

Least-square Support Vector Machine for Financial Crisis Forecast Based on Particle Swarm Optimization

Xinli Wang

Economics and Management Department, North China Electric Power University, Baoding City, China
Email: wxli0704@163.com

Abstract—Whether listed companies run soundly or not has direct impact on development of capital market, therefore, how to forecast financial crisis of listed companies accurately has been a widespread topic. Essentially financial crisis of listed companies is mainly about model pattern classification. Considering that Particle Swarm Optimization (PSO) and Support Vector Machine (SVM) have great performance and strength on classification and regression analysis, this paper puts forward the hybrid forecast thought by combination of the two methods above as research focus. Firstly via building performance indicators, the forecast model based on classification is established and related parameters are optimized by PSO. Then empirical financial crisis analysis will be conducted on this method using financial data of listed companies. The simulation results indicate that the forecast model established in this paper combines the strength of artificial intelligence and statistics, and can avoid phenomenon of over fitting and under fitting compared with traditional models. Moreover, with strong generalization ability, the model is accurate and universal, hence having high application value.

Index Terms—particle swarm optimization, least squares support vector machine, pattern classification, financial crisis

I. INTRODUCTION

Since 1990's, securities market in China has developed dramatically. However overall speaking, performance of listed companies is under expectation, the main problems lie in the fact that serious financial crisis rises in some of the listed companies, and deficit scale and amount tends to increase annually. In current international circumstances, financial integration has become the general trend and grown rapidly. With relations of each economy closer and closer, financial safety of each nation will have great impact on economic development of others. Financial crisis in 2008 reflects this very well. Therefore, it is crucial for listed companies to strengthen risk management awareness, improve risk prevention policies, structure and detail financial management, and detect financial crisis in a timely, effective, and accurate manner.

Traditional financial crisis forecast models that are widely applied mainly include statistics model and

artificial intelligence model. However, these models have drawbacks in application and accuracy, easy to get constrained by strict assumption and sample selection, and consequently these problems lead to model generalization ability decrease and low forecast accuracy. In recent years, one machine learning method based on statistics theory – Support Vector Machine (SVM) has drawn more and more attention, but its accuracy to a large extent relies on selection of operation parameters whose process is very difficult and results are often not like what are expected. Particle Swarm Optimization (PSO) is an emerging technology in artificial intelligence field and is a branch of evolutionary algorithm. Having advantages of simple solving, few parameter adjustment, and strong ability of searching global optimum, PSO has gained wide acknowledgement. In order to improve forecast accuracy, this paper will use PSO and SVM to forecast financial crisis of listed companies.

This paper selects listed companies in Shanghai and Shenzhen as research samples to build PSO-SVM model, and conducts forecast analysis on financial crisis of those companies. In this model LS-SVM is applied. To validate the generalization ability of this model, this article uses different indicators to conduct forecast analysis on built model. The result indicates that the forecast model built in this paper has higher accuracy.

II. LEAST SQUARES SUPPORT VECTOR MACHINE

Least Squares Support Vector Machine (LS-SVM) is an evolution of SVM. It changes inequality constraints in traditional SVM to equality constraints, and treats Sum Squares Error loss function as experience loss of training set, thus transforming quadratic programming to solving linear equations and further increasing speed of solving problems and convergence accuracy.

Training data having N groups of samples is represented as $\{x_k, y_k\}$, $k=1,2,\dots, N$. In it, x_k is Dimension n input vector; y_k is a group of output vector. In feature space, classification model of SVM is

$$y(x) = \text{sign}[w^T \varphi(x) + b] \quad (1)$$

In it, nonlinear reflection $\phi(\bullet)$ reflects input data to high dimension feature space. Used for classification, optimization problem of LS-SVM is

$$\min_{w,e} J(w, e) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{k=1}^N e_k^2 \quad (2)$$

Constraint condition is

$$y[w^T \phi(x) + b] = 1 - e_k (k = 1, 2, \dots, N) \quad (3)$$

The equivalent Lagrangian function to build LS-SVM optimization problem is

$$L(w, b, e, \alpha) = J(w, e) - \sum_{k=1}^N \alpha_k \times \{y_k[w^T \phi(x_k) + b] - 1 + e_k\} \quad (4)$$

In it, α_k is Lagrange multiplier.

Condition of the optimum is

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \\ \frac{\partial L}{\partial b} = 0 \\ \frac{\partial L}{\partial \xi} = 0 \\ \frac{\partial L}{\partial a} = 0 \end{cases} \quad \left\{ \begin{array}{l} \omega = \sum_{i=1}^I a_i \bullet \phi(x_i) \\ \sum_{i=1}^I a_i = 0 \end{array} \right.$$

Lagrangian derivation of w, e_k, b, α_k , the result below can be got

$$Z^T = [\phi(x_1)^T y_1; \phi(x_1)^T y_2; \dots; \phi(x_N)^T y_N];$$

$$y = [y_1; y_2; \dots; y_N]; \quad 1_N = [1; 2; \dots; 1];$$

$$e = [y_1; y_2; \dots; e_n];$$

$$\alpha = [\alpha_1; \alpha_2; \dots; \alpha_n] .$$

Thus the linear equality below can be got

$$\left[\begin{array}{c} 0 \\ y \end{array} \middle| \frac{y^T}{\Omega + I_N/\gamma} \right] \left[\begin{array}{c} b \\ \alpha \end{array} \right] = \left[\begin{array}{c} 0 \\ 1_N \end{array} \right] \quad (5)$$

In it,

$$\Omega \in Z^T Z; \quad \Omega_{ij} = y_i y_j \phi(x_i)^T \phi(x_j) =$$

$$y_i y_j K(x_i, x_j), \quad i, j = 1, 2, \dots, N;$$

$$I_N \in R^{N \times N}$$

is the identity matrix.

Therefore, for a new point X^* , classification function

$$\hat{f}$$
 is:

$$\hat{f}(x^*) = \text{sign} \left[\sum_{i=1}^N \hat{\alpha}_i y_i K(x_i, x^*) + \hat{b} \right] \quad (6)$$

Step function is

$$\text{sign}(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (7)$$

In it, $\hat{\alpha}, \hat{b}$, can be got uniquely from formula (9), and radial basis function can be taken as kernel function, shown as follows:

$$K(x_i, x_j) = \exp \left(- \frac{\|x_i - x_j\|_2^2}{\sigma^2} \right) \quad (8)$$

In it, as the width of radial basis function, σ is an undetermined parameter.

III. PARAMETER Γ AND Σ SELECTION OF PSO

This paper selects Radial Basis Function (RBF) as kernel function of SVM. Parameter C and Y in RBF determine classification accuracy to a large extent. C represents punishment cost, whose value will have influence on accuracy of classification results. If C is big, the model's classification accuracy in the training will be high; if small, then low accuracy and it even makes the model useless. However, Y has greater impact on classification results than C, because its value will directly influence data distribution in feature space. Big value of Y will lead to over fitting, and small value to under fitting. So the best values for both C and Y have to be set up. While unfortunately it is very difficult to specify the best values manually. Therefore, in the model built in this paper, in order to improve model accuracy PSO is used to optimize SVM and further parameter C and Y.

A. Validation of Performance Indicators

With large amount of samples, in order to optimize γ and σ , n validation samples $\{x_j^y, y_j^y\}$ selected without replacement from original sample set I including N samples constitute validation sample set V, and remaining samples are used as training sample P for validation of performance indicators. Normally $n=1/3N$, and the validation performance indicator can use the formula below.

$$\begin{aligned} \min_{\gamma, \sigma} \left\{ \sum_{j=1}^n (y_j^y - \hat{f}_{\gamma, \sigma}(x_j^y))^2 + \sum_{i=1}^{N-n} (y_i^p - \hat{f}_{\gamma, \sigma}(x_i^p))^2 \right\} = \\ \sum_{j=1}^n \left(y_j^y - \left[\frac{1}{\Omega} \right]^T \left[\begin{array}{c} 0 \\ y \end{array} \middle| \frac{y^T}{\Omega + I_{N-n}/\gamma} \right]^{-1} \left[\begin{array}{c} 0 \\ 1_{N-n} \end{array} \right] \right)^2 + \\ \sum_{j=1}^{N-n} \left(y_j^p - \left[\frac{1}{\Omega^p} \right]^T \left[\begin{array}{c} 0 \\ y \end{array} \middle| \frac{y^T}{\Omega + I_{N-n}/\gamma} \right]^{-1} \left[\begin{array}{c} 0 \\ 1_{N-n} \end{array} \right] \right)^2 \end{aligned} \quad (9)$$

In it,
 $\Omega^v \in R^{(N-n) \times n}$; $\Omega^v_{i,j} = y_i y_j K(x_i, x_j)$;
 $i=1,2,\dots,N-n; j=1,2,\dots,n$

Part I in the formula above shows the classification error of sample P as support vector to sample V, and Part II shows the classification error of sample P as support vector to sample P. Regarding binary classification, equation (9) divided by 2 is the amount of misclassification of original sample set I, and further divided by N, will be misclassification percentage of I. Generalization ability of the model is taken into consideration when the performance indicator function is used to obtain LS-SVM parameters. Parameter γ and σ can be gotten through optimization of the equation above.

B. PSO Algorithm

In each iteration of PSO, a particle updates itself by tracking two extremums. One is the optimal solution found by an individual particle, called personal best (pbest); the other is the optimal solution found by the whole swarm, called global best (gbest). When finding the two optimal solutions, every particle updates its own speed and position according to the formula below.

$$v_{k+1} = \omega v_k + c_1(pbest_k - x_k) + c_2(gbest_k - x_k) \tag{10}$$

$$x_{k+1} = x_k + v_{k+1} \tag{11}$$

In the formula, v_k represents particle speed vector; x_k represents position of current particle; $pbest_k$ represents the position of optimal solution found by an individual particle; $gbest_k$ represents the position of optimal solution found by the whole swarm; ω represents inertia weight, whose value determines how much speed of current particle will be carried on, and proper selection will endow particles with balanced search ability (wide area search ability) and development ability (local search ability); c_1, c_2 represents cognitive factor, also known as learning factor. ω is some random number between (0, 1), c_1, c_2 are random numbers between (0,2). In each dimension every particle's speed will be limited within a maximum speed v_{max} ($v_{max} > 0$). If the updated speed in some dimension exceeds v_{max} set by customers, i.e. $v_k > v_{max}$, then $v_k = v_{max}$; when $v_k < -v_{max}$, then $v_k = -v_{max}$.

错误!未找到引用源。 . ANALYSIS

A. Selection of Model Variables and Samples

For every listed company, 4 financial indicators of operations condition are mainly taken into consideration,

that is, x 1= earning per share, x 2=net asset per share, x 3= Rate of return on net assets, x 4= cash flow per share. Each sample has 4 kinds of properties.

At present, domestic studies all use ST companies as financial distress samples. ST companies are ones that accept special treatment due to financial deficit for two consecutive years or net asset per share lower than face value of share. From S companies announced by Shanghai Stock Exchange in 2010, 32 companies that have been STed in the past four years are selected as financial distress samples; 122 companies that have not been STed in 2010 and the four years backwards are selected as non-financial distress samples. So the total sample amount is 154.

This article uses 21 ST companies and 80 non-ST companies, in total 101, as training samples, taken as sample I, and remaining 53 as testing samples, taken as sample T. So there are 101 training samples and 53 testing samples. Financial data respectively in 2007 and 2009 (source: www.stockstar.com) is used to forecast financial situation in 2010. Firstly sample data is processed, LS-SVM output value under financial distress is set as 0, and output value as 1 when not under distress.

B. Steps that PSO optimizes LS-SVM Parameters

- 1) Training sample I mentioned above has 21 ST companies and 80 non-ST companies, in total 101; 7 ST companies and 27 non-ST companies selected at random, in total 34 are used as validation samples in equality (9) validation performance indicators, and remaining 67 companies are used as training samples.
- 2) Selection of swarm m. When m is small, possibility of going into local optimum is big; when m is large, optimization ability of PSO is great but with increasing calculation complexity. In it here m is set as 50.
- 3) Selection of learning factor c1, c2. Learning factor endows particles with abilities of self summarizing and learning from the best individual in the swarm, thus approaching the optimal point in the swarm or neighborhood. c1, c2 are normally set as 2, but literatures may have other values. c1 is equal to c2 and both are in the range of 0~4. This article sets c1, c2 as 2.
- 4) Selection of inertia weight ω . Larger inertia weight makes the particle gain a faster speed in its own original direction, hence flying further and having better exploration ability; smaller inertia weight makes the particle inherit a slower speed in original direction, hence flying nearer and having better development ability. This article sets ω as 0.5.
- 5) Selection of maximum velocity v_{max} . Maximum velocity determines the longest moving distance in iteration. Max v is set as changing domain of each dimension. Parameters of LV-SVM $\gamma \in (0,100)$, $\sigma^2 \in (0,100)$.
- 6) Selection of neighborhood topology. Global version PSO regards the whole swarm as particle

- 7) neighborhood, with faster speed but easy to fall into local optimum sometimes; local version PSO bears lower convergence speed, but hard to fall into local optimum. Global version neighborhood topology is selected in this article.
- 8) Stopping criterion. The largest iteration times, 100 are used.
- 9) Fitness function is shown in formula (9), and iteration is conducted according to formula (10) and (11).
- 10) Take parameter σ_2 into function (8), and radical basis function kernel can be obtained; then parameter γ , radical basis function kernel, and training sample I data are taken into equality (5), and parameter α and b can be obtained; lastly take related parameters and training sample I into function (6), the new LS-SVM classification function can be obtained.
- 11) Take testing sample T into LS-SVM classification function, and financial forecast results can be obtained.

C. Simulation Training and Financial Crisis Forecast

Based on model variables and samples selected above, and using LS-SVM model, this paper compares the results by particle swarm optimization LS-SVM parameter method, and mesh searching algorithm, and 10-fold Cross-Validation. Each method is operated for 10 times, and initial parameter value is generated within defined range at random.

The restrained conditions for the optimization process mentioned above could be either that mean squared error of model forecast result is zero or that particle iterative time is the set value. Endow the optimal particle swarm position (c, σ) obtained from optimization with LS-SVM, use testing samples to rebuild regression model, and then model forecast results of testing samples are obtained. Finally the system model parameters that have been optimized are gotten, based on which empirical model of financial forecast can be further built for actual forecast. Regarding the financial forecast model built via LS-SVM, PSO is used for optimal solution searching. Selected parameters are as follows: equation (8) for kernel function of LS-SVM, population size as 50, real number for coding, global mode for particle swarm algorithm, learning factor as $C1=C2=2$, inertia weight ω as 0.5, iterative time of algorithm termination as 100.

The calculation process of defining financial forecast of LS-SVM parameters by PSO algorithm is as follows.

- 1) To initialize algorithm
- 2) To train LS-SVM, and to determine if fitness function is the smallest. If yes, turn to step 4
- 3) To optimize LS-SVM via PSO, and to turn to Step 2
- 4) To specify LS-SVM parameters, and to train LS-SVM again
- 5) To forecast

MATLAB is applied to realize PSO and LS-SVM algorithms.

(1) Financial Data Training in year 2009

Regarding financial data in 2009, parameters of optimizing LS-SVM by PSO are used to conduct simulation trainings for 10 times, and equivalent parameters to the optimal results are $\gamma=68.4158$, $\sigma_2=2.4952$. Optimal and average results in financial forecast of year 2009 are shown in Table 1. Mesh searching algorithm and 10-fold Cross-Validation are used to get $\gamma=2.382$, $\sigma_2=2.2303$. Financial crisis results in 2010 are shown in Table I.

TABLE I
TRAINING AND TESTING RESULTS OF 2009 FINANCIAL DATA

Algorithm	Training results Error amount	Correct identification rate	Testing results Error amount	Correct identification rate
Algorithm of this paper (optimal)	0.0	100.00	1.0	98.12
Algorithm of this paper (average)	0.3	99.7	1.4	97.36
Mesh searching algorithm	1.0	98.68	2.0	96.23

CONTINUED TABLE

Algorithm	Total correct identification rate
Algorithm of this paper (optimal)	99.35
Algorithm of this paper (average)	98.90
Mesh searching algorithm	98.06

(2) Financial data training in year 2008

Regarding financial data in 2008, parameters of optimizing LS-SVM by PSO are used to conduct simulation trainings for 10 times, and equivalent parameters to the optimal results are $\gamma=96.5099$, $\sigma_2=1.7415$. Optimal and average results in 2010 financial forecast are shown in Table 2. 10-fold Cross-Validation method is used to get $\gamma=91.99932$, $\sigma=5.40038$.

TABLE II
TRAINING AND TESTING RESULTS OF 2008 FINANCIAL DATA

Algorithm	Trainin g results Error amount	Correct identification rate	Testing results Error amount	Correct identification rate
Algorithm of this paper (optimal)	0.0	100.00	5.0	90.57
Algorithm of this paper (average)	0.8	99.22	6.1	88.30
Mesh searching algorithm	4.0	96.04	7.0	86.79

CONTINUED TABLE

Algorithm	Total correct identification rate
Algorithm of this paper (optimal)	96.75
Algorithm of this paper (average)	95.45
Mesh searching algorithm	92.86

(3) Financial data training in year 2007

Regarding financial data in 2007, parameters of optimizing LS-SVM by PSO are used to conduct simulation trainings for 10 times, and equivalent parameters to the optimal results are $\gamma=80.4051$, $\sigma_2=0.0061$. Optimal and average results in 2010 financial forecast are shown in Table 3. 10-fold Cross-Validation method is used to get $\gamma=3.0524$, $2\sigma=0.9254$.

TABLE III
TRAINING AND TESTING RESULTS OF 2007 FINANCIAL DATA

Algorithm	Training results Error amount	Correct identification rate	Testing results Error amount	Correct identification rate
Algorithm of this paper (optimal)	0	100.00	11.0	79.25
Algorithm of this paper (average)	4	96.04	11.9	77.55
Mesh searching algorithm	8	92.08	13.0	75.47

CONTINUED TABLE

Algorithm	Total correct identification rate
Algorithm of this paper (optimal)	92.86
Algorithm of this paper (average)	89.68
Mesh searching algorithm	86.36

10-fold cross-validation method divides training samples into 10 parts, and each part is treated as validation sample for every calculation. Parameters of LS-SVM are obtained via mesh searching algorithm. This method takes account of the model's generalization, and is one of the traditional methods for LS-SVM parameter selection.

It can be seen from the training results above that the correct identification rate of using 2007 financial data to forecast 2010 financial situation is 92.86%, rate of using 2008 data is 96.75%, and rate of using 2009 data is 99.35%. The correct rate of forecast is respectively 79.25%, 90.57%, and 98.11%.

Comparing either total correct identification rate or forecast correct rate, it shows that the method applied in

this paper is better than mesh searching-based cross validation method. Financial forecast model built via optimization of LS-SVM parameters by PSO has higher accuracy rate to forecast financial situation of listed companies in short-term, medium-term, and long-term manner.

V. CONCLUSION

This paper conducts forecast analysis on financial crisis of listed companies by using PSO to specify the optimal parameters of LS-SVM, and gains satisfying results. It enhances searching ability and accuracy of the model, breaks through the restriction of manual selection of optimal parameters by SVM, and improves parameter quality, thus increasing accuracy of SVM classification, effectively avoiding over fitting and under fitting, and having strong generalization ability.

Differing from neural network forecast, SVM does not need to specify hidden node amount, which has always been a tricky problem for neural network. After LS-SVM specifies input and output data, and adjustment parameters via PSO, its weight and threshold are obtained by solving linear equality, and the solution is unique. LS-SVM applies structure risk minimization principle and takes account of sample error and complexity, while most of neural network merely takes minimization of sample error into consideration. Therefore, if relevant adjustment parameters are selected properly, LS-SVM has great generalization ability and high value of application.

ACKNOWLEDGMENT

This research was supported by "the Fundamental Research Funds for the Central Universities". Fund No: 12MS139.

REFERENCES

- [1] Gao Yanqing, Luan Fugui. Enterprise financial crisis forecast model based on fuzzy comprehensive evaluation [J]. Inquiry on Economic Issues, Jan, 2005. PP: 56-57.
- [2] Zhang Genming, Xiang Xiaoji, Sun Jingyi. Financial crisis forecast of listed manufactory companies based on BP neural network[J]. Journal of Shandong Institute of Business and Technology, Vol.20, No.4. 2006. PP: 46-48.
- [3] Ding Lin. Research on BP neural network financial forecast model of listed companies based on corporate governance[J]. Contemporary Economics, vol. 22, No. 4, 2008.PP: 58-59.
- [4] Chu Xiao. Research on financial crisis forecast of small and medium enterprise based on logistic regression model [D]. Thesis of China University of Geosciences, 2010.
- [5] Hu Yanjie, Xia Guoping. Empirical research on financial crisis forecast based on BP neural network[J]. Journal of Beijing University of Aeronautics and Astronautics, vol. 22, No. 4, 2009. PP:25-27.
- [6] Zhang Xiaoqi. Study on application of SVM in financial crisis forecast model of high-tech companies[J]. Science and Technology Management Research, No. 6, 2010, PP:147-149.
- [7] Liu Li, Luo Hui. Research on financial crisis forecast analysis of listed companies based on data mining[J].

- [8] Application of Statistics and Management, vol.23, No.3, 2004. PP: 56-58.
- [9] Charles E. Mossman, Geoffrey G. Bell, L. Miek Swartz, Harry Turtle. An Empirical Comparison of Bankruptcy Models [J]. The Financial Review, vol.33, No.2, Mas, 1998, PP.35-54.
- [10] Arindam Chaudhuri, Kajal De. Fuzzy Support Vector Machine for Bankruptcy Prediction [J]. Applied SoftCozn Puting, Oct, 2010. PP.1-32.
- [11] Ming Chang Lee, Chang To. Comparison of Support Vector Machine and Bank propagation Neural Network in Evaluating the Enterprise Financial Distress [J]. International Journal of Artificial Intelligence & Applications (IJAIA), vol.1, No.3, July, 2010.PP.31-43.
- [12] Luo Yong, Chang Hua, Wang Fangjun. Analysis on reasons that financial crisis of listed companies in China occurs [J]. Industrial Economics Forum. Vol.6, No. 11, 2007.PP.21-23.
- [13] Wang Qiang, Lan Hui. Analysis on reasons that financial crisis of listed companies in China occurs [J]. Friends of Accounting . No.12, 2006.PP.45-46.
- [14] Guo Ruiying. Research on financial crisis forecast of listed companies based on data mining [D]. North China Electric Power University graduate thesis, Mar 2009.
- [15] Zhang Jiafei. Research on financial crisis of listed companies [D]. Southwest University of Finance and Economics graduate thesis, 2004.
- [16] Li Ling, Tan Tao. Analysis on theory of building enterprise financial crisis forecast models[J]. Science-technology and Management. Vol. 10, No. 1, Jan, 2008.
- [17] Shang Huijie. Empirical analysis on influencing factor model and forecast model of listed company financial crisis in China [D]. Jinan University graduate thesis. 2007.
- [18] Guojun Ding, Lide Wang, Peng Yang, Ping Shen, Shuping Dang.Diagnosis Model Based on Least Squares Support Vector Machine Optimized by Multi-swarm Cooperative Chaos Particle Swarm Optimization and Its Application [J].Journal of Computers. Vol, 8, No 4 , 2013.pp:975-982.
- [19] YingGuang Yang, Shanxiao Yang .Study of Emotion Recognition Based on Surface Electromyography and Improved Least Squares Support Vector Machine [J]. Journal of Computers. Vol, 8, No 4 , 2013.pp:1707-1704.
- [20] Sijin Zhu, Guixian Li, Junwei Han.An Improved PSO Algorithm with Object-Oriented Performance Database for Flight Trajectory Optimization [J]. Journal of Computers. Vol 7, No 7 , 2012.pp:1555-1563.
- [21] Tay F E H, Cao L. Application of support vector machines in financial time series forecasting [J].Ome-ga, 2001, pp:309-317.
- [22] Vapnik V N. The nature of statistical learning theory [M]. NY: Springer-Verlag, 1995.

Xinli Wang, born in Baoding City, China, graduated from Agricultural University of Hebei in 2005, and gained the master's degree of management. The author's major field of study is the information management. Since 2005, she has been working at the North China Electric Power University, Baoding City, China, and she has published more than 10 papers.