International Journal of Scientific Research in Computer Science, Engineering and Information Technology

Soft Computing as a tool for Classification of Cardiovascular Abnormalities

Dilip Kumar S*1, Akshaya Yadhav2, Archana Sankar3

*1Assistant Professor, Instrumentation and Control Engineering, Sri Krishna College of Technology, Coimbatore, Tamilnadu, India
2Student, Instrumentation and Control Engineering, Sri Krishna College of Technology, Coimbatore, Tamilnadu, India
3Student, Instrumentation and Control Engineering, Sri Krishna College of Technology, Coimbatore, Tamilnadu, India

ABSTRACT

Classification of Electrocardiogram (ECG) for Cardio-Vascular Abnormalities (CVA) in the process of diagnosis is inevitable. In this paper, we propose a scheme to integrate Principal Component Analysis (PCA) with Neural Networks (NN) for classification of ECG Signals. A Neural Network (NN) with Back Propagation Algorithm is deployed as classifier. ECG samples consisting of Normal signals and three abnormal signals are taken from physionet arrhythmias database for our experiments. The PCA is used to minimize ECG signals into weighted sum of basic components that are statistically mutual independent. Thus, PCA is used for dimensionality reduction of data. Here a comparison of performance of Neural Network (NN) and Principal Component Analysis (PCA) with Neural Network (NN) are investigated. Principal Component Analysis (PCA) eliminates the least considerable data values, hence helps in improving the performance in classification of ECG signals. The results obtained suggest that Principal Component Analysis (PCA) with Neural Network (NN) performance is faster and better than Neural Network (NN) Classifier alone.

Keywords: Principal Component Analysis (PCA), Electrocardiogram (ECG), Neural Network (NN). Cardio-Vascular Abnormalities (CVA)

I. INTRODUCTION

Cardiac surgeons to efficiently diagnose the Cardiovascular Diseases (CVD) for the last seven decades have intensively used ECG signal. Traditionally the automatic analysis of ECG signals, including delineation, was taking place online on bulky, high performance beside cardiac monitors, or performed offline during a pre-processing stage after ambulatory ECG recording using wearable, yet obtrusive, ECG data loggers. Maintaining and updating such a system for every new abnormality is intrinsically complex. This introduces a problem of finding a simple and fast solution toward heart disease classification from ECG that raises alert to cardiac specialist as soon as a cardiac disease is recognized. However, the issues faced in ECG analysis of one patient differ with other patients ECG waveforms; due to this, the performance of classifiers will be low during training of data.

II. METHODS AND MATERIAL

It is known that from years, researchers have proposed various methods for ECG beat Classification using Neural Network (NN) classifier [1][2][3]. By convention back propagation, Neural Networks (BPNN) is used. The important feature of BPNN is its ability to recognize and classify ECG signals; the shortcoming with this method is slow convergence to local and global minima. To outcome this problem, researchers proposed Hybrid Neural Networks.

In [4] Dipti Patra *et al* had applied Integration of fuzzy c-means (FCM), Principal Component Analysis (PCA) and Neural Networks for Classification of ECG data, in which training the fuzzy layer by fuzzy c-means algorithm. Fuzzy c-means clustering algorithm divides the data into fuzzy functions that with overlap with one another. Applying PCA to the obtained data, results in dimensionality reduction of the data, which results in elimination of insignificant data values present. Hence the reduced matrix is the input for NN and the classification for ECG Arrhythmias are obtained.

In [6] Atena Sajedin *et al* had applied a trainable Neural Network model for ECG beat classification, in which topologies of multilayer perceptrons neural networks are designed. Comparative analyses of combination of different topologies are performed. In [8] Wei Jiang *et al* had applied Block-based Neural Networks (BbNN) for ECG signal Classification, in which BbNNs are utilized for personalized health monitoring.

In [5] Dayong GAO *et al* had applied ECG Arrhythmia Identification using a Neural Network based on a Bayesian Framework, in which Bayesian framework is based on logistic regression model and the back propagation algorithm. Here a dual threshold method is applied to determine false alarm signals. In [7] Philip Langley *et al* had applied Principal Component Analysis (PCA) for analysing Beat-to-beat changes in ECG features, in which Coherence and correlation are obtained for the ECG features.

In this paper, we evaluate the performance of Neural Network and the integration of Principal Component Analysis (PCA) with Neural Network for ECG features. The proposed structure consists of layer of feature extraction with Principal Component Analysis (PCA) and classification by Neural Networks using Back Propagation Neural Network (BPNN) Algorithm. Principal Component Analysis (PCA) performs the extraction of Principal Components from the raw data and the multilayer perceptron works as a final classifier. Initially the raw data for ECG is trained for Neural Network by varying the sigmoid function, number of hidden layers, training function.

However, in Principal Component Analysis (PCA), the raw data is reduced and the principal components obtained are given as input to Neural Network. It is observed from the results obtained that performance of Principal Component Analysis (PCA) with Neural Network (NN) is more generalized and faster in computation than Neural Network (NN) alone.

A. ECG Signal Classification Methods

For classification of ECG signals based on their arrhythmias, various solutions were presented in the literature. We present integration of Principal Component Analysis (PCA) with Neural Network (NN) and compare the performance of the model with Neural Network (NN). The data set is the prerequisite for the Neural Network (NN); the data set is obtained by taking the ECG values from four different subjects for Normal wave, Arrhythmia Wave, Ventricular Tachyarrhythmia, Supra Ventricular Arrhythmia from physionet database www.physionet.org [9]. A set of thousand values are taken from subjects the values are put in columns and output classifiers [0 0 0 1] , [0 0 1 0], [0 0 1 1], [0 1 0 0] are marked for corresponding input values. The input values and the output classifiers are shuffled and the data set can be utilized for Neural Network (NN). Similarly, this raw data is used in Principal Component Analysis (PCA) for dimensionality reduction and the output obtained is the input for Neural Network (NN) training, thus classification of ECG signals based on cardiovascular abnormalities is done

B. Neural Network Classifier

The classifier implemented for this work is a standard, feed forward, Neural Network (NN) with error back propagation algorithm with two or more hidden layers and output layer. The activation function for all units is the asymmetric sigmoid function. Training the network is accomplished by initializing all weights to small, random values and then performing a gradient-descent search in the network's weight space for a minimum of a squared error function of the network's output. The error obtained will be the difference of the network's output and the target value for each input vector. For the experiments, the target values were set to [0 0 0 1] for the Normal ECG wave, [0 0 1 0] for the Arrhythmia Wave, [0 0 1 1] for the Ventricular Tachyarrhythmia wave, and [0 1 0 0] for the Supra ventricular Arrhythmia.

The steps for performing NN are,

[1] Load the data set.

[2] Specify the sigmoid function, training function, Number of hidden layers.

[3] Specify data for training and testing.

[4] Mean Square Error for training and testing are obtained.

Classification of cardiovascular abnormalities is obtained using Neural Network

C. Principal Component Analysis

Principal Component Analysis (PCA) is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance (i.e., accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to (i.e., uncorrelated with) the preceding components. Since Principal Component Analysis (PCA) is known for its dimensionality reduction technique gives the linearly correlated variables where its matrix dimensions are reduced, which gives high accuracy, and quick computation.

The steps for performing PCA are,

[1] Load the input raw data set.

[2] Find the mean from the data set.

[3]Subtract the mean from individual components.

[4] Compute the Covariance from the data.

[5] Determine the Eigen vectors and values of the Covariance matrix.

[6] Principal Components are chosen and the matrix is formed.

The most assumption made in PCA for dimensionality reduction is to obtain the Principal Components from the principal axes, which consists of relevant information. Thus, computational time will be reduced. To this model-reduced matrix Neural Network with Back Propagation is performed.

III. PROPOSED METHODOLOGY

In this paper, the proposed method is divided into three steps: (A) ECG dataset formation, (B) Dimensionality reduction, (C) Classification by Neural Networks.

A. ECG Dataset Formation

For our experiment, ECG samples such as Normal wave, Arrhythmia Wave, Ventricular Tachyarrhythmia, Supra Ventricular Arrhythmia are obtained from physionet database www.physionet.org [9]. The data values are taken from signal before and after R peak, since this region consists of vital information values of Heart. Data are obtained from four subjects each for four beat types. Output classifier is marked and the data are shuffled. To this dataset, 75% of data is trained and 25% of data will be tested.

B. Dimensionality Reduction

The dataset formed will be of raw data and it is required to obtain the principal components from the dataset. Hence, by using PCA to the dataset the dimension of the dataset is reduced without compromising the feature vector. The most assumption made in PCA for dimensionality reduction is to obtain the Principal Components from the principal axes that consists of relevant information. Thus, computational time of the process will be reduced.

C. Classification by Neural Networks

For our experiment back propagation, Neural Network (NN) is used. In which the Neural Network (NN) structure consists of three-layer structure. The three layers are input layer, hidden layer and output layer. The data obtained because of PCA is given as input to the input layer of Neural Network. The no of hidden layers can be varied from 3-20. The output layer consists of four neurons, where ECG signals of four types are to be classified. In our study, tansig is sigmoid function and training type is Scaled Conjugate Gradient. The weight and bias values are updated with a learning rate of 0.01.

IV. RESULTS AND DISCUSSION

We got the samples from different subjects for four types of ECG Beats from physionet database. The obtained samples comprises of 4000x8 data matrix where 8 columns represents the input and output classifiers. By applying PCA to the data matrix, model reduced matrix is obtained. Neural Network (NN) training is done to the model-reduced matrix. This PCA-NN structure when compared with NN depicts the variation in classification of ECG Beats. Moreover, the numbers of hidden layers are less in PCA-NN than NN structure alone. The error obtained during training and testing are less in PCA-NN. It is observed that PCA-NN structure is performing well than NN structure.

The comparison of results is depicted in table 1. The Fig 1, 2 represents the regression plots of Neural Network and Principal Component Analysis with Neural Network. Fig 3, 4 represents the performance plots of Neural Network and Principal Component Analysis with Neural Network. Fig 5, 6 represents the classification of Neural Network and Principal Component Analysis with Neural Network. The table depicts the results which are ought to be well performing structure of NN and PCA-NN. The value obtained is concluded by obtaining the values by varying the parameters (via) No of Hidden layers, sigmoid function, Training Type.

TABLE 1. Tabulation of NN and PCA-NN Results

S.NO	Parameters	NN	PCA-NN
1	No of Hidden Layers	19	12
2	No of iterations	1000	1000
3	Sigmoid Function	tansig	tansig
4	Training Type	Trainscg	Trainscg
5	MSE training	0.0487	0.0933
6	MSE testing	0.018	0.092
7	Computation time (sec)	64.4908	71.8385
8	% Correctly Classified	90%	100%

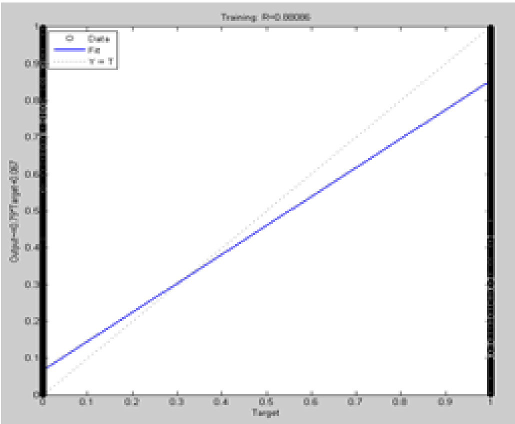

Figure 1. Regression Plot of NN

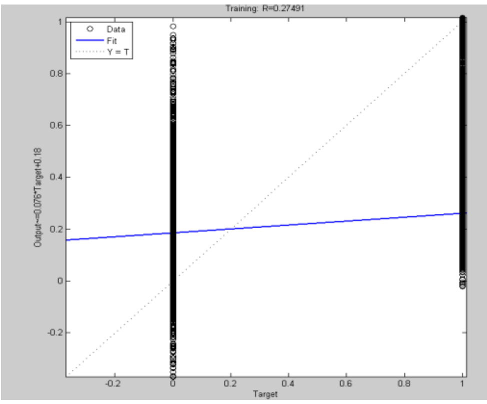

Figure 2. Regression plot of PCA-NN

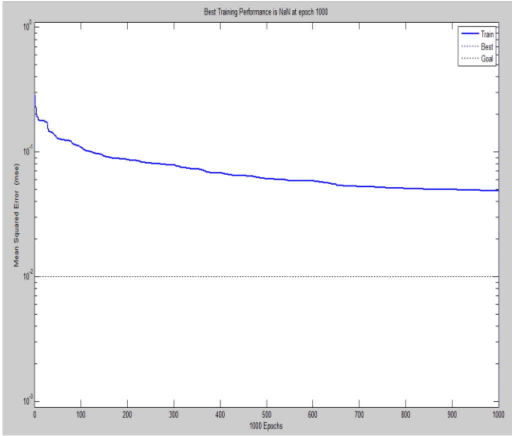

Figure 3. Performance plot of NN

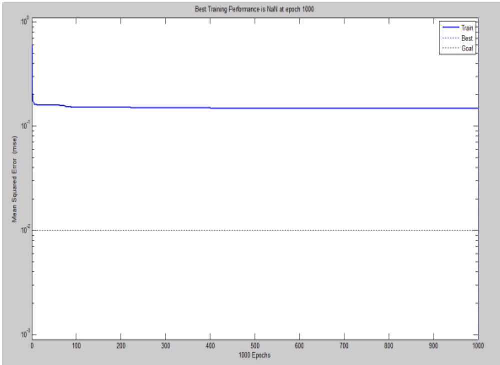

Figure 4. Performance plot of PCA- NN

V. CONCLUSION AND FUTURE WORK

In this study, the integration of PCA-NN structure performs well for the cardiovascular abnormalities classification than the conventional Neural Network (NN) classifier alone. Previously using Neural Network (NN) classifier alone for the data increases the computational time. This is curbed and the performance of the overall system is improvised in PCA-NN structure. By making a comparative analysis of PCA-NN and NN structure, it is observed that the Performance of PCA-NN structure is well served in recognizing and classification of ECG waves with better accuracy and higher computational rate. Further as a part of future work, other soft computing techniques can be employed for the classification of cardiovascular abnormalities.

VI. REFERENCES

[1]. T Acharya, R., Bhat, P. S., Iyengar, S.S., Roo, A. & Dua, S., "Classification of heart rate using aritificial neural network and fuzzy equivalence

relation," The Journal of the pattern Recognition Society, 2002.

[2]. De Chazal, P., & Reilly, R. B., "Automatic classification of ECG beats using waveform shape and heart beat interval features," In IEEE international conference on acoustic, speech and signal processing (ICASSP '03), vol. 2, pp. 269-272, Hong Kong, China, 2003

[3]. Osowski, S & Linh, T. H., "ECG beat recognition using fuzzy hybrid neural network," IEEE Transaction on Biomedical Engineering, vol. 48, no. 11, pp, 1265-1271, 2001.

[4]. Dipti Patra, Manab Kumar Das,Smita Pradhan "Integration of FCM, PCM and Neural Networks for Classification of ECG Arrhythmias" IAENG International Journal of Computer Science, 36:3,IJCS_36_3_05.

[5]. Dayong Gao, Micheal Madden, Micheal Schukat, Des Chambers, and Gerard Lyons "Arrhythmia Identification from ECG Signals with a Neural Network Classifier Based on a Bayesian Framework" In the Twenty-fourth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence, December 2004.

[6]. Atena Sajedin, Shokoufeh Zakernejad, Soheil Faridi, Mehrdad Javadi and Reza Ebrahimpour "A Trainable Neural Network Ensemble for ECG Beat Classification" Proceedings of the International Conference on Neural Networks (ICNN2010), Amsterdam, Netherland, Publication year 2010, Page no 28-30

[7]. Philip Langley, Emma J.Bowers, and Alan Murray "Principal Component Analysis as a Tool for Analyzing Beat-to-Beat Changes in ECG Features: Application to ECG Derived respiration" IEEE Trans Biomed Eng. 2010 Apr; 57(4):821-9. Epub 2009 Apr

[8]. Wei Jiang, Seong G.Kong, and Gregory D.Peterson "ECG Signal Classification using Block-based Neural Networks" Neural Networks, 2005. IJCNN '05. Proceedings. 2005 IEEE International Joint Conference on 31 July-4 Aug. 2005, page (s) 326 - 331 vol. 1.