

Arrhythmia classification using SVM with selected features

Narendra Kohli* and Nishchal K. Verma

Department of Electrical Engineering, Indian Institute of Technology Kanpur, INDIA
**Corresponding Author: e-mail: kohl.hbti@gmail.com, Tel +91-512-2582426, Fax. +91-512-2533412*

Abstract

This paper presents support vector machine based methods for arrhythmia classification in ECG datasets with selected features. Among various existing SVM methods, four well-known and widely used algorithms One Against One (OAO), One Against All (OAA), Fuzzy Decision Function (FDF) and Decision Directed Acyclic Graph (DDAG) are used here to distinguish between the presence and absence of cardiac arrhythmia and classifying them into one of the arrhythmia groups. The various types of arrhythmias in the cardiac arrhythmias ECG database chosen from University of California at Irvine (UCI) to train SVM include ischemic changes (coronary artery disease), old inferior myocardial infarction, sinus bradycardia, right bundle branch block, and others. ECG arrhythmia datasets are of generally complex nature and the results obtained through implementation of four well known methods are thus compared as per their accuracy rate in percentages and the performance of the SVM classifier using One Against All (OAA) technique was found to be of vital importance for classification based diagnosing diseases pertaining to abnormal heart beats. We have also used feature selection on the ECG datasets using Principal Component Analysis (PCA) method. It has been observed that after applying feature selection and then performing classification of ECG datasets through SVM based methods, One Against All (OAA) gives better results than classification without feature selection.

Keywords: Electrocardiogram, Support Vector Machine, Classification, Principal Component Analysis, Arrhythmia.

DOI: <http://dx.doi.org/10.4314/ijest.v3i8.10>

I. Introduction

Among various illness human being has suffered from, heart disease is still one of the important problems in the society. Early detection and proper medical treatment of diseases pertaining to heart can save lives of patients in cases of sudden death (Ozbay and Karlik, 2001, Anuradha and Reddy, 2001). Heart beat is stimulated by electric signal which basically originates from the Sino Atrial node (SA) known as heart's "natural pacemaker" and located at the top of the right chamber or Atrium (RA) in the heart. Any disruption of this pacemaker may cause the heart to beat at an abnormal rate, thereby adversely impacting the circulation of blood throughout the body (Kim et. al., 2008). Arrhythmia is a kind of heart conduction system disease that causes an inefficient heart beat (Song et. al., 2005). It represents abrupt and abnormal ECG beats. In general, arrhythmia is diagnosed by an electrocardiogram procedure. ECG signals are comprised of P wave, QRS complex, and T wave. They are designated by capital letters P, Q, R, S, and T. A typical normal ECG signal is shown in figure 1.

The main parameters included for inspection in heart-patients are the shape, the duration, and the relationship with each other of P wave (during normal atrial depolarization, the main electrical vector is directed from the SA node towards the AV node, and spreads from the right atrium to the left atrium. This turns into the P wave on the ECG), QRS complex (The QRS complex reflects the rapid depolarization of the right and left ventricles. They have a large muscle mass compared to the atria and so the QRS complex usually has a much larger amplitude than the P-wave), and T wave components (QRS complex to the apex of the T wave is referred to as the absolute refractory period. The last half of the T wave is referred to as the relative refractory period) and also R-R interval (The interval between an R wave and the next R wave. Normal resting heart rate is between 60 and 100 bpm). Any changes in these parameters signify an illness of the heart. The entire irregular beat phases are commonly called arrhythmia and some arrhythmias are very dangerous for a patient (Ozbay and Karlik, 2001).

Support Vector Machine based methods have been used to classify the electrocardiogram (ECG) arrhythmias (Kohli et. al., 2010). In classification applications, selecting a proper subset of features is an important task. The goal of feature selection is to reduce the number of features and retain the characteristics necessary for classification. Principal Component Analysis (PCA) is a statistical method used to select features for classification application. Its purpose is to condense the information in a large set of correlated original features (Zhang et. al., 2010). This paper compares the results of support vector machine based methods for arrhythmias classification in ECG datasets with and without feature selection method.

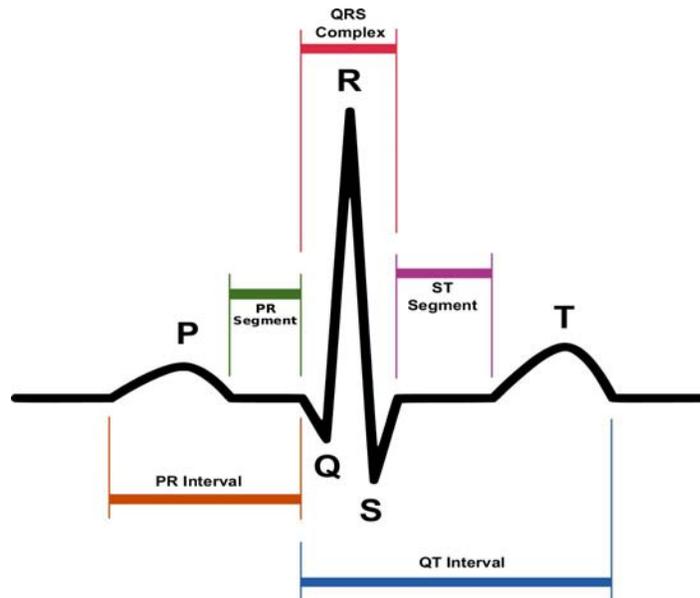


Figure 1: Schematic Representation of Normal ECG

This paper is organized into seven sections. Literature review has been discussed in section two. A brief discussion on the multiclass support vector machine methods is presented in section three. For feature selection, principle component analysis based approach (PCA) has been discussed in section four. Section five presents the experimental evaluations. Section six presents the results of all the classification methods including the results of comparison for arrhythmia classification in ECG on the basis of percent accuracy rate obtained. Finally, the conclusion and future work are drawn in section seven.

2. Literature Review

An ECG facilitates two major kinds of information; firstly, if the time intervals on the ECG are measured, it helps in determining the duration of the electrical wave crossing the heart and consequently we can determine whether the electrical activity is normal or slow, fast or irregular. Secondly, if the amount of electrical activity passing through the heart muscle is measured, it enables a pediatric cardiologist to find out if parts of the heart are too large or are overworked (Anuradha and Reddy, 2001). Thus, physicians diagnose arrhythmia based on long-term ECG data using an ECG recording system. Physicians interpret the morphology of the ECG waveform and decide whether the heartbeat belongs to the normal sinus rhythm or to the class of arrhythmia (Song et. al., 2005). With the various remote and mobile healthcare systems adapting ECG recorders, are being increased in number these days, the importance of a better and robust automatic arrhythmia classification algorithm is being increasingly acknowledged.

The analysis of ECG is basically recognizing its' pattern and classifying arrhythmia in real-time. To date, several researchers have made attempts to use SVM and various other classifiers to classify electrocardiograph (ECG) beats. Many algorithms have been proposed over previous years for developing the automated systems to accurately classify the electrocardiographic signals. (Devine and Macfarlane, 1993) used ANN for the detection of left ventricular strain by classifying ST-T abnormalities of the ECG. In the same year, (Hu et al., 1993) investigated the use of ANN for QRS detection and beat classification. They modeled the non-linear background noise using adaptive multilayer perception structure in order to augment the QRS complexities for its more reliable detection. (Silipo et al., 1996) presented a comparison work for ECG classification using two classification techniques; one with supervised; and other with unsupervised learning. (Yu et al., 1998 and Guyon et al.2003) presented the use of feature selection methods for choosing a number of features among the original features. An obvious advantage of using feature selection is reduction in the time and cost of feature acquisition as well as reduction in classifier training and testing time. Feature selection is also helpful in improving classifier accuracy, provided that noisy, irrelevant or redundant features are eliminated. (Song et al.,

1993) proposed SVM based arrhythmia classification with the reduction of feature dimensions by linear discriminant analysis (LDA). Since, a SVM is known to have the advantage of offering remarkable performance of classification; in this study we have chosen four well-known and most widely used SVM based methods optimized by feature selection for classification of standard arrhythmia dataset and thereby comparing their accuracy rates obtained for best results.

3. Multiclass Classification

Support Vector Machines (SVMs) were originally designed for binary classification. How to effectively extend it for multiclass classification is still an ongoing research issue (Hsu and Lin, 2002, Gunn, 1998). Currently there are two types of approaches for multiclass SVM. One is by constructing and combining several binary classifiers while the other is by directly considering all data in one optimization formulation. The formulation to solve multiclass SVM problems in one step has variables proportional to the number of classes. Therefore, for multiclass SVM methods, either several binary classifiers have to be constructed or a larger optimization problem is needed. Hence in general it is computationally more expensive to solve a multiclass problem than a binary problem with the same number of data. There are four methods for multiclass classification based on binary classification: One Against One (OAO), One Against All (OAA), Fuzzy Decision Function (FDF) and Decision Directed Acyclic Graph (DDAG) SVM.

Assume $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i)\}$ is a training set, where $x_i \in R^m$ and $y_i \in (1, 2, \dots, k)$. For the One Against-One method (Debnath et al., 2004), one needs to determine $k(k-1)/2$ classifiers for the k -classes problems. The optimal hyperplane with SVMs for class i against class j is

$$D_{ij}(x) = w_{ij}^T \phi(x) + b_{ij} = 0, \quad i < j, 1 < j \leq k, 1 \leq i < k \quad (1)$$

where w_{ij}^T is a vector in the feature space, $\phi(x)$ is mapping function, and b_{ij} is a scalar. Here the orientation of the optimal hyper plane is defined as per the following equation:

$$D_{ij}(x) = -D_{ji}(x) \quad (2)$$

One-Against-One Method: for the input vector, one computes

$$D_i(x) = \sum_{j \neq i, j=1}^k \text{sgn}(D_{ij}(x)) \quad (3)$$

and classifies x into the class

$$\arg \max_{i=1, \dots, k} (D_i(x))$$

One-Against-All Method: For a k class problem, the One-Against-All method constructs k SVM models. The i th SVM is trained with all of the training examples in the i th class with positive labels and all other examples with negative labels. The final output of the one-against-all method is the class that corresponds to the SVM with the highest output value (Liu et al. 2008). Thus, by solving the optimization problem of SVM using all the training samples in the dataset, the decision function of the i th SVM is

$$D_i(x) = w_i^T \phi(x) + b_i \quad (4)$$

The input vector x will be assigned to the class that corresponds to the largest value of the decision functions.

Sample x is classified into the class

$$\mathbf{x} = \arg \max_{i=1, \dots, k} (D_i(x)). \quad (5)$$

Fuzzy Decision Function Method: In the FDF method (Liu et al., 2008) for the input vector x , the 1-D membership function $m_{ij}(x)$ ($i, j = 1, 2, \dots, k$) in the directions orthogonal to the optimal separating hyperplanes $D_{ij}(x)=0$ is defined as follows:

$$m_{ij}(x) = \begin{cases} 1, & 1 \leq D_{ij}(x) \\ D_{ij}(x), & \text{otherwise} \end{cases}$$

In (Tsujinishi and Abe, 2003) the membership functions $m_i(x)$ are given by

$$m_i(x) = \min_{j=1, \dots, k} (m_{ij}(x)) \quad (6)$$

Using (6), sample x is classified into the class

$$\arg \max_{i=1, \dots, k} (m_i(x)) \quad (7)$$

Using above three SVM methods few of the instances remain unclassified at the end, thereby creating unclassifiable region and then DDAG method was developed as a solution to this issue.

Decision Directed Acyclic Graph (DDAG): To resolve the unclassifiable region, the DDAG method developed based on the One Against One scheme. At the top-level classification, one can choose any pair of class, and except for the leaf node, if $D_{ij}(x) > 0$, then one can regard x to be not belonging to class 'j'. If $D_{12}(x) > 0$, it means 'x' does not belong to class 2. Thus it belongs to either class 1 or class 3 and the next classification pair is class 1 and class 3. Thus, the unclassifiable region is resolved but it depends on the tree information. DDAG method is as per figures 1 and 2.

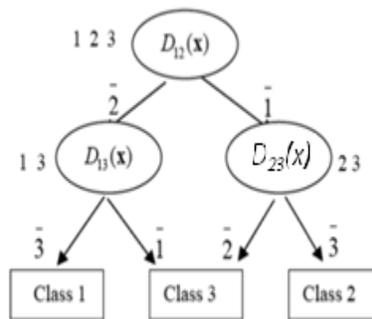


Figure 2: One case of DDAG decision function method based on One-Against-One Scheme (Liu et al., 2008)

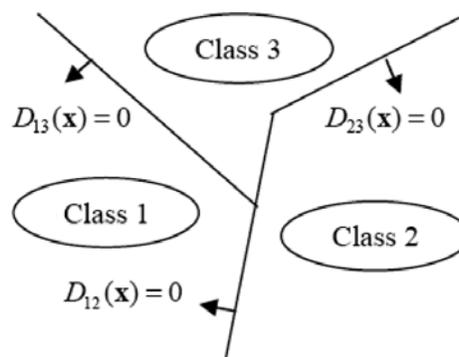


Figure 3: Resolved unclassifiable region based on DDAG Decision Function Method (Liu et al., 2008)

4. Feature Selection

Feature selection is the process of selecting a subset of the features occurring in the training set and using only this subset as features in classification. Feature selection serves two main purposes. First, it makes training and applying a classifier more efficient by decreasing the size of the effective features. Second, feature selection often increases classification accuracy by eliminating noisy features. A noisy feature is one that when added to the document representation, increases the classification error on new data. In our proposed classification technique, we have used Principal Component Analysis (PCA) method for feature selection (Polat and Gunes, 2007, Kohli and Verma, 2010).

Principle Component Analysis based Approach (PCA): PCA is a well-established technique for feature extraction and dimensionality reduction. It is based on the assumption that most information about classes is contained in the directions along which the variations are the largest. These directions are called principal components. Figure 3 shows an example of principal components in case of 2 dimensions where first principal component is denoted by pc1 whereas second one is denoted by pc2. The most common derivation of PCA is in terms of a standardized linear projection which maximizes the variance in the projected space. Principle Component Analysis method is useful for data compression, by reducing the number of dimensions, without much loss of information.

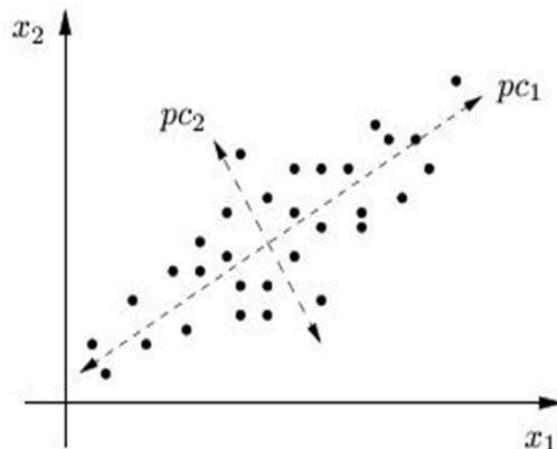


Figure 3: Principle Component Analysis Method (Polat and Gunes, 2007)

5. Experimental Evaluations

The standard multivariate ECG datasets taken here is chosen from the University of California at Irvine (UCI) Cardiac Arrhythmias Database (Murphy and Aha, 2010). The experiments are conducted on a personal computer with 1.5 GHz CPU and 1 GB of RAM. Initially this database contained 452 instances and 279 attributes. But owing to the presence of many missing values and also zero valued columns, it became imperative to preprocess and resize the datasets while maintaining the reliability and relevance of the datasets. Thus, columns containing all zeroes or all missing values were removed first from the datasets, followed by the removal of rows having missing values in the dataset and removal of classes having insignificant number of instances. This resized the datasets to total 377 instances and 166 attributes, distributed into 6 Classes with Class 1 referring to “normal” ECG, Classes 2 to 5 referring to different classes of arrhythmia and Class 6 referring to the rest of unclassified ones. Table 1 shows different arrhythmia classes with number of instances belonging to each of the classes in the datasets.

Table 1: Arrhythmia classes with corresponding number of instances in the dataset

Class	Class name	No. of instances
1	Normal	237
2	Ischemic changes (Coronary Artery Disease)	36
3	Old Inferior Myocardial Infarction	14
4	Sinus bradycardy	24
5	Right bundle branch block	48
6	Others	18

All the SVM based methods taken here were trained by half of the total datasets chosen fairly from the main datasets ensuring representation of all classes present in the required percentage. The remaining half of the main datasets was used for testing and analysis purpose. Table 2 shows the representation of each of the classes in the training and the testing datasets.

Table 2: Representation of each of the classes in the training and the testing datasets

Data	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Total
Initial	237	36	14	24	48	18	377
Train	119	18	7	12	24	9	189
Test	118	18	7	12	24	9	188

6. Results

(a) Without Feature Selection on ECG datasets:

We applied the SVM based methods using Gaussian kernel (Genton et al., 2001, Hsu and Lin, 2002). The kernel parameter σ and the regularization parameter λ were empirically optimized and shown in Table III and Table IV respectively, by minimizing the error rate on the validation datasets in order to obtain the best accuracy rate in terms of percentage and thus ensuring better classification of the arrhythmia datasets. A radial basis function (RBF) is a real-valued function whose value depends only on the distance from the origin, so that $\phi(x) = \phi\|x\|$; or alternatively on the distance from some other point c , called a center, so that $\phi(x, c) = \phi\|x - c\|$. Any function ϕ that satisfies the property $\phi(x) = \phi\|x\|$ is a radial function.

Gaussian function named after Johann Carl Friedrich Gauss is a function of the form $f(x) = ae^{-\frac{(x-b)^2}{2c^2}}$ for some real constants a , b , $c > 0$, and $e \approx 2.718281828$.

For Arrhythmia classification in ECG data with 166 attributes and 377 instances is taken from the Cardiac Arrhythmias Database of University of California at Irvine (UCI). The experiments are conducted on a personal computer with 1.5 GHz CPU and 1 GB of RAM. Initially this database contained 452 instances and 279 attributes. But owing to the presence of many missing values and also zero valued columns, it became imperative to preprocess and resize the datasets while maintaining the reliability and relevance of the datasets. Thus, columns containing all zeroes or all missing values were removed first from the datasets, followed by the removal of rows having missing values in the dataset and removal of classes having insignificant number of instances. This resized the datasets to total 377 instances and 166 attributes, distributed into 6 Classes with Class 1 referring to normal ECG, Classes 2 to 5 referring to different Classes of arrhythmia and Class 6 referring to the rest of unclassified ones.

Table 3: The accuracy rate (in %) wrt σ value with all Six classes

No.	σ	OAA	FDf	DDAG	OAO
1	2^{-4}	62.7660	62.7660	62.7660	62.7660
2	2^{-3}	62.7660	62.7660	62.7660	62.7660
3	2^{-2}	62.7660	62.7660	62.7660	62.7660
4	2^{-1}	62.7660	62.7660	62.7660	62.7660
5	2^0	62.7660	62.7660	62.7660	62.7660
6	2^1	62.7660	62.7660	62.7660	62.7660
7	2^2	62.7660	62.7660	62.7660	62.7660
8	2^3	62.7660	62.7660	62.7660	62.7660
9	2^4	66.4894	62.7660	62.7660	71.2766
10	2^5	73.9362	62.7660	62.7660	68.4894
11	2^6	68.6170	62.7660	62.7660	63.2979
12	2^7	63.8298	62.7660	62.7660	62.7660
13	2^8	62.7660	62.7660	62.7660	62.7660

Table 4: The accuracy rate (in %) wrt λ value at best σ value with all Six classes

No.	λ	OAA	FDf	DDAG	OAO
1	2^{-4}	73.9362	62.7660	62.7660	3.7234
2	2^{-3}	73.9362	62.7660	62.7660	3.7234
3	2^{-2}	73.9362	62.7660	62.7660	3.7234
4	2^{-1}	73.9362	62.7660	62.7660	18.6170
5	2^0	75	62.7660	62.7660	73.9362
6	2^1	74.4681	62.7660	62.7660	71.2766
7	2^2	73.9362	62.7660	62.7660	71.2766
8	2^3	73.9362	62.7660	62.7660	71.2766
9	2^4	73.9362	62.7660	62.7660	71.2766
10	2^5	73.9362	62.7660	62.7660	71.2766
11	2^6	73.9362	62.7660	62.7660	71.2766
12	2^7	73.9362	62.7660	62.7660	71.2766
13	2^8	73.9362	62.7660	62.7660	71.2766

Table 5: The accuracy rate (in %) wrt σ value with 1, 2 and 5 classes

No.	σ	OAA	FDF	DDAG	OA0
1	2^{-4}	73.7500	73.7500	73.7500	73.7500
2	2^{-3}	73.7500	73.7500	73.7500	73.7500
3	2^{-2}	73.7500	73.7500	73.7500	73.7500
4	2^{-1}	73.7500	73.7500	73.7500	73.7500
5	2^0	73.7500	73.7500	73.7500	73.7500
6	2^1	73.7500	73.7500	73.7500	73.7500
7	2^2	73.7500	73.7500	73.7500	73.7500
8	2^3	75	73.7500	73.7500	73.7500
9	2^4	78.1250	73.7500	73.7500	77.5000
10	2^5	76.8750	73.7500	73.7500	75
11	2^6	73.7500	73.7500	73.7500	73.7500
12	2^7	73.7500	73.7500	73.7500	73.7500
13	2^8	73.7500	73.7500	73.7500	73.7500

Table 6: The accuracy rate (in %) wrt λ value at best σ value with 1, 2 and 5 classes

No.	λ	OAA	FDF	DDAG	OA0
1	2^{-4}	75.6250	73.7500	73.7500	11.2500
2	2^{-3}	75.6250	73.7500	73.7500	11.2500
3	2^{-2}	75.6250	73.7500	73.7500	11.2500
4	2^{-1}	78.1250	73.7500	73.7500	74.3750
5	2^0	77.5000	73.7500	73.7500	78.1250
6	2^1	78.1250	73.7500	73.7500	77.5000
7	2^2	78.1250	73.7500	73.7500	77.5000
8	2^3	78.1250	73.7500	73.7500	77.5000
9	2^4	78.1250	73.7500	73.7500	77.5000
10	2^5	78.1250	73.7500	73.7500	77.5000
11	2^6	78.1250	73.7500	73.7500	77.5000
12	2^7	78.1250	73.7500	73.7500	77.5000
13	2^8	78.1250	73.7500	73.7500	77.5000

(b) With Feature Selection on ECG Datasets:

We have performed classification with feature selection for data composed of only Classes 1, 2 and 5. Since number of data samples in Classes 3, 4 and 6 were very less and were not sufficient for training the PCA, we have not considered them in the experiments. After feature selection using PCA of ECG dataset of Cardiac Arrhythmias Database of University of California at Irvine (UCI), we have selected dataset with 321 instances and 70 attributes of classes 1, 2 and 5 for Arrhythmia classification in ECG. Table VII and Table VIII show the classification results when feature selection is used.

Table 7: The accuracy rate (in %) wrt σ value

No.	σ	OAA	FDF	DDAG	OA0
1	2^{-4}	74.6032	74.6032	74.6032	74.6032
2	2^{-3}	74.6032	74.6032	74.6032	74.6032
3	2^{-2}	74.6032	74.6032	74.6032	74.6032
4	2^{-1}	74.6032	74.6032	74.6032	74.6032
5	2^0	76.1905	74.6032	74.6032	82.5397
6	2^1	82.5397	74.6032	74.6032	82.5397
7	2^2	85.7143	74.6032	74.6032	79.3697
8	2^3	85.7143	82.5440	82.5397	74.6032
9	2^4	85.7143	79.3657	82.5397	74.6032
10	2^5	85.7143	74.6032	74.6032	74.6032
11	2^6	85.7143	74.6032	74.6032	74.6032
12	2^7	85.7143	74.6032	74.6032	74.6032
13	2^8	85.7143	74.6032	74.6032	74.6032

Table 8: The accuracy rate (in %) wrt λ value at best σ value

No.	λ	OAA	FDG	DDAG	OAQ
1	2^{-4}	74.6032	74.6032	74.6032	11.1111
2	2^{-3}	74.6032	74.6032	74.6032	11.1111
3	2^{-2}	74.6032	74.6032	74.6032	11.1111
4	2^{-1}	82.5397	76.1902	74.6032	72.4286
5	2^0	83.7125	82.5397	82.5397	82.5397
6	2^1	83.7125	82.5397	82.5397	82.5397
7	2^2	83.7125	82.5397	82.5397	82.5397
8	2^3	83.7125	82.5397	82.5397	82.5397
9	2^4	74.6032	74.6032	74.6032	82.5397
10	2^5	74.6032	74.6032	74.6032	82.5397
11	2^6	74.6032	74.6032	74.6032	82.5397
12	2^7	74.6032	74.6032	74.6032	82.5397
13	2^8	74.6032	74.6032	74.6032	82.5397

Table 3 to 8 presenting the results of various classification methods that are as follows:

For Arrhythmias classification in ECG datasets using SVM methods such as One Against One, One Against All, Fuzzy Decision Function and Decision Directed Acyclic Graph without feature selection, the system was trained and optimized by classifying the datasets at various σ values i.e. $\sigma = [2^{-2}, \dots, 2^3, 2^4]$. After that σ value was chosen and fixed, which gives highest percentages of accuracy rate i.e. $\sigma = 32$ (using 1, 2, 3,4,5,6 classes) as per Table 5.3 and $\sigma = 16$ (using 1, 2, 5 classes) as per Table 5.5. After that the system was converged by classifying it again at various λ values i.e. $\lambda = [2^{-2}, \dots, 2^3, 2^4]$. As per Table 5.4 and Table 5.6, it is found that performance wise One Against All method in ECG classification gives better classification results i.e. $\lambda = 1$, accuracy rate is 75 percent (with all 6 classes) and $\lambda = 16$, accuracy rate is 78.1250 percent (with 3 classes i.e. 1, 2 and 5) than One Against One, Fuzzy Decision Function and Decision Directed Acyclic Graph methods.

With feature selection on ECG datasets, we have done the Arrhythmias classification using SVM methods such as One Against One, One Against All, Fuzzy Decision Function and Decision Directed Acyclic Graph of ECG datasets of 3 classes (1, 2 and 5) with 70 attributes and 321 instances. The system was trained and optimized by classifying the datasets at various σ values i.e. $\sigma = [2^{-2}, \dots, 2^3, 2^4]$. After that σ value was chosen and fixed, which gives highest percentages of accuracy rate i.e. as per Table 5.7, $\sigma = 4$ and accuracy rate is 85.7143 and next the system was converged by classifying it again at various λ values i.e. $\lambda = [2^{-2}, \dots, 2^3, 2^4]$. As per Table 5.8, it is found that performance wise One Against All method in ECG classification gives better classification results i.e. $\lambda = 4$, accuracy rate is 83.7125 percent, than One Against One, Fuzzy Decision Function and Decision Directed Acyclic Graph methods.

(c) Comparison of performance:

The classification results of arrhythmia in ECG database indicate that among various SVM based methods, one-against-all algorithm shows the highest percentage of accuracy rate. The system was trained and optimized by classifying the dataset at various σ values. That σ value was chosen and fixed, which gives highest percentage of accuracy rate; and next, the system was converged by classifying it again at various λ values. The range of values taken for σ and λ values are, $\sigma = [2^{-4}, 2^{-3}, 2^{-2}, \dots, 2^7, 2^8]$ and $\lambda = [2^{-4}, 2^{-3}, 2^{-2}, \dots, 2^7, 2^8]$. It is found that performance wise One Against One algorithm trails behind One Against All method in ECG classification but its results could be competitive at times. FDF and DDAG methods showed poor and constant classification results on ECG dataset. To remove ambiguity over having any potential to improve its performance, DDAG and FDF methods were used for classification with σ value ranging from 1 to 5000 but no improvement in accurate rate was observed. Clearly, OAO method resulted in the highest accuracy rate but in general, the very high accuracy rate is difficult to obtain. This could be due to the presence of a particular class sweeping maximum share of number of instances in the total datasets. This reflects the further potential of OAO method to give even higher results in cases of ECG datasets with more uniform distribution, thereby ensuring better training of the system.

Results are further improved when PCA method is used for feature selection in ECG datasets. Feature selection experiment considers only three classes (namely, Classes 1, 2 and 5), for experimentation as these classes contain the sufficient number of data samples required for the training of the PCA. Table VII and Table VIII shows the classification results when feature selection is used. To fairly compare the performance of classification with feature selection to without feature selection case, we also performed classification experiments by only considering Classes 1, 2 and 5 in without feature selection case. Obtained

classification results for this are shown in Table V and Table VI. It is evident from Tables V, VI, VII and VIII that classification accuracy is much better when feature selection is used.

7. Conclusion and future work

This paper presents classification of arrhythmia in ECG datasets with various popular and widely used SVM based methods. We have applied some well known classification techniques, viz., One Against One, Decision Directed Acyclic Graph, Fuzzy Decision Function and One Against All to distinguish between the presence and absence of cardiac arrhythmia and classify it in one of the six groups. We have also used feature selection method i.e. Principle Component Analysis method (PCA) on ECG datasets (Classes 1, 2 and 5). We found better classification results with feature selection. SVM is a strong classifier and using various methods it can deliver high performance even with smaller learning data (Song et al., 2005). The obtained results clearly indicate the One Against All method to be the most powerful method for arrhythmia classification in ECG.

Acknowledgement

We acknowledge Mr. Abhishek K. Roy, Accenture, Gurgoan for his help in the preparation of paper.

References

- Ozbay Y. and Karlik B. 2001. A Recognition of ECG Arrhythmia Using Artificial Neural Networks, Proceedings – 23rd Annual conference, IEEE/EMBS, Istanbul, Turkey.
- Anuradha B. and Veera V.C. Reddy. Cardiac Arrhythmia Classification Using Fuzzy Classifier. *Journal of Theoretical and Applied Information Technology*.
- Kim J., Shin H.S., Shin K., and Lee M. 2009. Robust Algorithm for Arrhythmia Classification in ECG Using Extreme Learning Machine, *BioMedical Engineering OnLine*, Vol. 8. p. 31.
- Song M.H., Lee J., Cho S.P., Lee K.J., and Yoo S.K. 2005. Support Vector Machine Based Arrhythmia Classification Using Reduced Features, *International Journal of Control, Automation, and Systems*, Vol. 3, No. 4, pp. 571-579.
- Devine B. and Macfarlane P.W. 1993. Detection of Electrocardiographic Left Ventricular Strain using Neural Nets, *Journal of Medical Biological Engineering and Computing*, Vol. 31. No. 4, pp. 343-348.
- Hu Y.H., Tompkins W.J., Urristi J.L., and Valtino X.A. 1993. Application of Artificial Neural Networks for ECG Signal Detection and Classification, *Journal of Electrocardiology*, Vol. 26, pp. 66-73.
- Silipo R., Bortolan G., and Marchesi C. 1996. Supervised and Unsupervised Learning for Diagnostic ECG Classification, *IEEE Eng. in Medicine and Biology*, Vol. 31, pp.1054-59, Amsterdam, The Netherlands.
- Sugiura T., Hirata H., Harada, Y. and Kazui, T. 1998. Automatic discrimination of arrhythmia waveforms using fuzzy logic, *Proceedings of 20th Annual International Conference of IEEE Engineering in Medicine and Biology Society*, Vol. 20, No. 1..
- Acharya U R, Subbanna Bhat P, Iyengar S S, Rao A and Dua S. 2003. Classification of Heart Rate using Artificial Neural Network and Fuzzy Equivalence Relation, *Pattern Recognit.*, Vol. 36, pp. 61–68.
- Liu B., Hao Z., and Tsang E.C.C. 2008. Nesting One-Against-One Algorithm based on SVM's for Pattern Classification, *IEEE Transactions on Neural Networks*, Vol. 19, No. 12, pp. 2044-2052.
- Vapnik V.N.1995. *The Nature of Statistical Learning Theory*. London, U.K.: Springer-Verlag.
- Aizeman M., Braverman E., and Rozonoer L. 1964. Theoretical Foundations of Potential Function Method in Pattern Recognition Learning, *Journal of Automation Remote Control*, Vol. 25, pp. 821–837.
- S. Saitoh S. 1988. *Theory of Reproducing Kernels and Its Applications*. Harlow, U.K.: Longman.
- Abe S., *Support Vector Machines for Pattern Classification*. London, U.K.: Springer-Verlag, 2006.
- Kreßel U.H.G. 1998. *Pairwise Classification and Support Vector Machines in Advances in Kernel Methods- Support Vector Learning*, Cambridge, MA: MIT Press, pp. 255–268.
- Tsujinishi D. and Abe S. 2003. Fuzzy Least Squares Support Vector Machines for Multiclass Problems, *IEEE Transactions of Neural Networks*, Vol. 16, pp. 785–792.
- Debnath R., Takahide N., and Takahashi H. 2004. A Decision based One-Against-One Method for Multi-Class Support Vector Machine, *Pattern Anal. Appl.*, Vol. 7, pp. 164-175.
- Murphy P.M. and Aha D.W. 2010. UCI Machine Learning Repository Database, [Online]. Available: [<http://archive.ics.uci.edu/ml>], Irvine, CA: University of California, School of Information and Computer Science.
- Electrocardiography. <http://www.en.wikipedia.org/wiki/Electrocardiography>.
- Genton Marc G., Nello Cristianini Nello, John Shawe-taylor, and Robert Williamson. 2001. Classes of kernels for machine learning: A statistics perspective. *Journal of Machine Learning Research*, Vol. 2, No. 1, pp. 299-312, 2001.
- Kohli Narendra and Verma Nishchal K. 2010. SVM based Methods for Arrhythmia Classification in ECG. *Proceedings of IEEE International Conference*, Vol. 2, pp. 119-130.
- Polat Kemal and Gunes Salih. 2007. Detection of ECG arrhythmia using a differential expert system approach based on principal

- component analysis and least square support vector machine. *Journal of Applied Mathematics and Computation*, Vol. 186, No. 1, pp.898- 906.
- Chih-Wei Hsu and Chih-Jen Lin. 2002. A comparison of Methods for Multiclass Support Vector Machines. *IEEE Transaction on Neural Networks*, Vol. 13, No. 2, pp. 415- 425.
- Steve R. Gunn. 1998. Support vector machine for classification and regression. Technical report, Image Speech and Intelligent Systems Group, University of Southampton.
- Kohli Narendra and Verma Nishchal K. 2010. SVM based methods optimized by feature selection for arrhythmia classification in ECG. *Proceedings of International Conference Biomedical Engineering and Assistive Technologies, India*.
- Zhang Zhen-Xiay, Lee Sang-Hong and Lim S Joon. 2010. Comparison of Feature Selection Methods in ECG Signal classifications, *ACM*.
- Gaussian function, http://www.en.wikipedia.org/wiki/Gaussian_function.
- Radial basis function, http://www.en.wikipedia.org/wiki/Radial_basis_function.

Biographical notes

Dr. Narendra Kohli was born in India on April 13, 1963. He received M. Sc. Engg. from Kiev university Kiev USSR (earlier) in 1988 and the Ph.D. degree from Indian Institute of Technology Kanpur, India, in 2012. He is currently an Associate Professor and Head with the Department of Computer Science & Engineering, Harcourt Butler Technological Institute Kanpur, India. His research interests include data mining, telemedicine and support vector machine. He is a Member of IE (India), Fellow of IETE (India), and senior member of ISE.

Dr. Nishchal K. Verma was born in India on September 9, 1973. He received the B.E. degree from the Faculty of Engineering, Dayalbagh Educational Institute, Agra, India, in 1996, the M.Tech. degree from the Indian Institute of Technology (IIT) Roorkee, Roorkee, India, in 2003, and the Ph.D. degree from IIT Delhi, New Delhi, India, in 2007, all in electrical engineering. He is currently an Assistant Professor with the Department of Electrical Engineering, IIT Kanpur, India. His research interests include fuzzy systems, neural networks, data mining, fault diagnosis, bioinformatics, color segmentation, video clip or image sequence modeling, machine learning, and computational intelligence. Dr. Verma is a reviewer for several reputed national and international journals and conferences, including the *IEEE TRANSACTIONS ON FUZZY SYSTEMS*, the *IEEE TRANSACTIONS ON SYSTEMS, MAN, and CYBERNETICS: PARTS A, B, AND C*, the *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, and *Pattern Recognition*.

Received December 2011

Accepted January 2012

Final acceptance in revised form March 2012