

Training of Multilayer Perceptrons with Improved Particle Swarm Optimization for the Heart Diseases Prediction

Kelwade JP^{1*} and Salankar SS²

¹Bapurao Deshmukh College of Engineering, RTM Nagpur University, India

²G. H. Rasoni College of Engineering, RTM Nagpur University, India

Abstract

The study of Heart rate variability is recently gained momentum for an estimation of heart health. This paper suggests a new approach for enhancement of the prediction accuracy of Multi-Layer Perceptrons (MLP) neural network using improved Particle Swarm Optimization (IPSO) technique. The IPSO computes the weights and biases of MLP for the more accurate prediction of the cardiac arrhythmia classes. This study for heart condition prediction involves selection of Three types of heart signals including Left Bundle Branch Block (LBBB), Normal Sinus Rhythm (NSR), Right Bundle Branch Block (RBBB) from MIT-BIH arrhythmia database, formation of heart rate time series, extraction of features from RR interval time series, implementation of training algorithm and prediction of arrhythmia classes. Several experiments on the proposed training method are carried out to superior the convergence ability of MLP. The experimental results gives comparably better evaluation over gradient based Back-Propagation (BP) learning algorithm.

Keywords: Cardiac arrhythmia; Particle swarm optimization; Multilayer perceptron; Heart rate variability; Back propagation learning; Training algorithm

Introduction

The graphical depiction of the heart beats in the form of electrical signals is known as electrocardiogram (ECG). The abnormal rhythms of the heart beats are termed as cardiac arrhythmias. Some arrhythmias are life threatening. Therefore there is need to identify the heart conditions of cardiac patients. The identification of cardiac arrhythmias in early stage can save the patients from sudden cardiac arrest.

A variation in the consecutive cardiac beats is referred to as heart rate variability (HRV). By means of HRV analysis technique cardiac health can usually be computed. An estimation of HRV is recently being adopted as investigation tool for recognition of heart abnormalities in cardiology.

Some methods have been suggested in reported literature for recognition, classification or prediction of cardiac arrhythmia. Özbay and Karlik presented Artificial Neural Network (ANN) and built up ECGWin Software to interpret and classify more number of cardiac arrhythmias [1]. Habboush et al. compared neural networks with Karhunen-LoGve transform for compression and classification [2]. Saini and Saini used multilayer perceptron (MLP) feedforward neural network to classify four arrhythmias [3]. Franklin and Wallcave utilized ANN to categorize heartbeat into 6 types with 85% correct identification rate [4].

Deshmukh and Patil proposed Empirical Mode Decomposition and feed-forward propagation neural network to classify different types of Abnormal beats [5]. Ozbay et al. studied MLP and fuzzy clustering Neural Network for classification of 10 different arrhythmias [6]. Wang et al. suggested methods to distinguish eight types of ECG using principal component analysis (PCA), linear discriminant analysis (LDA) and a probabilistic neural network (PNN) [7]. The generalized linear model (GLM) algorithm to discriminate arrhythmias classes using AR coefficients of normal and abnormal ECG signals [8,9].

These techniques are generally based on extraction of morphological and temporal features from processing of ECG signals. The main disadvantages of ECG signal processing for features detection

are 1) requirement of large computation time and 2) Introduction of noise in the ECG at the time of processing. An alternative approach is to extract HRV signals from RR time intervals of ECG signal. The main advantages of HRV analysis for features detection include 1) RR time intervals are less prone to the noise and 2) HRV signals signify the function of autonomic nervous system (ANS) and cardiovascular system [10].

HRV signal is useful tool for estimation of overall cardiac health and condition of the ANS. Therefore, HRV analysis can be treated as valuable investigation tool in detection, prediction of arrhythmias classes in the medical field of cardiology [11].

Some of proposed approaches for analyzing the HRV signal in detection and prediction of cardiac arrhythmia classes are reported in the given literature.

Yaghouby et al. investigated four cardiac arrhythmias such as left bundle branch block, first degree heart block, Supraventricular tachyarrhythmia and ventricular trigeminy based on the Generalized Discriminant Analysis (GDA)/feature reduction technique and MLP [10]. Acharya et al. proposed ANN and Fuzzy equivalence relationship for classification of eight types of cardiac arrhythmias [12]. Anuradha and Reddy employed non-linear methods such as Spectral entropy, Poincaré plot geometry, Largest Lyapunov exponent and detrended fluctuation analysis to extract features from HRV signal for fuzzy classifier [13]. Asl et al. studied four types of cardiac arrhythmias for adaptive-learning-rate neural network classifier adopting linear, nonlinear, and chaotic features of the RR interval signals [14]. Dallali et al. presented combined approach using fuzzy c-means (FCM)

***Corresponding author:** Kelwade JP, Bapurao Deshmukh College of Engineering, RTM Nagpur University, India, Tel: +917755903397; E-mail: kelwade2007@gmail.com

Received May 02, 2017; **Accepted** May 26, 2017; **Published** Jun 02, 2017

Citation: Kelwade JP, Salankar SS (2017) Training of Multilayer Perceptrons with Improved Particle Swarm Optimization for the Heart Diseases Prediction. Int J Swarm Intel Evol Comput 6: 156. doi: [10.4172/2090-4908.1000156](https://doi.org/10.4172/2090-4908.1000156)

Copyright: © 2017 Kelwade JP, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

clustering, wavelet transform and PCA to classify four kinds of heart diseases [15]. Asl and Setarehdan proposed automatic detection and classification method using ANN classifier for five classes of arrhythmia [16].

Kelwade and Salankar predicted classes of cardiac arrhythmia with MLP and radial basis function Neural (RBFN) network [17,18]. Goshvarpour focused on the Lyapunov Exponents and Entropy features to train Quadratic classifier and compare the result with Fisher and k-Nearest Neighbor (k-NN) classifiers [19]. Kampouraki et al. used statistical methods and signal analysis techniques to extract features of heartbeat time series [20]. Rawther and Cheriyan investigated support vector machine (SVM) for Life threatening arrhythmias such as Ventricular Tachycardia (VT) and Ventricular Fibrillation (VF) to detect and classify by make use of temporal and wavelet features [21]. Asl et al. developed an effective algorithm based on GDA and SVM classifier using HRV [22].

This paper for prediction of arrhythmia classes is organized in remaining sections as follows. Materials and methods section presents the overall methods such as MLP and PSO. Followed by experimental results and discussions of the proposed algorithm. Finally, the paper ends with significant remarks in conclusion section.

Materials and Methods

The ECG records for analysis and prediction are captured from standard MIT-BIH arrhythmia database. The filtering of ECG signals is performed with bandpass filter to remove powerline interferences. Pan and Tompkins algorithm for detection of QRS complexes and then R peaks is utilized in this study. Three kinds of heart rhythms including LBBB, NSR and RBBB are selected. The Records 109, 233 and 118 of ECG signals possessing LBBB, NSR and RBBB rhythms are particularly selected. The RR interval time series (RRITS) signals from the records to estimate the HRV is detected. The segments of the RRITS signals to extract the features are formed. The features such as normalized Low Frequency (nLF) and High Frequency (nHF) power components, SD1/SD2 ratio, Spectral Entropy (SE), Largest Lyapunov Exponent (LLE) and Hurst exponent (HE) extracted from HRV signal using linear and nonlinear methods are presented to train MLP for better prediction accuracy [12,13,16-18].

Training of multi-layer perceptrons

The MLP is a most popular multilayer feed-forward neural network. The adopted neurons configuration, shown in Figure 1, includes 6 neurons, 10 neurons and 3 neurons in input, hidden and output layers, respectively.

For prediction problems, The MLP is generally trained with a Back Propagation (BP) learning algorithm by computing the connection weights and biases. The BP learning algorithm, which is largely depends on selection of initial values of weights for faster convergence and a minimum generalization error, is an extension of the Least Mean Square (LMS) rule.

Many training algorithms have been reported in the literature to optimize the generalization errors and convergence speed of the MLP. Recently, extensive research and significant progress have been made in the area of nonlinear system. However, when a neural system is used to handle unlimited examples, including training and testing data, an important issue is how well it generalizes to patterns of the testing data, which is known as generalization ability. Many algorithms have been proposed so far to deal with the problem of appropriate weight-update

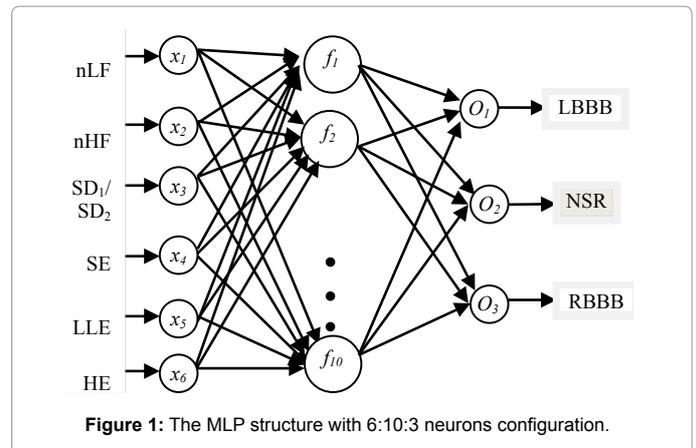


Figure 1: The MLP structure with 6:10:3 neurons configuration.

by doing some sort of parameter adaptation during learning. Singhal and Wu illustrated the application of extended Kalman algorithm which converged quickly compared to BP algorithm but required more computation [23]. Sarkalehm and Shahbahrami suggested several training algorithms such as Gradient Decent algorithm (GDA) with adaptive Learning Rate, Resilient and Levenberg-Marquardt algorithms for MLP to classify Paced Beat (PB), Atrial Premature Beat (APB) and NSR [24]. Suykens and Vandewalle determined output weights of single hidden layer MLP classifier using SVM method [25].

Tzikas and Likas effectively utilized incremental Bayesian learning method for linear models to train the MLP [26]. Ni and Song provided online learning algorithm for the neural tracking control system [27]. Battiti reviewed first and second order optimization methods for feed forward neural networks learning [28]. Riedmiller and Braun proposed new learning algorithm-Resilient backPropagation (RPROP) for MLP to improve generalization error a gradient-descent algorithm [29].

Moller introduced new supervised learning algorithm, scaled conjugate gradient (SCG) to speed-up convergence rate than BP, conjugate gradient algorithm with line search (CGL), Broyden-fletcher-Goldfarb-Shanno (BFGS) memory less quasi-Newton algorithm [30]. Nasir et al. demonstrated ability of Bayesian Regulation algorithm and LM to train MLP and Simplified Fuzzy ARTMAP (SFAM) for classifying the acute leukemia cells in blood sample [31]. Sut and Celik predicted mortality in stroke patients using MLP trained with algorithms namely quick propagation (QP), LM, BP, quasi-Newton (QN), delta bar delta (DBD), and CGD [32]. Abid et al. proposed learning algorithm based on combination of Least Square (LS) and Least Fourth (LF) criterion [33].

In fact the gradient-based training algorithms often require large iterations so as to evade from being spellbound in local optima and tuning of learning rate. Numerous modifications have been suggested to overcome the limitations of the gradient-based algorithm.

The evolutionary approaches such as Genetic algorithm, Ant Colony Optimization, artificial bee colony, Cuckoo search, PSO are usually being used in avoiding local minima and improving convergence rate of training algorithm [34-40].

In latest years, swarm intelligence algorithm such as PSO has been applied to solve real life problems in the area of optimization [40].

Particle swarm optimization

The PSO algorithm replicates social intelligence of particle swarm

namely flock of birds and school of fishes. PSO is most popular among other evolutionary algorithms because of ease of implementation and requirement of tuning of few parameters. PSO has recently been employed in the field of an optimization problem such as training of neural network [41].

In many literatures, the PSO has been proposed as an effective tool for training neural networks [42]. The basic PSO often get trapped in local optima and resulted in poor convergence. To efficiently control the local search and convergence to the global optimum solution, time varying acceleration coefficients (TVAC) are introduced in addition to the time varying inertia weight factor in PSO to estimate the new velocity of each particle and particles are reinitialized whenever they are stagnated in the search space [43].

The particle of the PSO possesses two characteristics namely position and velocity. A solution of any optimization problem contains updating of personal position and velocity in response with cognitive and social experience [40].

At current iteration time (t), current velocity v_{ij} and new position x_{ij} are modified using equation (1) and equation (2) respectively.

$$v_{ij}(t+1) = w(t)v_{ij}(t) + c_1r_1(p_{ij}(t) - x_{ij}(t)) + c_2r_2(p_{gj}(t) - x_{ij}(t)) \quad (1)$$

$$x_{ij}(t+1) = v_{ij}(t+1) + x_{ij}(t) \quad (2)$$

Where $w(t)$, c_1 and c_2 , r_1 and r_2 , $p_{ij}(t)$ and $p_{gj}(t)$, $x_{ij}(t)$ represent inertia weight, cognitive and social acceleration coefficients, random variables, personal best position, global best position and previous personal best position respectively.

The inertia weight may be randomly chosen. In improved PSO, inertia weight can be computed using time linear decreasing method. Equation (3) gives inertia weight as follows [44]:

$$w(t) = w_{max} - \frac{w_{max} - w_{min}}{T} * t \quad (3)$$

Where w_{max} , w_{min} , T and t represent maximum and minimum value

w_{min}	w_{max}	\dot{c}_1	\dot{c}_2	\dot{c}_1''	\dot{c}_2''	No. of particles	Maximum iteration
0.8	2	0.8	0.8	3.5	3.5	50	50

Table 1: Parameters setting of IPSO.

of inertia weights, maximum iteration and current iteration respectively.

The performance of PSO is dependent to the proper tuned parameters that results in the optimum solutions. Normally, cognitive (c_1) and social acceleration coefficients (c_2) are randomly selected to constant values. If value of c_2 is selected higher than value of c_1 , the PSO will converge prematurely [45].

Initially choosing high c_2 and small c_1 will make particles to move towards optimum solution. As optimization progresses, the values of c_1 and c_2 will get modified, which direct the particles to the global solution [45]. The acceleration coefficients are determined according the following equations (4) and (5) [46].

$$c_1(t) = (c_1' - c_1'') * \frac{i}{i_{max}} + c_1'' \quad (4)$$

$$c_2(t) = (c_2' - c_2'') * \frac{i}{i_{max}} + c_2'' \quad (5)$$

Where c_1 and c_1' , c_2 and c_2' are minimum and maximum values of cognitive coefficients, minimum and maximum values of social coefficients respectively.

In the simulation, the parameters of PSO algorithm are set initially as shown in Table 1.

Each particle possesses fitness value and fitness of the particle is measured by a fitness function. In this approach, mean squared error is used as the fitness function to test the performance of individual particle.

The fitness of k^{th} particle at t^{th} iteration is assessed using equation (6).

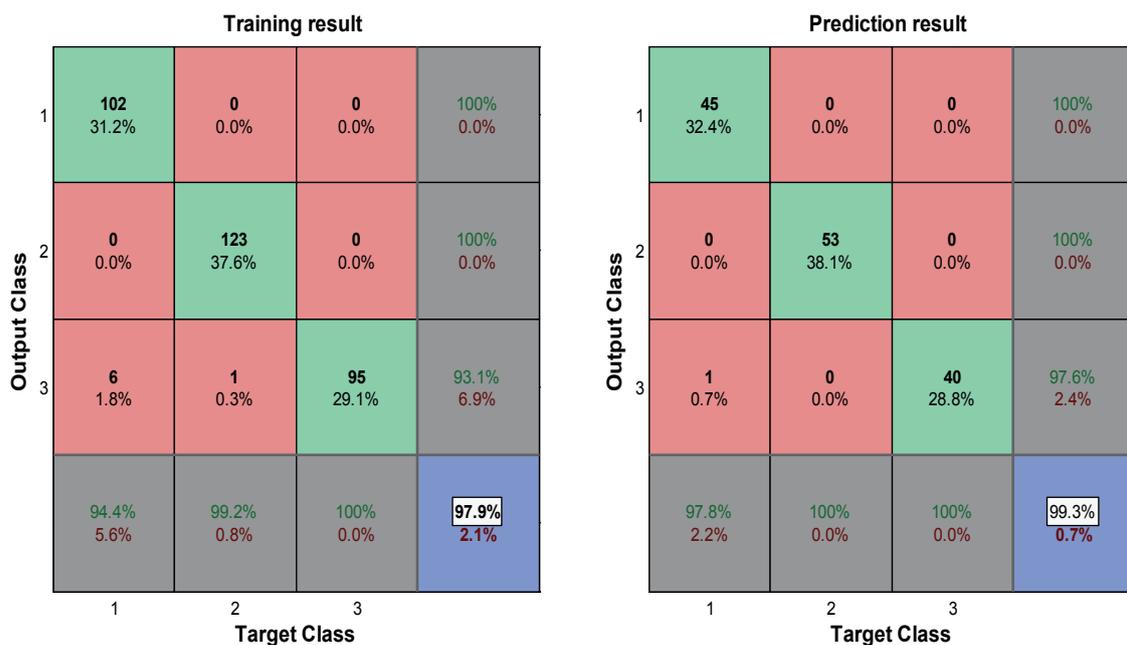


Figure 2: The performance of MLP-PSO.

$$f_k(t) = \frac{1}{P} \sum_{n=1}^P [O_n - O'_n]^2 \quad (6)$$

Where P, O_n and O'_n denote number of training datasets, desired output and actual network output, respectively.

A personal best position $pbest_k$ and a global best position $gbest_k$ of k^{th} particle will be adapted in t^{th} iteration using equation (6).

$$pbest_k(t) = \begin{cases} p_k(t) & f_k(t) \leq f_k(t-1) \\ pbest_k(t) & f_k(t) > f_k(t-1) \end{cases} \quad (7)$$

$$gbest_k(t) = best \{ pbest_k(t) \}_{k=1} \quad (8)$$

Results and Discussion

In this study, our aim was to train the MLP using the IPSO to

enhance the performance of MLP to predict the classes of cardiac arrhythmia. The experiments are carried out on arrhythmia data segments to make the MLP to effectively evolve the weights and biases with the help of IPSO. The experimental result of IPSO is compared with standard gradient based learning algorithms namely gradient decent algorithm with adaptive learning rate (GDX), RPROP, SCG and one-step secant method (OSS). Figures depict the training and Prediction results in confusion matrices. The Training performance of IPSO, GDX, RPROP, SCG, OSS learning algorithm is measured by plotting MSE against iterations. The plotted result is shown in figures. The IPSO enables the MLP to dynamically evolve weights and biases effectively. The performance of IPSO algorithm is found to be quite competitive in comparison with other algorithms (Figures 2-11).

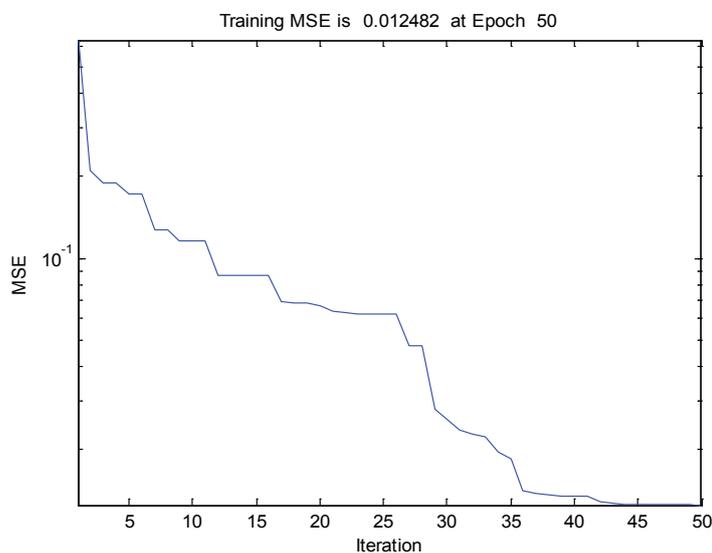


Figure 3: Training result of PSO.

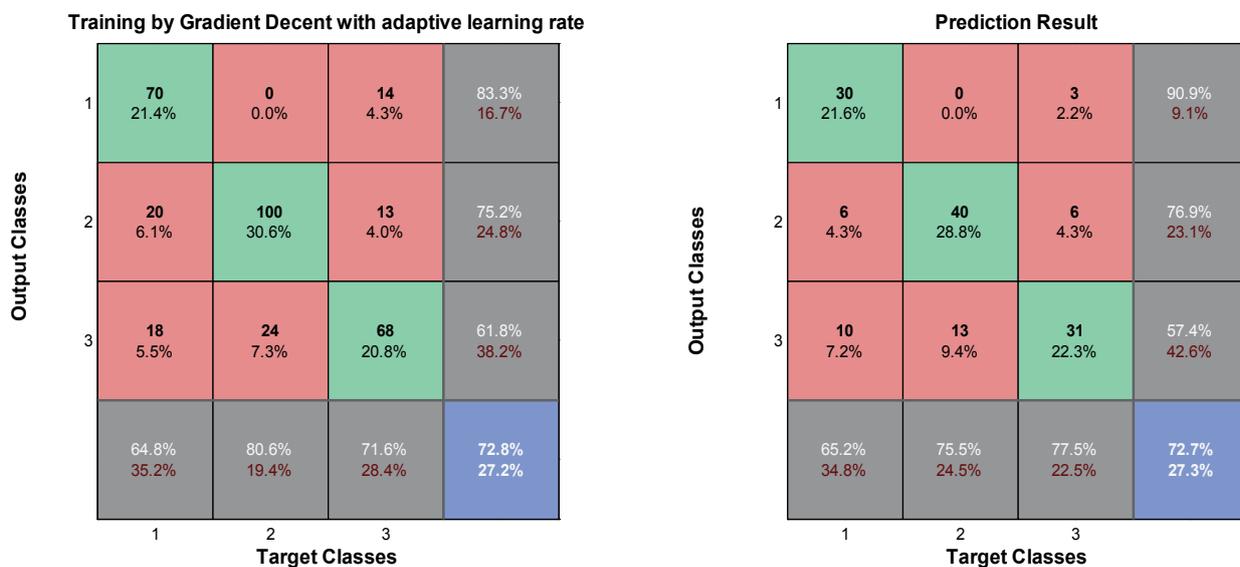


Figure 4: Training and prediction result using gradient decent with adaptive learning rate.

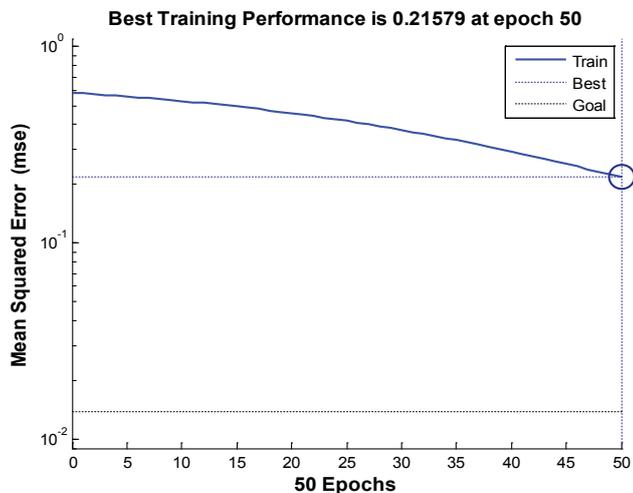


Figure 5: Experimental training results using gradient decent with adaptive learning rate.

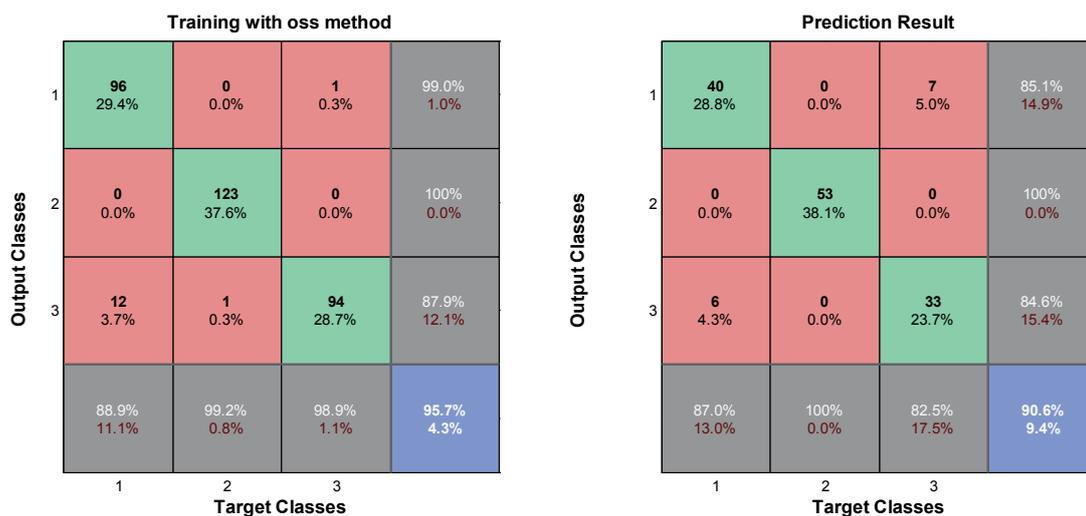


Figure 6: The performance of MLP with OSS training method and prediction of arrhythmia classes.

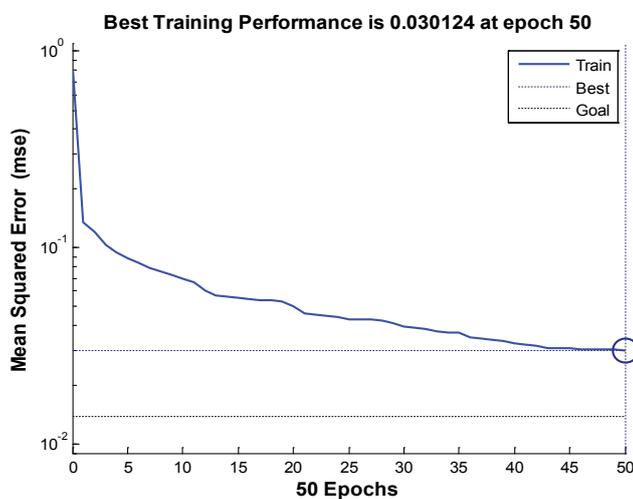
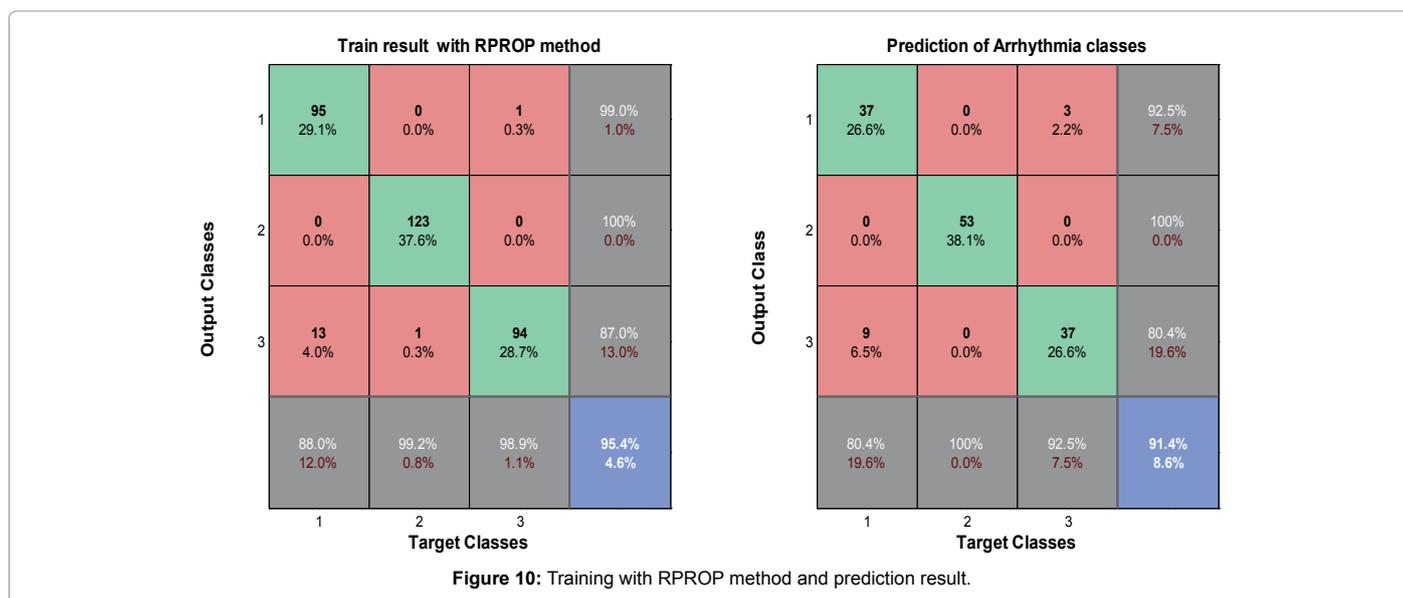
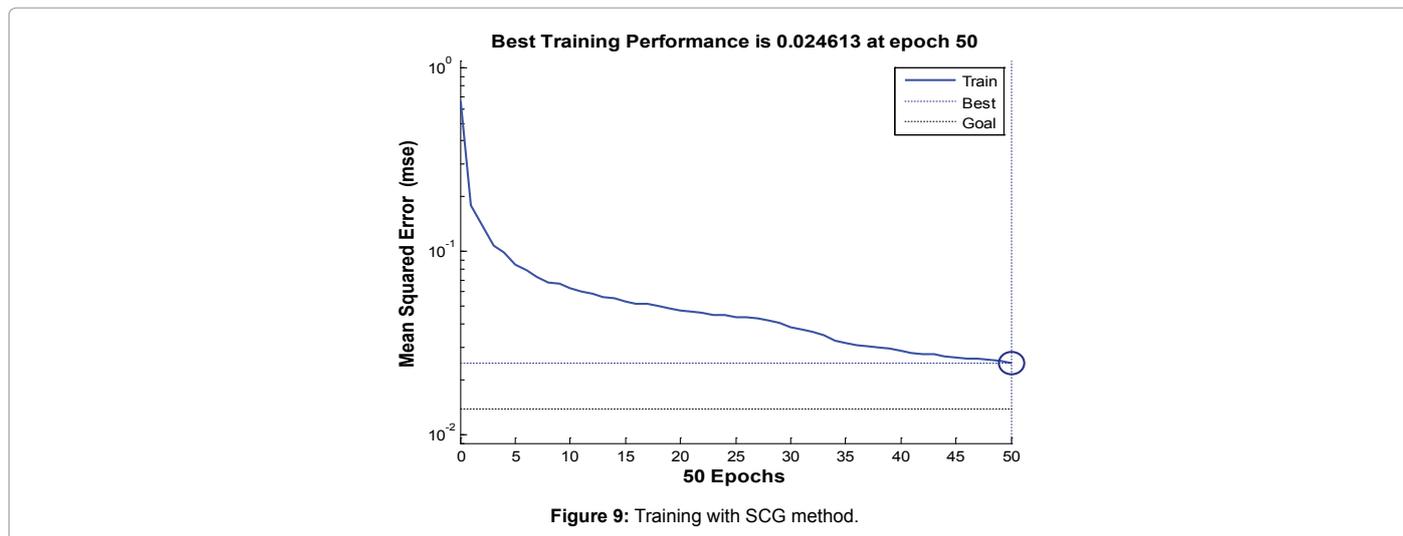
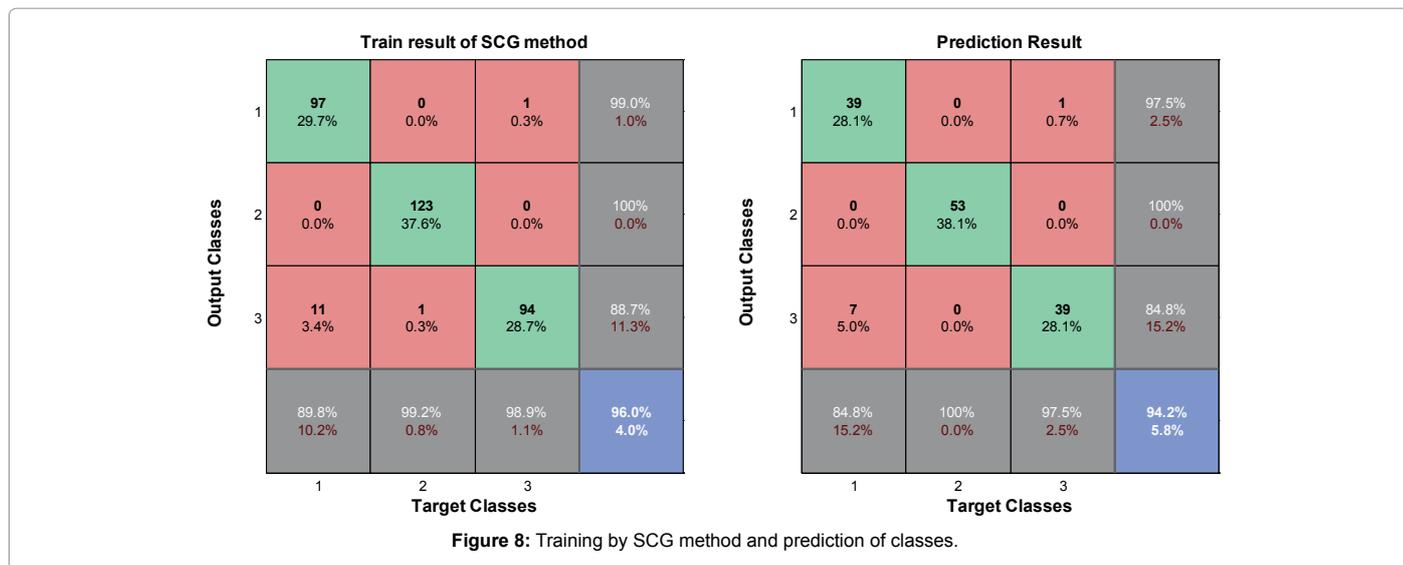


Figure 7: Training performance using OSS method.



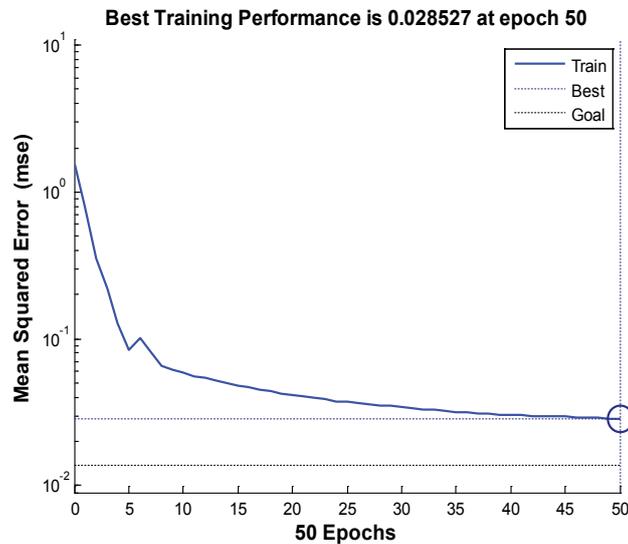


Figure 11: Training with RPROP method.

Conclusion

The several experiments are carried out on datasets of three classes of cardiac arrhythmia. The learning of MLP is performed using IPSO, GDX, RPROP, SCG, OSS learning algorithm. The experimental results are compared with the existing learning algorithms and it shows that evolutionary method such as IPSO outperform GDX, RPROP, SCG, OSS learning algorithm. The IPSO makes MLP to converge faster in very little iterations as compared with other training methods. The IPSO is proved to be used as another alternative learning algorithm for enhancing the ability of MLP to predict the arrhythmia classes. The presented technique increases the prediction capability of MLP using few linear and nonlinear parameters, which are obtained from datasets of normal and abnormal HRV signals. The condition of heart health can be assessed using the proposed hybrid approach.

References

- Özbay Y, Karlik B (2001) A recognition of ECG arrhythmias using artificial neural networks. *Proceedings of the 23rd Annual Conference, IEEE/EMBS, Istanbul, Turkey.*
- Habboush I, Moody GB, Mark RG (1992) Neural networks for ECG compression and classification, pp: 185-188.
- Saini I, Saini BS (2012) Cardiac arrhythmia classification using error back propagation method. *Int J Comput Theor Eng* 4: 462-464.
- Franklin S, Wallcave J (2013) Cardiac condition detection using artificial neural networks. Ph.D. Thesis, California Polytechnic State University.
- Deshmukh R, Patil AJ (2012) Layered approach for ECG beat classification utilizing neural network. *Int J Eng Res Appl* 2: 1495-1500.
- Ozbay Y, Ceylan R, Karlik B (2006) A fuzzy clustering neural network architecture for classification of ECG arrhythmias. *Comput Biol Med* 36: 376-388.
- Wang JS, Chiang WC, Hsu YL, Yang YTC (2013) ECG arrhythmia classification using a probabilistic neural network with a feature reduction method. *Neurocomputing* 116: 38-45.
- Dingfei G, Srinivasan N, Krishnan SM (2002) Cardiac arrhythmia classification using autoregressive modeling. *Biomed Eng Online*, pp: 1-12.
- Vuksanovic B, Alhamdi M (2013) AR-based method for ECG classification and patient recognition. *Int J Biom Bioinformatics* 7: 74-92.
- Yaghoubi F, Ayatollahi A, Soleimani R (2009) Classification of cardiac abnormalities using reduced features of heart rate variability signal. *World Appl Sci J* 6: 1547-1554.
- Acharya UR, Joseph KP, Kannathal N, Lim CM, Suri JS (2006) Heart rate variability: A review. *Med Biol Eng Comput* 44: 1031-1051.
- Acharya R, Kumar A, Bhat PS, Lim CM, Iyengar SS (2004) Classification of cardiac abnormalities using heart rate signals. *Med Biol Eng Comput* 42: 288-293.
- Anuradha B, Reddy VCV (2008) Cardiac arrhythmia classification using fuzzy classifiers. *J Theor Appl Inf Technol* 353-359.
- Asl BM, Sharafat AR, Setarehdan SK (2012) An adaptive back propagation neural network for arrhythmia classification using R-R interval signal. *Neural Network World* 6: 535-548.
- Dallali A, Kachouri A, Samet M (2011) Classification of cardiac arrhythmia using WT, HRV and Fuzzy C-means clustering. *Signal Process* 5: 101-109.
- Asl BM, Setarehdan SK (2006) Neural network based arrhythmia classification using heart rate variability signal. *European Signal Processing Conference.*
- Kelwade JP, Salankar SS (2015) Prediction of cardiac arrhythmia using artificial neural network. *Int J Comput Appl* 115: 30-35.
- Kelwade JP, Salankar SS (2016) Radial basis function neural network for prediction of cardiac arrhythmias based on heart rate time series. *IEEE Int conference on control, measurement and instrumentation.* 454-458.
- Goshvarpour A (2012) Classification of heart rate signals during meditation using Lyapunov exponents and entropy. *Int J Intell Syst Appl* 2: 35-41.
- Kampouraki A, Manis G, Nikou C (2009) Heartbeat time series classification with support vector machines. *IEEE Trans Inf Technol Biomed* 13:512-518.
- Rawther NN, Cheriyan J (2015) Detection and classification of cardiac arrhythmias based on ECG and PCG using temporal and wavelet features. *Int J Adv Res Comput Commun Eng* 4: 474-479.
- Asl BM, Setarehdan SK, Mohebbi M (2008) Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. *Artif Intell Med* 44: 51-64.
- Singhal S, Wu L (1989) Training multilayer perceptrons with the extended kalman algorithm. *Bell Commun Res Inc.*, pp: 133-140.
- Sarkalehm MK, Shahbahrami A (2012) Classification of ECG arrhythmias using discrete wavelet transform and neural networks. *Int J Comput Sci Eng Appl* 2: 1-13.
- Suykens JAK, Vandewalle J (1999) Training multilayer perceptron classifiers based on a modified support vector method. *IEEE T Neural Networ* 10: 907-911.

26. Tzikas D, Likas A (2010) An incremental Bayesian approach for training multilayer perceptrons. *ICANN 2010, Part I, LNCS 6352*, pp: 87-96.
27. Ni J, Song Q (2006) Dynamic pruning algorithm for multilayer perceptron based neural control systems. *Neurocomputing* 69: 2097-2111.
28. Battiti R (1992) First and second order methods for learning: Between steepest descent and Newton's method. *Neural Comput* 4: 141-166.
29. Riedmiller M, Braun H (1993) A direct adaptive method for faster back propagation learning: The RPROP algorithm. *Proceedings of IEEE International Conference on Neural Networks* 1: 586-591.
30. Moller MF (1993) A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks* 6: 525-533.
31. Nasir A, Mashor MY, Hassan R (2013) Classification of acute leukemia cells using multilayer perceptron and simplified Fuzzy ARTMAP neural networks. *Int Arab J Inf Technol* 10: 356-364.
32. Sut N, Celik Y (2012) Prediction of mortality in stroke patients using multilayer perceptron neural networks. *Turk J Med Sci* 42: 886-893.
33. Abid S, Fnaiech F, Jervis BW, Cheriet M (2005) Fast training of multilayer perceptron's with a mixed norm algorithm. *Int Joint Conference on Neural Networks*, pp: 1018-1022.
34. Hagan MT, Demuth HB, Beale MH (1996) *Neural Network Design*, Boston, MA: PWS Publishing.
35. Seiffert U (2001) Multiple layer perceptron training using genetic algorithms. *Proceedings of European symposium on artificial neural networks*, pp: 159-164.
36. Joy U (2011) Comparing the performance of back propagation algorithm and genetic algorithms in pattern recognition problems. *Int J Comput Inf Syst* 2: 7-11.
37. Ghanou Y, Bencheikh G (2016) Architecture optimization and training for the multilayer perceptron using ant system. *Int J Comput Sci* 43.
38. Shah H (2014) An improved artificial bee colony algorithm for training multilayer perceptron in time series prediction. PhD Thesis, University Tun Hussein Onn Malaysia.
39. Chen JF, Do QH, Hsieh HN (2015) Training artificial neural networks by a hybrid PSO-CS algorithm. *Algorithms* 8: 292-308.
40. Kennedy J, Eberhart R (1995) Particle swarm optimization. *IEEE International Conference on Neural Networks Proceedings*, Perth, Australia.
41. Karwowski J, Okulewicz M, Legierski J (2013) Application of particle swarm optimization algorithm to neural network training process in the localization of the mobile terminal. *Proceedings of International Conference on Engineering Applications in Neural Networks*, pp: 122-131.
42. Pu X, Fang Z, Liu Y (2007) Multilayer perceptron networks training using particle swarm optimization with minimum velocity constraints. *Proceedings of International Symposium on Neural Networks*, pp: 237-245.
43. Ratnaweera, Halgamuge SK, Watson HC (2004) Self-organizing hierarchical particle swarm optimizer with time varying acceleration coefficients. *IEEE Trans on Evolutionary Computing*. 8: 240-255.
44. Sun Y, Lang M, Wang D, Liu L (2014) A PSO-GRNN model for railway freight volume prediction: Empirical study from China. *J Ind Eng Manage* 7: 413-433.
45. Abdullah MN, Bakar AH, Rahim NA, Mokhlis H, Illias HA, et al. (2014) Modified particle swarm optimization with time varying acceleration coefficients for economic load dispatch with generator constraints. *J Electr Eng Technol* 9: 15-26.
46. Khokhar B, Parmar KPS, Dahiya S (2012) An efficient particle swarm optimization with time varying acceleration coefficients to solve economic dispatch problem with valve point loading. *Energy* 43: 74-80.