Support Vector Machine for Assistant Clinical Diagnosis of Cardiac Disease

Gong Wei Wang Shoubin

Department of Electron and Information Engineering, Tianjin Institute of Urban Construction, Tianjin , China

gongwei@tjuci.edu.cn sbin800@126.com

Abstract

Support Vector Machine (SVM) is a novel powerful machine learning method based on statistical learning theory, which is powerful for the characterization of small sample (nonlinearity, high dimension and local minima). Cardiac Diseases are very harmful to the human health. The application of electrocardiogram (ECG) is essential for the clinical diagnosis of cardiac diseases. The use of computers for accurately and quickly cardiac disease diagnosis has been a subject fervently pursued by both internal and external researchers. Therefore, it is great significant to explore even more accurate and higher-speed automatic ECG analysis method. In this paper, Clinical diagnosis of cardiac disease based on SVM is proposed. There are two input patterns of samples: 8-lead ECG in series and in parallel. All experiments are implemented on Pentium 350 MHz with 512 MB RAM. Matlab 6.5 is employ to solve the quadratic programming. Comparison in two input patterns based on SVM, the result shows that SVM method in parallel is highly reliable and accurate. It will have great potential application in clinical diagnosis.

1. Introduction

Support Vector Machine (SVM) is a novel powerful machine learning method based on Statistical Learning Theory (SLT), which is a small sample statistical theory introduced by Vapnik, et al.[1]. SVM is powerful for the characterized by small sample, nonlinearity, high dimension and local minima. SVM implements well trade off between the quality of the approximation of the given data and the complexity of the approximating function by Structural Risk Minimization (SRM) principle. Currently SVM is an active field in artificial intelligent technology, and has been applied to pattern recognition, function estimation, signal processing, control and other fields[2][3] .

The paper is organized as follows. SVM nonlinear classification algorithm is reviewed in Section II and parallel decision model based on SVM is introduced in Section III. In Section IV, SVM is applied to Clinical Diagnosis of Cardiac Disease. Conclusions are given in Section V.

2.SVM Nonlinear Classification Algorithm

Given a set of training data

$$(x_1, y_1), \cdots, (x_l, y_l) \in R^n \times \{+1, -1\}$$

The nonlinear function $\Psi(\bullet)$ is employed to map original input space R^n to high dimensional feature space:

$$\Psi(x) = (\phi_1(x), \phi_2(x), \cdots, \phi_N(x)).$$

Classifier is constructed in this high dimensional feature space. Classifier take the from below:

$$\gamma(x) = \mathrm{sgn}[w \bullet \phi(x) + b]$$

To obtain the classifier, one minimizes ||w|| subjected to

$$\gamma_i [\phi(x_i) \bullet \omega + b] \geq 1 - \xi_i \ ,$$

where variables ξ_i are slack variables which are needed in order to allow misclassification.

The optimization is formulate as

$$\min J(w, \xi) = \frac{1}{2} w \bullet w + c \sum_{i=1}^{l} \xi_i$$

which is subject to constraints

$$\gamma_i [\phi(x_i) \bullet \omega + b] \geq 1 - \xi_i$$
$$\xi_i \geq 0, \quad i = 1, \cdots, l.$$

978-0-7695-3571-5/09 $25.00 © 2009 IEEE
DOI 10.1109/GCIS.2009.186

where c is a constant. The parameter c can be regarded as a regularization parameter.

The optimization problem can be solved by constructing a Lagrangian. By conditions for optimality, the optimization problem yields the following dual problem:

$$\max W(a) = -\frac{1}{2}\sum_{i,j=1}^{l} a_i a_j y_i y_j K(x_i, x_j) + \sum_{i=1}^{l} a_i$$

which is subject to constraints

$$\sum_{i=1}^{l} a_i y_i = 0,\ \ 0 \le a_i \le c,\ \ i = 1, \cdots, l.$$

where $(a_i \ge 0, \quad y_i \ge 0 (i = 1, \cdots, l))$ are Lagrange multipliers and $K(x_i, x_j)$ are kernels function that satisfy Mercer condition of original space R^n [4].

Finally, we have the following nonlinear classifier:

$$f(x) = \mathrm{sgn}\left[\sum a_i y_i K(x, x_i) + b\right]$$

3. Parallel Decision Model Based on SVM

The structure of SVM is shown in Fig. 1 [4]. In SVM classifier, the structure of SVM consists of three layers.

The first layer is input layer. The second layer is support vector, i.e. hidden layer, and the third layer is output layer,

i.e. decision layer. By applying indicator function to decision function:

$$D_k = \sum a_i^k y_i K(x, x_i) + b_k$$

and a classifier is constructed:

$$f(x) = \mathrm{sgn}\left[\sum a_i y_i K(x, x_i) + b\right]$$

4. Application to Clinical Diagnosis of Cardiac Disease

The electrocardiogram is today used worldwide as a relatively simple way of diagnosing heart conditions. An ECG is a recording of the small electric waves being generated during heart activity. An ECG intervals and waves are shown in Fig. 2. This diagram illustrates ECG waves and intervals as well as standard time and voltage measures on the ECG paper.

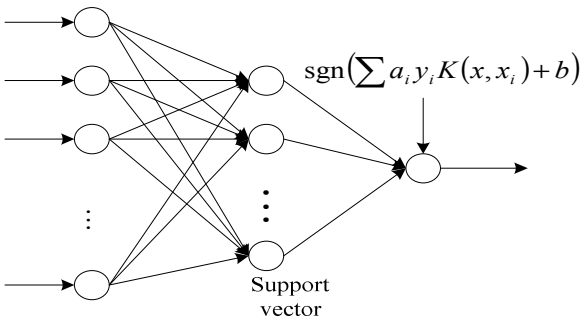

Fig. 1 Structure of SVM

The P wave represents atrial activation; the PR interval is the time from onset of atrial activation to onset of ventricular activation. The QRS complex represents ventricular activation; the QRS duration is the duration of ventricular activation. The ST-T wave represents ventricular repolarization. The QT interval is the duration of ventricular activation and recovery. The U wave probably represents "after depolarizations" in the ventricles.

ECG examination is one of means in clinical diagnosis, which is important to some diseases especially cardiac disease. Using computer to identify and classify ECG, first divide ECG exactly, and extract every wave, finally analyse data. This "method" is recommended when reading all 12-lead ECG's. Like the physical examination, it is desirable to follow a standardized sequence of steps in order to avoid missing subtle abnormalities in the ECG tracing, some of which may have clinical importance [5].

The data are from LDS Hospital in Salt Lake City provided by Frank G.Yanowitz of University of Utah School of ECG Department. Most of the 12- and 6-lead ECGs were recorded at LDS Hospital in Salt Lake City, Utah. Marquette Electronics has also given permission to use ECG rhythms and diagrams from their educational posters. Each of the ECGs has an interpretation and many have additional explanations that help explain the diagnosis [5].

There are 8 leads after the transform of ordinary 12-lead ECG. Every one notes the signal of ECG in 10 seconds. In the original data, every lead include 5000 samples, so every swatch has (8×5000) data. For the difference of individual and heartbeat frequency, the extracted full heartbeat samples are difference from one another. For the effective pattern classification, the original input data must be transformed to get the classified information.

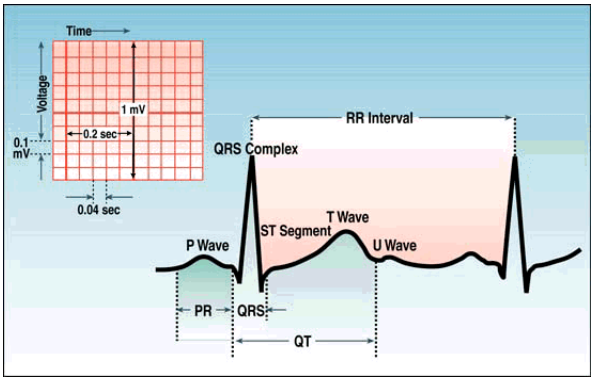

Fig. 2 ECG Intervals and Waves

There are two input patterns of samples:

1). 8-lead ECG in series: every sample can be regarded as (1,344) dimension data for training and testing.

2). 8-lead ECG in parallel: every sample can be regarded as (8, 43) dimension data, every 43 dimension data in one group are corresponding to one-lead heartbeat of the ECG. The samples make train and test for the corresponding lead data every time. 8 leads test one by one to get 8 output data, we can decide the final classification result by integrating 8 leads estimation values.

Pre-processing of Data

Each attribute is normalized:

$$\bar{x}_i = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)}, \bar{x}_i \in [0,1]$$

Where $\bar{x}_i$ are attributes after normalization.

Normal sample is replaced by +1, and abnormal sample is replaced by -1.

Data Kernels

Kernels $K(x, x_i)$ can be any symmetric function satisfying Mercer condition. Different kernels $K(x, x_i)$ can be selected to construct different SVM [4].

This paper employs radial basic function as kernels:

$$K(x, x_i) = \exp\left[-\frac{|x - x_i|^2}{\sigma^2}\right]$$

Where $|x\text{-}x_i|$ are calculated by formula $|x - x| = \sqrt{\sum_{k=1}^{n} (x^k - x_i^k)^2}$ and σ is kernels width.

Table 1 Comparison in Two Input Patterns Based on SVM

RBF Function	Classification precision of Normal and abnormal ECG (%)	
	8-lead ECG in series	8-lead ECG in parallel
σ^2=10	58.00	89.75
σ^2=50	78.00	91.00
σ^2=100	86.75	91.25
σ^2=200	89.25	90.75
σ^2=300	86.25	90.50
σ^2=400	84.50	90.25
σ^2=500	86.00	89.50
σ^2=1000	82.00	87.00
σ^2=5000	81.25	88.25

Result and Discussion

796 samples act as a training set, and 400 samples act as a testing set to evaluate learning result from LDS Hospital database. All experiments are implemented on Pentium 350 MHz with 512 MB RAM. Matlab 6.5 is employ to solve the quadratic programming.

The simulation of testing results in clinical diagnosis of cardiac disease is in Table 1, Fig.3. The Fig.3 shows the performances of nonlinear Classification precision in 8-lead ECG in series based on SVM, and it shows the performances of nonlinear Classification precision in 8-lead ECG in parallel based on SVM. In simulation figures, X-coordinate represents samples number, and Y-coordinate represents Classification precision.

From the results, we can draw two conclusions from the experiments:

1). It is found that 8-lead ECG in parallel connection is better than in series in precision promotion.

2). parameter selection is important and difficult in Kernels, take table 1 for example, it is great difference in classification capability of SVM when selecting different parameters. In parameter selection, 8-lead in parallel is more stable than in series.

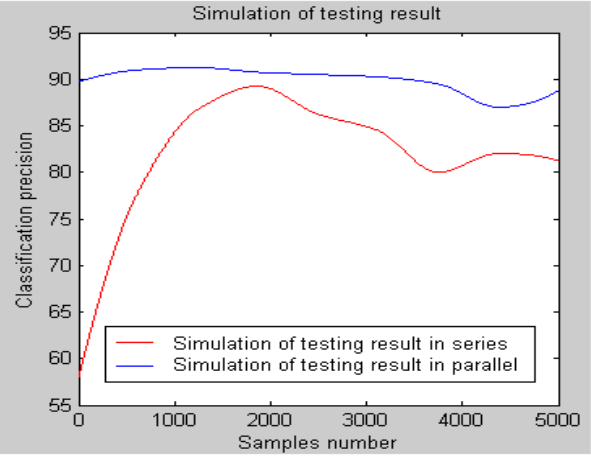

Fig. 3 Simulation of testing result in series and parallel

5. Conclusions

In the paper, clinical diagnosis of cardiac disease based on SVM is introduced. 8-lead ECG in parallel is highly reliable and accurate through comparison. SVM is powerful for the characterization of small sample (nonlinearity, high dimension and local minima). It will have great potential application in clinical diagnosis.

6. References

[1] Vapnik V N. The nature of statistical learning theory. New York: Springer-Verlag, 1995.

[2] Vapnik V N. IEEE Trans on Neural Network,1999,10(5):988.

[3] Suykens J A K, Vandewalle J, De MoorB. Neural Network, 2001, 14:23.

[4] Vapnik V N. The nature of statistical learning theory. New York: Springer-Verlag, 1999.

[5] http://www.ifmbe.org.

[6] L.Prechelt. Probenl. A set of neural network benchmark problem and benchmark rules [J]. Fakulitat fur informat ik, Univ. Karlsruhe,Germany, Tech. Rep. 21/94, sep 1994.

[7] Li Xiangyang, Zhu Xuefeng. Research and application of Kappa Number Prediction Model to Kraft Process, Guangdong papermaking 2000, (5): 23-27.

[8] Zonghai Sun, Youxian Sun. “Fuzzy support vector machine for regression estimation,” IEEE SMC’03, Proc. of international Conf. onSystems Man, and Cybernetics, vol4, Washington, DC, ctober, 2003, pp. 3336-3341.

[9] Ming-Hsuan Yang, Narendra Alruja, “A geometric approach to train support vector machines”, IEEE CVPR’2000, Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, vol1, Hilton Head Island June, 2000,pp. 430-437.