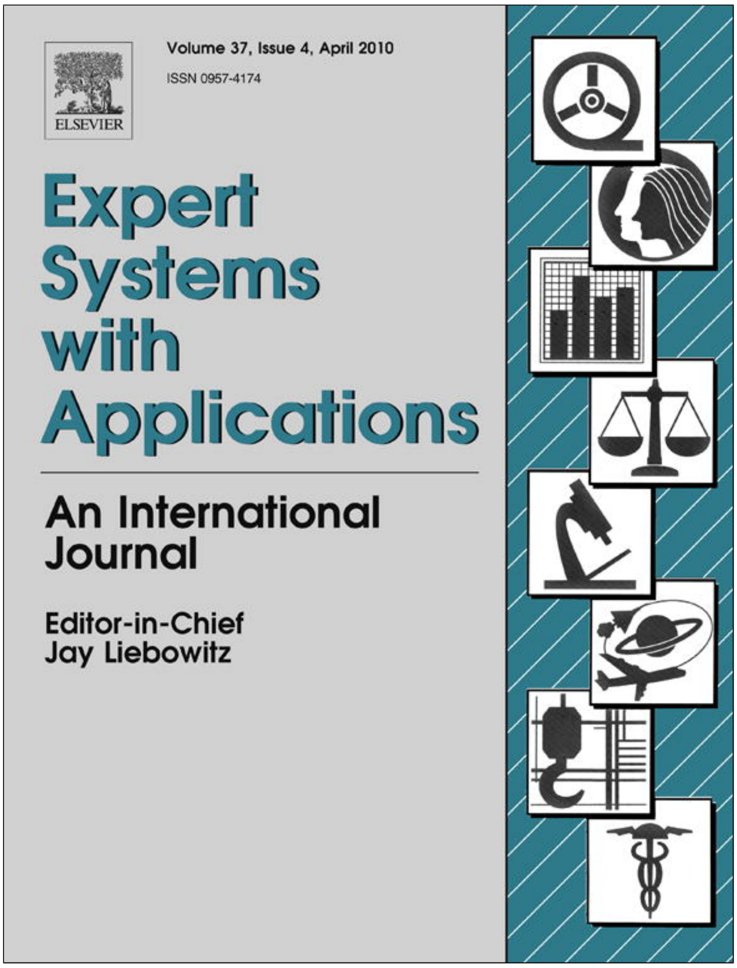
Volume 37, Issue 4, April 2010
ISSN 0957-4174
ELSEVIER
Expert Systems with Applications
An International Journal
Editor-in-Chief
Jay Liebowitz

Expert Systems with Applications 37 (2010) 3088–3093

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A qualitative comparison of Artificial Neural Networks and Support Vector Machines in ECG arrhythmias classification

Majid Moavenian, Hamid Khorrami *

Department of Mechanical Engineering, Ferdowsi University of Mashhad, Iran

ARTICLE INFO

Keywords:
ANN
SVM
ECG

ABSTRACT

In this paper, a novel use of Kernel–Adatron (K–A) learning algorithm to aid SVM (Support Vector Machine) for ECG arrhythmias classification is proposed. The proposed pattern classifier is compared with MLP (multi-layered perceptron) using back propagation (BP) learning algorithm. The ECG signals taken from MIT-BIH arrhythmia database are used in training to classify 6 different arrhythmia, plus normal ECG. The MLP and SVM training and testing stages were carried out twice. They were first trained only with one ECG lead signal and then a second ECG lead signal was added to the training and testing datasets. The aim was to investigate its influence on training and testing performance (generalization ability) plus time of training for both classifiers. Implementation of these three criteria for evaluation of ECG signals classification will ease the problem of structural comparisons, which has not been given attention in previous research works. The results indicate that SVM in comparison to MLP is much faster in training stage and nearly seven times higher in performance, but MLP generalization ability in terms of mean square error is more than three times less. The proposed SVM method shows considerable improvement in comparison to recently reported results obtained by Osowski et al. (2008).

1. Introduction

Accurate diagnosis of heart disease has obviously always been of high importance. The use of ElectroCardioGram (ECG) aids the diagnosis of patient's problems. ECG signals is the electrical record of heart beats in relation to time (see Fig. 1), detecting any arrhythmia. One of the problems though with ECG signal classification is that in some cases patients with identical defects may not have completely similar ECG wave form signal or as in other cases two various diseases may have nearly the same effect on ECG signals, hence complicating the defects diagnosis. This has encouraged many researchers to continue their efforts to obtain a more precise diagnostic system for contributing to the clinical applications (Acir, 2006). Several computer based diagnostic systems have been designed using Artificial Neural Network (ANN) with various structural designs. Improved ANN techniques using data reduction and feature extraction methods (Ceylan & Ozbay, 2007, 2009; Dokur & Olmez, 2001; Engin, 2004; Gholam Hosseini & Luo, 2006; Guler & Ubeyli, 2005a, 2005b; Mohammadzadeh Asl & Setarehdan, 2008; Yu & Chen, 2007; Yu & Chou, 2007, 2009) have been presented in recent years (see Table 1).

It is stated that MLP is capable of recognizing and classifying ECG signals more accurately than other methods of ANN. However, MLP with back propagation (BP) training algorithm suffers from slow convergence to local and global minima and from random settings of initial values of weights (Ozbay, Ceylan & Karlik, 2006).

In order to solve this problem we have proposed use of SVM classifier with (K–A) training algorithm. SVM classifiers do not trap in local minima points and need less training input; therefore they are faster than ANN (Abe, 2005). Some researches have used various methods of SVM for ECG signals classification, some of them have improved SVM classification results by using both data reduction and feature extraction methods (Acir, 2006; Mohammadzadeh Asl & Setarehdan, 2008; Osowski et al., 2008; Polat & Akdemir, 2008; Polat & Gunes, 2007; Ubeyli, 2007, 2008; Yu & Chou, 2009; Zhang & Zhang, 2005) (see Table 2).

In this paper we first use one lead ECG signals (II) that contains six different arrhythmias together with normal ECG signal for training the multi-layered perceptron using BP algorithm. We also use this ECG dataset for training support vector machine supported by (K–A) algorithm. In the second part we use two lead ECG signals II and V1 (see Fig. 2) for MLP and SVM training in order to find out if there is any advantage when using extra lead. Finally we compare the application of MLP and SVM tests carried out in this work and also with recently reported SVM test results in the same area. It is worth mention that comparison is carried out considering three criteria (Training and Testing Performance (Tr.P and Te.P)

* Corresponding author. Tel.: +98 9155515832; fax: +98 5118763304.
E-mail addresses: moaven@ferdowsi.um.ac.ir (M. Moavenian), khorrami.hamid@gmail.com (H. Khorrami).

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.09.021

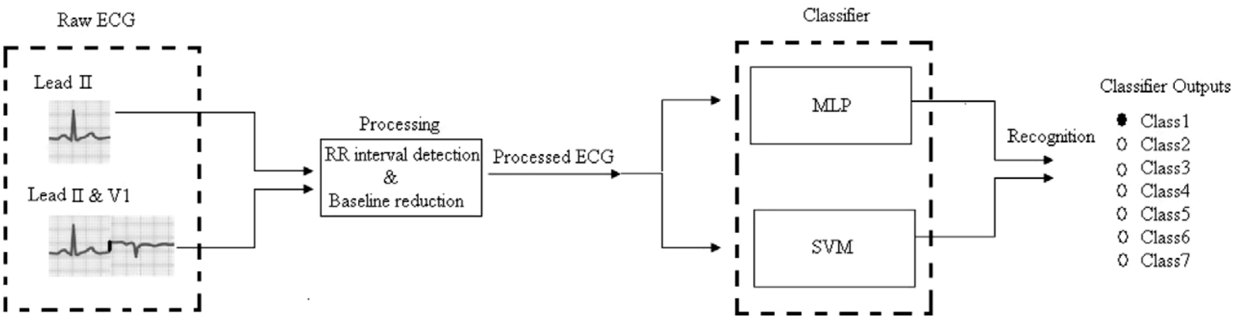

Fig. 1. Classification stages.

Table 1
Researches used MLP for ECG arrhythmias classification.

Structure	# of arrhythmias	Training accuracy (%)	Researcher	
DFT-MLP	10	78	Dokur and Ölmez	1
DWT-MLP	10	96	Dokur and Ölmez	2
(DWT+ARM+high-order Cumulant)-Fuzzy Hybrid	4	98	Engin	3
DWT-MLP	4	96.94	Guler and Ubeyli	4
(DWT+Lyapunov Exponent)-modified mixture of experts	5	97.98	Guler and Ubeyli	5
FCM-PCA-MLP	10	99	Ceylan and Ozbay	6
NET-BST	6	93	Gholam Hosseini, Luo and Reynolds	7
ICA-MDC(Edm)	6	98.29	Yu and Chou	8
ICA-MDC(Mdm)	6	99.42	Yu and Chou	9
ICA-BmC	6	99.51	Yu and Chou	10
DWT-PNN	6	99.65	Yu and Chen	11
T2FCM-MLP	10	99	Ceylan, Ozbay and Karlik	12
ICA-PNN	8	98.7	Yu and Chou	13
MLP	6	98.22	Mohammadzadeh Asl and Setarehdan	14
PCA-MLP	6	96.93	Mohammadzadeh Asl and Setarehdan	15
LDA-MLP	6	98.10	Mohammadzadeh Asl and Setarehdan	16
GDA-MLP	6	98.49	Mohammadzadeh Asl and Setarehdan	17

Table 2
Researches used SVM for ECG arrhythmias classification.

Structure	# of arrhythmias	Training accuracy (%)	Researcher	
PCA-SVM	4	99.17	Zang and Zang	1
SVM	4	89.1	Acir	2
DCT-SVM	4	96.5	Acir	3
DWT-SVM	4	94	Acir	4
PCA-LSSVM	15	100	Polat and Gunes	5
ICA-SVM	8	98.7	Yu and Chou	6
SVM	6	98.49	Mohammadzadeh Asl and Setarehdan	7
PCA-SVM	6	97.65	Mohammadzadeh Asl and Setarehdan	8
LDA-SVM	6	98.06	Mohammadzadeh Asl and Setarehdan	9
GDA-			SVM(OAA)	6
99.16			Mohammadzadeh Asl and Setarehdan	10
SVM	2	99.44	Ubeyli	11
SVM	7	98.04	Osowski and Markeiwicz	12

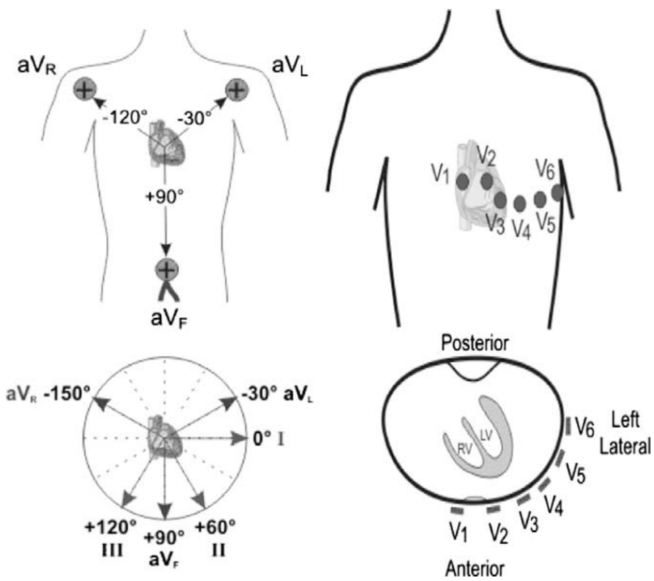

Fig. 2. Record position of ECG leads on patients body.

plus Training Time (Tr.T)) where as most previous experiments only focused on Tr.P. Addition of Tr.P and Tr.T, increases the capability of qualitative evaluation for the selection of diagnostic system.

2. Materials and preprocessing

The ECG signals for training and testing datasets are taken from MIT-BIH arrhythmia database which contained two lead ECG signals of 48 patients. The selected arrhythmias are LBBB (Left Bundle Branch Block), RBBB (Right Bundle Branch Block), PAB (Premature Atrial Beat), PVB (Premature Ventricular Beat), PB (Paced Beat) and FB (Fusion of paced and normal Beat). Ninety beats were chosen for each arrhythmia and normal ECG divided into three groups of: training (50 beats), validation (30 beats) and testing (10 beats) data (see Table 3). Each ECG beat is a matrix (334×1) when one

Table 3
Number of training, testing and validation data in first and second types of datasets.

	LBBB	RBBB	Normal	PVB	PAB	FB	PB	Total
# Of training data beats	50	50	50	50	50	50	50	350
# Of validation data beats for MLP	30	30	30	30	30	30	30	210
# Of testing data beats	10	10	10	10	10	10	10	70
Total	90	90	90	90	90	90	90	630
MIT-BIH data file	111–207–214–109	118–207–212–231	101–105–209–234	107–108–109–119–200–203–207–223–233	118–200–201–202–207–209–	217	107	

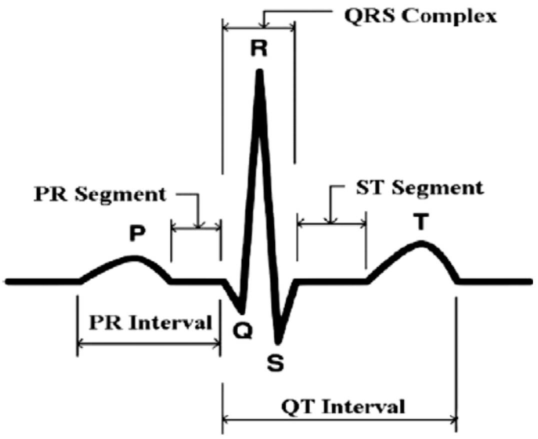

Fig. 3. Standard ECG beat.

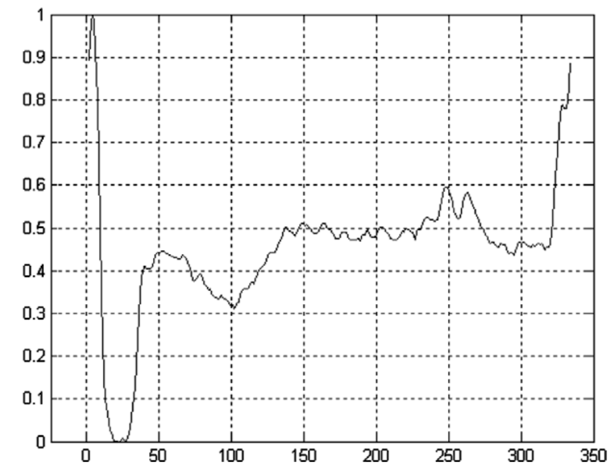

Fig. 4. An ECG signal of RBBB arrhythmia which its baseline is rejected.

ECG lead (II) is used and a matrix (668×1) when two ECG leads (II and V1) are used. Every ECG signal has five distinct points (P, Q, R, S and T) used for the interpretation of the ECG (Fig. 3). Every R–R, interval duration was considered as a beat in the study. Because no-fixed ECG base line exists for individual patients, we located every beat from zero to one vertical scale for better arrhythmias classification (see Fig. 4).

3. Multi-layered perceptron

In our study, a three-layered feed-forward neural network was trained, using (BP) algorithm. The (BP) training algorithm with generalized delta learning rule is an iterative gradient algorithm designed to minimize the mean square error between the actual output of a multi-layered feed-forward neural network and a desired output. Each layer is fully connected to the previous layer, and has no other connection.

3.1. Backpropagation algorithm (summary)

Given a finite length input patterns $x_1(k), x_2(k), \ldots, x_n(k) \in \Re$, $(1 \leqslant k \leqslant K)$ and the desired patterns $x_1(k), x_2(k), \ldots, x_m(k) \in \Re$,

Step 1: Select the total number of layers M, the number n_i ($i = 1, 2, \ldots, M-1$) of the neurons in each hidden layer, and an error tolerance parameter $\varepsilon > 0$.
Step 2: Randomly select the initial values of the weight vectors $w_{aj}^{(i)}$ for $i = 1, 2, \ldots, n_i$.
Step 3: Initialization:

$$w_{aj}^{(i)} \leftarrow w_{aj}^{(i)}(0), \quad E \leftarrow 0, \quad k \leftarrow 1$$

Step 4: Calculate the neural outputs

$$\begin{cases} s_j^{(i)} = (w_{aj}^{(i)})^T x_a^{(i-1)} \\ x_j^{(i)} = \sigma(s_j^{(i)}) \end{cases} \tag{1}$$

For $i = 1, 2, \ldots, M$ and $j = 1, 2, \ldots, n_i$.

Step 5: Calculate the output error

$$e_j = d_j - x_j^{(M)} \tag{2}$$

for $j = 1, 2, \ldots, m$

Step 6: Calculate the output deltas

$$\delta_j^{(M)} = e_j \sigma'\left(s_j^{(M)}\right) \tag{3}$$

Step 7: Recursively calculate the propagation errors of the hidden neurons

$$e_j^{(i)} = \sum_{l=1}^{n_i+1} \delta_l^{(i+1)} w_{lj}^{(i+1)} \tag{4}$$

From the layer $M-1, M-2, \ldots,$ to layer 1.

Step 8: Recursively calculate the hidden neural delta values:

$$\delta_j^{(i)} = e_j \sigma'\left(s_j^{(i)}\right) \tag{5}$$

Step 9: Update weight vectors

$$w_{aj}^{(i)} = w_{aj}^{(i)} + \eta \delta_j^{(i)} x_a^{(i-1)} \tag{6}$$

Step 10: calculate the error function

$$E = E + \frac{1}{k}\sum_{j=1}^{m} e_j^2 \tag{7}$$

Step 11: if $k = K$ then go to step 12; otherwise, $k \leftarrow k+1$ and go to step 4.
Step 12: if $E \leqslant \varepsilon$ then go to step 13; otherwise go to step 3.
Step 13: learning is completed. Output the weights (Gupta, Jin, & Homma, 2003).

After completing the training procedure of the neural network, the weights of MLP are frozen and MLP is made ready for testing stage. MATLAB software is employed to run structures using MLP with BP algorithm.

4. Support vector machines

A special forms of ANNs are SVMs, introduced by Boser, Guyon and Vapnik in 1992. The SVM performs classification by non-linearly mapping their n-dimensional input into a high dimensional feature space. In this high dimensional feature space a linear classifier is constructed. Doing the explicit mapping would be computationally unreasonable, and the algorithm avoids that by introducing the kernel, which is possible since the algorithm only uses the scalar product of the inputs. From this the classification problem is translated into a convex quadratic optimization problem, which due to its convexity has a unique solution.

The simplest version of a SVM is the so-called Maximal Margin Classifier. It works only for data which are linearly separable. It is a good start for understanding the basic ideas behind more sophisticated SVMs. Consider a linearly separable dataset $\{(X_i, d_i)\}$, where X_i is the input pattern for the ith example and d_i is the corresponding desired output $\{-1, 1\}$. The assumption, the dataset is linearly separable, means there exist a hyperplane working as the decision surface. We can write:

$$\begin{aligned} &W^T X_i + b \geqslant 0 \quad then, \quad d_i = +1 \\ &W^T X_i + b \leqslant 0 \quad then, \quad d_i = -1 \end{aligned} \tag{8}$$

where $W^T X + b$ is the output function. The distance from the hyperplane to the closest point is called the geometric margin. The idea is, in order to have a good machine, we want the geometric margin to be maximized. To get that, we first introduce the functional marginal $W^T X + b$. Because the dataset is linearly separable we can rewrite Eq. (8) as follow:

$$\begin{aligned} &W^T X^+ + b = +1 \\ &W^T X^- + b = -1 \end{aligned} \tag{9}$$

where $X^+(X^-)$ is the closest data point on the positive (negative) side of the hyperplane. Now it is straight forward to compute the geometric margin

$$\begin{aligned} \gamma &= \frac{1}{2}\left(\frac{W^T X^+ + b}{|w|} - \frac{W^T X^- + b}{|w|}\right) \\ &= \frac{1}{2|w|}\left(W^T X^+ + b - W^T X^- - b\right) = \frac{1}{2|w|}(1-(-1)) = \frac{1}{|w|} \end{aligned} \tag{10}$$

Hence, equivalent to maximize the geometric margin is fixing the functional margin to one and minimizing the norm of the weight vector, $|w|$.

This can be formulated as a quadratic (ww^T) problem with inequality constraints $d_i(w^T x_i + b) \geqslant 1$.

$$\begin{aligned} &\min : \frac{1}{2} W^T W \ (quadratic - problem) \\ &\text{subject to} : d_i\left(w^T x_i + b\right) \geqslant 1 \end{aligned} \tag{11}$$

By the use of Lagrange multipliers $\alpha_i \geqslant 0$ the original problem is transformed into the dual problem. From the Kuhan–Tuker theory we have the following condition

$$\alpha_i[d_i(W^T x_i + b) - 1] = 0 \tag{12}$$

Which means only the points with functional margin unity contributes to the output function. These points are called the support vectors. Since they are supporting, the separating hyperplane. For more information about SVM classifying, non-separable datasets and classifying more than two classes, see (Abe, 2005).

4.1. Kernel–Adatron algorithm (summary)

Support Vector Machines work by mapping training data for classification tasks into a high dimensional feature space. In the feature space they then find a maximal margin hyperplane which separates the data. This hyperplane is usually found using a quadratic programming routine which is computationally intensive and non trivial to implement. In this section we briefly explain the (K–A) algorithm for SVM classification. The algorithm is simple and can find rapid solution for SVM classification with an exponentially fast rate of convergence (in the number of iterations) towards the optimal solution as follows:

Step 1: Initialize Lagrangian parameters $\alpha_i = 1$.
Step 2: Starting from pattern $i = 1$, for labeled points $\{(x_i, y_i)\}$ calculates.

$$z_i = y_i \sum_{j=1}^{p} \alpha_i y_i K(x_i, x_j) \tag{13}$$

Step 3: For all patterns i calculate

$$\gamma_i = y_i z_i \tag{14}$$

And execute steps 4–5 below.

Step 4: Let

$$\delta\alpha^i = \eta(1 - \gamma^i) \tag{15}$$

Be the proposed change to the multipliers α^i.

Step 5.1: If $(\alpha^i + \delta\alpha^i) \leqslant 0$ then the proposed change to the multipliers would result in a negative α^i.

Consequently to avoid this problem we set $\alpha^i = 0$.

Step 5.2: If $(\alpha^i + \delta\alpha^i) > 0$ then the multipliers are updated through the addition of the $\delta\alpha^i$ i.e. $\alpha^i \leftarrow \alpha^i + \delta\alpha^i$.
Step 6: Calculate the bias b from

$$b = \frac{1}{2}\left(\min(z_i^+) + \max(z_i^-)\right) \tag{16}$$

where z_i^+ are those patterns i with class label +1 and z_i^- are those with class label -1.

Step 7: If a maximum number of presentations of the pattern set has been exceeded then stop, otherwise return to step 2.

Table 4
Some conventional kernels.

Kernel function	Type of classifier
$K(x, x_i) = \exp(-\gamma\|x - x_i\|^2)$	Gaussian radial basis function (RBF)
$K(x, x_i) = (x^T x_i + 1)^d$	Polynomial of degree d
$K(x, x_i) = \tanh(x^T x_i - \theta)$	Multi layer perceptron

The kernel $K(x, x')$ can be any function satisfying Mercer's condition; in particular it is possible to use RBF or polynomial kernels (Abe, 2005). Some conventional kernels are introduced in Table 4.

5. Structure

In this study, two different structures were formed for classification of ECG arrhythmias given in Table 3 as follows:

A. In the first structure MLP with BP training algorithm classifier has been trained and tested using dataset designations shown in Table 3. In training a classifier, the aim is to maximize classification performance for the training data. But if the classifier is too fit for the training data, the classification (generalization) ability for test data is degraded. This phenomenon is called *overfitting* (Abe, 2005). MLP *overfitting* problem during learning is avoided using 210 out of 630 selected ECG beats as validation data in this research.

Learning or training of MLP has been done for two types of datasets, (1 and 2) assigning 350 beats for training, 210 for validation and 70 for testing. In order to find the best structure for utilization of MLP we have calculated Tr.P with respect to the number of neurons in the middle layer (see Fig. 5). MLP with tree layers containing 11, 80 and 7 neurons, respectively, was found to be the best.

B. In second structure SVM with (K–A) training algorithm classifier have been trained with two types of training datasets according to Table 3. In this structure we have utilized RBF kernel for mapping datasets into a high dimensional feature space. Since the number of classes (arrhythmias) for classification are more than two (seven classes used), we have used one-against-all method, see (Abe, 2005), for SVM classification. Also, SVMs are motivated by the concept of training and using only those inputs that are near the decision surface (This provides the most information concerning the classification).

6. Training and test performances

Training and test performances are calculated and presented in Table 5 and 6 using Eq. (17)

$$MSE = \left(\frac{\sum_{j=1}^{P} \sum_{i=1}^{N} (d_{ij} - y_{ij})^2}{NP} \right) \tag{17}$$

where P = number of sample points in each beat, N = number of beats in input matrix, d_{ij} = desired output of classifier for jth sample point and ith beat, y_{ij} = real output of classifier for jth sample point and ith beat. MSE = mean square error.

7. Numerical experiments

All tests carried out in this work were organized in three parts. In the first part, tests were carried out with use of MLP dealing with two sets of data (dataset 1 and 2). The results show that Te.P and Tr.T of MLP dealing with dataset 2 is about 50% and 20%, respectively more than MLP dealing with dataset 1 (This means that use of dataset 1 is preferred in these cases to classify arrhythmias). However, the Tr.T in MLP dealing with dataset 1 is 52% more than MLP dealing with dataset 2 (This means that use of dataset 1 is not in this case preferred to classify arrhythmias).

In the second part, tests were carried out, using SVM dealing with two sets of data (dataset 1 and 2) .The procedure was:

First keep Tr.P constant, as it was for the test when using MLP (see Table 6, rows 1 and 3).

Second keep the Tr.T constant, respectively (see Table 6, rows 2 and 4). Considering Table 6(row 1) for SVM,Tr.P in a few number of iterations (only 8) reaches the same level as shown in Table 5 (row 1), having Tr.T more than 80 times less than MLP for the same condition. Also when Tr.T is kept constant (Table 6 row 2), then Tr.P reduces about seven times with only 217 iterations, but Te.P increases. Table 6 (rows 3 and 4) shows that similar results are obtained, when comparing SVM and MLP using dataset type 2. The results show that Te.P and Tr.T of SVM dealing with dataset 2 is about 6% and 20%, respectively more than MLP dealing with dataset 1. However, the Tr.T in MLP dealing with dataset 1 is 7% more than MLP dealing with dataset 2.

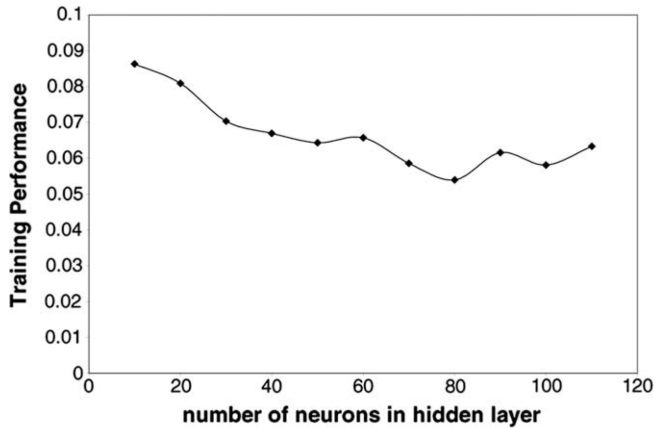

Fig. 5. Tr.P of MLP with respect to the number of neurons in the middle layer of MLP.

Table 5
Performance of training, testing and time of training for ANN.

Structure	Training performance	Training time (min:s)	Number of iterations	Testing performance
MLP (dataset type 1)	0.0539517	03:58	5000	0.0414
MLP (dataset type 2)	0.0355928	04:45	5000	0.0596

Table 6
Performance of training, testing and time of training for SVM.

Structure	Training performance	Training time (min:s)	Number of iterations	Testing performance
SVM (dataset type 1)	0.0539517	00:03	8	–
SVM (dataset type 1)	0.0082264	03:58	217	0.1444
SVM (dataset type 2)	0.0355928	00:18	17	–
SVM (dataset type 2)	0.007656	04:45	207	0.1539

In the third part a comparison is made between MLP and SVM structures using datasets 1 and 2. Table 5 and 6 clearly show that SVM has the best Tr.T and Tr.P, but use of MLP is only suggested when dealing with Te.P.

8. Conclusion

Classification of ECG arrhythmias taken from different and numerous patients (selected from MIT-BIH Arrhythmia Database) due to non-stationary inherent nature of ECG signals, is an applicable way for predicting the existence of arrhythmia in an ECG signal. This paper qualitatively compares two classifiers, MLP with (BP) training algorithm and SVM with (K–A) training algorithm, without employing any data reduction or feature extraction methods with regard to training performance, testing performance (generalization ability) and training time. Tables 5 and 6 strongly suggest that the selected SVM classifier generally could be used when Tr.T and Tr.P were examined and MLP is only preferred for Te.P examination. Introducing three criteria for evaluation of ECG signals will ease the problem of structural comparisons which has not been given attention in previous research works. It is also clarified that use of a second lead could only improve Tr.P which the improvement is about 33% when dealing with MLP and 7% with SVM. This improvement is appropriate when time spent on signal analysis is not of high importance and emergency.

References

Abe, S. (2005). *Support vector machines for pattern classification*. London, NJ: Springer-Verlag.

Acir, N. (2006). A support vector machine classifier algorithm based on perturbation method and its application to ECG beat recognition systems. *Expert Systems with Applications, 31*, 150–158.

Ceylan, R., & Ozbay, Y. (2007). Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network. *Expert Systems with Applications, 33*, 286–295.

Ceylan, R., Ozbay, Y., & Karlik, B. (2009). A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network. *Expert Systems with Applications, 36*, 6721–6726.

Dokur, Z., & Ölmez, T. (2001). ECG beat classification by a novel hybrid neural network. *Computer Methods and Programs in Biomedicine, 66*, 167–181.

Engin, M. (2004). ECG beat classification using neuro-fuzzy network. *Pattern Recognition Letters, 25*, 1715–1722.

Gholam Hosseini, H., Luo, D., & Reynolds, K. J. (2006). The comparison of different feed forward neural network architectures for ECG signal diagnosis. *Medical Engineering and Physics, 28*, 372–378.

Guler, I., & Ubeyli, E. (2005a). A modified mixture of experts network structure for ECG beats classification with diverse features. *Engineering Applications of Artificial Intelligence, 18*, 845–856.

Guler, I., & Ubeyli, E. (2005b). ECG beat classifier designed by combined neural network mode. *Pattern Recognition, 38*, 199–208.

Gupta, M., Jin, L., & Homma, N. (2003). *Static and dynamic neural networks from fundamentals to advanced theory*. NJ: John Wiley & Sons.

MIT-BIH Arrhythmia Database, <http://www.physionet.org./physiobank/database/mitdb/>.

Mohammadzadeh Asl, B., Setarehdan, S. K., & Mohebbi, M. (2008). Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. *Artificial Intelligence in Medicine, 44*, 51–64.

Osowski, S., Markiewicz, T., & Tran Hoai, L. (2008). Recognition and classification system of arrhythmia using ensemble of neural networks. *Measurement, 41*, 610–617.

Özbay, Y., Ceylan, R., & Karlik, B. (2006). A fuzzy clustering neural network architecture for classification of ECG arrhythmias. *Computers in Biology and Medicine, 36*, 376–388.

Polat, K., Akdemir, B., & Güneş, S. (2008). Computer aided diagnosis of ECG data on the least square support vector machine. *Digital Signal Processing, 18*, 25–32.

Polat, K., & Gunes, S. (2007). Detection of ECG Arrhythmia using a differential expert system approach based on principal component analysis and least square support vector machine. *Applied Mathematics and Computing, 186*, 898–906.

Ubeyli, E. (2008). Support vector machines for detection of electrocardiographic changes in partial epileptic patients. *Engineering Applications of Artificial Intelligence, 21*, 1196–1203.

Ubeyli, E. (2007). ECG beats classification using multiclass support vector machines with error correcting output codes. *Digital Signal Processing, 17*, 675–684.

Yu, S., & Chen, Y. (2007). Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network. *Pattern Recognition Letters, 28*, 1142–1150.

Yu, S., & Chou, K. (2007). A switchable scheme for ECG beat classification based on independent component analysis. *Expert Systems with Applications, 33*, 824–829.

Yu, S., & Chou, K. (2009). Selection of significant independent components for ECG beat classification. *Expert Systems with Applications, 36*, 2088–2096.

Zhang, H., & Zhang, L. (2005). ECG analysis based on PCA support vector machine. *International Conference on Neural Networks and Brain, ICNN&B'05, 2*, 743–747.