Financial distress prediction model of dual constraint LS-SVM based on neighborhood rough set index optimization

Guanhua Zhao[1], Tanglu[2], Lin Gao[3]*

[1]*Accounting Institute of Shandong Finance University, Shandong Accounting Science Research Center, Ji'nan, China*

[2]*Accounting Institute of Shandong Finance University, Ji'nan, China*

[3]*Qingdao Vocational and Technical College of Hotel Management, Qingdao, China*

Received 1 October 2014, www.cmnt.lv

Abstract

In order to improve the accuracy of financial distress prediction and the effect of model forecast, and make the neighborhood rough set and genetic algorithm apply to the dual constrained least squares support vector machine. This study proposes a dual constrained least squares support vector machine prediction model based on the neighborhood rough set attribute reduction. At the same time, this study gives the steps to improve this model. The empirical results show that, after the pre-treatment of neighborhood rough set index and optimization of parameters of genetic algorithm. It not only improves the model prediction accuracy, but also reduces the model run time, therefore it confirmed that the application of the improved model to the financial distress prediction is effective.

Keywords: financial distress prediction, least squares support vector machine, dual constraint, the neighborhood rough set, genetic algorithm

1 Introduction

Traditional statistical prediction model has some limits in practical application. The method itself also has some limitations [1-3]. In order to solve these problems, some artificial intelligent prediction models are widely used in financial distress prediction and thy have achieved good results.

In the early 90's, some researchers tried to apply the artificial neural networks (Artificial Neural Networks, ANN) model to the study of financial distress prediction, and compared it with the traditional statistical model [4-6]. Compared with the traditional statistical model, ANN is superiority. It can also deal with the qualitative variables and quantitative variables, and it need not to consider the statistical relationship between variables. Its disadvantage is the definition of topology model is hard to achieve, the amount of calculation is very large, and discriminant ability is not strong. Fan A etc. [7] established the earliest support vector machine (Support Vector Machine, SVM) prediction model of financial distress. Some researchers used various algorithms to optimize the model itself and the kernel function, and established the improved SVM model. The prediction is not only superior to the effect of statistical model, but also better than that of ANN model [8-10]. Based on empirical risk minimization, SVM adopts the principle of structural risk minimization. It can improve the generalization ability of the model [11] to a large extent. But the traditional SVM model also has problems that misclassification leads to increased risk of experience, especially when the sample point and the optimal hyper plane are closed to the experience risk of misclassification significantly increased [12]. In addition, the traditional SVM requires solving a quadratic programming problem which is very difficult and occupy more running time and space.

Suykens [13] proposed the least squares support vector machine (Least Squares Support Vector Machine, LS-SVM), its core idea is using straits to replace the inequality constraints in SVM. Thus, quadratic programming problem is transformed into a system of linear equations to solve, greatly reduces the difficulty of solving, and the convergence speed is faster than that of SVM model. On the basis of this, the researchers at home and abroad put forward some improved LS-SVM models [14-16], which can improve the prediction accuracy and stability.

This paper made the following points on the basis of the above research: 1) tried using neighborhood rough set attribute reduction algorithm to optimize financial variables, and compared it with the classical rough set and factor analysis method. 2) in the constrained condition of the traditional LS-SVM model, tried to improve the prediction results of model by increasing the dual constraints and improving kernel function. 3) in the selection of the parameters of LS-SVM model, genetic algorithm is used to automatic optimization of the parameters of the model.

2 Dual constraint type LS-SVM principle

General description of the LS-SVM is as follows: given L of training samples cc, the input data of i is $x_i \in R^n$, the output data of i, $y_i \in \{+1,-1\}$, is a dichotomic variable. A classification function is constructed by LS-SVM:

$$f(x) = \operatorname{sgn}\left(w^T \phi(x) + b\right). \tag{1}$$

***Corresponding author's** e-mail: jiaoshigaolin@sina.cn

So the sample of x can be correctly classified by $f(x)$, that is to solve the following optimization problem:

$$\min \cdot \frac{1}{2} w^T w + \frac{1}{2} C \sum_{i=1}^{l} \xi_i^2 , \tag{2}$$

$$s.t. \; y_i (w^T \phi(x_i) + b) = 1 - \xi_i , \; i = 1, 2, ..., l , \tag{3}$$

among them, $w = (w_1, w_2, ..., w_l)^T$ is a weight vector, which is perpendicular to the hyperplane, *C* is the penalty factor, ξ_i is a slack variable greater than zero, $\phi(x_i)$ called the mapping function, *b* is a constant error.

Now, on the basis of the Equation (3), adding a constraint condition:

$$w^T \phi(x_i) = a w^T \phi(-x_i) , \; i = 1, 2, ..., l , \tag{4}$$

among them, when the nonlinear function is an odd function, $a = 1$, when the nonlinear function is a even function, $a = -1$.(2)~(4) dual problem is Lagrange polynomials:

$$L(w, b, \xi_i, \alpha_i, \beta_i) = \frac{1}{2} w^T w + \frac{1}{2} C \sum_{i=1}^{l} \xi_i^2 - \sum_{i=1}^{l} \alpha_i \left(w^T \phi(x_i) + b + \xi_i - y_i \right) - \sum_{i=1}^{l} \beta_i \left(w^T \phi(x_i) - a w^T \phi(-x_i) \right) , \tag{5}$$

among them, α_i is the Lagrange multiplier, we respectively get the partial derivative of w, b, ξ_i, α_i, β_i from Equation (5) and let them equal to zero. According to the Karush-Kuhn-Tucker (KKT) complementary conditions, and take $y_i \in \{+1, -1\}$ into account at the same time, there is:

$$\begin{cases} w = \sum_{i=1}^{l} (\alpha_i + \beta_i) \phi(x_i) - a \sum_{i=1}^{l} \beta_i \phi(-x_i) \\ \sum_{i=1}^{l} \alpha_i = 0 \\ C \xi_i = \alpha_i \\ y_i = w^T \phi(x_i) + b + \xi_i \\ w^T \phi(x_i) = a w^T \phi(-x_i) \end{cases} . \tag{6}$$

After the elimination of ξ_i and w, we can obtain linear Equations as follows:

$$\begin{vmatrix} 0 & \underline{1}^T \\ \underline{1} & (\Omega + a\Omega^*)/2 + I/C \end{vmatrix} \begin{vmatrix} b \\ \alpha \end{vmatrix} = \begin{vmatrix} 0 \\ Y \end{vmatrix} . \tag{7}$$

Among them, $\underline{1} = (1, 1, \cdots, 1)^T$, $\Omega_{ij} = y_i y_j K(x_i, x_j)$, $\Omega^*_{i,j} = y_i y_j K(-x_i, x_j)$, $Y = (y_1, y_2, ..., y_l)^T$, $\alpha = (\alpha_1, \alpha_2, ..., \alpha_l)^T$, *I* is the unit matrix, $\forall i, j = 1.2, ..., l$.

To solve the linear Equations, and replace the inner product operation to the symmetric kernel function $K(x_i, x_j) = \phi(x_i)\phi(x_j)$ which satisfying the Mercer conditions, so as to obtain the prediction model of dual constrained LS-SVM:

$$y_i = \frac{1}{2} \sum_{i=1}^{l} \left(w^T \varphi(x_i) + a w^T \varphi(-x_i) \right) + b + \frac{1}{C} \alpha_i = \frac{1}{2} \sum_{i=1}^{l} \alpha_i \left(K(x_i, x_j) + aK(-x_i, x_j) \right) + b + \frac{1}{C} \alpha_i K_{equ}(x_i, x) = \frac{1}{2} \left(K(x_i, x) + aK(-x_i, x) \right) . \tag{8}$$

Let $K_{equ}(x_i, x) = \frac{1}{2} \left(K(x_i, x) + aK(-x_i, x) \right)$, and introduce into the symbol function, Equation (8) is equivalent to:

$$f(x) = \mathrm{sgn} \left(\sum_{i=1}^{l} \alpha_i y_i K_{equ}(x_i, x) + b \right) , \tag{9}$$

where b, α in Equation (9) can be obtained from the Equation (7), y_i is determined according to the training sample properties. When the sample is ST, $y_i = -1$, otherwise, $y_i = +1$. $K_{equ}(x_i, x)$ is the kernel function, and this paper used the Gauss kernel. Its expression is $K_{equ}(x_i, x) = \exp\left(-|x_i - x|^2 / 2\sigma^2 \right)$.

σ^2 is the parameter of the Gauss kernel. σ^2 and the penalty factor *C* have a direct influence on the prediction results of model. The paper used genetic algorithm to search the optimal σ^2 and the *C* value automatically.

3 Neighborhood rough set and attribute reduction algorithm

3.1 NEIGHBORHOOD ROUGH SET

Rough set was proposed by Professor Pawlak in 1982. It used the rough approximation method to screen tap the information which is potential, valuable and indispensable from the massive rambling and strong interference data [13]. Therefore, rough set is widely used in attribute reduction, decision rule extraction, classical prediction and other fields. But the classical rough set is only suitable for the treatment of nominal variables. For the numerical type data, we need to use various methods to make them discretization, while discrete process inevitably brings the loss of effective information [17].

T.Y.Lin [18] first proposed the concept of neighborhood model. This model used the neighborhood space points to granulate the domain universe. It took the neighborhood as the basic information particle, and used it to describe the other concept of space. Hu Qinghua [19] improved the classical rough set theory based on neighborhood model, and put forward the concept of neighborhood rough set model. The neighborhood rough set model forms a δ neighbourhood, which is based on each point in real space. δ neighborhood group constitutes the basic information particles which can descript any concept of space, its basic description is as follows:

For information system $IS = < U, A, V, f >$, among them $U = \{x_1, x_2, ..., x_n\}$ represents the non-empty finite set called the domain. For the financial distress prediction, the domain consists of sample space set. A is the set of attributes, it refers to the predictor set. V is range. $\mathrm{f{:}U \times \vec{A}V}$ is an information function, which presents the corresponding mapping relationship between sample and the attribute value. If $\mathrm{A} = C \cup D$, C is a conditional attribute set, D is the decision attribute set, and that requires $\mathrm{C \cap D = \varnothing}$,then $IS = < U, A, V, f >$ is a decision table. Given $\mathrm{x}_i \subseteq \mathrm{U}$, $\mathrm{B \subseteq C}$, then the neighborhood of x_i in the attributes B can be defined as:

$$\delta_B(x_i) = \{x_i \,|\, x_i \in U, \Delta_B(x_1, x_2) \le \delta\}, \tag{10}$$

Among them, Δ is The distance function,
for $\forall x_1, x_2, x_3 \in U$, Δ satisfies the following relations:

$$\begin{cases} \Delta(x_1,x_2) \ge 0, \Delta(x_1,x_2) = 0, only x_1 = x_2 \\ \Delta(x_1,x_2) = \Delta(x_2,x_1) \\ \Delta(x_1,x_3) \le \Delta(x_1,x_2) + \Delta(x_2,x_3) \end{cases}. \tag{11}$$

For N attributes sample sets, the distance can be expressed by P norm:

$$\Delta_{\mathrm{p}}(x_1,x_2) = \left(\sum_{i=1}^{N} \left| f(x_1,a_i) - f(x_2,a_i) \right|^P \right)^{1/P}, \tag{12}$$

among them, $f(x_i, a_i)$ is the value of sample x_i on the attribute a_i. $\Delta_{\mathrm{p}}(x_1, x_2)$ is for the numerical attribute set, but the neighborhood model is easy to make the distance calculation extend to the data with symbols and numeric, the symbolic attributes a_i, can be defined as follows:

$$\begin{cases} f(x_1,a_i) - f(x_2,a_i) = 0, when\ x_1 = x_2 on a_i \\ f(x_1,a_i) - f(x_2,a_i) = 1, when\ x_1 \ne x_2 on a_i \end{cases}. \tag{13}$$

Thus, the lower and upper approximation of neighborhood rough set is defined as:

$$\begin{cases} \underline{\mathrm{N}}\mathrm{X} = \left\{ x_i \,|\, \delta(x_i) \subseteq X, x_i \in U \right\} \\ \overline{\mathrm{N}}\mathrm{X} = \{ x_i \,|\, \delta(x_i) \cap X \ne \varnothing, x_i \in U \} \end{cases}. \tag{14}$$

Then the approximate boundary corresponding for X is $\mathrm{BN(X)} = \overline{\mathrm{N}}\mathrm{X}\text{-}\underline{\mathrm{N}}\mathrm{X}$.

For a neighborhood decision system
$\mathrm{NDT} = < \mathrm{U,C \cup D,V,}f >$,

D divided U into N equivalence class: $X_1, X_2, \ldots, X_N$, $\forall B \subseteq C$,

the lower approximation, upper approximation and the decision boundary of decision D on B are defined respectively:

$$\underline{\mathrm{N_B}}\mathrm{D} = \bigcup_{i=1}^{N} \underline{\mathrm{N_B}}\mathrm{X}_i, \tag{15}$$

$$\overline{\mathrm{N_B}}D = \bigcup_{i=1}^{N} \overline{\mathrm{N_B}}\mathrm{X}_i, \tag{16}$$

$$\mathrm{BN(D)} = \overline{\mathrm{N_B}}D\text{-}\underline{\mathrm{N_B}}\mathrm{D}. \tag{17}$$

Lower approximation of decision D called the decision positive region, denoted as $\mathrm{POS_B(D)}$. The size of the positive region reflect the separable degree of classification problem in a given attribute space. The greater positive region presents the less overlap region or boundary. Therefore, defining the dependence of decision attribute D on condition attributes B is:

$$\gamma_{\mathrm{B}}(\mathrm{D}) = \frac{|\mathrm{POS_B(D)}|}{|\mathrm{U}|}. \tag{18}$$

Equation (18) presents a ratio that a sample which can be completely contained by a kind of decision in all samples, according to condition attribute description of B in sample set.

3.2 NEIGHBORHOOD ROUGH SET ATTRIBUTE REDUCTION ALGORITHM

The index which has influence on financial distress may up to several dozen. These data often has a great deal of information overlap. They not only affect the SVM generalization ability, but also make the structure of SVM more complex. Therefore, with the help of the neighborhood rough sets theory. We can make the numerous alternative index (the condition attribute) attribute reduct, that is eliminating the redundant attributes, reducing the cost of financial distress prediction while reducing noise and improving the prediction accuracy on the premise of keeping the ability of classification. On the basis of paper [19-23] combined with the actual of financial distress prediction, this paper presents a rapid algorithm of attribute reduction, measures the dependence of the decision attribute with the conditional attribute, and weights each of the attributes. The algorithm is programmed in MATLAB R2007, the specific steps are as follows:

Input: decision table $<$ U,C,D,V,f $>$

Output: reduction red.

Step 1: for each value in the condition attribute set of *C*, using the Equation of $x=(x-x_{\min})/(x_{\max}-x_{\min})$ to make all numerical attributes normalized to [0,1].

Step 2: let $red=\varnothing$, namely the initialization for testing sample set. $sam\_chk=U$

Step 3: flag=1.

Step 4: while $sam\_chk\neq\varnothing$.

Step 5: for each $k_i\in(C-red)$.

Step 6: $DT_i=<U,red\bigcup k_i,D,V,f>$.

Step 7: initialization $POS_i=\varnothing$.

Step 8: for each $a_j\in smp\_chk$.

Step 9: calculation the neighborhood $\delta(a_j)$ of a_j in DT_i .

Step 10: if each of the attribute value *D* of $\delta(a_j)$ has the same value.

Step 11: $POS_i=POS_i\bigcup a_j$.

Step 12: end if

Step 13: end for

Step 14: if flag=1

Step 15: $\gamma_i=\dfrac{POS_i}{smp\_chk}$

Step 16: end for

Step 17: flag=0

Step 18: find out the maximum of POS_i and the corresponding k_i .

Step 19: if $POS_i\neq\varnothing$.

Step 20: $red=red\bigcup k_i$.

Step 21: $smp\_chk=smp\_chk-POS_i$

Step 22: else

Step 23: exit the while loop.

Step 24: end if

Step 25: end while

Step 26: return red

Step 27: after the reduction of the training sample set, we attribute weighted the training sample set by multiplying the corresponding attribute importance degree γ_i .

4 Empirical study

4.1 THE SOURCE AND SELECTION OF SAMPLE

The data was from the CSMAR database which was developed by Shenzhen Tai'an Information Technology Co., Ltd. 384 A-share listed companies in 2004~2009 were randomly selected as the research sample from the Shanghai and Shenzhen stock markets, excluding financial and incomplete data samples, the remaining 352 A-share listed companies as the final research sample. Among them, 120 were the test samples, the remaining 232 were used as training samples. Pair the training and test samples respectively in accordance with the principle of 1:1. That is the training samples have 116 non ST companies and 116 ST companies. The test samples have 60 non ST companies and 60 ST companies.

4.2 THE SELECTION OF INITIAL SET OF VARIABLES

The factors causing financial distress are not only the financial aspects, but also the non-financial aspects, therefore, based on domestic and foreign scholars' choice of the index. This paper makes a new attempt, not only chooses the financial index, but also introduces the non-financial indicators, which can be shown in Table 1.

TABLE 1 The financial indicators and non-financial indicators

Index properties	Index name	Index properties	Index name
The short-term debt paying	The current ratio (X_1)	Growth ability	The growth rate of fixed assets (X_{17})
	The quick ratio (X_2)		The growth rate of total assets (X_{18})
	Operating funds to total assets ratio (X_3)		The growth rate of net profit (X_{19})
The long-term debt paying	The asset-liability ratio (X_4)	Cash flow	The main business income of cash ratio (X_{20})
	Long term liabilities ratio (X_5)		Per - share operating net cash flow (X_{21})
	Debt to tangible assets ratio (X_6)		Per - share net cash flow (X_{22})
Operation ability	Accounts receivable turnover ratio (X_7)		Sales cash ratio (X_{23})
	Inventory turnover ratio (X_8)	Equity structure	The proportion of executive stockholding (X_{24})
	Fixed assets turnover ratio (X_9)		The proportion of state-owned shares (X_{25})
	Total assets turnover ratio (X_{10})	Corporate governance	The part-time situation of chairman and general manager (X_{26})
Profitability	The main business profit ratio (X_{11})		The number of directors (X_{27})
	Ratio of return on assets (X_{12})		The total size of the board of supervisors (X_{28})
	Total assets profit ratio(X_{13})		The number of executives(X_{29})
	Fixed assets profit ratio(X_{14})	The scale of the corporate	Log of assets(X_{30})
	Ratio of return on net assets(X_{15})		
	The growth ratio of main business(X_{16})		

4.3 THE PRETREATMENT OF INDEX DATA

This paper uses factor analysis method and the neighborhood rough set method for index pre-conditioning. So we can compare the forecast effect of financial distress between the two methods.

1) Factor analysis. First of all, using the Wilcoxon signed rank method to test the significant difference of 31 indexes of the sample. Then put the variables which have significant difference into factor analysis, in order to eliminate the multicollinearity between variables. After the factor analysis, determine 8 factors ultimately, which include 16 primitive indexes. 8 factors are the input variables of these models. The 8 factors and the indexes can be seen in Table 2.

TABLE 2 Public factors and their financial indicators after the factor analysis

Public factors	Indicators	Public factors	Indicators
F_1	The main business profit ratio(X_{11})、Ratio of return on assets（X_{12}）、Ratio of return on net assets(X_{13})、Fixed assets profit ratio(X_{14})、Ratio of return on net assets(X_{15})	F_5	Sales cash ratio(X_{23})
F_2	The current ratio(X_1)、Operating funds to total assets ratio (X_3)、The asset - liability ratio(X_4)	F_6	The proportion of state-owned shares(X_{25})
F_3	Total assets turnover ratio（X_{10}）、The growth ratio of main business(X_{16})	F_7	Accounts receivable turnover ratio(X_7)
F_4	The main business income of cash ratio(X_{20})、Per - share operating net cash flow(X_{21})	F_8	The growth rate of net profit(X_{19})

2) The neighborhood rough set reduction method. Based on the standard of 31 alternative index data, using the neighborhood rough set reduction algorithm for index selection. In order to compare with the classical rough set approach, CART classification learning algorithm is introduced in the experiment. The specific steps are as follows: for the attribute reduction of classical rough set, first, use Equal frequency binning algorithm in Rosetta software to discretize the data, then use the Johnson reducer algorithm to reduct the discreted data. For the neighborhood rough set attribute reduction, use the proposed algorithm which is programmed in MATLAB R2007. As the number of attribute in this algorithm is affected by the size of the neighbourhood δ. This paper sets the step length of δ to 0.1 and δ ranges from 0.1 to 1. The results show that, when δ=0.9, classification accuracy is the best (to achieve the highest classification accuracy with the least number of attributes). Table 3 gives the results of attribute reduction by two methods.

TABLE 3 Index reduction methods comparison between classical rough sets and neighborhood rough sets

The index preconditioning method	Index after the reduction
Classical rough set attribute reduction	The current ratio (X_1). The quick ratio (X_2). The asset-liability ratio (X_4). Accounts receivable turnover ratio (X_7). Inventory turnover ratio (X_8). The main business profit ratio (X_{11}). The growth ratio of main business (X_{16}). The growth rate of total assets (X_{18}). The growth rate of net profit (X_{19}). The main business income of cash ratio (X_{20}). Per-share operating net cash flow (X_{21}). The number of executives (X_{29}).
The neighborhood rough set attribute reduction	Accounts receivable turnover ratio (X_7). The growth ratio of main business (X_{16}). The growth rate of net profit (X_{19}). Per - share operating net cash flow (X_{21}). Per-share net cash flow (X_{22}).

Comprehensive Table 2 and Table 3, we can see the factor analysis method extracted 8 public factors. These 8 factors include 16 indexes. Classical rough set reduction method extracted 12 indicators, and the neighborhood rough set reduction method extracted 5 indexes.

4.4 OPTIMIZATION OF THE MODEL PARAMETERS

The value of the penalty factor C in dual constraint LS-SVM model and Gauss kernel parameter σ^2 will directly affect the prediction effect of the model. In order to improve the model prediction accuracy, this paper uses genetic algorithm (Genetic Algorithm, GA) to optimize the parameters automatically. The core problem of parameter optimization with GA is the individual encoding, selection, crossover and mutation for specific problems.

This paper adopts binary coding, for the binary code is easy to realize selection, crossover and mutation operation. The genetic algorithm evaluates the quality of the individual by fitness. Here we use the model prediction accuracy as the fitness function, and use "roulette" method for operator selection with the optimal preservation strategy (using the parent optimal individuals to replace the offspring worst individuals).

Two point crossovers are used in the crossover operation of the crossover operator. Two intersection points K1, K2 $\in [1,l]$ are set randomly in two individual code string of σ^2,and then exchange part of genes between the intersection points of two chromosomes according to the crossover probability P_c . The mutation operator replace one or several loci original gene values which was random designated by the random number (0,1) in individual code string according to mutation probability P_m.Genetic algorithm is used for the optimization of process parameters, the process is realized by MATLAB R2007 programming with paper [17].

4.5 ANALYSIS OF THE EMPIRICAL RESULTS.

The forecasting accuracy of each model under differrent index pre-treatment methods is shown in Table 4. We can see, the prediction accuracy of LS-SVM model and dual constraint LS-SVM model is not only better than factor analysis method but also than the classical rough set method after using neighborhood rough set attribute reduction method to treat primary indicators. In addition, use genetic algorithm to optimize the parameters can significantly improve the prediction accuracy of the model.

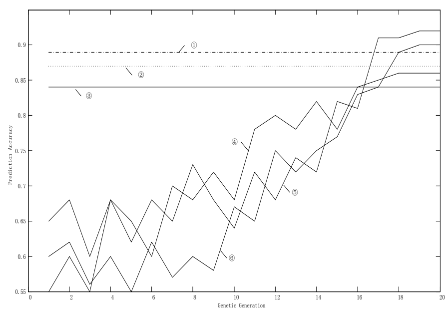
TABLE 4 The comparison of forecasting accuracy of each model under different index pretreatment methods

Model name	The index preconditioning method	The penalty factor C	The parameter of the Gauss kernel σ^2	Prediction accuracy
LS-SVM	Factor analysis	5	1	84.0%
	Classical rough set attribute reduction	5	1	86.5%
	The neighborhood rough set attribute reduction	5	1	88.6%
LS-SVM based on GA	Factor analysis	4.7613	1.0362	86.0%
	Classical rough set attribute reduction	4.7613	1.0362	90.4%
	The neighborhood rough set attribute reduction	4.7613	1.0362	92.2%
Dual constraint LS-SVM	Factor analysis	5	1	86.0%
	Classical rough set attribute reduction	5	1	88.2%
	The neighborhood rough set attribute reduction	5	1	89.8%
Dual constraint LS-SVM based on GA	Factor analysis	4.7613	1.0362	88.0%
	Classical rough set attribute reduction	4.7613	1.0362	90.6%
	The neighborhood rough set attribute reduction	4.7613	1.0362	94.0%

Figure 1 shows comparison of LS-SVM model and LS-SVM model prediction accuracy under different index pre-treatment methods.

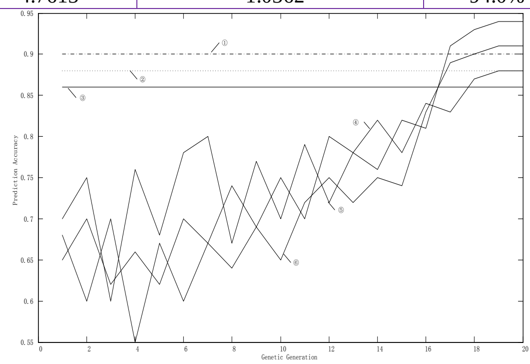
We can see from Table 4 and Figure 1, through 20 genetic manipulation, the prediction accuracy of LS-SVM model based on GA converges to a stable value. But the model under the neighborhood rough set index pre-treatment is the best one, which reached 92.2%. The following are the classical rough set and factor analysis index pre-treatment, respectively reached 90.4% and 86%; prediction accuracy of LS-SVM model without three kinds of index pre-treatment of GA parameter optimization were respectively reached 88.6%, 86.5%, 84%, though the neighborhood rough sets index preconditioning is also the best. They are lower than that predicted by model with GA parameter optimization.

Figure 2 shows comparison of prediction accuracy between dual constraint LS-SVM model and dual constraint LS-SVM model based on GA under different index pre-treatment methods. The conclusion is similar to Figure 1. But the overall prediction effect of dual constraint LS-SVM model is better than that of LS-SVM model.



① The LS-SVM model based on neighborhood rough set
② The LS-SVM model based on classical rough set
③ The LS-SVM model based on factor analysis
④ The LS-SVM model based on factor analysis and GA
⑤ The LS-SVM model based on neighborhood rough set and GA
⑥ The LS-SVM model based on classical rough set and GA

FIGURE 1 Comparison of LS-SVM model and LS-SVM model based on GA prediction accuracy under different index pretreatment methods



① The dual constraint LS-SVM model based on neighborhood rough set
② The dual constraint LS-SVM model based on classical rough set
③ The dual constraint LS-SVM model based on factor analysis
④ The dual constraint LS-SVM model based on factor analysis and GA
⑤ The dual constraint LS-SVM model based on neighborhood rough set and GA
⑥ The dual constraint LS-SVM model based on classical rough set and GA

FIGURE 2 Comparison of prediction accuracy between dual constraint LS-SVM model and dual constraint LS-SVM model based on GA under different index pretreatment methods

5 Conclusions

In order to improve the support vector machine financial distress prediction accuracy, and reduce the misjudgement rate, this paper has drawn the following conclusions from three perspectives through empirical research. 1) for the pre-treatment method of alternative indicators, use the index pre-treatment method based on neighborhood rough sets, and compare with the classical rough set and factor analysis method. The empirical results show that by using neighborhood rough set reduction method can significantly improve the prediction accuracy of the models with lose index number but large amount of effective information. 2) the overall prediction effect of dual constraint LS-SVM model is better than that of LS-SVM model. This confirms that the method, which improve the kernel function by adding the dual constraints in the LS-SVM model can improve the prediction accuracy of LS-

SVM model. 3) to optimize the kernel parameter of LS-SVM model with GA through 20 genetic manipulation, the prediction accuracy of model converges to a stable value. GA can significantly improve the model prediction accuracy. To sum up, this paper uses the neighborhood rough set attribute reduction method for index pre-treatment, uses GA for the optimization of dual constraint LS-SVM parameter, financial distress prediction effect is improved significantly, the results show that the improved model proposed in this paper is effective.

References

[1] Beaver W 1966 Financial ratios as predictors of failure, empirical research in accounting: Selected studies *Journal of accounting research* **19**(5) 71-111

[2] Altman E 1968 Financial ratios,discriminant analysis and the prediction of corporate bankruptcy *The journal of finance* **23**(4) 589-609

[3] Ohlson J 1980 Financial ratios and the probabilistic prediction of bankruptcy *Journal of accounting research* **18**(1) 109-31

[4] Odom M,Sharda R A 1990 *Proceedings of the IEEE 1990 international conference on neural network* 163-8

[5] Yang S, Li H 2005 listed Corporation financial distress Prediction modelbased on BP neural network *Systems engineering theory and practice* (1)12-8

[6] Li B 2004 Nonlinear combination of enterprise financial crisis prediction method and the empirical *Systems engineering theory and practice* (6) 222-5

[7] Fan A, Palaniswami M 2000 *Proceeding of the international joint conference on neural network* 354–9

[8] Shi H 2012 Applying Principal Component Analysis, Genetic Algorithm and Support Vector Machine for Risk Forecasting of General Contracting *Journal of Computers* **7**(1) 301-7

[9] Kan B, Liu Z 2012 Applications of improved support vector machine in the false financial report recognition *Management review* (5) 35-9

[10] Ding D 2011 An Empirical Study on the financial risk early warning model of hybrid HOGA-SVM *Management journal* (4) 46-50

[11] Tan J, Chen X, Du Min 2012 An Internet Traffic Identification Approach Based on GA and PSO-SVM *Journal of Computers* **7**(1) 19-29

[12] Ding S, Wu F, Ru N, Yu J 2013 Twin Support Vector Machines Based on Quantum Particle Swarm Optimization *Journal of Software* **8**(7) 1743-50

[13] Suykens J A K, Vandewalle J.1999 Least squares support vector machine classifiers *Neural processing letter* (9) 293-300

[14] Zhao G 2010 Reduced memory type LS-SVM study on the financial distress prediction model based on genetic algorithm and entropy *Operations research and management science* (5) 71-7

[15] Lin H 2006 The study of Least squares support vector machine method on listing Corporation financial crisis forecasting *Market modernization* (10) 346-8

[16] Pawlak Z 1982 Rough sets *International journal of information and computer sciences* (11) 341-56

[17] Hu Q, Yu D 2008 Neighborhood roush set based heteroge-neous feature subset selection *Information sciences* (8) 357-59

[18] Lin T Y 1998 Granular computing on binary relations:data mining and neighborhood system *Processing of rough sets in knowledge discovery* (10) 107-21

[19] Hu Q, Zhao H, Yu D 2007 Seventh China rough set and Soft Computing Conference *Shanxi, Taiyuan*

[20] Du T 2012 Evaluation model of personal credit support based on rough set vector machine Statistics and decision (1) 26-30

[21] Zhou X,Wu K 2012 Study of listing Corporation early warning of financial crisis based on rough sets and ANN *Enterprise economy* (2) 56-9

[22] Bao X, Yang Y 2013 Enterprise financial crisis early warning system based onclustering-rough set neural network *Journal of systems management* (5) 34-9

[23] Zhang Y, Yang X 2012 Intuitionistic Fuzzy Dominance–based Rough Set Approach: Model and Attribute Reductions *Journal of Software* **7**(3) 551-63

Acknowledgements

Thanks to Tang Lu, Zhao Juan, Hao Min, Wang Fang's study of part the research work. This research is supported by the National Natural Science Fund Project (70840018), Shandong Provincial Natural Science Foundation Project (ZR2012GM009) and the Shandong Provincial Science and Technology Research Project (2008GG30009005).

Authors

Zhao Guanhua, December 1962, Jiangsu Xuzhou, China.

Current position, grades: working in Shandong University of Finance and Economics, PhD in Management.
University studies: MsC in Computing Science in the University of Geosciences (Beijing).
Scientific interest: early warning of financial risks, financial data mining technology, machine learning and optimization theory
Publications 30 papers.
Experience: 1 project, 1 teaching material.

Tanglu, April 23, 1990, Shandong Zibo, China.

Current position, grades: a graduate student in Shandong University of Finance and Economics, major in accounting information system.
University studies: MsC in Shandong University of Finance and Economics.
Scientific interest: early warning of financial risk, the study of accounting theory.
Publications: 2 papers.

Gao Lin, October 1988, Jinan, Shandong, China.

Current position, grades Qingdao Vocational and Technical College of Hotel Management, Master of management.
University studies: MsC in Shandong University of Finance and Economics and accounting professional learning.
Scientific interest: financial risk early warning, financial management.