Performance Evaluation of Artificial Neural Network Classifier Based On Receiver Operating Characteristic (ROC) Curves

Ms. J.D. Dhande and Dr. S. M. Gulhane

Abstract: - Classification is one of the most active research and application area of neural networks. Artificial Neural Networks have been applied as a classifier to find one "best" detection rate using Receiver operating characteristic (ROC) curve. This research work has been proposed the design of multilayered neural network (MLP NN) classifier for classification of Echocardiogram dataset as a two class problem. The performance of MLP NN classifiers are examined by generating the ROC curves for training and testing dataset. ROC gives better results in terms of Area under the ROC curve (AUC) is greater and in the sense of being composed of a better distribution of operating points. Simulation results shown that the classifier can achieve more than 90% classification accuracy using ROC curves.

Keywords: Artificial Neural Network (ANN), Classifier, MLPNN, Receiver operating characteristic (ROC)

I. INTRODUCTION

Classification is one of the most frequently used decision making tasks of human activity. A classification problem occurs when an object needs to be assigned into a predefined class or group based on a number of observed attributes related to that object. Many problems in business, science, industry, and medicine can be treated as classification problems. Examples, includes bankruptcy, prediction, credit scoring, medical diagnosis, Quality control, hand written character recognition and speech recognition.

Ms. J.D. Dhande and Dr. S. M. Gulhane, Emails : jayshridhande@rediffmail.com , mgulhane67@rediffmail.com

Traditional statistical classification procedure such as discriminate analysis is built on the Bayesian decision theory [1][5]. The effectiveness of these methods depends on a large extent on the various assumptions. Artificial neural networks have emerged as an important tool for classification. The recent vast research activities in neural network classification have established that artificial neural networks are a promising alternative to various traditional classification methods. The neural networks are nonlinear models, which makes them flexible in modeling real world complex relationships. In spite of the successful application of ANN's to several pattern recognition problems [1][2]. ANN also provide a powerful tool to help doctors to analyze, model and make sense of complex clinical data across a broad range of medical applications [3]. Most applications of artificial neural networks to medicine are classification problems; that is, the task is on the basis of the measured features to assign the patient to one of a small set of classes [4][5].The performance of classifier is measured in terms of classification accuracy and another method of specifying the performance of classifier for generating the Receiver Operating Characteristic (ROC) Curve.

Receiver Operating Characteristic (ROC) analysis is an established method of measuring diagnostic performance of in medical imaging and signaling studies. Artificial neural networks (ANN's) have been applied as a classifier to find one "best" detection rate. The current standard method of generating ROC curves for an ANN is to vary the output node threshold for classification [6].

This paper is organized as follow. Section II gives some background on Receiver Operating Characteristic (ROC) Curve for classifier. Section III describes the basic ANN, in particular, multilayered back propagation network. Section IV presents the methodology and simulation results .Section V discuss the conclusion of work.

II. BACKGROUND OF RECEIVER OPERATING CHARACTERISTIC (ROC) CURVE

The Receiver Operating Characteristic (ROC) curve is the method of specifying the performance of a classifier is to note its true positive (TP) rate and false positive (FP) rate for a dataset. The True positive (TP) rate is the percentage of target samples that are correctly classified as target samples. The false positive (FP) rate is the percentage of non target samples that are incorrectly classified as target samples. For particularly applications, we may require the

classifier to operate at some point other than the one to which it naturally trained.

An ROC curve is a plot of operating points showing the possible trade off between a classifier's TP rates versus its FP rate. The TP rate is commonly referred to as "sensitivity" and (1-FP) rate is called "specificity". The ROC curves are a useful way to interpret sensitivity and specificity levels and to determine related cut scores .ROC curves are a generalization of the set of potential combinations of sensitivity and specificity possible for predictors. ROC curve analyze not only provide information about cut scores , but also provide a natural common scale for comparing different predictors that are measured in different units. An overall indication of the diagnostic accuracy of a ROC curve is the area under the curve (AUC). Area under the curve values are closer to 1 indicate the screening measure reliably distinguishes among satisfactory and unsatisfactory performance, whereas values at .5 indicate the predictor is no better than chance. Figure 1 shows the two typical ROC curves for performance evaluation

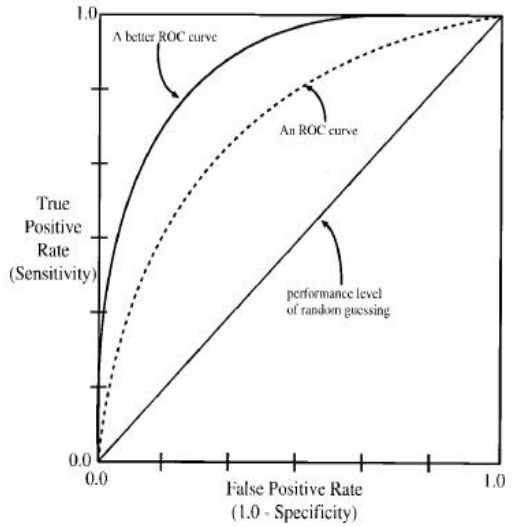

Figure 1 Two typical ROC Curves

Receiver Operating Characteristic (ROC) curve analysis is a method of measuring diagnostic performance in medical application. This work is examined the method of performance evaluation of ANN classifier by generating ROC curve for two class problem.

III. ARTIFICIAL NEURAL NETWORK CLASSIFIER

An Artificial Neural Network (ANN) is a computational model that attempts to account for the parallel nature of human brain. An ANN is a network of highly interconnecting processing elements (neurons) operating in parallel. These elements are inspired by biological nervous systems. As in nature, the connections between elements largely determine the network function. A subgroup of processing elements is called a layer in the network. The first layer is the input layer and the last layer is the output layer. Between the input and output layer, there may be additional layers of units called hidden layers. Figure 2 represents the typical neural network. The neural network is train to perform a particular function by adjusting the values of the connection weights between elements.

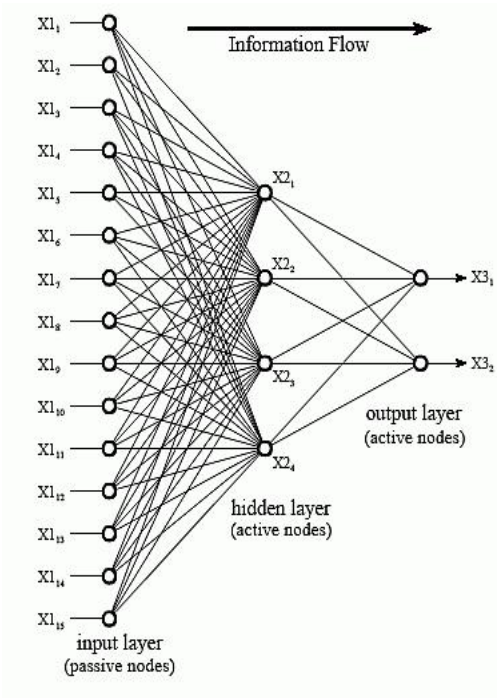

Figure 2 A typical Neural network

The most popular neural network is Multilayer feed forward network (MLP NN) with back- propagation learning algorithm. Although many types of neural networks can be used for classification purposes [8-9].The main focus on multilayer perceptron neural network which are the most widely studied and used neural network classifier. The back propagation neural network architecture is the most popular, effective and easy-to-learn model for complex multilayer networks. To this network, training inputs are applied to the input layer of the network, and desired outputs are compared to the output layer. During learning process, a feed forward sweep is made through the network, grid the output of each element is computed layer by layer. The difference between the output of the final layer and the desired output is back-propagated to the previous layers, usually modified by the derivative of the transfer function, and the connection weights are normally adjusted .This process proceeds for the previous layers until the input layer is reached [5][9].

IV. METHODOLOGY AND SIMULATION RESULTS

The echocardiogram dataset is used for design the ANN based classifier. This section gives Overview of echocardiogram dataset, training of neural network, and

simulation result of neural network design with performance curves.

A) Dataset Description

In this experiment the medical data related to echocardiogram dataset of patients for survival analysis after the heart attacks is considered. This dataset has been taken from publically available UCI repository machine learning [7]. The dataset concerns classification of patient's survival at least one year that is death or life after a heart attack. The dataset size is reduced by ignoring one attribute values from given dataset. Table 1 gives attributes description of echocardiogram dataset. The dataset is partitioned into training set having 100 samples and testing set having 32 samples for design and test the classifier.

Data Representation:

Number of instances: 132

Number of attributes: 11 and a class attribute

Class:

Class 0: A patient was either dead after 1 year or had been followed for less than 1 year.

Class 1: Patients was alive at 1 year

Table 1 Attributes Description

Sr. no	Attributes	Description
1	Survival	Number of months patients survived
2	Still-alive	A binary variable: 0 = dead at end of survival period, 1 means still alive
3	Age _at_ heart attack	Age in years when heart attack occurred
4	Pericardial effusion	Binary: 0 means no fluid, 1=fluid
5	Fractional – shortening	A measure of contractility around the heart lower numbers are increasingly abnormal
6	Epss	E-point septal separation, another measure of contractibility
7	Lvdd	A measure of the size of the heart at end –diastole
8	Wall-motionscore	A measure of how the segments of the left ventricle are moving
9	Wall-motion-index	Equals wall motion score divided by number of segments seen.
10	Mult	A derivate var
11	Name	Name of patients

A) Training the Neural Network

In this experimentation, Multilayer feed forward neural network is created by using the training set and the network is trained on given data and used back propagation learning algorithm. Multilayer feed forward neural network is three layers model. The input layer network consists of 11 neurons to represent each attributes as the database attributes. The numbers of classes are two: class 0 and class 1. The output layer consists of one neuron to represent these classes. The hidden layer consists of 20 neurons in MLP neural network. The design model of MLPNN classifier is 11-20-1.The optimal value of hidden node is obtained based on trail and error experimentation. This model is trained using back propagation learning algorithm with more runs. Figure 3 shows the relation between the number of epochs and the performance in term of Mean square error (MSE) during training process for MLP NN network model. For training the MLPNN 11 epochs required with minimum MSE value.

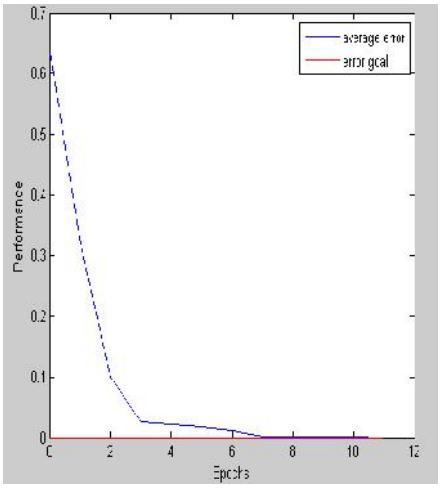

Figure 3 Training curve for MLPNN

B) Performance of Neural Network

MLPNN model is trained on 11 epochs and this network is validated on testing set. The testing set with 32 samples is tested on trained network. For performance evaluation of network, Receiver operating characteristic (ROC) curves is plotted for training and testing dataset. ROC curve for training data with better classification curve with best TP and FP rates has shown in figure 4.

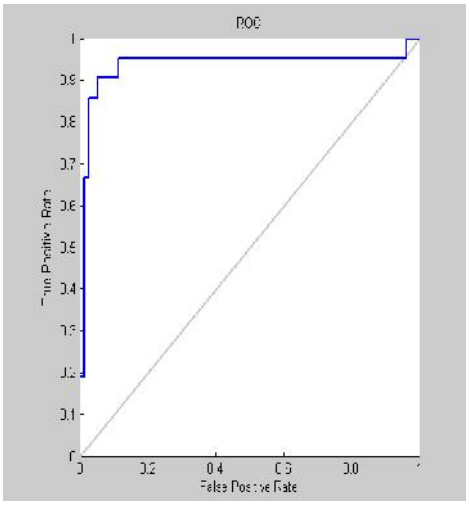

Figure 4 ROC curve for MLPNN on training set

For generalization of trained network, the performance evaluated on testing set as a unknown input samples as input to trained network and output of network is calculated with the adjusted weights. The output of network is compared with the target output to study the learning ability of network for classification of echocardiogram dataset with better accuracy based on trail and error experimentation. ROC curve is plotted for testing dataset and is shown in figure 5.

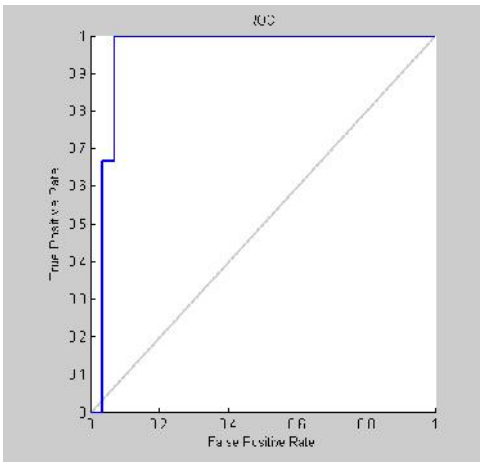

Figure 5 ROC curve for MLPNN testing set

V. CONCLUSION

For wide range of applicability of ANN and their ability to learn complex and non linear relationships including noisy or less precise information , neural networks are well suited to solve problems in biomedical field. In this study, neural network technique is adopted for classification of medical dataset with back propagation learning algorithm. The performance of network is analyzed on training and testing dataset with ROC curve generating from trained classifier. The manipulation of classifier parameters will in turn move a decision boundary in feature space between the two classes and result in a new sensitivity/ specificity tradeoff. We examined that the performance of design neural network classifier by generating the better ROC curve in sense of having greater area under the ROC curve (AUC) and achieved more than 90% classification accuracy on generalization.

REFERENCES

[1] P.O. Duda, P.E.Hart, pattern classification and scene analysis. New York: Wiley, 1973

[2] S.Haykin, Neural Networks: A comprehensive foundation. Prentice Hall, 1999.

[3] R.Dybowski, V.Gant, clinical applications of artificial neural networks, Cambridge university press, 2007.

[4] R.P Lippman,"Pattern classification using neural networks (1989)"IEEE communication Mag., pp 47-64.

[5] G.P. Zhang "Neural Networks for classification: A survey IEEE transaction on systems, Man and cybernetics part-c: Applications and review; Vol. 30, No. 4 November 2000.

[6] K.Woods, K.W.Bowyer," Generating ROC curves for Artificial neural networks IEEE Transactions on medical imaging, vol. 16, no. 3, June 1987"

[7] Murphy P.M. and Aha D.W., UCI Machine Learning Databases repository Irvine C.A: University of California Department of Information and computer science, ftp://ftp.ics.uci.edu/pub/machine-learning databases/echocardiogram.

[8] Zhou J.H., Simopoulos S.K. Otsuki C., Comaniciu J.D., Automatic Cardiac View Classification of Echocardiogram ,Park computer vision , 2007, IEEE 11th International conference, ICCV 2007,ISSN: 1550-5499,publication Date (14-21 Oct.2007), 1-8.

[9] J.D.Dhande, S.M.Gulhane,"Design of classifier using artificial neural network for patient's survival analysis", International journal of Engineering science and innovative technology (IJESIT) volume 1, Issue 2, page no. 278-282, November 2012.