Review of Economics & Finance
Submitted on 03/12/2017
Article ID: 1923-7529-2019-01-57-10
Siyang Jiang, Xiuyan Yao, Qianju Long, Jiali Chen, and Hui Jiang

Fund Investment Decision in Support Vector Classification Based on Information Entropy[1]

Siyang Jiang
School of Mathematics and Information Science, Shaanxi Normal University,
Chang'an Avenue, Chang'an District, Xi'an, Shaanxi 710119, CHINA
Tel: +8615619263398 E-mail: jiangsy@snnu.edu.cn

Xiuyan Yao, ***Qianju Long***, ***Jiali Chen***
School of Mathematics, Huizhou University,
Yan Da Avenue, Huicheng District, Huizhou, Guangdong 516007, CHINA
Tel: +8615767973182 E-mail: 1174094357@qq.com

Prof. Dr. ***Hui Jiang*** (Correspondence author)
School of Information Science and Technology, Huizhou University,
Yan Da Avenue, Huicheng District, Huizhou, Guangdong 516007, CHINA
Tel: +8615322143789 E-mail: jianghui200207@163.com Fax: +867522527017

Abstract: As to the complex investment decision with big data, how to portray the essential characteristics to answer its evolving complexity of the mechanism, and how to make risk identification, assessment and measurement have become problems which urgently need to be addressed by investors. In this paper, the support vector classification based on information entropy (IE-SVC) is put forward to improve the accuracy in the field of capital investment decisions. Two classic methods, the K-Nearest Neighbors algorithm (K-NN) and the Radius Basis Function Neural Network (RBFNN), are applied to compare the performance. In the experiment of Gates foundation investment decision, its results show that the IE-SVC can be faster and higher accuracy than those of other methods.

Keywords: Information entropy; Support vector classification; Radius basis function neural network; K-Nearest Neighbors algorithm

JEL Classifications: C63, C81, C89

1. Introduction

In the current information era of big data, financial data are mostly non-linear and unstructured. They show some characteristics such as large-scale, complex relationship, and changeable regularity. Especially, in the field of financial investment research, a large quantity of data has many uncertainties. Some models based on traditional statistical techniques have many assumptions

[1] This research is funded by the national social science fund in China, No. 15BTJ024, and by the '12.5' planning project of common construction subject for philosophical and social sciences in Guangdong , No. GD12XGL21.

and can't be applied to the practical application. We are in great need of finding more efficient algorithms for the investment fund to improve the quality and reduce the risk.

In recent years, some methods are applied to fund investment decision. For example, Jane and Ligia (2014) have proposed a spatial analysis methodology to identify optimal areas for the development of new affordable housing in Polk County Jane and Ligia (2014). Another method, the K-Nearest Neighbors (K-NN), has often been applied to assess investment projects and selection by experts, e.g. Stone, *et al*. (2013). K-NN is one of the statistical methods with a more mature theory, and a non-parametric method used for classification and regression in pattern recognition, stated by Hall, *et al*. (2008) and Altman, *et al*. (1992). It can improve the performance among business processes, planning process, potential performance forecasting. In addition, Deshpande and Karypis (2014) recommended it as a cutting-edge research in the field of the international academic community, artificial neural networks have a very wide range of applications in the control and optimization, prediction and management, credit investment, communications and other aspects, and are mainly used for stock market transactions in the financial sector, economic forecasting and analysis of credit risk. Due to the excellent characteristic about fault-tolerant computing, artificial neural networks had been applied to evaluate all aspects of investment decision.

In fact, support vector classification (SVC) and neural networks have a lot of similarities. Compared to a feasible straight line found by single-layer perceptron in neural network, SVC can find the best hyperplane of classification. Han, *et al*.(2010) showed that SVC has an outstanding ability about solving curse of dimensionality. With kernel method, it can provide an effective learning way to financial data. In recent years, many scholars have designed effective support vector algorithms for online learning sample storage control. Wang, *et al*.(2002) have designed a new geometric algorithm of support vector classification and proposed a new method based on fixed point theorem. It can handle large-scale mapping problems. Zhu, *et al*.(2004) have proposed projected support vector classification to put the training and test data from the front projection into the subspace structure, and finally applied the linear SVC to the projected data classification. SVC can not only speed up the computing speed, but also reduce the interference of noise. In practice, Bottou, *et al*. (2007) endorsed it is widely used, particularly in the volatile data, non-bulk samples to obtain training and limited computer storage space under the classification conditions since it can effectively improve the learning performance. Ye, *et al*. (2006) have introduced entropy support vector classification modeling and focused on the data analysis and the relationship between the distribution of information entropy and support vectors, contributing to the improvement of the classification speed. Some scholars have combined k-means clustering algorithm and the method of support vector classification to reduce the size of the training set. It can effectively save time in SVC training and prediction while ensuring the generalization, according to Yao, *et al*. (2013).

However, there is no such an application that combines information entropy and SVC in the area of the investment classification, especially for large-scale data of fund investment decisions. Inspired by these references above, we will use information entropy theory to filter out data from a large-scale data set, then adopt SVC to train the selected data for an optimal classifier. Ultimately, we obtain a method of controlling costs and storage capacity of classification in learning theory. Experiment is carried out in the fund investment decisions in high-dimensional data. Compared with two classical methods, the final results of the proposed method demonstrate the better performance.

The rest of this paper is organized as follows. Support vector classification based on information entropy is provided in section 2, two classic approaches in section 3, experimental results in fund investment decision in section 4, and the overall conclusion in the last section.

2. Support Vector Classification Based on Information Entropy

2.1 Support vector classification

Based on statistical learning theory, SVC was proposed by Vapnik. Its important foundations are VC dimension and structural risk minimization theory, according to Shalev-Shwartz and Srebro (2008). SVC was originally proposed by the linearly separable problem, which can be classified without mapping in the original input space. But in practice, as the vast majority of problems are nonlinear, the linear separable SVC is powerless. Yu, *et al*. (2015) solved the linearly inseparable problem. The common approach is that the original samples of the input space are mapped into high-dimensional feature space by nonlinear mapping to construct the optimal hyperplane. The details of the basic idea and principle of linearly inseparable SVC are described as follows.

Let $T=\{(x_i,y_i)\}_{i=1}^{n}$ be the training set, which consists of two categories of component. If the input x_i belongs to positive class, then the output $y_i=1$; Otherwise, $y_i=-1$. Our goal is to map the input into a higher dimensional feature space F via a nonlinear mapping $\phi(x)$ and find a function f that has an ϵ-deviation from the actually obtained target y_i for all training set at the same time is flat as possible. When using nonlinear kernel SVC, we can obtain the border discriminant function as follows.

$$f(x)=sgn(\theta\cdot(\phi(x)\cdot\phi(x_i)+b)) \tag{1}$$

where $\phi(x)$, θ and b denote a nonlinear mapping, a weighted value and a bias, respectively. Considering the soft margin formulation, we have to solve the following problem with slack variables.

$$\min \quad \frac{||\theta||^2}{2}+C\sum_{i=1}^{n}\xi_i \tag{2}$$

subject to

$$\begin{aligned} & y_i(\theta\cdot\phi(x)+b)\geq 1-\xi_i \\ & \xi_i>0, i=1,2,\ldots,n \end{aligned} \tag{3}$$

Here, ξ_i is a slack variable. C is a penalty factor and plays an important role in controlling the punishment degree of misclassified. Solution of the formula is also converted to its dual problem. To solve the dual problem, we need to calculate dot product of the sample point vector. When the samples map into the high dimensional feature space, it also needs to calculate the dot product, resulting in the increase of amount of computation. A suitable kernel function $\Phi(x_i,x_j)$ should be found to replace dot product operation.

$$\Phi(x_i,x_j)=\phi(x_i)\phi(x_j) \tag{4}$$

In this paper, we choose the sigmoid kernel function with parameters η and μ

$$\Phi(x,x_i)=tanh(\eta(xx_i)+\mu) \tag{5}$$

Therefore, Eq.(2) can be changed to the following dual problem.

$$\max \quad \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\, \alpha_j y_i y_j \Phi(x_i, x_j) \tag{6}$$

subject to

$$\begin{aligned} &\sum_{i=1}^{n} \alpha_i\, y_i = 0 \\ &0 \le \alpha_i \le C, i = 1,2, \dots, n \end{aligned} \tag{7}$$

where $\alpha = [\alpha_1, \alpha_2, \cdots, \alpha_n]^T$ is the vector of Lagrangian multiplier. Let $\boldsymbol{\alpha}^* = [\alpha_1^*, \alpha_2^*, \cdots, \alpha_n^*]^T$ be the solution of the dual problem, then the optimum solution $[\theta^*, b^*]^T$ can be obtained.

$$\begin{cases} \theta^* = \sum_{i=1}^{n} \alpha_i^*\, y_i \phi(x_i) \\ \\ b^* = -\frac{1}{2}\theta^*(\phi(x_r) + \phi(x_s)) \end{cases} \tag{8}$$

Here, the instance $\{(x_i, y_i)\}(i = 1,2, \cdots, n)$ corresponding to $\alpha_i^* \neq 0(i = 1,2, \cdots, n)$ names a support vector. $\phi(x_r)$ and $\phi(x_s)$ are two random categories in any pair of support vectors. Ultimately, the optimal classification function is Eq.(9).

$$f(x) = sgn(\sum_{i=1}^{n} \alpha_i^*\, y_i \Phi(x_i, x_j) + b^*) \tag{9}$$

2.2 Support vector determined with information entropy

Information Entropy (IE) is defined as the occurrence probability of a discrete random event, which can be understood as the probability of occurrence of a particular information. It can help us to solve the problem of quantitative measurement of information, which represents the uncertainty of a random event or the measure of information. IE is the theory foundation of modern information theory, and greatly promotes the development of the information theory. Its principle is summarized as follows (Babaie and Lucas, 2009).

Let $\boldsymbol{x} = [x_1, x_2, \cdots, x_N]$ be a finite set and $\boldsymbol{p}(\boldsymbol{x}) = [p_1, p_2, \cdots, p_n]$ be a corresponding proper probability mass function. The amount of information needed to fully characterize all of the elements of this set consisting of N discrete elements is defined by $\boldsymbol{I}(\boldsymbol{x}_N) = \log_2 N$ (Hartley's formula). Shannon built on this formula to develop his information entropy as Eq.(10).

$$\boldsymbol{IE}(\boldsymbol{x}) = \mathrm{E}[-\log_2 p(\boldsymbol{x})] = -\sum_{i=1}^{N} p_i \log_2 p_i \tag{10}$$

Here, with $p\log_2 p$ tending to zero as p tends to zero. $\mathrm{E}(\cdot)$ denotes the expectation. This information entropy measures the uncertainty or informational content that is implied by $\boldsymbol{p}(\boldsymbol{x})$. The entropy-uncertainty measure $\boldsymbol{IE}(\boldsymbol{x})$ reaches a maximum when $p_1 = p_2 = \cdots = p_n = \frac{1}{n}$ and a minimum with a point mass function.IE measures the degree of system chaos and data smoothness. Generally, more chaos data system is, greater entropy there are.

In SVC, we should find a hyperplane with positive and negative category in the training sample set (Figure.1). This hyperplane can split the training data into two parts ($H+$ and $H-$). The perpendicular distance to the hyperplane between two domain boundaries is the largest. Namely the interval between $H+$ and $H-$ is the maximum. From Eq.(8), we know the support vectors are crucial to determine the hyperplane, and most of instances are not support vectors and have no effect on the result of the classification. Therefore, we can get rid of these useless vectors to accelerate the training process of SVC.

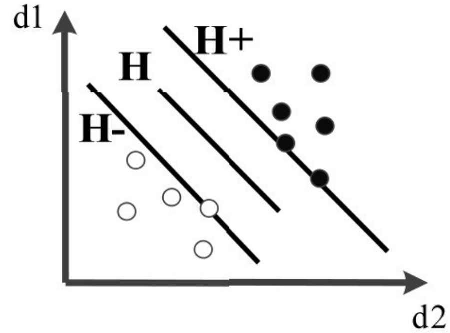


Figure 1. Support vector classification

Ye, *et al*. (2006) found that the high entropy points are concentrated on the boundary. Experimental results demonstrated that support vectors contained high IEs. Based on these points, we proposed SVC with information entropy (IE-SVC) to make a decision of the fund investment. At first, we can derive an optimal hyperplane from the preliminary training set by SVC and calculate IE of each training instance. Then the threshold of IE is set to choose the key instances. Finally, we carry out the SVC with the optimal training sample.

The calculation of information entropy steps is described as follows.

(1) Change the original data into the standardization of pretreatment.

(2) Obtain the preliminary random hyperplane H.

(3) Compute the relative Euclidean distance $d(x+)$ and $d(x-)$, $d(x+)$ is the distance between the positive instance $(x+)$ and H; $d(x-)$ is the distance between the negative instance $(x-)$ and H.

(4) Use the distance information from the instance x to the hyperplane H and then introduce the parameter g of the Sigmoid function to compute the posterior probability.

$$\boldsymbol{p}(+|\boldsymbol{x}) = \frac{1}{1+\exp(-g \cdot d(x+))} \tag{11}$$

$$\boldsymbol{p}(-|\boldsymbol{x}) = \frac{1}{1+\exp(-g \cdot d(x-))} \tag{12}$$

where g is the parameter of sigmoid function and $g > 0$.

(5) Calculate the positive and negative instance IE values, respectively.

$$\mathbf{IE}+= -\frac{1}{1+\exp(-g \cdot d(x+))}\log_2^{\frac{1}{1+\exp(-g \cdot d(x+))}} \tag{13}$$

$$\mathbf{IE}-= -\frac{1}{1+\exp(-g \cdot d(x-))}\log_2^{\frac{1}{1+\exp(-g \cdot d(x-))}} \tag{14}$$

(6) Repeat the above steps until the IEs of all positive and negative instances are obtained.

After then, a sample is selected for the final training data set depending on whether its IE value is larger than the threshold or not. In this way, we can more effectively partition the large-scale data, and further improve the classification speed. Meanwhile, the accuracy of classification is higher.

3. Two Classic Approaches

3.1 Radial basis function neural networks

Radial basis function neural network (RBFNN) is an artificial network that uses radial basis functions as activation functions. Formulated by Broomhead and Lowe firstly, RBFNN has been applied to many applications, such as pattern classification, system identification, nonlinear function approximation, adaptive control, speech recognition, time-series prediction, and so on. It is a commonly used three layers feed forward neural network and has an input, hidden and output layer. The input layer is composed of input vectors. The hidden layer consists of radial basis activation function as networks neuron. The net input to the radial basis activation function is the vector distance between its weight and the input vectors, multiplied by the bias. It has not only solid theoretical basis, a concise mathematical form, but also some advantages of an easy design, good generalization, and high tolerance of input noises and ability of online learning. In recent years, many experts, say, Ahmed, *et al*. (2015) and Komijani, *et al*. (2016), have applied RBFNN to lots of practical problems.

Let the input vector $\mathbf{x} = [x_1, x_2, \cdots, x_n]$ and the output layer vector $\boldsymbol{y} = [y_1, y_2, \cdots, y_m]$. $\boldsymbol{w} = [\omega_1, \omega_2, \cdots, \omega_k]$ is the hidden-to-output weight vector, and ν denotes the total number of hidden nodes. Figure.2 presents the structure of a single output RBF network. The network output can be obtained by Eq.(15).

$$\boldsymbol{y} = \Gamma(\boldsymbol{x}) = \sum_{i=1}^{\nu} \omega_i \, \psi_i(\boldsymbol{x}) \tag{15}$$

where $\Gamma(\boldsymbol{x})$ is the final output, $\psi_i(\cdot)$ denotes the radial basis function of the i-th hidden node.

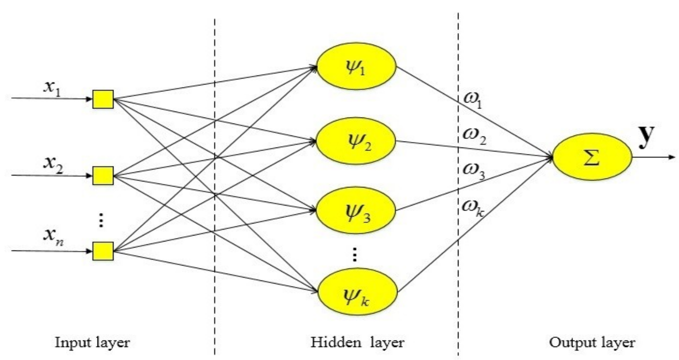


Figure 2. The structure of a radial basis function neural network

The radial basis functions can have different forms. The most popular among them is Gaussian function:

$$\psi_i(\boldsymbol{x}) = \exp(-\frac{1}{2\sigma_i^2} ||\boldsymbol{x} - c_i||^2) \tag{16}$$

where c_i and σ_i denote the center and spread width of the i-th node respectively.

Generally, there are two steps for the RBFNN training. The first step is to determine the parameters of radial basis functions, and the second to determine the output weight $\boldsymbol{w}$ by supervised learning method. The detailed process is available in the reference Sun, *et al*. (2009).

3.2 K-Nearest Neighbors classification

K-Nearest Neighbors (K-NN) classification is also one of the well-known machine learning algorithms. It finds a group of k instances in the training set that are the nearest neighbor to the test instance, and bases the assignment of a label on the performance of a particular class in this neighborhood. As a lazy learning algorithm based on examples, it compares the similarity of training tuples and test tuples by learning analogy. It is easy to implement and support incremental learning, but also make ultra-polygon modeling in complex decision space. K-NN is an online technology where new data can be added directly to a data set without retraining and widely used in such areas as text classification, pattern recognition, image processing and spatiotemporal classification. Zhang, *et al*. (2016) conducted sets of experiments on big data and found that the proposed K-NN algorithm classification works well in terms of accuracy and efficiency. Niu, *et al.* (2013) showed that, its reformed algorithm also used in massive data acquisition and storage and showed excellent ability of classification and low computation cost.

Let the training set $T = \{(x_i, y_i)\}_{i=1}^{n}$ and the test set $D = \{(x_j, y_j)\}_{j=1}^{l}$, where $x_i, x_j \in X \subseteq \mathrm{R}^m$, $y_i, y_j \in Y \subseteq \{0,1\}$, n and l are the training and test sample sizes respectively. Given a test instance $z = (x_q, y_q) \in D$, the algorithm computes the distance between z and all the training instances to determine its nearest neighbor list, T_z. Once T_z is obtained, the test instance is classified based on the majority class of its nearest neighbors, according to Wu, *et al*. (2008):

$$y_q = \underset{\tau}{\operatorname{argmax}} \sum\nolimits_{(x_i, y_i) \in T_z} F(\tau = y_i) \tag{17}$$

where τ is a class label, y_i is the class label for the i −th nearest neighbors, and $F\ (\cdot)$ is an indicator function that returns the value 1 if its argument is true and 0 otherwise. The K-NN classification pseudo codes are available in the reference Zhao and Chen (2016).

4. Fund Investment Decision

In this section, experiments on Goodgrant Foundation will be carried out to assess the performance of our methods. Sigmoid kernel is adopted here and all calculations are performed with programs developed in MATLAB2016a. In order to speed up the global learning, we exploit the optimizer based on MATLAB support vector machine toolbox (Libsvm-3.21). All experiments are carried out in the same personal computer with Intel(R)Core(TM)i7-6500U CPU@2.50GHz 2.50GHz.

In American mathematical modeling contest of 2016, problem C intends to determine the best investment strategy for Goodgrant Foundation (https://github.com/772922440/ProblemCDATA). Large-scale data provided have 7804 schools' information with 114 dimensions including geographical distribution, race property, tuition fees, specialization ratio, SAT and ACT, graduate employment and so on. In addition, there are some characteristics such as complex relationship, no significant distribution and non-linear complex structures. Generally, we should give some assumptions, as the following, to this experiment.

(1) Results of the investing decision are decided by the sample data, not by human factors.

(2) The data set is reliable and truthful and there is no human revising in it.

(3) The missing values which are supplemented are close to the real values.

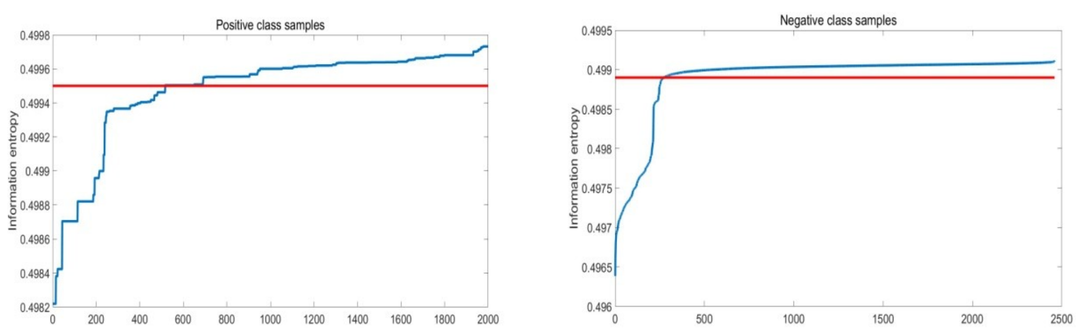
Based on the mature investment models of large investment institutions, we analyze the relationship between IE value and the support vector in line with the data features. A new information entropy support vector machine model is carried out to make investment decision.

We resort to the perfect investment system of large foundation, the Gates foundation, to determine the sample. From the foundation's website, the list of investment schools is downloaded. According to the characteristics of data attached to the topic, we firstly fill the missing value by the mode and quantify some data to satisfy the input of SVC. Then the schools fall into two categories. The positive class is composed of the schools invested by Gates foundation, and the others belong to the negative class. Finally, we divide them into the training and test set. Synthetic minority over-sampling technique, proposed by Chawla, *et al*. (2002), is used to balance the training set in order to eliminating the bias. In addition, the K-NN classifier and RBFNN are used to solve this problem, comparing the results with the IE-SVM.

From part 2.2, an initial hyperplane should be obtained to calculate the IE of each object. In this paper, we balance the training set and obtain the primary sample. In this process, C-SVC is carried out. Penalty factor C and the kernel function parameters are optimized by cross-validation. The Libsvm-3.21 based on MATLAB2016a is used to find the initial hyperplane. The experimental result (hyperplane H_0) based on the given data is as follows.

$$f_0(x) = \theta_0^* \cdot \phi(x) + b_0^* \tag{18}$$

Here, $\theta_0^* =$ [-1.7664, -0.2011, -0.5404, ... , -0.7310, -1.1255, 0.0423]^T, which is a vector of 114 dimensions. Because the appearance is limited, we just write some values. Then, b_0^* =-0.0142.



(a) (b)

Figure 3. Information entropy of positive and negative class samples

Euclidean distance between each instance and H_0 is calculated to determine the posterior probability in view of Eq.(11) and Eq.(12). Then, the IE of each instance can be calculated by means of Eq.(13) and Eq.(14). Meanwhile, the IE thresholds of positive and negative classes are determined by their mean values respectively. The results are described in Figure 3a and Figure 3b.

Thus, we filter out the instances whose IEs are less than the threshold. Namely, we extract the corresponding instances in the upper part of the red line in Figure.3a and Figure 3b. Once again, we use the Libsvm-3.21 toolbox to train the small-scale SVC to get the new hyperplane H_1.

$$f(x) = \theta^* \cdot \phi(x) + b^* \tag{19}$$

Here, $\theta^* =$ [-1.1368, 0.1875, 0.8868, ⋯, -0.5512, 0.3789, 0.1781]^T, and b^*= -0.0127.

We use the test and validation sets to verify the performance of IE-SVM. In order to reflect the efficient and reliable classification performance of this model in the field of financial investment, we carry out K-NN and RBFNN to solve the same problem. All experiments are conducted on the same personal computer. The results are shown in table 1. In this table, F1 Score is used to measure the correction of classification. It is an indicator used to measure statistics dichotomous model accuracy. F1 scores are regarded as a weighted average model precision and recall rates. Its maximum value is 1, the minimum value is 0.

Table1. The results of various classification models

		KNN	RBFNN	SVC	IE-SVC
The test set	F1 score	0.0309	0.0290	1	1
	Runtime(s)	34.456	63.751	63.751	26.318
The validation set	F1 score	0.0542	0.0161	1	1
	Runtime(s)	33.758	45.838	62.934	26.254

As can be seen from Table 1, compared with the results of the two classical algorithms, the accuracy of SVC is the highest but the running time is longer than RBFNN. However, the classification performance IE-SVC is the best whether the accuracy or the running time. These experiments are conducted in Matlab2016a. In Python, these results are the same. By analysis of the fund investment data which are large scale and high dimensional, the IE-SVC can significantly reduce the space and time complexity of classification and improve the accuracy.

5. Conclusion

Nowadays, many fund investments become more complex in the current big data background. As an important part of the financial sector, how to dig out valuable investment decision-making is of great significance. In order to improve the efficiency of investment decisions, we apply IE-SVC to the field of investment funds in this paper. Experiments on Goodgrant Foundation are carried out to assess the performance of our methods. Compared with the results of K-NN classification and other classifications, IE-SVC classification has higher performance of reducing the space and time complexity and improving the accuracy. IE-SVC classification can quickly and accurately identify a large number of investment candidates. This method can provide an efficient way to solve high-dimensional funds investment decisions.

Two main issues need to be pointed out in the experimental section. On one hand, different types of sigmoid kernel, such as Gaussian kernel, maybe impact our experimental results. We leave it as future work. On the other hand, we adopt cross-validation algorithm to optimize the parameters and resort to the Libsvm-3.21. Exploiting other algorithms to accelerate the computation and improve performance will be a highly challenging yet interesting topic.

References

[1] Ahmed, M. H., Hasan, S. and Ali. A. (2015). "Learning enhancement of radial basis function neural network with harmony search algorithm", *International Journal of Advances in Soft Computing & Its Applications*, 7(1): 78-103.

[2] Altman, N.S. (1992). "An introduction to kernel and nearest-neighbor nonparametric regression", *The American Statistician*, 46 (3): 175-185.

[3] Babaie, T., and Lucas, C. (2009). "Variable selection using information entropy in time series prediction", In: *Proceedings of the Seventh International Conference on Computer Science and Information Technologies*, Yerevan, Armenia: 118-121.

[4] Bottou, L., and Bousquet, O. (2007). "The tradeoffs of large scale learning", *International Conference on Neural Information Processing Systems*, Curran Associates Inc.: 161-168.

[5] Chawla, N.V., Bowyer, K.W., and Hall, L.O., *et al*. (2002). "SMOTE: synthetic minority over-sampling technique", *Journal of Artificial Intelligence Research*, 16(1): 321-357.

[6] Deshpande, M., and Karypis, G. (2014). "Item-based top-N recommendation algorithms", *ACM Transactions on Information Systems*, 22 (1): 143-177.

[7] Hall, P., Park, B.U., and Samworth, R.J. (2008). "Choice of neighbor order in nearest-neighbor classification", *Annals of Statistics*, 36(5): 2135-2152.

[8] Han, Z. H., and Zhu, X. (2010). "Selection of training sample length in support vector regression based on information entropy", *Proceedings of the CSEE*, 30(20): 112-116.

[9] Jane, M.R., and Ligia, D.O. (2014). "Where to invest affordable housing dollars in Polk County? A spatial analysis of opportunity areas", *Resilient Neighborhoods Technical Reports and White Papers*. [Online]Available at: http://works.bepress.com/jane_rongerude/4/

[10] Komijani, H., Rezaeihassanabadi, S., and Parsaei, M.R. *et al*. (2016). "Radial basis function neural network for electrochemical impedance prediction at presence of corrosion inhibitor", *Periodica Polytechnica Chemical Engineering*, 61(2):1-5.

[11] Niu, K., Zhao, F., and Zhang, S. (2013). "A fast classification algorithm for big data based on KNN", *Journal of Applied Sciences*, 13(12): 2208-2212.

[12] Shalev-Shwartz, S. and Srebro, N. (2008). "SVM optimization: inverse dependence on training set size", *International Conference on Machine Learning*, 28(3): 928-935.

[13] Stone, T., Zhang, W.N., and Zhao, X. (2013). "An empirical study of top-N recommendation for venture finance", *Acm International Conference on Conference on Information and Knowledge Management*: 1865-1868.

[14] Sun, T.Y., Liu, C., and Lin, C.L. *et al*. (2009). "A radial basis function neural network with adaptive structure via particle swarm optimization", In: *Particle Swarm Optimization*, Intech Publisher: 423-434.

[15] Wang, J.Q., Tao, Q., and Wang, J. (2002). "Kernel projection algorithm for large-scale SVM problems", *Journal of Computer Science and Technology*, 17(5): 556-564.

[16] Wu, X.D., Kumar, V., and Quinlan, J., *et al*. (2008). "Top 10 algorithms in data mining", *Knowledge and Information Systems*, 14(1): 1-37.

[17] Yao, Y.K., Liu, Y., and Yu, Y.Q., *et al*. (2013). "K-SVM: An effective SVM algorithm based on K-means clustering", *Journal of Computers*, 8(10): 2632-2639.

[18] Ye, N., Wang, D., and Dou, L. J. (2006). "Relation of information entropy and support vector", *Journal of Guangxi Normal University (Natural Science Edition)*, 24(4): 127-130.

[19] Yu, L., Shi, F., and Wang, H. (2015). *Thirty case analysis of MATLAB intelligent algorithm (Second Edition)*, Beijing University of Aeronautics and Astronautics Press: 240-279.

[20] Zhang, S., Deng, Z., and Cheng, D., *et al*. (2016). "Efficient KNN classification algorithm for big data", *Neurocomputing*, 195(C): 143-148.

[21] Zhao, M. and Chen, J.C. (2016). "Improvement and comparison of weighted k nearest neighbors classifiers for model selection", *Journal of software Engineering*, 10(1): 109-118.

[22] Zhu, Y. and Zhang, Y. (2004). "A new type SVM—projected SVM", *Science China (Series G:Physics,Mechanics and Astronomy)*, 47(S1): 21-28.