Development Trend of Support Vector Machine and Applications on the Field of Computer Science

Hainan Wang[1], Haoming Liu[2] and Xiaoliang Zhang[3]

[1]College of Computer Science and Technology, Jilin University, Changchun, 130000, China;

[2]Indiana University Bloomington, USA;

[3]Tianjin University, China.

Abstract. In the field of machine learning, support vector machine is a supervised learning model which can analyze the data and identify patterns, and this theory is used for classification and regression analysis which is related to the learning algorithm. Support vector machine is a new machine learning method based on statistical learning theory, and has become a hot research topic in the machine learning field because of its excellent learning performance. Support vector machine is also a learning machine based on kernel function, and its generalization ability depends on the chosen kernel function to a great extent. Nevertheless, traditional support vector machine cannot achieve the desired results when training large-scale data set. Therefore, in order to improve training efficiency and generalization performance of support vector machine, we have to improve this algorithm. Although support vector machine has great advantages in theory, the research on its application is comparatively delayed. In this paper, we focus on novel support vector machine generated from multi-subjects amalgamation, such as fuzzy support vector machine, granular support vector machine, twin support vector machine. We analyze a large number of literature review about development trend of support vector machine and its applications in computer field, laying foundation for the systematic research on the correlation algorithm based on support vector machine.

Keywords: Support vector machine (SVM); Development; Machine learning; Computer field; Applications.

1. Introduction

The 21st century is the age of information and big data, in this age, with the production of huge amounts of data, how to excavate the useful knowledge or rules from massive data is a task of data mining. Data mining is an important step of Knowledge-Discovery in Database (KDD), and is also a nontrivial process to reveal hidden, unknown and potentially valuable information from large amounts of data [1].

Support vector machine (SVM) [2, 3, 11, 12, 13] is a kind of machine learning method based on statistical learning theory [4, 5] and structural risk minimization principle. SVM shows many unique advantages in solving small sample, nonlinear and high dimensional pattern recognition problems, and overcomes many problems of "dimension disaster" and "over learning" to a great extent. In addition, because of its substantial theoretical foundation and simple mathematical model, SVM is widely used in pattern recognition, regression analysis, function estimation, time series prediction, text recognition [6], handwriting font recognition [7], face recognition [8], gene classification [9], time series prediction [10], et al.

Standard support vector machine learning algorithm problem can be reduced to a quadratic programming (QP) problem. For small-scale quadratic programming problem, we can easily solve this problem by using Newton method, interior point method and other classical optimization algorithm. However, when facing large-scale data set, the training speed becomes slow, the algorithm is complex, and the efficiency is pretty low. At present, some traditional training algorithm is to divide original large-scale QP problem into a series of small-scale QP problems, and repeatedly solves the small-scale QP problem according to iterative strategy, then generate an approximate solution of the original large-scale QP problem, finally figure out the optimum solution. In addition,

existed large-scale problem training algorithm cannot completely solve the problems, therefore, we have to improve original algorithms and figure out an optimal algorithm.

In this paper, we firstly introduce the fundamental theory of SVM, and then summarize the current SVM training algorithm, we also prospect the future research direction.

2. Support Vector Machine Theory

2.1 Statistical Learning Theory

Statistical learning theory (SLT) is a kind of special research on the theory of machine learning rules under the condition of small-scale sample. The theory establishes a set of new theoretical system for the small-scale sample statistical problem, and the statistical inference rules under this system not only take the requirements of asymptotic performance into account, but also obtain the optimal results under the condition of limited information.

A core concept of statistical learning theory is Vapnik-Chervonenkis (VC) dimension. The intuitive definition of VC dimension in pattern recognition method is: for an indicator function set, if there have h samples that can be separated into 2^h different forms, then we say function set can break up h samples, and the number of VC dimension is the maximum number of samples. VC dimension reflects the learning ability of function set, namely, the larger the dimension of VC, the more complex the learning machine (the stronger the learning ability).

SLT systematically studies the relationship between empirical risk (training error) and actual risk (expected risk) of function sets with various types, i.e. the bound of extending.

For all functions existed in indicator function set, the relationship of empirical risk and actual risk is denoted as follows:

$$R(w) \leq R_{emp}(w) + \sqrt{\frac{h(\ln \frac{2l}{h} + 1) - \ln \frac{\eta}{4}}{l}} \tag{1}$$

Where h is VC dimension of function set, and l is the number of samples.

The actual risk is related to the VC dimension of the learning machine and the number of training samples, and can be denoted as follows:

$$R(w) \leq R_{emp}(w) + \phi(h/n) \tag{2}$$

This means, under the condition of limited training samples, the higher the VC dimension, the greater the confidence range, and leads to greater difference between actual risk and empirical risk. This is the reason why there will be a phenomenon of over learning. In the process of machine learning, we need to keep lower empirical risk and VC dimension in order to reduce confidence range, in this way can we obtain less actual risk.

SLT puts forward a new method to solve this problem, which is, dividing a function set $S = \{f(x,a), a \in \Omega\}$ into function subset sequence denoted as follows:

$$S_1 \subset S_2 \subset ... \subset S_k \subset ...S \tag{3}$$

$$h_1 \subset h_2 \subset ... \subset h_k \subset ... \tag{4}$$

The confidence range is the same in one subset. In each sub focus on finding the minimum empirical risk and confidence range to achieve the minimum value of the actual risk is called structural risk minimization (SRM).

2.2 Support Vector Machines

Support vector machine(SVM) is a data mining method based on statistical learning theory which can successfully deal with problems such as regression analysis and pattern recognition, and it can be extended to the fields and subjects of prediction and comprehensive evaluation[14, 15].

On the basis of SLT, SVM is proposed which is also an excellent general machine learning method [16-20]. The main idea is separating two kinds of training samples to the largest limit, and constructing a classification hyper plane for training samples of binary classification problem. The most important advantage is to overcome the curse of dimensionality, thus, SVM is widely used in text classification, bioinformatics and other classification problems [21].

Figure 1 explains the linear SVM intuitively from geometry, and figure 2 shows the linear twin support vector machine (TWSVM).
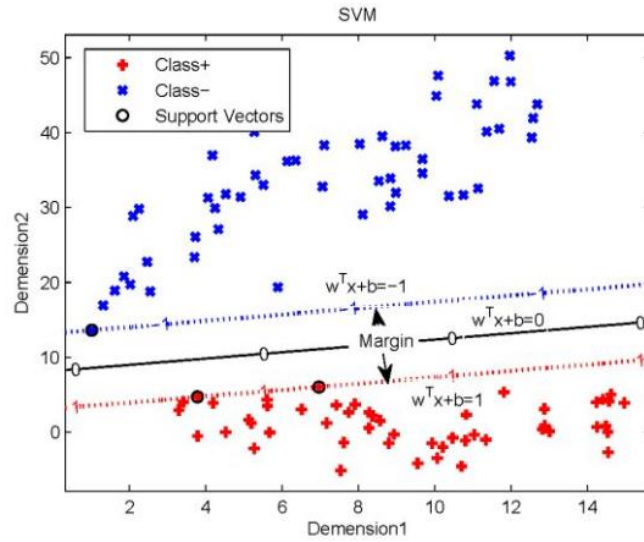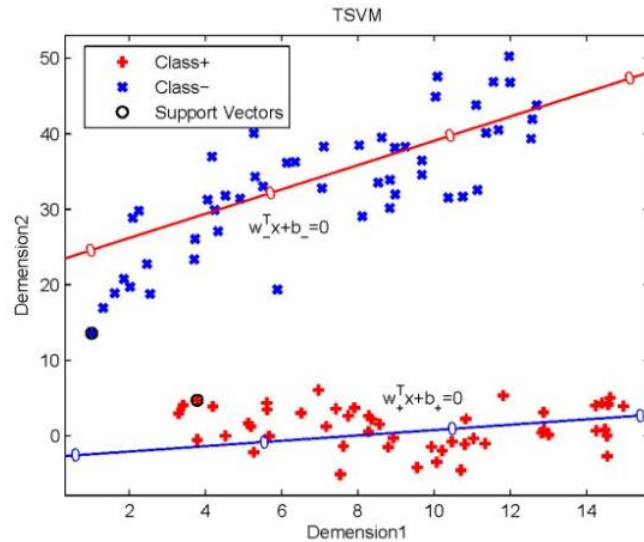

Fig.1 Linear SVM

Fig.2 Linear TWSVM

In order to analyze the optimization effect of SVMs, we use artificial data to carry out simple verification. Figure 3 and figure 4 show the effect diagram of RBF parameter optimization.
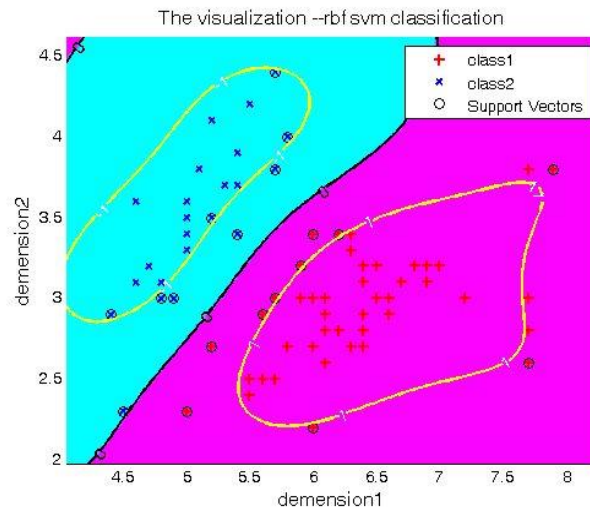

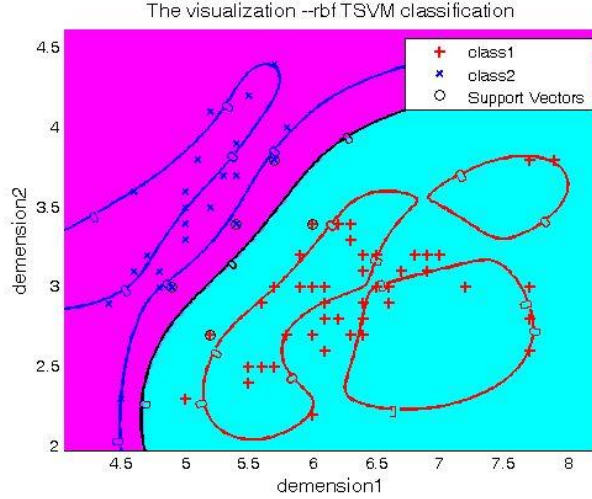Fig.3 The Visualization of RBF paraeter based on SVM

Fig.4 The Visualization of RBF paraeter based on TWSVM

SVM is to figure out an optimal classification hyper plane which satisfies the classification requirements, so that the hyper plane can ensure the classification accuracy and maximize the blank area on both sides of the hyper plane. Theoretically, SVM can achieve the optimal classification of linear separable data.

Assume we have a training sample set denoted as:

$$(x_i, y_i), i = 1, 2, ...l, x \in R^n, y \in \{\pm 1\} \tag{5}$$

And the hyper plane is denoted as:

$$(w \bullet x) + b = 0 \tag{6}$$

In order to classify all samples correctly, we have the following constraint condition:

$$y_i[(w \bullet x_i) + b] \geq 1, i = 1, 2, ..., l \tag{7}$$

Thus we can compute the classification intervalis $2/\parallel w \parallel$, then the problem of constructing the optimal hyper plane can be transformed to calculate Eq. (6) under constraint condition:

$$\min \phi(w) = \frac{1}{2} \parallel w \parallel^2 = \frac{1}{2}(w^{'} \bullet w) \tag{8}$$

To solve this constrained optimization problem, we introduce Lagrange function:

$$L(w, b, a) = \frac{1}{2} \parallel w \parallel - a(y((w \bullet x) + b) - 1) \tag{9}$$

Where $a_i > 0$ is Lagrange multiplier. The solution of constrained optimization problem is determined by saddle point of Lagrange function. We can transform this QP problem to dual problem, i.e.:

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

$$s.t. \sum_{j=1}^{l} a_j y_j = 0, j = 1, 2, ..., l, a_j \geq 0, j = 1, 2, ..., l \tag{10}$$

The optimum solution is: $a^* = (a_1^*, a_2^*, ..., a_l^*)^T$.

Compute optimal weight vector w^* and optimal allocation b^*:

$$w^* = \sum_{j=1}^{l} a_j^* y_j x_j \tag{11}$$

$$b^* = y_i - \sum_{j=1}^{l} y_j a_j^* (x_j \bullet x_i) \tag{12}$$

Where $j \in \{j \mid a_j^* > 0\}$. Thus we obtain optimal classification hyper plane denoted as $(w^* \bullet x) + b^* = 0$. Optimal classification function is:

$$f(x) = \text{sgn}\{(w^* \bullet x) + b^*\} = \text{sgn}\left\{\left(\sum_{j=1}^{l} a_j^* y_j (x_j \bullet x_l)\right) + b^*\right\}$$

$$x \in R^n \tag{13}$$

For linearly inseparable case, the main idea of SVM is mapping the input vector into a high dimensional feature vector space, and the optimal classification surface is constructed in the feature space:

$$x \to \phi(x) = (\phi_1(x), \phi_2(x), ..., \phi_l(x))^T \tag{14}$$

The feature vector $\phi(x)$ replaces input vector x, and we can obtain optimal classification function denoted as:

$$f(x) = \text{sgn}\{(w \bullet \phi(x) + b\} = \text{sgn}\left(\sum_{i=1}^{l} a_i y_i \phi(x_i) \bullet \phi(x) + b\right) \tag{15}$$

In the above dual problem, both of the objective function and decision function are only related to the inner product between training samples. In the high dimensional space, avoiding high dimensional complex operations, and we only need to implement inner product.

3. Research Direction and Hot Spot

With the deep research of SVMs, the research mainly focuses on the following aspects.

3.1 Factorization Algorithm

Due to the appearance of large-scale and large amount of data, thus, it takes us a lot of time to process the data comprehensively. Researchers introduce factorization algorithms in order to improve the efficiency of data processing. The key of the factorization algorithm is the determination of the size of the working set. Commonly used factorization algorithms are: Chunking algorithm proposed by Cortes and Vapnik [22]; fixed working sample set algorithm proposed by Osuna [23]; SMO algorithm proposed by Platt [24].

3.2 Online Learning Problems

Online learning problems are online incremental learning and online reduced learning [25-28].When new data is added to the training data set (or reduce from training data set), we do not need to re-train the SVM model, this makes training process simple. Incremental learning concept is first proposed by Syed et al. [26] and they put forward incremental learning strategy, however, they did not improve the algorithm with slow training process, and the solution of their method is approximate solution. Cauwenberghs et al. [27] proposed online incremental learning method which can obtain exact solution. In [27], the researchers analyzed the effect of adding a sample or reducing a sample, and the results showed that incremental learning algorithm obtained the same solution as non-incremental learning algorithm did.

3.3 Multi-class classification problem

Multi-class classification problem is an important topic in reality, and is a generalization of 2-classification question.

One-against-one approach (OVO) constructs pairwise combination with different types. Assume there are K different classes, we need to construct $K(K-1)/2$ different classifiers. The main idea of this method is that: arbitrarily extract two classes among K different classes, and implement training learning process based on SVM, then we can obtain one classifier. In this way, we construct $K(K-1)/2$ different classifiers in all. The training time of this method is pretty long because of too many classifiers.

One-against-several approach (OVS) constructs K two-class classification machines. Assume given data have k classes. We regard 1 class and other k-1 classes as one class, thus the question can be transformed into a 2-class classification problem, and we obtain a support vector classifier. Totally, we need construct k support vector classifiers. Nevertheless, there are two problems to be solved in

this method. Problem of training time is too long due to a large amount of training sample data, and data imbalance problem.

Binomial tree method (BT): First of all, all classes are divided into two sub-classes, and each sub-class is divided into two sub-classes. This method can effectively reduce the amount of calculation and test. But when the father node is wrong, the end result must be wrong.

3.4 Privacy Preserving Support Vector Machine [29-32]

With the emphasis on privacy issues in recent years, the researchers propose privacy preserving SVM (PPSVM). PPSVM is divided into vertical data based partition and horizontal data based partition. HwanjoYu [29, 30] et al. applied secure multi-party computation (SMC) and cryptography into SVMs, and constructed kernel function which can obtain PPSVM without information leakage. Olvi L. Mangasarian [31] used reduced support vector machine (RSVM) into the process of calculating the kernel matrix.

4. Some Novel SVMs

4.1 Granular Support Vector Machines

Granular support vector machine (GSVM) is to construct granular space by means of frequently-used granular partition method in order to obtain a series of information granules [33-36], and then learn the obtained information granules, finally, the method generates SVM decision function through aggregating the information. This learning mechanism can transform a linear non-separable problem into a series of linear separable problems according to the granulation of data, and can obtain multiple decision functions. GSVM also enhances the generalization performance of data that means, the method can obtain the hyper plane with a wider spacing during the process of SVM training.

In [37], the researchers proposed GSVM based on association rules. The main idea of their method is, mapping the sample in original space to the feature space in terms of RBF kernel function, and expanding into MacLaurin series, then mining the vital association rules from the obtained expansion. Thus, the association rule classifier is obtained from the classification model of SVM [38].

In [39, 40], the authors proposed GSVM based on clustering theory. The main idea of their method is: Firstly, the original data is divided into a number of granules according to the commonly-used clustering method combined with some evaluation rules. Next, select the granules which contain more information, and use these granules to implement the process of classification and regression. Thus SVM can get high efficiency and generalization ability in data set training.

In [41, 42], the researchers proposed GSVM based on fuzzy set theory. The main idea of this method is to preprocess the original data according to fuzzy set theory, in order to reduce redundant data and compress the data. In this way, the researchers can improve the classification speed of SVM.

In [43, 44], the authors proposed a novel GSVM based on kernel space (KGSVM). KGSVM can mapping the original data to the kernel space, and then SVM training in the same kernel space is carried out. Compared with the traditional GSVM model, KGSVM model overcomes the problem of inconsistency of kernel spatial data distribution. Because traditional GSVM implements the process of information granules division in original space, thus, we may lose a lot of useful information. However, KGSVM implements the process of information granules division in kernel space, so we can get better classification accuracy and generalization performance by using radius rule. After that, in [45], the authors improved the KGSVM model, and they proposed a kernel granular SVM based on association rules.

4.2 Twin Support Vector Machines

In [46], the researchers proposed a novel classifier based on binary data that is twin support vector machine (TWSVM). TWSVMs are formally similar to the conventional support vector machines [47, 48], which not only has the advantages of traditional support vector machines, but also has a better processing ability for large-scale data training. TWSVMs constructs a classification hyper plane by

optimizing a pair of planes [49, 50], namely, TWSVMs needs to solve a pair of QP problems, and one class of data is the constraint condition of another QP problem.

4.3 Ranking Support Vector Machine

Rank learning is a hot issue in the field of information retrieval and machine learning, and this method is widely used in many fields, such as, document retrieval, collaborative filtering, keyword extraction, definition discovery and so on. Rank learning problems is divided into two categories: regression-based rank learning and classification-based rank learning [51].

RSVM is widely applied in information retrieval [52] and recommendation system [53]. One of the most important problems is how to set up the damage function which is suitable for the specific application. When RSVM uses cost function which is different from traditional cost function used in SVM, we can obtain precise ranking. SVM regression model can also be used in recommend system, compared with the traditional recommendation system using heuristic, this method will obtain better performance under the condition of large-scale sample.

The traditional RSVM has its limitation in application, namely, the model is too simple to deal with complex ranking process, and it is difficult to add the priori knowledge into SVM model. In order to overcome these problems, a new RSVM is proposed [54]. This novel RSVM can map the scoring function to a probability sigmoid function which uses cross-examination to implement the process of data training.

5. Conclusion

The problem of machine learning is systematically studied by statistical learning theory, especially in the case of finite sample, and SVM is a universal machine learning method under the framework of this theory. Due to the superiority of SVMs, these methods are widely used in the fields of pattern recognition, regression analysis, function estimation, time series prediction, et al. The research of SVM theory and algorithm has made great progress, however, in the practical application of large-scale training data, there are still some issues such as the calculation speed and storage capacity. The field needs to be further developed and improved, including the following research direction:

(1) Further improve the SVM algorithms, and choose different functions to construct SVM classifiers.

(2) Exploring the integration of SVM with other subjects. In recent years, there are many novel SVMs generated from the combination of traditional SVMs and other disciplines, such as FSVM, GSVM, TWSVM, et al. Nevertheless, these SVMs have their own disadvantages. Thus, we have to explore more methods to improve SVMs.

(3) Continuously explore new application areas of SVMs. Although SVM has a very prominent advantage in theory, compared with its theoretical research, the application of research is still relatively backward. Therefore, in the future research work, how to use SVM in people's daily life is an important topic.

References

[1] JiaweiHan, MiehelineKmabe. Data Mining: Concept and Technology [M]. 2001.

[2] Cristianini N, Taylor J S. An introduction ti support vector machines and other kernel-based learning methods [M]. Translated by Li Guozheng, Wang Meng, and ZengHuajun. Beijing: Publishing House of Electronics Industry, 2004.

[3] Zhang Xuegong. Statistical learning theory and support vector machines [J]. ActaAutomaticalSinica, 2000, 26 (1): 32-41.

[4] Vapnik V N. The nature of statistical learning theory [M]. Translated by Zhang Xuegong. Beijing: Tsinghua University Press, 2000.

[5] Vapnik V N. Statistical learning theory [M]. Translated by Xu Jianhua, Zhuang Xuegong. Beijing: Publishing House of Electronics Industry, 2004.

[6] Liu Xiaoliang, Ding Shifei. Appropriateness in applying SVMs to text classification [J]. Computer Engineering and Science, 2010, 32(6): 106-108.

[7] Lin Kaibiao, Wang Zhoujing. The method of fax receiver's name recognition based on SV, [J]. Computer Engineering and Applications, 2006, 42(7):156-158.

[8] Xie Saiqin, Shen Fuming, Qiu Xuena. Face recognition using SVMs [J]. Computer Engineering, 2009, 35(16): 186-188.

[9] Li Yingxin, Ruan Xiaogang. Feature selection for cancer classification based on SVM [J]. Journal of Computer Research and Development, 2005, 42(10): 1796-1801.

[10] Gao Wei, Wang Ning. Prediction of shallow-water reverberation time series using SVM [J]. Computer Engineering, 200834(6):25-27.

[11] V. N. Vapnik. An overview of statistical learning theory [J]. IEEE Transaction on Neural Networks, 1999, 10(5): 988-999.

[12] C. Nello, S T John. An introduction to SVM and other kernel-based learning methods [M]. Canbridge: Cambridge University Press, 2000.

[13] C.J.C. Burges. A tutorial on support vector machines for pattern recognition [J]. Data Mining and Knowledge Discovery, 1998(2):121-167.

[14] Deng Naiyang, Tian Yingjie. New method in Data Mining: Support vector machine [M]. 2003.

[15] Yuan Yaxiang, Sun Wenyu. Optimization theory and method [M]. 1997.

[16] Su Jinshu, Zhang Bofeng, Xu Xin. Research progress of text categorization technology based on machine learning [J], 2006, 17(9): 1848-1859.

[17] Zhang Xuegong. The essence of statistical learning theory [M]. Beijing: Tsinghua University Press, 2000.

[18] Sommer R, Paxson V. Outside the closed world: On using machine learning for network intrusion detection[C]. IEEE Symposium on Security and Privacy (SP), 2010:305-316.

[19] Sangkatsanee P, Wattanapongsakorn N, Charnsripinyo C. Practical real-time intrusiondetection using machine learning approaches [J]. Computer Communications, 2011, 34(18): 2227-2235.

[20] Suthaharan S. Big data classification: problems and challenges in network intrusionprediction with machine learning [J].ACM SIGMETRICS Performance EvaluationReview, 2014, 41(4): 70-73.

[21] Wang Guosheng. Research on theory and algorithm of support vector machine [D].Beijing: Beijing University of Posts and Telecommunications, 2007.

[22] B. Boser, I. Guyon, V. Vapnik. Atraining algorithm for optimal margin classifiers[C]. Pittsburgh: ACM Press, 1992.

[23] E. E. Osuna. Support vector machines: training and applications [D]. MA: MIT University, 1998.

[24] J. C. Platt. Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines [D]. MA: MIT Press, 199.

[25] Duan Hua. Research on incremental learning algorithm of support vector machine [D]. Shanghai: Shanghai Jiao Tong University, 2008.

[26] N. A. SYED, H. Liu, K. K. SuNG. Incremental learning with support vector machines [J]. In Proc. Int. Joint Conf. on Artifical Intelligence, 1999: 196-215.

[27] G. Cauwenderghs, T. Poggio. Invremental and decremental support vector machine learning [M]. MA: MIT Press, 2001.

[28] Hou Weizhen. Research on several optimization algorithms for solving support vector machines [D]. Qingdao: Shandong University of Science and Technology, 2007.

[29] Hwanjo Yu, Xiaoqian Jiang, Jaideep Vaidya. Privacy-preserving SVM using nonlinear kernels on horizontally partitioned data[C]. New York: ACM, 2006.

[30] O.L. Mangasarian, E. W. Wild, G. M. Fung. Privacy-preserving classification of vertically partitioned data via random kernels [J]. ACM Transactions on Knowledge Discovery from Data, 2008, 2(3): 473-479.

[31] Shao Fubo. Research on privacy preserving support vector machine model and algorithm [D]. Qingdao: Shandong University of Science and Technology, 2009.

[32] Hu Yunhong. Research on some algorithms of support vector machine [D]. Qingdao: Shandong University of Science and Technology, 2011.

[33] Tang Yuchun, Jin Bo, Zhang Yanqing et al. Granular support vector machines for medical binart classification problems[C]. Proceedings of the IEEE CIBIB. Piscataway, HJ: IEEE Computational Intelligence Society, 2004: 73-78.

[34] Li Daoguo, Miao Duoqian, Zhang Dongxing et al. An overview of granular computing [J] Computer Science, 2005, 32(9): 1-12.

[35] Cheng Wei, Shi Yang, Zhong Yanping. Three primary methods of granular computing [J]. Computer Technology and Development, 2007, 17(3): 91-94.

[36] Ding Shifei, Xu Li, Zhu Hong et al. Research and progress of cluster algorithms based on granular computing [J]. International Journal of Digital Content Technology and its Applications, 2010, 4(5): 96-104.

[37] Tang Yuchun, Jin Bo, Zhuang Yanqing. Granular support vector machines with association rules mining for protein homology prediction [J].

[38] Zhuang Wenhan, Wang Wenjian. A kernel granular support vector machine based on association rules [J]. Journal of Guangxi Normal University, 2009, 27(3): 89-92.

[39] Wang Wenjian, Xu Zongben. A heuristic training in support vector regression [J]. Neurocomputing, 2004, 61: 259-275.

[40] Zhang Xin, et al. A support vector machine learning strategy based on granular [J]. Computer Science, 2008, 35(8): 101-106.

[41] Duan Danqing, Chen Songqiao, Yang Weijun, et al. Detect intrusion using rough set and support vector machine [J]. Journal of Chinese Computer Systems, 2008, 29(4): 627-630.

[42] Li Ye, Cai Yunze, Li Yuangui, et al. Rough sets method for SVM data preprecessing[C]. IEEE Conference on Cybernetics and Intelligent Systems. Singapore: [s. n], 2004: 1039-1042.

[43] Guo Husheng, Wang Wenjian, Men Chngqian. A novel learning model-kernel granular support vector machine[C]. Preceeding of the 8th international conference on machine learning and cybernetic. Baoding: [s. n], 2009: 930-935.

[44] Wang Wenjian et al. Granular support vector machine learning model [J]. Journal of Shanxi Unversity: Natural Science Edition, 2009, 32(4): 535-540.

[45] Zhang Wenhao, et al. A kernel granular SVM based on association rules [J]. Journal of Guangxi Normal University: Natural Science Edition, 2009, 27(3): 89-92.

[46] Jayadcva R, Khemchandani S.C. Twin support vector machines for pattern classification [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007, 29(5):905-910.

[47] Arun K.M, Gopal M. Least squares twin support vector machines for pattern classification [J], 2009, 36: 7535-7543.

[48] Zhang Xinsheng. Boosting twin SVM approach for MCs detection[C]. Proceedings of Asia-Pacific Conference on Information Processing. Shenzhen, 2009, 46:149-1502.

[49] Peng Xinjun. A v-twin support vector machine classifier and its geometric algorithms [J]. Information Sciences, 2010, 180 (20): 3863-3875.

[50] Chen Jing, Ji Guangrong. Weighted least squares twin support vector machines for pattern classification[C]. Proceedings of the 2dn Internation Conference on Computer and Automation Engineering. Singapore, 2010, 2: 242-246.

[51] Wang Yang, Huang Yanlo, et al. Algorithm of active learning to rank based on Rank Algorithm [J]. Computer Engineering, 2008, 34(21): 38-39, 47.

[52] Ding Shifei, Liu Xiaoliang. Research on ranking support vector machine and prospects[C]. Proceeding of the 29th Chinese Control Conference. Beijing, 2010: 2829-2831.

[53] Yanjima Y, Kuo Tien-fang. Efficient formulations for RSVM and its application to recommendation tasks [J]. Journal of Computers, 2006, 1 (3): 27-34.

[54] Nguyen T, Ngo A.V, Nguyen H.V, et al. Probabilistic ranking support vector machine. Advances in Neural Networks, Beilin: Springer, 2009: 345-353