

# Electrocardiogram Classification Method Based on SVM

Xiao Tang Zhiwen Mo

College of mathematics and software science, Sichuan normal university, Chengdu 610066, P. R. China

## Abstract

Heart disease is one of the main diseases threatening human beings health, and electrocardiogram is the important basis of diagnosing cardiovascular disease. Because of distinctiveness and variability of QRS wave, many present ECG classification techniques are difficult to realize. Although many methods could work successfully in recognizing certain types of ECG signals, the recognition rate usually can not be substantially promoted throughout all kinds of ECG signals. In this paper, 1-vs-rest algorithm of SVM is used for ECG classification. The algorithm for ECG classification is tested with the data of MIT-BIH. Finally a high recognize rate is obtained.

**Keywords:** Support vector machine, Feature extraction, Electrocardiogram classification, CAD

## 1. Introduction

Electrocardiography is an important tool in diagnosing the condition of the heart. The electrocardiogram (ECG) is the record of variation of bioelectric potential with respect to time as the human heart beats. It provides valuable information about the functional aspects of the heart and cardiovascular system. Early detection of heart diseases can prolong life and enhance the quality of living through appropriate treatment. Therefore, there are numerous research and work analyzing the ECG signals. The state of cardiac health is generally reflected in the shape of ECG waveform and heart rate. It may contain important pointers to the nature of diseases afflicting the heart. Since the biosignals are nonstationary signals, this reflection may occur at random in the time scale. In this situation, the disease symptoms may not show up all the time, but would manifest at certain irregular intervals during the day. The recognition of ECG is the main process for the cardiovascular diagnosis. Therefore, the computer auto analysis of ECG remains the research hotspot in medical Engineering. There mainly have two major types to the classification of

QRS wave based on feature and waveforms methods. So, many methods and algorithms have been proposed for the classification of ECG signal (CWA, direct ECG characteristic, Fourier transform, Power spectrum analysis, Complexity measure, TLS(Total Least Squares), Adaptive kalman filter, TLS parameter and Fuzzy ARTMAP neural network classification)[1]-[9].

Although the methods described above could work successfully in recognizing certain types of ECG signals, the recognition rate usually can not be substantially promoted throughout all kinds of ECG signals. In fact, the major of these methods still need artificial assistance at the present. For improving the accuracy, many scholars are trying to look for a more excellent method. Zhou k led the SVM methods in finding methods to ECG classification [10]. However, he didn't gain the recognition rate of ECG classification. In this paper, 1-vs-rest algorithm of SVM is used for ECG classification. The algorithm for ECG classification is tested with the data of MIT-BIH. Finally a high recognize rate is obtained. It is better than normal way in constructing algorithm model and classification speed.

## 2. ECG signal analysis

The electrocardiogram (ECG) is noninvasive in nature and abundant in diagnostic information, which is one of the most important tools in the diagnosis of heart diseases. Due to the high mortality rate of heart diseases, faithful detection and classification of ECG arrhythmias is essential for the treatment of patient in the clinics. Many researches have been conducted to explore effective signal analysis and pattern recognition techniques for computer-aided diagnosis (CAD) based on ECG signals. Most of the ECG CAD systems have two functional units: feature extraction and pattern classification. The ECG features can be extracted in time domain, in frequency domain, by multi-scale decomposition, or represented as statistical measures.

At present, the methods based on syntax analysis and vector measure are the mainly methods for ECG analysis. Here follow is flow chart of cardiogram computer auto analysis and representative electrocardiogram.

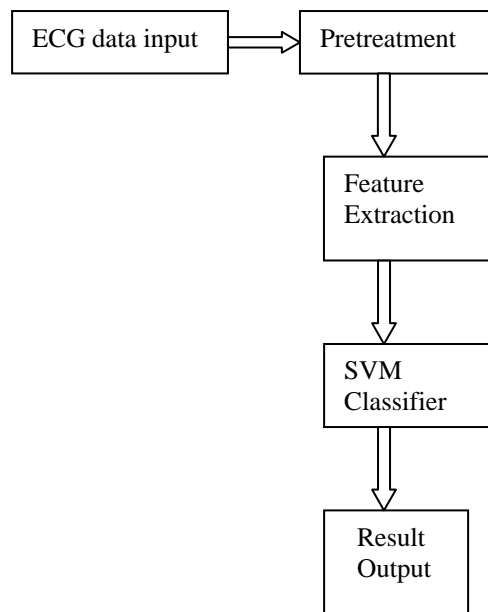


Fig.1: Flow chart of cardiogram computer auto analysis.

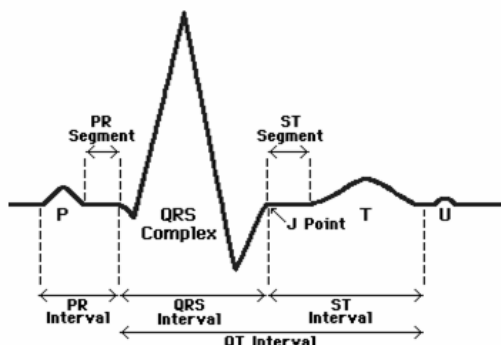


Fig.2: Representative electrocardiogram.

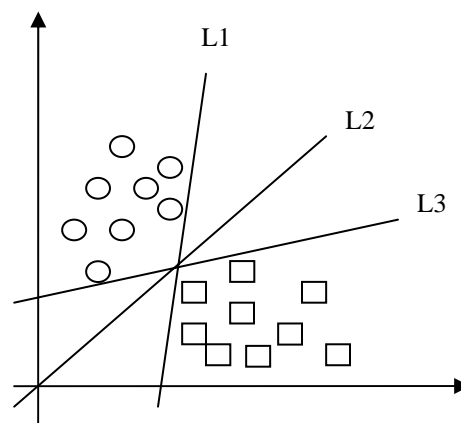
### 3. SVM theory summarize

Support Vector Machine (SVM) is the statistical learning theory developed on the basis of a new generation of learning algorithm. It is applied to text classification, handwriting recognition, image classification. It is a good application in weather forecast. SVM is also a data mining of a new technology, is using the most optimal solution to the problem of machine learning new tools. It was originally put forward in the 1990s last century by Vapnik. Recent years, its theoretical research and the algorithms have made breakthrough progress, and started to become overcome the "dimension of the

disaster" and to "learn" traditional difficulties powerful tool. Support vector machine maps input vectors to a higher dimensional space where a maximal separating hyperplane is constructed. Two parallel hyperplanes are constructed on each side of the hyperplane that separates the data. The separating hyperplane is the hyperplane that maximizes the distance between the two parallel hyperplanes. An assumption is made that the larger the margin between these parallel hyperplanes is, the better the impact of the classifier will be.

Often we are interested in classifying data as a part of a machine-learning process. Each data point will be represented by a p-dimensional vector (a list of p numbers). Each of these data points belongs to only one of two classes. We are interested in whether we can separate them with an "p minus 1" dimensional hyperplane. This is a typical form of linear classifier. There are many linear classifiers that might satisfy this property. However, we are additionally interested in finding out if we can achieve maximum separation (margin) between the two classes. By this we mean that we pick the hyperplane so that the distance from the hyperplane to the nearest data point is maximized. That is to say that the nearest distance between a point in one separated hyperplane and a point in the other separated hyperplane is maximized. Now, if such a hyperplane exists, it is clearly of interest and is known as the maximum-margin hyperplane and such a linear classifier is known as a maximum margin classifier.

Many linear classifiers (hyperplanes) separate the data. However, only one achieves maximum separation. Maximum-margin hyperplanes for a SVM trained with samples from two classes. Samples along the hyperplanes are called the support vectors. (see figure 3)



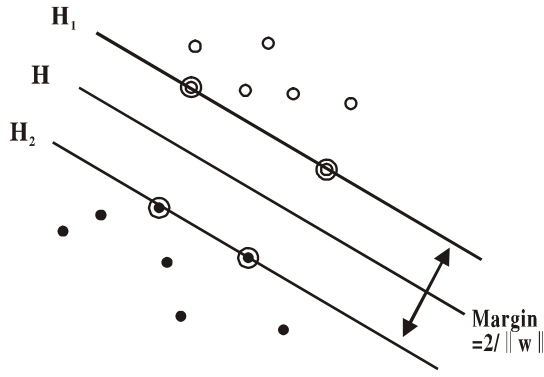


Fig.3: Sketch map of optimization classification plane.

An assumption is made that there are two types of points in input pattern set  $\{x_i\} \in R^n$ . If  $x_i$  belongs to type 1, then  $y_i = 1$ ; if  $x_i$  belongs to type 2, then  $y_i = -1$ . There exist a training sample set  $\{x_i, y_i\}$ ,  $i = 1, 2, 3, \dots, n$ . The aim of SVM is to construct a target function to separate the two classes by Structure Risk Minimization. Usually it is divided into two kinds of circumstances to discuss: 1). The linear classification, 2). The Non-linear classification.

### 3.1. The linear classification

In such situation, we can view this as training samples, which denotes the correct classification which we would like the SVM to eventually distinguish, by means of the dividing (or separating) hyperplane, which takes the form:

$$w * x + b = 0 \quad (1)$$

where “\*” is dot matrix,  $w$  is n-dimensional vector,  $b$  is offset.

The optimal hyperplane can gain from solving the quadratic optimization problem.

$$\min \phi(w) = \frac{1}{2} \|w\|^2 \quad (2)$$

Such that

$$y_i (w * x_i + b) \geq 1, i = 1, 2, 3, \dots, n. \quad (3)$$

When the feature number is very big, the quadratic programming problems may convert into its dual problem.

$$\max W(a) = \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i,j=1}^n a_i a_j y_i y_j (x_i * y_j) \quad (4)$$

$$w^* = \sum_{i=1}^n a_i y_i x_i \quad (5)$$

$$b^* = y_i - w * x_i \quad (6)$$

Such that:

$$\sum_{i=1}^n y_i a_i = 0, a_i \geq 0, i = 1, 2, 3, \dots, n. \quad (7)$$

Where  $a = (a_1, a_2, \dots, a_n)$  is Lagrange arithmetic operators,  $w^*$  is the normal vector of the optimal hyperplane,  $b^*$  is the offset of the optimal hyperplane. The KKT condition will play the important role in the solution and analysis of the optimal problem.

At (7), the solution must hold in:

$$a_i \{y_i (w * x + b) - 1\} = 0, i = 1, 2, 3, \dots, n. \quad (8)$$

From (5) we may see that these samples of  $a_i = 0$  are no effect for classification. While those samples of  $a_i > 0$  get function for classification. So these samples are called support vector. Then we may gain the final classification function

$$f(x) = \sum_{i=1}^n y_i a_i (x * x_i) + b^* \quad (9)$$

Where, the symbol of  $f(x)$  determine the attribute of  $x$ .

### 3.2. The non-linear classification

Support Vector Machines non-linearly map their n-dimensional input space into a high dimensional feature space. In this high dimensional feature space a linear classifier is constructed. In the feature space, the function of original space achieves scalar product operation. According to functional analysis theory, if one of the kernel functions satisfies Mercer condition, it corresponds to scalar product. Taken together, this leads to the decision function

$$\max W(a) = \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i,j=1}^n a_i a_j y_i y_j K(x_i * y_j) \quad (10)$$

Its correspondence classification function is

$$f(x) = \sum_{i=1}^n y_i a_i K(x * x_i) + b^* \quad (11)$$

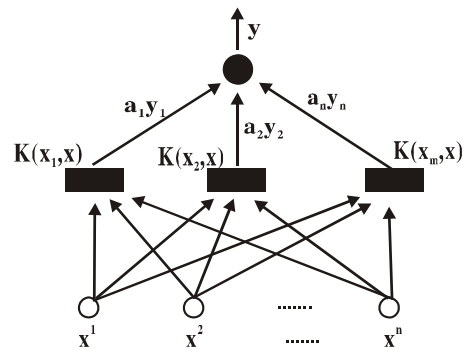


Fig.4: sketch map of SVM rationale

## 4. Electrocardiogram classification method based on SVM 1-vs-rest algorithmic

1-vs-rest classification means that we are dealing with a dataset consisting of more than two classes and that we are trying to distinguishing members of one class from non-members of this class (“the rest”). Prediction accuracy is the number of test points that were correctly classified divided by the number of all test points. Sensitivity is the percentage of test points with “+” label that were classified to be “+”. Specificity is the percentage of test points with “-” label that were classified to be “-”. The “+” predictive value is the percentage of test points with “+” labels among all test

points that were predicted to have “+” labels. The “-” predictive value is the percentage of test points with “-” label among all test points that were predicted to have “-” labels. We will report these classification quality measures as percentages throughout the thesis.

The ECG data sample is the file T101.dat with 12th—14th seconds data segment from MIT-BIH data-base .Its sampling frequency is 250Hz.(see Tab 1)

Because the conditional attribute of ECG is successive value, it have to be discretized at first. (see Tab 2)

NO	Attribute	Values						
		0	1	2	3	4	5	6
1	First-order	<-44	[-44,-10)	[-10, 0)	[0, 10)	[10, 30)	[30, 80)	>=80
2	Second-order	<-37.6	[-37.6,0)	[0,41.2)	[41.2,80)	>=80	-----	-----
3								

Table1: Discretization method of First-order and Second –order difference of ECG.

U	Conditional Attribute											Class
	1	2	3	4	5	6	7	8	9	10	11	
1	2	2	1	2	2	1	2	2	3	2	1	1
2	2	1	2	2	1	2	2	2	2	1	1	1
3	3	2	2	2	1	2	2	1	3	2	1	1
4	3	1	2	2	1	2	3	1	2	2	1	1
5	4	0	6	0	3	4	1	2	0	4	6	2
6	4	0	6	0	4	4	1	2	0	4	6	2
7	5	0	6	0	2	4	1	2	0	4	6	2
8	4	0	6	0	4	4	1	2	0	4	6	2
9	5	0	6	0	2	4	1	2	0	4	6	2
10	3	1	4	0	4	2	2	2	1	2	5	3
...												...
...												...

Table2: Decision table.

## 5. Conclusions

This text leads support vector machine (SVM) 1-vs-rest arithmetic in finding method for ECG classification. The algorithm for ECG classification is tested with the data of MIT-BIH. Finally a high recognize rate is obtained. It is better than normal way in constructing algorithm model and classification speed. We adopt MATLAB and CMSVM1. 0 for modeling programs, and select radial basis function as SVM kernel function. The parameter of the optimal model  $c=100$ , kernel parameter  $g=0.001$ . The 75% of the samples data are training samples. The 20% of the samples data are classification experiment samples, and the 5% of the samples data are classification inspection samples. The success ratio of prediction is 98.7%.How to extract the ECG feature is the next work.

## References

- [1] Caswells A, kluge K S, Chiang C M J, Pattern recognition of cardiac arrhythmia using two intra-cardiac channels. *Proc Comp Cardio*, 181, 1993.
- [2] S.H. Zhou, Rautaharju P M. Calhoun H P, Selection of a reduced set of parameters for classification of ventricular conduction defects by cluster analysis. *Proc Comp Cardiol*, 879, 1993.
- [3] Minam I K C, Nakajima H, Toyoshma T, Real-time discrimination of ventricular tachy arrhythmia with fourier-transform neural network. *IEEE Trans Biomed Eng*, 46:179-187, 1999.

- [4] Afonso V X, Tompkins W J, Detecting ventricular fibrillation: Selecting the appropriate time-frequency analysis tool for the application. *IEEE Eng Biomed Mag*, 14:152-161, 1995.
- [5] X..S. Zhang, Y.S. Zhu, Thakor NV, etal, Detecting ventricular tachycardia and fibrillation by complex its measure. *IEEE Trans Biomed Eng*, 46:548-556, 1999.
- [6] S.W. Chen, Two-stage discrimination of cardiac arrhythmia using a total least squares-based prony modeling algorithm. *IEEE Trans Biomed Eng*, 47:1317-1325, 2000.
- [7] M. Arnold, W. H. R .Miltner, H. Witte, Adaptive modeling of nonstationary time series by means of Kalman filtering. *IEEE Trans Biomed Eng*, 45:533-541, 1998.
- [8] F.M. Ham , S. Han, Classification of cardiac arrhythmia using fuzzy ARTMAP. *IEEE Trans Biomed Eng*, 43:425, 1996.
- [9] J.F. Ren, L. Guo, G. Li, Improving Scene Image Classification with Multi-Class SVMs. *Journal of Northwestern Polytechnical University*, 23(3):295-304,2005.
- [10] K. Zhou, H. Peng, J.S. Hu, SVM Applied in ECG Classification Diagnostic. *Control & Automation*, 22(3):237-239, 2006.
- [11] H.Y. Zhang, L.P. Xing, Support Vector Machine Classification Method based on Kernel Function. *Computer Development & Applications*, 18(7):2-11, 2005.
- [12] F.L. Huang, Z .Zhong, Support Vector Machine Used in Classification. *Journal of GuangXi Teachers College*, 21(3):75-82, 2004.