Mathematical Biosciences and Engineering

Research article

An improved least squares SVM with adaptive PSO for the prediction of coal spontaneous combustion

Qian Zhang and Haigang Li *

School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, Jiangsu, 221116, China

* Correspondence: Email: haigangli@cumt.edu.cn; Tel: +1-395-219-4041.

Abstract: The problem of coal spontaneous combustion prediction is very complex, and there are many factors that affect the prediction results. In order to solve the issues of high dimension and redundancy among features and limited samples in the prediction of coal spontaneous combustion, this paper proposes a prediction algorithm of coal spontaneous combustion based on least squares support vector machine and adaptive particle swarm optimization (APSO-LSSVM). The adaptive PSO algorithm is used to solve the problems such as high computational complexity and slow calculation speed of the LS-SVM model for large-scale samples, so that it can always obtain the optimal solution, and its training speed and accuracy are improved. This method adjusts the inertia weight based on the convergence degree of group and the adaptive value of an individual for accelerating the training speed of swarm. After that, the improved PSO is used to iteratively solve the matrix equations in LS-SVM. APSO-LSSVM avoids the matrix inversion, saves the internal memory and obtains the optimum solution. The experiment results show that this method simplifies the training sample, accelerates the training speed, and also offers superior classification accuracy, fast convergence speed and good generalization ability.

Keywords: LS-SVM; adaptive; particle swarm optimization; prediction; coal spontaneous combustion

1. Introduction

The coal spontaneous combustion disaster in China was very serious and caused nearly 4,000 fire hazards due to coal spontaneous combustion. The direct economic losses of combusted coal amounted to billions [1]. Furthermore, the coal spontaneous combustion also caused secondary

disasters of gas and dust explosions, air pollution and other disasters. This helped us realize that effective prediction is the key to prevention and control of coal spontaneous combustion. Along with coal oxidation and temperature rising, it will release the corresponding indicator gases, such as CO, CO_2, CH_4, C_2H_6, C_2H_4, C_2H_2 and N_2. Therefore, there is a very complex nonlinear relationship between the degree of coal spontaneous combustion and gas products. By finding out the congruent relationships between these indicator gases and temperature of coal, monitoring the gas products of coal sample reaction, temperature, oxygen consumption and other indexes, the signs of coal spontaneous combustion could be detected so as to predict the development trend of spontaneous combustion [2]. However, in the coal spontaneous combustion process, due to the limitations of experimental conditions, the heating rates of coal temperature are different at different stages.

2. Related work

The most common method for prediction of coal spontaneous combustion is the comprehensive evaluation method, which is predicted by machine learning methods such as cluster analysis [3], neural network [4] and support vector machine [5]. Among them, the support vector machines are based on VC dimensional theory and structural risk minimization of statistical theory, and have a good generalization performance for small sample learning, rendering the possibility to effectively and accurately do predictions. Moreover, the traditional support vector machine does not work well on imbalanced sample classifications [6].

Through the proximate analysis of the composition in different coal samples, by means of the mathematical model of multiple regression analysis, Deng set up a regression equation to predict the system, the significance of the equation and its coefficient is tested, and residual analysis is carried out to identify the appropriateness of the equation [7]. Wang proposed a method to predict coal spontaneous combustion, which combines grey GM(1,1) model with Markov Model [8]. Based on the analysis on the seam spontaneous combustion in the coal mining face, the prediction technology with the support vector machine was applied to predict and analyze the spontaneous combustion danger of the residual coal in the goaf [9]. Paper [10] introduce fuzzy membership and least squares method, adopt neighborhood rough set method to reduce the dimensions of input vectors of the coal spontaneous combustion, and use PSO algorithm to optimize the parameters of SVM model. Then a fuzzy least square spherical support vector machine is presented, use sequential minimize optimization method to solve the quadratic programming problem, and establish a coal spontaneous combustion forecast model. Therefore, this paper uses the least square method, transferring learning and particle swarm algorithms in the support vector machine, and proposes a prediction method based on adaptive particle swarm optimization least squares support vector machine.

3. LS-SVM

3.1. LS-SVM basic theory

The design complexity of standard SVM algorithm is related to the number of training samples. When the number of training samples is too large, solving the corresponding quadratic programming problem becomes complex and the calculation speed will be slowed down accordingly. Therefore, the calculation speed of standard SVM in practical application is greatly restricted. The Least Square Support Vector Machine (LS-SVM) algorithm can solve this kind of problem. The main difference

between LS-SVM and standard SVM lies in the loss function and equality constraints.

In the structure of support vector machine, the input space of support vector machine is composed of original observation data, and the data of input space is mapped to high-dimensional feature space by kernel function. In feature space, support vector machine classifies or fits data by linear regression function. LS-SVM for classification is deduced as follows: Set the training set $S = \{(x_i, y_i)|i = 1,2,\cdots,m\}$, in which $x_i \in R^n$ and $y_i \in R$ are input data and output data; different from the classical SVM. LS-SVM uses SRM criterion to construct the minimum objective function and its constraint conditions as follows [11]:

$$\begin{aligned} \min J(w, e) &= \frac{1}{2}w^T w + \frac{\gamma}{2}\sum_{i=1}^{m} e_i^2 \\ s.t. \quad y_i &= w^T \Phi(x_i) + b + e_i \end{aligned} \tag{1}$$

Wherein, w is the weight vector, γ is the constant, b is the constant deviation.

In order to solve the optimization problem of equation (1), we translate it into the following system of linear equations.

$$\begin{bmatrix} 0 & L^T \\ L & Q + \gamma^{-1}I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{2}$$

Wherein, $L = [1,1,\cdots,1] \in R^m$, $\alpha = [\alpha_1, \alpha_2, \cdots, \alpha_m]^T \in R^m$, I is an identity matrix, $Q = y_i y_j \Phi(x_i)^T \Phi(x_i) = y_i y_j K(x_i, x_j), K(x_i, x_j) = \exp(-\|x - x_i\|^2/2\sigma^2)$, are the kernel functions satisfying the Mercer condition, $y = [y_1, y_2, \cdots, y_m]^T \in R^m$. Then the classification decision function of LS-SVM is [12]:

$$f(x) = \text{sgn}(\sum_{i=1}^{m} \alpha_i y_i K(x, x_i) + b) \tag{3}$$

Equation (2) is rewritten into a matrix equation $AX = z$, X is solved via the least squares method in the LS-SVM algorithm, so an A inversion is required. However, for large-scale problems in practical engineering, the dimensions of $A^T A$ are larger, so it is difficult to realize the process of matrix inversion [13]. Therefore, we can solve the matrix equation by the iterative computation of a PSO algorithm.

3.2. Multi-classification LS-SVM

The LS-SVM algorithm is initially designed for binary classification problems. In handling multi-class problems, it needs to construct a proper multi-class classifier. At present, the method for constructing the LS-SVM multi-class classifier is mainly to combine multiple binary classifiers to realize the construction of a multi-class classifier. The common methods are one to one and one to many [14]. The method of one to one is to design a LS-SVM between any two categories of samples. Thus, it needs k LS-SVM for $k(k-1)/2$ categories of samples. In classifying an unknown sample, the category with the most votes shall be the category for it. It needs to traverse all the classifiers for the method, which is complex in training and low in classification effectiveness. In the training of the method of one to many, it needs to organize the samples of some kind of category as one category, and the remaining samples are the other category. Therefore, samples of k categories will construct k LS-SVM. In the classification, the unknown sample is classified as the category with a maximum classification function value. Compared with the one to one method, the method greatly reduces the necessary classifiers for training and enhances the training speed.

4. LS-SVM algorithm based on an adaptive particle swarm optimization

4.1. Adaptive particle swarm optimization algorithm

The PSO Algorithm is an evolutionary computation suggested in 1995 by Doctor Eberhart and Doctor Kennedy of Purdue University as inspired by observing bird foraging behavior. The basic idea is to seek for the optimal solution through cooperation and information sharing between individuals among a group. That is, the system shall initialize a group of random particles, and find out the optimal solution by iteration. In every iteration, the particle updates itself by tracking two "extreme values". After finding the two optimal values, the particle will adjust its speed in every dimensional space and calculate its new position [15,16]. The particle evolution formula is

$$v_i^{k+1} = \omega v_i^k + c_1 r_1 \left(p_i^k - x_i^k \right) + c_2 r_2 \left(p_g^k - x_i^k \right)$$
$$x_i^{k+1} = x_i^k + v_i^{k+1} \tag{4}$$

In which the $r_1, r_2 \in U(0,1)$, v_i^k indicates the speed of the particle in the k time(s) of the iteration, x_i^k indicates the position of the particle in the k time(s) of iteration, and p_i^k indicates the individual extreme of the particle in the k time (s) iteration. p_g^k indicates the present global extreme point of the swarm in the k time(s) of iteration, ω is the inertia weight factor, and the constant c_1, c_2 is the acceleration factor.

From the model of PSO algorithm, if the acceleration factor c_1, c_2 or the inertia weight factor ω are with a bigger value, the particle swarms may miss the optimal solution, which will result in a non-convergence of the algorithm [17,18]. And even in convergence, if the tracking of the particle is the process of gradual convergence of the particle swarm, all the particles will tend to the optimal solution, which will cause an immature convergence, resulting in a rate of convergence in the later period of the algorithm to slow down and the decline in precision.

For this situation, the adjustment strategy is, if the search space dimension is large, in order to enhance the global search ability, the inertia weight should be increased appropriately; while the search space dimension is small, the inertia weight should be reduced appropriately to ensure the search efficiency of the algorithm. When adjusting the inertia weight, the balance between the search ability and the search efficiency should be achieved. To solve the problem, we can adjust the inertia weight factor ω and add the constraint factor α. We can introduce the source field factor and target domain factor by transferring learning [19], etc., to improve the basic model and obtain the self-adaptive PSO model:

$$v_i^{k+1} = \omega \cdot v_i^k + c_1 r_1 \left[\xi_q \left(p_i^k - x_i^k \right) + \xi_{q-1} \left(p_i^{k-1} - x_i^{k-1} \right) \right]$$

$$+ c_2 r_2 \left[\xi_q \left(p_g^k - x_i^k \right) + \xi_{q-1} \left(p_g^{k-1} - x_i^{k-1} \right) \right] \tag{5}$$

$$v_i = \begin{cases} v_{max} & v_i > v_{max} \\ -v_{max} & v_i < -v_{max} \end{cases} \tag{6}$$

In which the $\xi_q, \xi_{q-1} \in R^n$, $\sum_{i=0}^{q} \xi_{i-1} = 1$, ξ_q become the target domain factors, and the ξ_{q-1} is the source field factor. In particular, when $\xi_{q-1} = 0$ and $\xi_q = 1$, PSO is a special condition

of the APSO. From the perspective of psychology, the usage of knowledge of the source domain means the accumulation of the individual search experience of the particle, which is in favor of the quicker convergence of the algorithm.

The adjustment of the inertia weight factor ω can be done according to the advantages and disadvantages of the fitness of the particles. That is, in the initial stage of the algorithm, give ω a bigger positive value to obtain a better global searching ability; while in the later stage of the algorithm, give ω a smaller value to make the algorithm easier for convergence. The method suggested in the text may adjust the ω dynamically according to the degree of convergence and individual fitness of the swarm. The specific method is shown as follows:

When $f(x_i) < f(p_i)$, the particles meeting the condition are the ordinary particles in the swarm, which have good global optimizing capacity and local optimizing capacity. The inertia weight ω is changed as the following formula with the searching:

$$\omega = \omega_{min} + (\omega_{max} - \omega_{min})\frac{k_{max}-k}{k_{max}}$$

(7)

In which the ω_{max} is the maximum value ω in the beginning of search, which is 0.9. ω_{min} is the minimum value ω in the end of search, which is 0.2. k are the steps needed in the iteration, and k_{max} are the maximum times of iteration.

When $f(x_i) < f(p_g)$, the particles meeting the condition are the better particles in the swarm, which are close to the global optimal, so they shall be given with a smaller inertia weight to accelerate the convergence to the global optimum. The inertia weight ω is changed as depicted in the following formula with the searching:

$$\omega = \omega - (\omega - \omega_{min})\left|\frac{f(x_i)-f(p_i)}{f(p_g)-f(p_i)}\right|$$

(8)

In which the ω_{min} is the minimum value ω in the end of search, which is 0.2. The better the adaptive value of the particles, the smaller the inertia weight, which is beneficial to the local optimization.

4.2. Steps of the APSO-LSSVM algorithm

The text suggests that the LS-SVM algorithm is based on self-adaption, which may translate the solution of the least squares of the matrix equation $AX = z$ into the self-adaptive PSO algorithm for an iterative solution. In this way, a matrix inversion may be avoided, which ensures the convergence of the algorithm, accelerates the calculation speed, and enhances the solution accuracy. The process is as follows:

Step 1: Choose the proper training samples n_b and test the samples for pretreatment.

Step 2: Initialize the parameters of the particle swarm, including the speed and location of the particle. Set the parameters of the particle swarm, and R^n randomly produces m particles in space $(x_1, x_2, \cdots, x_m)$ to form the initial swarm X(k). The initial speed $(v_{i1}, v_{i2}, \cdots, v_{im})$ of the randomly produced particle forms the speed matrix V(k). The initial value of the optimal value of every particle is the initial value of x_i.

Step 3: Train LS-SVM with training samples to calculate the adaptive value f(x) of every

particle, and the formula of fitness function is: $f(x_i) = \frac{1}{l}\sum_{i=1}^{l}(y_i - x_i)^2$

In the formula, x_i is the actual value of the i sample, y_i is the predicted value of the i sample, and the l is the amount of the test sample. According to the fitness value of the particle, update p_{id} and p_{gd};

Step 4: for each particle, we compared the current fitness $f(x_i)$ with the fitness $f(p_i)$ of the best historical position. If $f(x_i) < f(p_i)$, then $p_i = x_i$, and we adjust ω according to Eq. (6). We compare the current fitness $f(x_i)$ of all the particles of the group with the fitness $f(p_g)$ of the best position of the group. If $f(x_i) < f(p_g)$, then $p_g = x_i$, and adjust ω according to Eq. (8).

Step 5: According to the improved PSO model $[v_i^{k+1}, x_i^{k+1}]$, update the velocity and position of the particle to produce a new population $X(k + 1)$ for $m = 1$ to M,

Step 6: Determine whether the velocity vector meets the constraints $-v_{max} \le v_i \le v_{max}$, and if not, adjust in accordance with the Eq. (6).

Step 7: Determine whether the fitness value meets the requirements or whether the maximum number of iterations has been reached. If the stop condition is satisfied, optimization ends, and the global optimal particle will be mapped to the optimized LS-SVM model parameters. Otherwise k = k + 1, and go to Step 3.

Step 8: The LS-SVM is solved by using the training sample data and the parameters obtained in Step 7 to obtain the least squares solution of the matrix equation. That is, the optimal parameters α_i and b are in the corresponding Eq.(2).

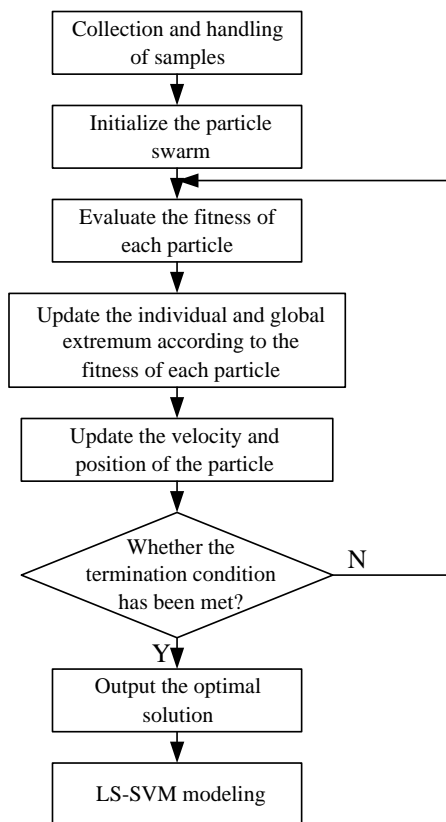The APSO-LSSVM algorithm flow chart is as follows:

Figure 1. Algorithm flow chart.

4.3. Convergence test

In order to verify the performance of APSO-LSSVM, three common Benchmark functions are selected as test functions. Rosenbrock is a single-peak function, search space is $[-100,100]^D$, Schwefel and Penalized are two-peak functions, and search space are $[-500,500]^D$ and $[-50,50]^D$, respectively. Figure 2 shows the convergence curves of three test functions optimized by different algorithms.
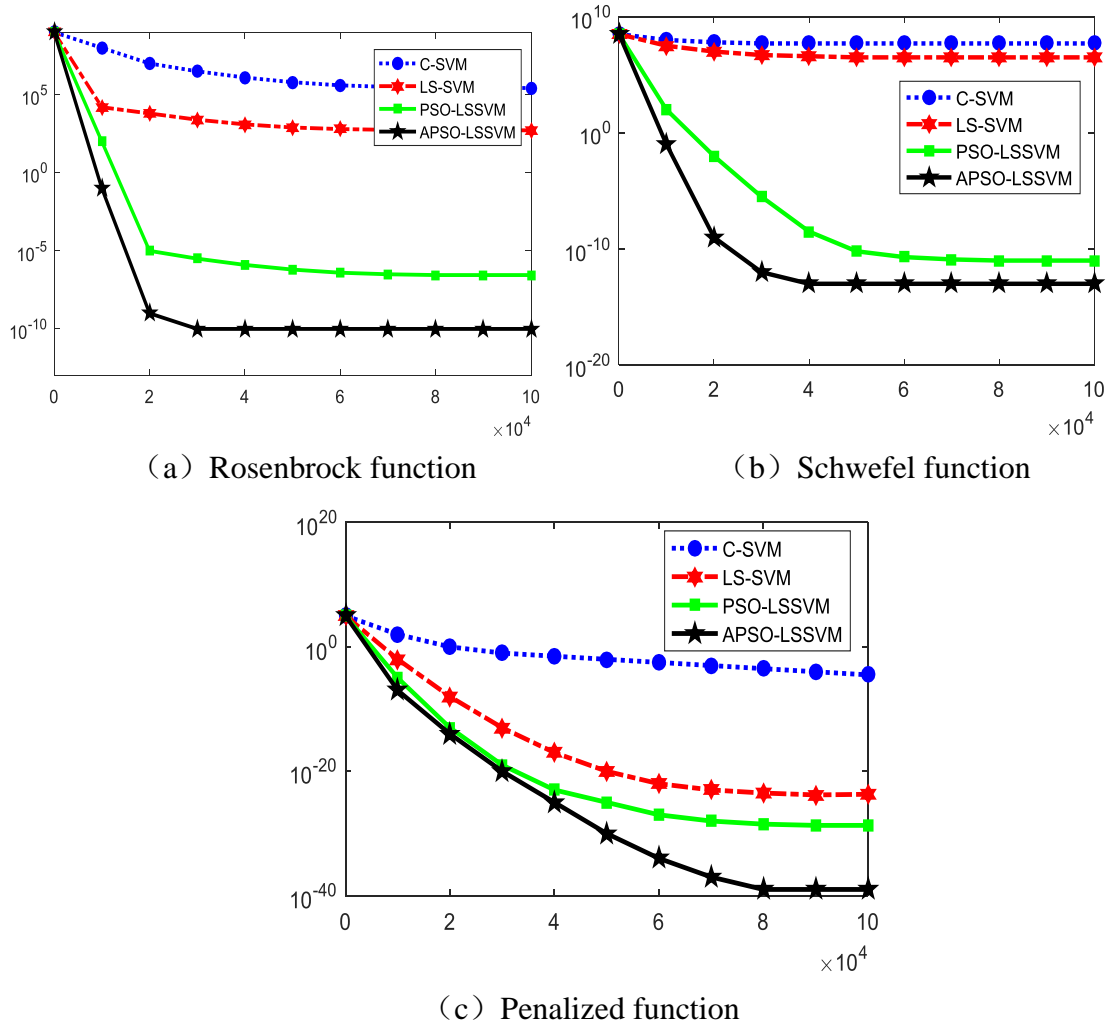

（a）Rosenbrock function   （b）Schwefel function

（c）Penalized function

Figure 2. Convergence contrast diagrams of three test functions.

It can be seen that APSO-LSSVM performs very well on all three functions. Since the adaptive strategy is adopted in the algorithm, the ability of the population to jump out of the local optimum and improve the accuracy of the solution can be improved while effectively maintaining the diversity of the population.

5. Experiments and analysis

5.1. Experimental setting

In 2013, coal samples from a certain coalmine in Hebi have been collected, and have carried out a spontaneous ignition test of the coal, and collected the sample data. We forecast the development trends of spontaneous ignition of coal by analyzing the characteristic parameters such as the concentration, ratio and occurrence rate of the indicator gas produced during the spontaneous ignition of coal. The process of coal spontaneous combustion can be generally divided into 3 stages [2]: the preparation period, the spontaneous heating period and the burning period. In order to better predict the state of coal spontaneous combustion, the spontaneous heating period can be further divided into 3 stages: early, mid-term and the later period of spontaneous heating.

C-SVM, LS-SVM, PSO-LSSVM and APSO-LSSVM algorithms were used to predict the degree of danger. The experiment uses the MATLAB 2010 version for programming with a 2.19GHz CPU and 2GB of memory.

To improve the accuracy of the sample, it will first be normalized to avoid the effects of singular point data on the performance of the Support Vector Machine. Set the size of the particle swarm to 25, the solution space to 350 dimensions, the maximum number of iterations to 1000, the acceleration factor $c_1 = c_2 = 2$, initial $\omega = 0.9$. The regularized parameter $\gamma = 1000$, the width parameter of the radial basis function $\sigma^2 = 0.15$, and establish 5 LS-SVM classifiers.

5.2. Experimental results and analysis

Test data samples with the prediction model obtained by iteration of the adaptive PSO, establish standard support vector model C-SVM, LS-SVM model and LS-SVM model of a standard PSO, and we compared the results with the predicted results proposed in this article. The C-SVM model adopts the radial basis function. The inertia weight ω in the LS-SVM model of the standard PSO is constant. The correlation curve of the training time is shown in Figure 3, and the range of the number of training sample data is [50,300]. The testing time of the prediction of accuracy is shown in Figure 4. The range of the number of testing samples is [50,300], and the number of testing samples is 300. The correlation curve of the prediction of accuracy is shown in Figure 5. The range of the number of training samples is [50,300]. The number of testing samples is 300.
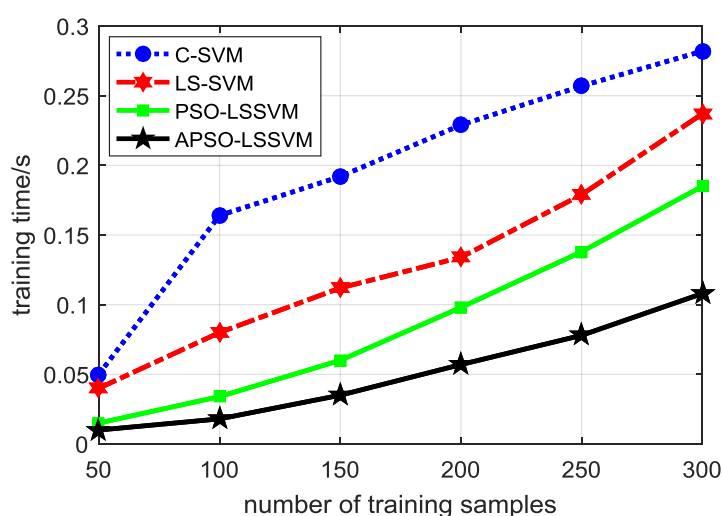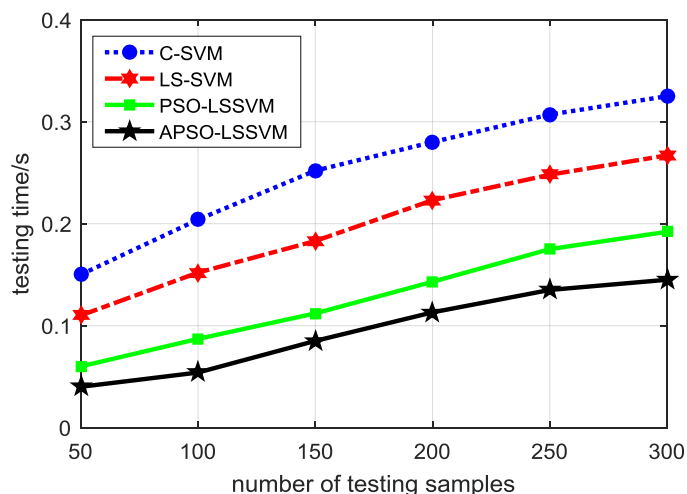

Figure 3. Correlation curve of training time.

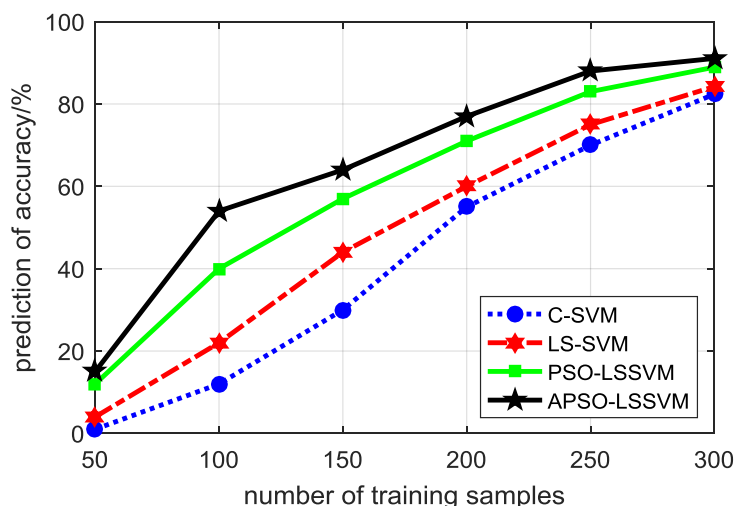**Figure 4.** Correlation curve of testing time.

Figure 5. Correlation curve of the prediction of accuracy.

It can be seen from the above response curves. As the number of training samples increases, the training time of 4 kinds of classification algorithms have been obviously increased. However, the training time of APSO-LSSVM is significantly shorter than that of C-SVM, LS-SVM and PSO-LSSVM, which proves that APSO-LSSVM has good adaptability to test conditions and the environment of the number of different samples and fast learning process. In terms of testing time, the processing time of 4 algorithms increases linearly with the increase of the number of test samples. However, the processing time of APSO-LSSVM is obviously shorter than that of C-SVM, LS-SVM and PSO-LSSVM and APSO-LSSVM, showing a good ability of real-time processing of APSO-LSSVM. Under the same conditions, the accuracy of classification of the 4 kinds of algorithms also increases with the increase of the number of training samples. However, the accuracy of APSO-LSSVM is slightly higher than that of C-SVM, LS-SVM and PSO-LSSVM. It can be seen that the adaptive PSO algorithm has obtained higher accuracy during the iteration process of the matrix.

Secondly, in order to test the performance of 4 kinds of algorithms and predict the accuracy of the results with the distribution of different samples, we selected coal samples of different coalmines for the test of spontaneous ignition to obtain the second group of sample data. We established the prediction model with different numbers of training samples. The test parameters respectively are as follows: training time, testing time, accuracy prediction and performance test, which are shown in Table 1.The test parameters respectively are as follows: training time, testing time, accuracy prediction and performance test, which are shown in Table 2.

Table 1. Experimental sample data.

	Data set	Class 1	CLASS 2	CLASS 3	Class 4	Class 5
Frist Group	Training Sample	55	67	59	67	62
	Test sample	33	48	31	43	45
Second group	Training Sample	35	40	36	31	36
	Test sample	21	31	24	19	25

Table 2. Comparison of algorithm performance.

	Algorithm	Training time (s)	Test time(s)	Forecast accuracy (%)
Frist group	C-SVM	0.282	0.325	82.38
	LS-SVM	0.237	0.267	84.23
	PSO-LSSVM	0.185	0.192	88.93
	APSO-LSSVM	0.108	0.145	91.07
Second group	C-SVM	0.188	0.253	83.57
	LS-SVM	0.145	0.226	85.76
	PSO-LSSVM	0.112	0.170	89.25
	APSO-LSSVM	0.073	0.104	92.12

As seen from the statistical results of the performance test of different coal sample data in two regions that the training and testing time consumption of APSO-LSSVM are significantly smaller than those of C-SVM, LS-SVM and PSO-LSSVM, which proves that APSO-LSSVM has competitive advantages in dealing with relatively complex issues and issues requiring higher real-time performance. The accuracy of the training set and testing set of APSO-LSSVM is slightly higher than that of C-SVM, LS-SVM and PSO-LSSVM, and APSO-LSSVM has a relatively small error, which shows it has a better classification effect.

Next, we consider the problem of time complexity. Assuming the complexity of training a classifier is C_h and updating a training sample is C_w, the time complexity of C-SVM, LS-SVM, PSO-LSSVM and APSO-LSSVM can be approximated to $C_h O(k_{max}) + C_w O(n_b k_{max})$,
$C_h O(k_{max}) + C_w O(lk_{max}), C_h O(Nk_{max}) + C_w O(lNk_{max})$ and $C_h O(Nk_{max}) + C_w O(lk_{max})$. The average training time of the four algorithms is compared in Figure 6.
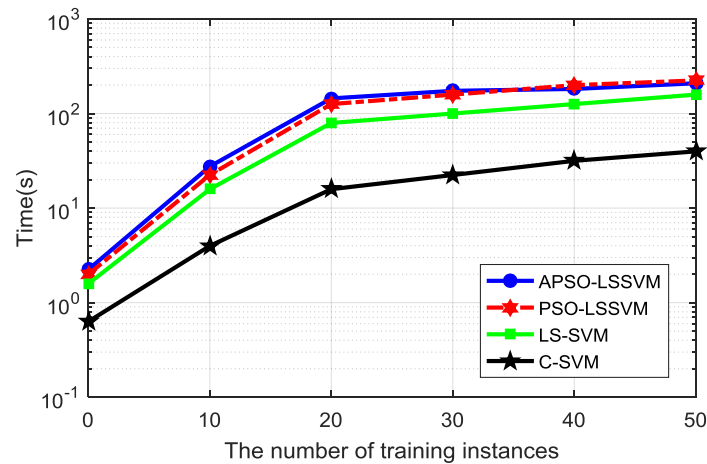
Figure 6. Time cost comparison on different methods.

In order to have objective and scientific comparison results, hypothesis testing is used on the experimental results. Let the variable X_1, X_2, X_3, X_4 denote the classification error rate of APSO-LSSVM, PSO-LSSVM, LS-SVM and C-SVM algorithms, respectively. Since the value of X_1, X_2, X_3, X_4 is subject to many random factors, we assume that they submit to normal distribution, $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1,2,3,4$. Now, we compare the random variable mean of these algorithms, $\mu_i, i = 1,2,3,4$. The smaller the μ_i is, the lower the expected classification error rate is, and the higher the efficiency is. Because the sample variance is the unbiased estimation of the overall variance, the sample variance value is used as an estimate of the generality variance. In this experiment the significance level α sets as 0.01.

Table 3. Hypothesis testing for experimental results.

Hypothesis	$H_0: \mu_1 \geq \mu_2$ $H_1: \mu_1 < \mu_2$	$H_0: \mu_1 \geq \mu_3$ $H_1: \mu_1 < \mu_3$	$H_0: \mu_1 \geq \mu_4$ $H_1: \mu_1 < \mu_4$
Statistics	$U_1 = \dfrac{\overline{X_1} - \overline{X_2}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$	$U_2 = \dfrac{\overline{X_1} - \overline{X_3}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_3^2}{n_3}}}$	$U_3 = \dfrac{\overline{X_1} - \overline{X_4}}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_4^2}{n_4}}}$
Rejection region	$U_1 \leq -Z_\alpha = -2.325$	$U_2 \leq -Z_\alpha = -2.325$	$U_3 \leq -Z_\alpha = -2.325$
Value of the statistic	$U_1 = -52.58$	$U_2 = -105.64$	$U_3 = -124.68$
Conclusion	$H_1: \mu_1 < \mu_2$	$H_1: \mu_1 < \mu_3$	$H_1: \mu_1 < \mu_4$

Table 3 shows the comparison process on μ_i and other parameters. We can see that, the expectations of prediction accuracy in APSO-LSSVM is far below than other algorithms.

5.3. Results of UCI database

In order to verify the effectiveness of the algorithm, the data is tested using the Wine dataset and the Iris dataset in the UCI database. Wine dataset contains 3 types of samples, a total of 178 samples. 89 are selected as training samples and 89 are used as test samples. The Iris data set contains 3 types

of samples, a total of 150 samples. This article selects 75 samples as training samples and 75 as test samples. In order to improve the accuracy of the sample, the sample was first normalized to [0, 1], support vector machine uses C-SVM, and kernel function uses Gauss radial base kernel function. The training process uses K-fold cross verification to test the accuracy of the judgment sample. Learning factors $c_1 = c_2 = 2$, and number of particles N = 40, maximum number of iterations M = 100, inertial weights $\omega = 1$.

Table 4. Data test results.

Dataset	Parameter C	Parameter σ	Time of parameter optimization	Training sample accuracy	Test sample accuracy	Range of parameter C	Range of parameter σ
Wine	68.94	0.01	20.36s	96.85%	95.34%	[0.1,1000]	[0.01,1000]
Iris	100	0.01	13.32s	97.80%	98.10%	[0.1,1000]	[0.01,1000]

6. Conclusion

The problem of coal spontaneous combustion prediction is very complex, and there are many factors that affect the prediction results. The LS-SVM algorithm based on adaptive PSO optimization proposed in this paper uses the LS-SVM algorithm to solve problems such as small samples, non-linear, high dimension and local very small points. The adaptive PSO algorithm is used to solve the problems such as high computational complexity and slow calculation speed of the LS-SVM model for large-scale samples, so that it can always obtain the optimal solution, and its training speed and accuracy are improved. In the coal spontaneous combustion experiment, the training and testing time consumption of the proposed method are significantly smaller than those of the remaining examined methods, which proves that APSO-LSSVM has competitive advantages in dealing with relatively complex issues and issues requiring higher real-time performance. The accuracy of the training set and testing set of APSO-LSSVM is slightly higher than that of the remaining examined methods, and APSO-LSSVM has a relatively small error, which shows it has a better classification effect.

7. Future work

On the basis of this article, there are still the following aspects to be studied. 1) How to more accurately solve the problem of unbalanced data in SVM requires further research. At the same time, the parameter selection problem of SVM is supported. Although the PSO is used to solve the problem to a certain extent, the parameters sought are only optimal for the relative training set. How to obtain the parameters in theory is a direction of research at present. 2) PSO will encounter a local optimal situation, so that the obtained parameters are not global optimal, but sub-optimal. How to obtain the global optimal solution more stably and efficiently for particles will be studied as the next step. 3) The prediction in this paper is only based on the index gas and does not combine with the surrounding environmental factors. In order to improve the effect of this model in practical application, the next step will comprehensively consider other factors, so that this model in practice to give full play to its advantages.

Acknowledgments

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. H. G. Li and Q. Zhang, Test system and automatic control of coal spontaneous combustion tendency oxidation kinetics, *Chin. Univ. Min. Techn. Press*, (2014), 25–49.
2. H. F. Zhang, X. G. Chen and X. L. Sun, Classification research on coal spontaneous combustion based on fuzzy comprehensive evaluation, *Coal Techn.*, **33**(2014), 209–210.
3. T. S. Li, C. J. Lin, P. H. Kuo, et al., Grasping posture control design for a home service robot using an ABC-based adaptive PSO algorithm, *Int. J. Adv. Robot. Syst.*, (2016), 1–15.
4. K. Meng, Z. Y. Dong and D. H. Wang, A self-adaptive RBF neural network classifier for transformer Fault analysis, *IEEE Trans. Power Appar. Syst.*, **25**(2010), 1350–1360.
5. W. C. Gu, B. R. Chai and Y. P. Teng, Research on support vector machine based on particle swarm optimization, *T. Beijing Inst. Techn.*, **34**(2014), 705–709.
6. H. G. Bargh and M. H. Sadr, Vibration Optimization of fiber-metal laminated composite shallow shell panels using an adaptive PSO algorithm, *Mech. Adv. Compos. Struct.*, **4**(2017), 99–110.
7. J. Deng, Z. Xing and L. Ma, Application of multiple regression analysis in coal spontaneous combustion prediction, *J. Xi'An Univ. Sci. Technol.*, **3**(2011), 645–648.
8. L. Wang, S.J. Wu and C. Q. Li, Coal spontaneous combustion prediction based on grey-markov model, *Comput. S.*, **31**(2014), 416–420.
9. Y. Gao, M. G. Qin and M. J. Li, Analysis on prediction of residual coal spontaneous combustion in goaf based on support vector machine, *Coal Sci. Technol.*, **38**(2010),50–54.
10. Y. P. Jin, B. Zhang and K. Gao, Application of spherical SVM in forecasting spontaneous combustion of coal, *Comput. Appl. Softw.*, **30**(2013), 57–60.
11. A. Sahu and S. Patnaik, Evolving neuro structure using adaptive PSO and modified TLBO for classification, *Procedia Computer Sci.*, **92** (2016), 450–454.
12. W. P. Ding, C. T. Lin, M. Prasad, et al., A layered coevolution based attribute-boosted reduction using adaptive quantum behavior PSO and its consistent segmentation for neonates brain tissue, *IEEE T. Fuzzy Syst.*, (2017),1–16.
13. K. W. Xia, Y. Dong and H. B. Du, Oil layer recognition model of LS-SVM based on improved PSO algorithm, *Control Decis.*, **22**(2007), 1385–1389.
14. S. R. Inkollu and V. R. Kota, Optimal setting of FACTS devices for voltage stability improvement using PSO adaptive GSA hybrid algorithm, *Eng. Sci. Techn.*, **19**(2016),1166–1176.
15. Y. Zhang, D. W. Gong and J. Cheng, Multi-objective particle swarm optimization approach for cost-based feature selection in classification, *IEEE/ACM T. Comput. Bi.*, **14**(2017), 64–75.
16. L. L. Shan, H. J. Zhang, J. Wang, et al., Parameters optimization and implementation of mixed kernels E-SVM based on improved PSO algorithm, *Appl. Res. Comp.*, **30**(2013), 1636–1639.

17. Y. Zhang, D. W. Gong and Z. H. Ding, Handling multi-objective optimization problems with a multi-swarm cooperative particle swarm optimizer, *Expert Syst. Appl.*, **38**(2011): 3933–3941.

18. L. S. Coelho, Gaussian quantum-behaved particle swarm optimization approaches for constrained engineering design problems, *Expert Syst. Appl.*, **37**(2010), 1676–1683.

19. Q. Zhang, M. Li, X. S. Wang, et al., Instance-based transfer learning for multi-source domains, *Acta Autom. Sin.*, **40**(2014), 1176–1183.