oswaldo Ludwig and Urbano Nunes, “Novel Maximum-Margin Training Algorithms for Supervised Neural Networks,” IEEE transactions on Neural Networks, vol. 21, no. 6,June 2010

[5] Ram Dayal Goyal, “Knowledge based Text Classification,” 0-7695-3032-X/07 © 2007 IEEE, 2007, IEEE International Conference on Granular Computing.

[6] K. Sookhanaphibarn, T. Raicharoen and C. LursSinsap, “A Supervised Neural Network Approach to Invariant Image Recognition,” 0-7803-8653-1/04/$20.00 Q 2004 IEEE, 2004, 8th International Conference on Control, Automation, Robotics and Vision, Kunming, China, 68th December 2004.

[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] Bin Tian, Mukhtiar A. Shaikh, Mahmood R. Azimi-Sadjadi, Senior Member, IEEE, Thomas H. Vonder Haar, and Donald L. Reinke, " A Study of Cloud Classification with Neural Networks Using Spectral and Textural Features**,"** IEEE transactions on Neural Networks, vol. 10, no. 1, January 1999

[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] Essarn Al-Daoud, "A Compraision between three Neural Network Models for Classification problems," Journal of Artificial Intelligence 2(2), 56-64, 2009.

[11] D. Ben Ayed Mezghani, S. Zribi Boujelbene, N. Ellouze, "Evaluation of SVM Kernels and Conventional Machine Learning Algorithms for Speaker Identification," International Journal of Hybrid Information Technology ,Vol.3, No.3, July, 2010

[12] Durgesh k. Srivastava, Lekha bhambhu, "Data classification using Support Vector Machine," Journal of Theoretical and Applied Information Technology© 2005 - 2009 JATIT.

[13] M. Bak, "Support Vector Classifier with Linguistic Interpretation of the Kernel Matrix in Speaker Verification", Man-Machine Interactions, Krzysztof A. Cyran,Stanislaw Kozielski, James F. Peters (eds.), ISSN 1867-5662, Springer, 2009, pp 399-406.

[14] Hiroyuki Mori 1, Wenjun Jiang, "An ANN-Based Risk Assessment Method for Carbon Pricing," 978-1-4244-1744-5/08©2008 IEEE.