TELKOMNIKA, Vol.14, No.2, June 2016, pp. 638~646
ISSN: 1693-6930, accredited A by DIKTI, Decree No: 58/DIKTI/Kep/2013
DOI: 10.12928/TELKOMNIKA.v14i1.2751

Semi-supervised Online Multiple Kernel Learning Algorithm for Big Data

Ning Liu[1]*, Jianhua Zhao[2]
[1]School of economics and management, Shangluo University, Shangluo 726000, China
[2]School of mathematics and computer applications, Shangluo University, Shangluo 726000, China
*Corresponding author, e-mail: liuning20122014@aliyun.com

Abstract

In order to improve the performance of machine learning in big data, online multiple kernel learning algorithms are proposed in this paper. First, a supervised online multiple kernel learning algorithm for big data (SOMK_bd) is proposed to reduce the computational workload during kernel modification. In SOMK_bd, the traditional kernel learning algorithm is improved and kernel integration is only carried out in the constructed kernel subset. Next, an unsupervised online multiple kernel learning algorithm for big data (UOMK_bd) is proposed. In UOMK_bd, the traditional kernel learning algorithm is improved to adapt to the online environment and data replacement strategy is used to modify the kernel function in unsupervised manner. Then, a semi-supervised online multiple kernel learning algorithm for big data (SSOMK_bd) is proposed. Based on incremental learning, SSOMK_bd makes full use of the abundant information of large scale incomplete labeled data, and uses SOMK_bd and UOMK_bd to update the current reading data. Finally, experiments are conducted on UCI data set and the results show that the proposed algorithms are effective.

***Keywords**: Semi-supervised Classification, Online Learning, Multiple Kernel, Big Data*

1. Introduction

In recent years, with the rapid development of the Internet, cloud computing, Internet of things, social media and other information technology, data from all walks of life are showing an explosive growth trend [1]. We have entered the era of big data, which has become an important strategic resource of the country, the management and analysis of big data has become a hot topic of academic and industrial attention [2, 3].

The purpose of data collection, data storage, data transmission and data management is to use big data effectively, where machine learning technology is essential [4, 5]. In recent years, machine learning in big data being one of the hotspots has attracted extensive attention, and new achievements are emerging [7, 8]. For example, Kleiner et al., [9] put forward a new data sampling method of BLB based Bagging learning thought, to solve the bottleneck problems of calculation in big data Bootstrap; Gonzalez et al., [10] presented distributed machine learning framework graph to realize the large-scale machine learning; Gao et al., [11] proposed the idea of "one pass learning" learning, trying to only scan again the data using constant storage to store intermediate results, which achieved good results in AUC optimization such a complex learning task.

However, there are still many problems to be solved in machine learning for big data due to it’s complexity [6, 7]. From the view of the algorithm, it mainly exists the following problems in machine learning and the analysis of big data mining. Because of the huge data size, it is not within the acceptable time to get results. So putting forward a new machine learning algorithm to meet the demand of high data processing and large data is one of the hot research points in machine learning [3, 7].

On this issue, the researchers generally solve the problem of big data machine learning through two methods. A method is to modify existing machine learning algorithm and transform it to be concurrent\parallel computing version[6,10]; the other method is to design a online learning version of existing machine learning and data mining algorithm [6, 7].

Because the online learning doesn't need to store samples of early data, or only needs to save the early data from a sample of a sufficient statistic, it is very suitable for big data

Received December 20, 2015; Revised March 18, 2015; Accepted April 8, 2016

analysis application scenarios. For very large data sets, online learning gets data in a sequential manner and synchronous update learning; for high speed data streams, online learning can be carried out while data is inputting. And the learning model can reflect a recent period of time input data rule and make effective prediction [12, 13]. In recent years, some researchers put online learning into the field of machine learning and produced better benefits. For example, Shalev-Shwartz S [14] proposed gradient ascent (descent) method based on the idea of online learning to implement a large-scale model of fast learning; Yang [15] proposed an incremental optimization fast decision tree algorithm for data with noise. Compared with the traditional big data mining decision tree algorithm, the main advantage of this algorithm was with real-time mining capabilities, which could store the complete data for training decision model when the mobile data stream was infinite.

As a kind of important online learning method, online multiple kernel learning has been widely used in different application fields [16-19]. However, due to the indirect influence of kernel learning problem on target analysis, online multiple kernel learning has not been fully researched in the field of big data analysis and mining [7].

In order to improve the performance of machine learning in big data environment, semi-supervised learning and online multiple kernel learning are introduced into the field of big data machine learning in this paper. First, a supervised online multiple kernel learning algorithm for big data (SOMK_bd) is proposed to reduce the computational workload during online learning kernel modification. Next, an unsupervised online multiple kernel learning algorithm (UOMK_bd) is proposed to adapt to the online learning environment. Then, an semi-supervised online multiple kernel learning algorithm for big data (SSOMK_bd) is proposed. Based on the online learning framework of incremental learning, SSOMK_bd makes full use of the abundant information of large scale labeled and unlabeled data, and uses the SOMK_bd algorithm and UOMK_bd algorithm to update the current reading data separately. Finally, experiments are conducted on the benchmark UCI data set to verify the effectiveness of the proposed algorithm.

The structure of this paper is as follows: the part I is the introduction; the part II introduces the proposed algorithms: SOMK_bd, UOMK_bd and SSOMK_bd; the part III is the experiment and analysis; the part IV is the conclusion; the part V is the acknowledgements.

2. Proposed Algorithms

In this paper, we propose three kinds of algorithms in big data environment, there are supervised online multiple kernel learning algorithm for big data (SOMK_bd), unsupervised online multiple kernel learning algorithm for big data (UOMK_bd) and semi-supervised online multiple kernel learning algorithm for big data (SSOMK_bd). The three algorithms are described in the following part of the paper.

2.1. Supervised Online Multiple Kernel Learning Algorithm for Big Data (SOMK_bd)

The main purpose of multiple kernel learning is to study kernel function with parametric or semi-parametric directly from the training data, which contain the information reflecting the data distribution. We take it as given a set of training data $D_L = \{ (x_i, y_i) | i = 1, 2, \ldots, n\}$, where x_i is feature set, $y_i \in \{-1, +1\}$ is the class label, $K_m = \{k_j(\cdot, \cdot) : X \times X \rightarrow R, j = 1, 2, \ldots, m\}$ is given a set containing m basic kernel functions, $u = \{ u_1, u_2, \ldots u_m, \sum u_i = 1\}$ is a set of non negative weights, minimizing the classification error of the kernel learning machine on the test set. Because of non negative weights, convex combination kernel function is still a valid kernel function. The problem can be formalized as the formula(1):

$$\min_{f \in H_k} \|f\|_{H_k}^2 + C \sum_{i=1}^{n} l(\mathrm{f}(\mathrm{x}_i), \mathrm{y}_i) \quad (1)$$

For the formula (1), it is difficult to directly compute the optimal value ,even it is basically impossible within acceptable time to find the optimal solution, especially under the big data environment .

In general, online multiple kernel learning will transform the optimization problem of formula (1) into the following problem: first, find out the optimal function f_i of each kernel K_i in their respective Hilbert space H_{ki} ; then, look for a set of weights u_i, which makes the f_i to be the best combination, and update weights u_i and f_i synchronously in the process of searching for the optimal value.

When the combination of kernel function is linear, the online multiple kernel learning problem can be solved by the following three steps: in the first step, train a classifier f_i using the basis of the training set for each kernel function. The second step is to perform online learning. After reading a training sample, different strategies are adopted to modify the kernel weight and the classifier according to different prediction results. In the third step, iterative until it meets certain condition, the optimal kernel function is a weighted combination of each kernel function with the optimal weight. During the second step and the third step, kernel weights and classifiers' updating strategy is as follows: first read a training sample, then determine whether the prediction of the sample is correct. If correct, do not perform any updating action; if not, update the corresponding kernel weight and classifier.

However, there is a particularly large number of kernel functions in the process of online multiple kernel learning. When each sample is input, all kernel function are used to predict and weight; once the forecast is not consistent with the correct label, the weight of all kernel functions are needed to change and modify. This will be a waste of computing resources, especially in big data environment, its efficiency is very low.

In order to improve the operating efficiency of online multiple kernel learning under the environment of big data and reduce its computational resource during modifying weight and classifier, we improve the traditional online multiple kernel learning algorithm , and put forward a supervised online multiple kernel learning algorithm for big data which is named as SOMK_bd.

In the integration process of online multiple kernel learning, SOMK_bd uses Bernoulli sampling to do a random sampling. Only the sample whose selection probability is 1 is selected to be constructed as a subset of kernel function and only the one in the subset is to carry out function prediction, weighted combination, kernel weight and classifier updating, reducing kernel calculation workload.

Algorithm 1. The Algorithm Description of SOMK_bd

Input:
– kernel function set: K_m= { k_1, k_2, …, k_m}
– The *t*-th labeled sample: (x_t ,y_t)
– Initialized classifier: F = { $f_1(t)$, $f_2(t)$, …, $f_m(t)$}
– Weight: $u_i(t)=1$, i=1,...,m
– Discount factor: $\beta \in (0,1)$
– Smoothing parameter: $\delta \in (0,1)$
Output: kernel weight $u_i(t+1)$ and classifier f_i (t+1), i=1,...,m
Procedure:
1) $q_{i(t)}=u_i(t)/[\max_{1\le j\le m} u_j(t)]$, i=1,...,m
/*Kernel weight probability $q_i(t)$*/
2) $p_i(t)=(1-\delta)q_i(t)+\delta/m$, i=1,...,m
/*Kernel selection probability $p_i(t)$*/
3) for i=1,2,...,m do
4) Sample $s_i(t)$=Bernoulli_Sampling(p_i)
/*$s_i(t)=1$ represents i-th kernel is selected, the one where $s_i(t)=1$construct a subset*/
5) end for
6) $y^*(t) = sign(\sum_{i=1}^{m} s_i(t) q_i(t) \, sign(f_i(t)))$
/*Weighting in the subset and predict the label of x_t*/
7) for i=1,2,...,m do
8) if $y^*(t)=y_t$ then /*correct, not update*/
9) $z_i(t)=0$
10) else /*not correct, upade*/
11) $z_i(t)=1$
12) end if
13) Update $u_i(t+1)=u_i(t)\beta^{z_i(t) m_i(t)}$
/*weight updating where $z_i(t)=1$ and $m_i(t)=1$*/
14) Update $f_i(t+1)= f_i(t) + z_i(t) m_i(t) y_t k_i(x_t ,\bullet)$
/*classifer updating where $z_i(t)=1$ and $m_i(t)=1$*/
15) end for

The detailed procedure of the algorithm SOMK_bd is shown in Algorithm 1. For the input labeled samples (x_t ,y_t), in step 1) the weighted probability $q_i(t)$ is calculated; in step 2) the i-th kernel selection probability p_i (t) is calculated. Here, $p_i(t)$ adds the smoothing parameter δ to ensure that each kernel is selected with the probability of δ/m, and avoid that the probability $p_i(t)$ is concentrated on a few kernel function; By step 4) Bernoulli sampling is carried out and the probability $s_i(t)$ is calculated. The one where s_i (t) = 1 represents that the i-th kernel is selected when the t-th sample is input. Then, the all kernel function selected by Bernoulli sampling(s_i (t) = 1) constitute a subset; In the step 6), the kernel is weighted in the subset and the predicted label of X is calculated; Step 13) and step 14) is the updating procedure of the kernel weight and kernel classifier. Here, the updating only focus on the subset of online multiple kernel learning, where $s_i(t)=1$; That is to say that we only update the kernel weight and classifier where $z_i(t)=1$ and $m_i(t)=1$ in SOMK_bd algorithm.

2.2. Unsupervised Online Multiple Kernel Learning Algorithm for Big Data (UOMK_bd)

Traditional kernel learning method based on data dependence is a commonly unsupervised kernel learning method, only considering the density of data distribution. In essence, it is to modify the kernel function in the training set. It can modify any existing kernel function based on the observed data samples , the essence is to modify the inner product on the Hilbert space induced by the kernel function. In general, the kernel function is modified by the formula (2), which makes the distribution of the D in the data set.

$$K_D(a, b) = k(a, b) - k_a^T(I + MK_D)^{-1}MK_b \tag{2}$$

Where D={x_1,x_2,…,x_n}; K is kernel function, k_{xi}=(k(x_i,x_1),k(x_i,x_2),…, k(x_i,x_n))), K represents the Gram matrix of k about D; a and b are 2 training samples; W_{ij}=RBF(x_i,x_j), Where x_i and x_j are the elements on the D , represents a symmetric distance matrix.

The calculation of formula (2) needs to be in the off-line batch, and the computational time complexity is high. At the same time, the kernel function’s updating of a and b need to compute K_a and K_b in the data samples. For large data streams, M and K_D is in changing, and not easy to calculate. In addition, direct calculating the whole data set M and K_D is not realistic in computing resources.

In order to make the data depend on kernel learning method suitable for online kernel modification in big data environment. We provide an unsupervised online multiple kernel learning algorithm for big data, which is named as UOMK_bd.

Algorithm 2. The Algorithm Description of UOMK_bd

Input:
–D={x_1,x_2,…,x_n}
–Current input sample: x_t
–Gram matrix and distance matrix: *K*, *M*
Output: Updated kernel matrix: *K*
procedure:
1) Initialization K_{x0}
2) K_{xi}=(k(x_i,x_1),k(x_i,x_2),…, k(x_i,x_n)))
3) for j=1,…,N do
4) K_2=K(j, •)
5) $K_{x_t} = k_1 - k_1^T(I + MK)^{-1}MK_2$
6) end for
7) Using K_{xt} to update matrix K in the last row and the last column
/*FIFO displacement*/
8) return *K*

The algorithm description of UOMK_bd is shown in Algorithm 2. UOMK_bd focuses on the online updating of M and K_D. In order to facilitate the calculation of M and K_D, the size of the M and K_D are restricted to be fixed N*N. In order to make full use of information of all of the data samples to perform kernel learning, the sample data replacement strategy in D is designed to ensure that the other samples can become the element of D. Because the data generation rules of big data may change with the change of time, we learn from the operating system “first in first out” (FIFO) page replacement policy to replace the data samples in D. Here, the timeliness of data is considered.

In UOMK_bd, it maintains a working set of M which can be used as a cache to reflect data distribution in a period time. The working set limiting strategy has a certain locality, but it is compromised under the limited computation and memory resources. In the realization of FIFO strategy, new entrants to the sample are put at the downward and the right column in matrix M, and the elements of the first row and first column are removed.

2.3. Semi-supervised Online Multiple Kernel Learning Algorithm for Big Data (SSOMK_bd)

Under the environment of big data, labeled data loss is very common and the big data set is a mixed data set including labeled data and unlabeled data. If you only used labeled data to learn by a supervised learning method, then a supervised learning model does not have good generalization ability and it can cause large waste of unlabeled data; If you only use a large amount of unlabeled data with implied information to learn by unsupervised learning method, the unsupervised learning will ignore the value of labeled data. Semi-supervised learning is a new machine learning method between the traditional supervised learning and unsupervised learning, its purpose is to make full use of a large number of unlabeled samples to make up for the lack of labeled samples and improve the learning performance effectively.

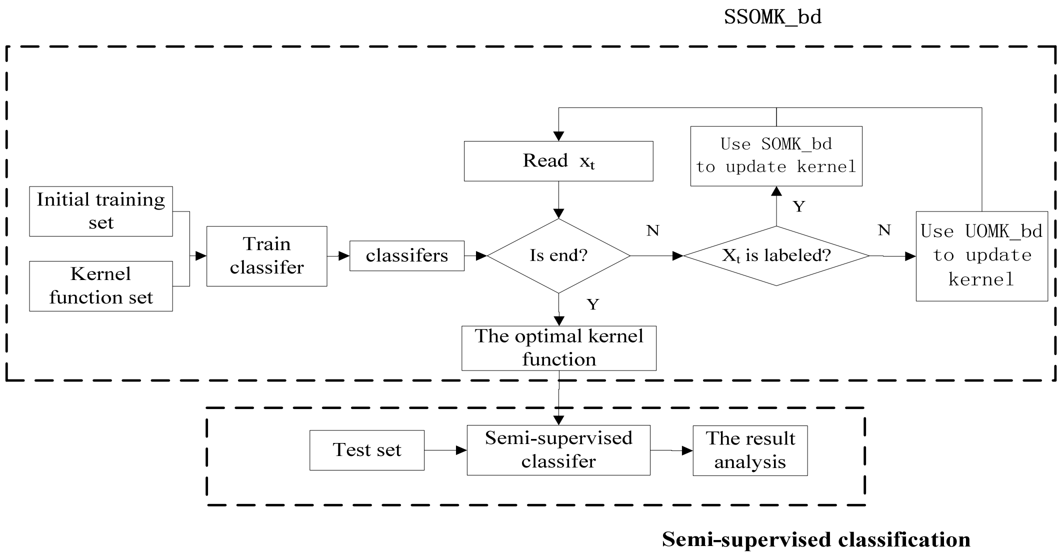

Figure 1. Semi-supervised Classification Flow Chart

Algorithm 3. The Algorithm Description of SSOMK_bd

Input:
–D={x_1,x_2,…,x_n}
–Sample: (x_t ,y_t)
Output: Updated kernel matrix K
Procedure:
1) Initialization *K*
2) Learn K from D_0
3) for each (x_t,y_t) in D
4) if y_t is not NULL then
5) call SOMK_bd to carry out updating
6) else
7) call UOMK_bd to carry out updating
8) end if
9) end for

Here, we proposed a semi-supervised online multiple kernel learning algorithm for big data which is named as SSOMK_bd, making full use of sample labeling information and unlabeled samples with implicit information to generate effective kernel functions and effective classifier. Specifically, when reading a sample, online learning method is based on the incremental learning to perform. First it determines whether the sample is marked. If marked, SSOMK_bd uses supervised online multiple kernel learning algorithm SOMK_bd presented above in part 2.1 to carry out kernel modification; otherwise, it uses the on-line unsupervised multiple kernel learning algorithm UOMK_bd presented above in part 2.2 to carry out kernel modifications. Iterative until they meet certain condition. Finally it generates the optimal kernel function to form a classifier and the test is carried out on the test data.

Among them, online multiple kernel learning and semi-supervised classification algorithm are shown in Figure 1, upper part within the dashed box represents online multiple kernel learning and the following part within the dashed box represents semi-supervised classification process. At the end of the online multiple kernel learning, it generates an optimal kernel function for semi-supervised classification to construct the classifier.

The description of algorithm SSOMK_bd is shown in Algorithm 3, where the kernel function is updated in step 5) and 7).

3. Experiment and Analysis

The experiments are carried out on UCI data sets and the proposed algorithms in the paper are compared with the existing online multiple kernel learning algorithm and batch processing of multiple kernel learning algorithm to make the effectiveness assessment.

In the experiment, SVM is used as a kernel function. RBF kernel and polynomial kernel , whose parameters are selected randomly, are used to construct kernel function [23, 24]. And 0-1 loss function is used to evaluate the error rate.

In order to cancel the orders of magnitude difference between the dimensions of data and avoid large prediction error caused by differences in input and output, data normalization function is used here. As shown in formula (3), the input feature value is normalized to [-1, 1] by data normalization function.

$$x_k = (x_k - x_{min}) / (x_k - x_{max}) \tag{3}$$

Where x_{min} represents the minimum value in data sequence, x_{max} represents the maximum value in data sequence.

Table 1. The Dateset1 of Experiment

Index	Dataset	Size	Dimensions
D1	Breast	683	9
D2	Splice	1000	60
D3	Dorothea	1150	100000
D4	Spambase	4601	57
D5	Mushrooms	8124	112

Table 2. The Dateset2 of Experiment

Index	Dataset	Size	Dimensions
D6	Forest CoverType	581012	54
D7	Poker-Hand	10^7	11
D8	Localization Data for Person Activity	164860	8

Such as Table 1 and Table 2, 8 kinds of UCI data sets are used in the experiment. Table 1 is used to verify the validity of the SOMK_bd algorithm and Table 2 is used to verify the validity of the SSOMK_bd algorithm.

For the sample selected in Table 2, the proportion of the training set and the test set is set to 1:1. The training set divided into labeled and unlabeled samples. Three types of semi-supervised classification experimental data sets are constructed according to the proportion of labeled samples to the number of training set samples. In first class, there are labeled samples accounted for 5% of the samples in the training set; in the second class, there are labeled samples accounted for 10% of the samples in the training set; in the third class, there are labeled samples accounted for 15% of the samples in the training set. We calculate the total classification rate by calculating the average of the 3 kinds of classification rates.

At the same time, the data in Table 2 is divided into 30 parts, which are to put in 4 files separately, and the data in the sequential read file is trained and tested on the online multiple kernel learning. All these data are used to evaluate the relationship between the data set size and the CPU processing time.

Table 3. The Experiment Result

	OM-2		SOMK_bd	
	mistake(%)	time(s)	mistake(%)	time(s)
D1	34.30±2.80	0.270±0.015	24.25±3.10	0.245±0.013
D2	30.80±1.41	0.420±0.005	25.56±1.21	0.380±0.020
D3	10.70±0.72	0.435±0.013	8.20±0.52	0.426±0.041
D4	58.16±1.50	5.184±0.239	23.30±1.10	3.30±0.68
D5	0.37±0.03	8.691±0.070	0.34±0.02	2.560±0.02

Table 4. The Correct Classification Rate Comparison (%)

	Algorithm 1	Algorithm 2	SSOMK_bd
D6	72.5±1.80	74.25±2.10	77.0±2.43
D7	78.7±2.01	78.18±1.15	81.70±1.37
D8	61.02±0.24	67.55±1.33	75.46±1.65

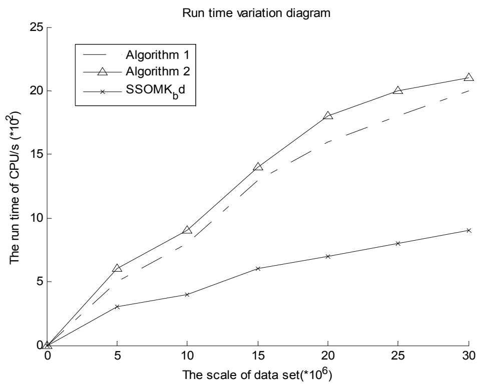

Figure 2. CPU Run Time Comparision

The first experiment is carried out to compare the proposed SOMK_bd with existing online multiple kernel learning algorithm om-2 [20] to verify its effectiveness in reducing the computational workload during kernel modification. As shown in Table 3, the experiment results show that SOMK_bd has effectiveness to reduce kernel scale, mainly in shortening the learning time and reducing the error rate. The reason is that SOMK_bd reduces calculation workload during kernel modification and selects some representative kernel to update.

The second experiment is carried out to compare the proposed semi-supervised online multiple kernel learning algorithm SSOMK_bd with the existing supervised online multiple kernel learning algorithm1 [21] and algorithm2 [22] to verify its effectiveness in improving the correct classification. As shown in Table 4, we can know that the proposed SSOMK_bd has higher correct classification rate than the existing algorithm1 [21] and algorithm2 [22]. Especially, SSOMK_bd is effective in dealing with large scale incomplete labeled data. The reason is that SSOMK_bd makes full use of the tag information of labeled samples and abundant information of unlabeled samples to train effective classifier, which can improve the performance of learning.

The data in Table 2 was divided into 30 parts to evaluate the relationship between the data set size and the CPU processing time. As shown in Figure 2, we can see SSOMK_bd has a good linear relationship between the growth rate of the running time and data set size. The experiment result indicates that SSOMK_bd is with good scalability and could be used in more regulations data analysis and application. The reason is that SSOMK_bd uses incremental learning framework to improve the classification performance and it is effective in the big data environment.

4. Conclusion

In this paper, the online multiple kernel learning algorithm under big data environment is studied deeply, and the semi-supervised learning is introduced into the field of big data machine learning. The traditional kernel learning algorithms is improved to reduce the computational workload during kernel modification. The incremental learning framework is used to improve the classification performance in big data environment. Based on the current reading of large data fragments in an online way, the algorithm makes full use of the rich information of labeled data and unlabeled data to achieve kernel updating and construct efficient kernel function. The experiment is conducted on the benchmark UCI large data set, the results show that the proposed algorithms are effective in shortening the learning time and reducing the error rate. Also the proposed algorithms could be used in big data analysis and application.

Acknowledgements

The Project Supported by Natural Science Basic Research Plan in Shaanxi Province of China (Program No.2015JM6347); The Project Supported by Science Technology Plan in Shangluo City of China (Program No. SK2014-01-15); The Project Supported by Science Research Plan of Shangluo University (Program No.14SKY026).

Reference

[1] LI Wujun, Zhou Zhihua. Learning to hash for big data: Current status and future trends. *Chinese Science Bulletin*. 2015; 60(5-6): 485-490.
[2] Mayer-Schönberger V, Cukier K. Big Data: A Revolution That Will Transform How We Live, Work, and Think. *American Journal of Epidemiology*. 2014; 17(1): 181-183.
[3] Li Zhijie, Li Yuanxiang, Wang Feng, et al. Online Learning Algorithms for Big Data Analytics: A Survey. *Journal of Computer Research and Development*. 2015; 52(8): 1707-1721.
[4] Zhou ZH, Chawla NV, Jin Y, et al. Big Data Opportunities and Challenges: Discussions from Data Analytics Perspectives. *IEEE Computational Intelligence Magazine*. 2014; 9(4): 62-74.
[5] Tan M, Tsang IW, Wang L. Towards Ultrahigh Dimensional Feature Selection for Big Data. *Journal of Machine Learning Research*. 2014; 15(2): 1371-1429.
[6] Yoo C, Ramirez L, Liuzzi J. Big data analysis using modern statistical and machine learning methods in medicine. *International Neurourology Journal*. 2014; 18(2): 50-57.
[7] HE Qing, LI Ning, LUO Wen-Juan, et al. A Survey of Machine Learning Algorithms for Big Data. *Pattern Recognition & Artificial Intelligence*. 2014; 27(4): 327-336.
[8] Hansen TJ, Mahoney M W. Semi-supervised Eigenvectors for Large-scale Locally-biased Learning. *Journal of Machine Learning Research*. 2014;15(11): 3691-3734.
[9] Kleiner A, Talwalkar A, Sarkar P, et al. *The big data bootstrap*. Proceedings of the 29th International Conference on Machine Learning (ICML). Edinburgh. 2012: 1759-1766.
[10] Gonzalez JE, Low Y, Gu H, et al. *PowerGraph: Distributed graph-parallel computation on natural graphs*. Proceedings of the 10th USENIX Symposium on Operating Systems Designand Implementation (OSDI). Hollywood. 2012: 17-30.
[11] Gao W, Jin R, Zhu S, et al. *One-pass AUC optimization*. Proceedings of the 30th International Conference on Machine Learning(ICML). Atlanta. 2013: 906-914.
[12] Hoi SCH, Wang J, Zhao P. LIBOL: A Library for Online Learning Algorithms. *Journal of Machine Learning Research*. 2014; 15(1): 495-499.
[13] Yang Haifeng, Liu Yuan, Xie Zhenping, et al. Efficiently Training Ball Vector Machine in Online Way. *Journal of Computer Research and Development*. 2013; 50(9): 1836-1842.
[14] Shalev-Shwartz S, Zhang T. *Accelerated proximal stochastic dual coordinate ascent for regularized loss minimization*. Proceedings of the 31st International Conference on Machine Learning (ICML). Beijing. 2014: 64-72.

[15] Yang H, Fong S. *Incrementally optimized decision tree for noisy big data*. Proceedings of the 1st International Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications. ACM. 2012.
[16] Motai Y. Kernel Association for Classification and Prediction: A Survey. *IEEE Trans Neural Netw Learn Syst*. 2014; 26(2): 208-223.
[17] Zheng H, Ye Q, Jin Z. A Novel Multiple Kernel Sparse Representation based Classification for Face Recognition. *Ksii Transactions on Internet & Information Systems*. 2014; 8(4): 1463-1480.
[18] Shrivastava A, Pillai JK, Patel VM. Multiple kernel-based dictionary learning for weakly supervised classification. *Pattern Recognition*. 2015; 48: 2667-2675.
[19] Liu KH, Lin YY, Chen CS. Linear Spectral Mixture Analysis via Multiple-Kernel Learning for Hyperspectral Image Classification. *IEEE Transactions on Geoscience & Remote Sensing*. 2015; 53(4): 2254-2269.
[20] Luo J, Orabona F, Fornoni M, et al. *OM-2: An Online Multi-class Multi-kernel Learning Algorithm*. Proceeding of CVPR 2010. 2010: 43-50.
[21] Orabona F, Luo J, Caputo B. Multi Kernel Learning with online-batch optimization. *Journal of Machine Learning Research*. 2012; 13(1): 227-253.
[22] Jin R, Hoi SCH, Yang T. Online Multiple Kernel Learning: Algorithms and Mistake Bounds. *Lecture Notes in Computer Science*. 2010; 6331(4): 390-404.
[23] Zhang W, Guo Z, Zhang L, et al. Adaptive Control for Robotic Manipulators base on RBF Neural Network. *TELKOMNIKA Telecommunication Computing Electronics and Control*. 2013; 8(3): 497-502.
[24] Pan LJ, Jin W, Wu J. A Novel Intrusion Detection Approach using Multi-Kernel Functions. *TELKOMNIKA Telecommunication Computing Electronics and Control*. 2014; 12(4): 1088-1095.