Application of Optimized SVM in Sample Classification

Xuemei Yao
School of Data Science and Information Engineering, Guizhou Minzu University, China

***Abstract*—Support vector machines (SVM) have unique advantages in solving problems with small samples, nonlinearity and high dimension. It has a relatively complete theory and has been widely used in various fields. The classification accuracy and generalization ability of SVMs are determined by the selected parameters, for which there is no solid theoretical guidance. To address this parameter optimization problem, we applied random selection, genetic algorithms (GA), particle swarm optimization (PSO) and K-fold cross validation (k-CV) method to optimize the parameters of SVMs. Taking the classification accuracy, mean squared error and squared correlation coefficient as the goal, the K-fold cross validation method is chosen as the best way to optimize SVM parameters. In order to further verify the best performance of the SVM whose parameters are optimized by the K-fold cross validation method, the back propagation neural network and decision tree are used as the contrast models. The experimental results show that the SVM-cross validation method has the highest classification accuracy in SVM parameter selection, which lead to SVM classifiers that outperform both BP neural networks and decision tree method.**

Keywords*—Support vector machine; parameter optimization; K-fold cross validation; sample classification

I. Introduction

Support Vector Machine (SVM) is a theory that studies the rule of machine learning in the case of limited sample based on the statistical learning theory including Vapnik-Chervonenkis Dimension (VCD) and Structural Risk Minimization (SRM). SVM has many unique advantages in solving the problem of high-dimensional pattern recognition. It can use the limited sample information to compromise the complexity and learning ability of the model, and avoid the problems caused by over-learning and under-learning as much as possible, so that the system has a better ability to extend. SVM has attracted the attention and research interests of experts and scholars in various fields, and made a lot of research results in practical application, which has promoted the development of various fields [1-3].

Although SVM has been widely used in text classification [4], image recognition[5], prediction[6] and so on, its performance is mainly dependent on the selection of penalty factor and kernel parameter. So far, SVM parameter selection still does not have a complete set of theory and standards, but in practical application, the choice of parameter directly determines the classification accuracy and generalization ability of SVM [7-8]. The common methods of optimizing parameters are as follows, experimental method, grid search and numerical method. The first way means that you need to select different parameters for multiple experiments, and then select a pair of parameters for the best results, which is not only time consuming and too random; the second method is feasible in the case of small samples, but the efficiency and feasibility is lower when the data is large; the third method is more sensitive to the selection of the initial value. With the development of artificial intelligence, some groups of intelligent optimization algorithm such as ant colony optimization algorithm (ACO), genetic algorithm (GA), particle swarm algorithm (PSO) and so on are used to optimize the parameters of SVM.

Rajeshwari et al [9] proposed a new Weighted-SVM kernel by applying a suitably transformed weight vector derived from particle swarm optimized neural networks. Yang et al [10] envisaged the analysis of the dissolved oxygen fault of the water quality monitoring system using the GA-SVM and the result exhibited a good accuracy. Faris et al [11] proposed a robust approach based on a recent nature-inspired metaheuristic called multi-verse optimizer for selecting optimal features and optimizing the parameters of SVM simultaneously and it can effectively reduce the number of features while maintaining a high prediction accuracy. John et al [12] proposed a detection technique using SVM with Grid search algorithm and recognized the disorder with an accuracy rate of 89%.

In the above research, it is for a specific field to select a suitable way to optimize the parameters of SVM, but it does not compare these parameter optimization algorithms in the same application to select the best way. To address this issue, based on the identification of Italian wine, random selection, GA, PSO and cross validation (CV) are used to optimize the parameters of SVM, and the best way was selected. In order to verify that the SVM which is optimized by the best way has a good classification accuracy, the neural network and decision tree are used as contrast model for analysis.

II. Methodology

A. Introduction of SVM

SVM is a statistical learning method proposed by Vapnik in 1992 based on VCD theory and SRM principle [13]. SVM maps the input space into a high dimensional kernel space by introducing the kernel function. In order to classify the sample, the optimal classification hyperplane with lower VCD is obtained in the high dimensional kernel space, so that one class is located on the side of the hyperplane and the other side of the hyperplane is another class. Take the Two-Dimensional (2D) space as an example showed in Fig. 1. There are many samples, some with five-pointed star and some with a circle.

This work is supported by the Natural Science Foundation of the Guizhou Province (No. [2020]1Y263).

Classification by SVM is to find a hyperplane (the red line in Fig. 1), according to the classification requirements, the sample will be divided into class 1 with five-pointed star and class 2 with circle. The point falling on the classification boundary is called support vector

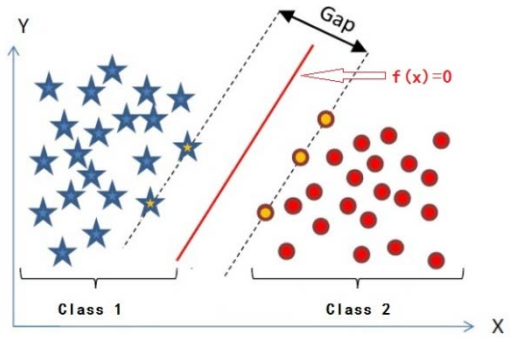


Fig. 1. Schematic Diagram of SVM in 2D Space.

The main idea of SVM is to establish a classification hyperplane *f(x) = 0* as a decision surface for a multidimensional space to be classified, for any positive case x satisfying *f(x) > 0*, and the counter example is satisfied with *f(x) < 0*. The purpose of SVM is to find the appropriate *f(x)* to maximize the isolation edge (Gap) and minimize the training error between the positive and negative examples. Through the decision-making surface to achieve the classification of the sample, the decision-making function can be written as:

$$f(x) = sign(\sum_{x_{i\in S_V}}^{n} a_i y_i k(x_i, x) + b) \quad (1)$$

where a_i: Lagrange multiplier, s.t. $0 < a_i < c$

S_V: Support vector.

$K(x_i,x)$: Kernel function.

x_i: The positive examples.

y_i: The negative examples.

b: Threshold.

n: The number of samples.

In practical application, SVM has a complete theoretical basis, but it needs to choose different parameters and kernel functions for different conditions in order to obtain the desired results. Different support vector classifiers can be generated by using different kernel functions. Commonly used kernel functions are linear kernel function, radial basis function (RBF), polynomial kernel function and so on. Since it is only necessary to determine a parameter in RBF, which is beneficial to parameter optimization, RBF is chosen as the kernel function of SVM in this article. The expression can now be written as:

$$K(x_i, x) = exp(\frac{1}{2g^2}||x_i - x||^2) \quad (2)$$

Where *g* is the nuclear parameter to be determined. The penalty factor *c* is an important parameter in SVM. It is used to control the weight of the loss and sorting interval. The bigger the value of *c*, the higher the fit degree of the model, but its generalization ability is reduced. According to SVM regression theory, *c* and *g* have a great influence on the classification results. Therefore, it is necessary to choose the best *c* and *g* to obtain the superior performance of SVM.

Definition 1

The mathematical model of SVM parameter will be written as: *P*= {*bestc*, *bestg*}.

Definition 2

The objective function of SVM parameter optimization is the accuracy of prediction model.

In this paper, the problem of SVM parameter optimization can be described mathematically:

For $\forall p=\{c,g\}$, satisfied *P*=max(Accuracy)

$$\text{s.t.}\begin{cases} P \subset p \\ c > 0 \\ g > 0 \end{cases} \quad (3)$$

SVM solves the practical puzzles such as small sample, nonlinearity and high dimension. It has strong versatility, robustness and effectiveness and it is widely used in classification and regression. The SVM is originally designed for two classification problems. When dealing with multiple classification, it is necessary to combine multiple two classifiers to construct the appropriate multi-classifier. In this paper, the tool used in constructing the forecasting model is the LibSVM software package developed by Chih-Jen Lin of Taiwan University.

B. SVM Parameter Optimiztion Method

1) Random selection: Random selection belongs to the experimental method. According to the theory of SVM, we randomly select different parameters to test and analyze the samples, and then choose a set of parameters with the highest classification accuracy. The method is too random and lacks of authoritative theoretical guidance.

2) Genetic Algorithm (GA): GA is proposed by Professor J. Holland of the United States in 1975. Its main feature is the direct operation of the structure of the object; there is no limit of derivation and continuity of the function; it has better global optimization ability [14]. GA is a computational model for simulating the natural selection and genetic mechanism of Darwin's theory of biological evolution. It is a method for searching the optimal solution by simulating natural evolutionary processes [15]. It introduces the principle of "survival of the fittest" in the process of coding of parameter optimization. According to the selected fitness function, the individuals are selected by the selection, crossover and mutation of the genetic process, so that the individuals with good fitness value are retained, the remaining are eliminated. The new group not only inherited the previous generation of information, but also better than it. It is repeatedly until the conditions are satisfied. Genetic algorithm has been widely used in combinatorial optimization, machine learning, adaptive control and other fields. The process of GA shows in Fig. 2.

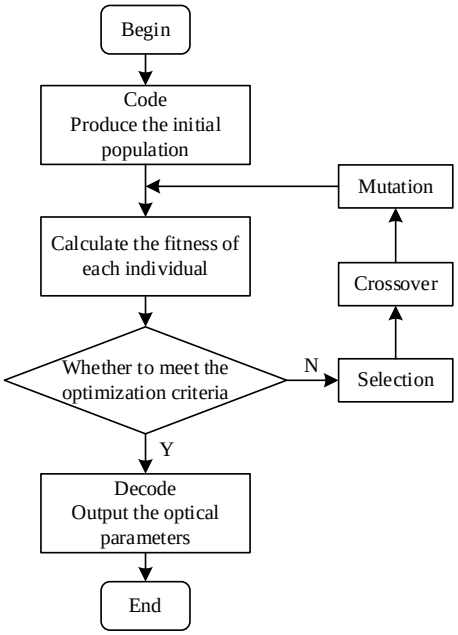


Fig. 2. Flow Chart of GA.

3) Particle Swarm Optimization (PSO): PSO is an optimization algorithm based on population intelligence in computing intelligence [16]. It is first proposed by Kennedy and Eberhart in 1995 and its basic concept comes from the study of predatory behavior of birds. PSO initializes a group of particles in the solvable space first. Each particle represents a potential optimal solution of the search space and the characteristics of particle are displayed by three indexes of position, velocity and fitness value. The particle moves in the solution space [17]. By tracking the trajectories of the particle, it updates the position and speed of individual constantly to follow the optimal output. The process of PSO algorithm shows in Fig. 3.

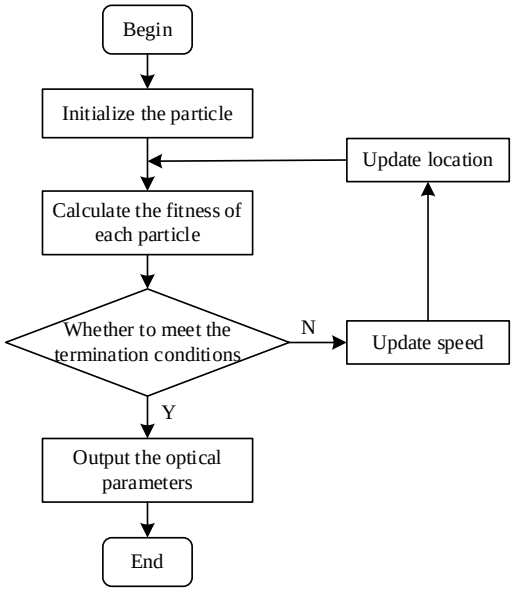


Fig. 3. Flow Chart of PSO.

4) Cross-Validation (CV): CV is a statistical analysis method used to verify the performance of a classifier. The basic idea is to divide the original data into two parts in a certain sense, one is the training set and the other is validation set [18-19]. Firstly, the training set is used to train the classifier, then the validation set is used to test the model to obtain the classification accuracy as the performance index of the classifier. The optimal parameters obtained by CV can effectively avoid the occurrence of over learning and less learning, and finally get the ideal prediction accuracy for the testing set. The Hold-out method, *K*-fold CV, Leave-one-out CV are commonly used methods. The second method is adopted in the article that is *K*-CV. Its basic idea is to divide the original data into *K* disjoint subsets, then it selects one of them as the validation set and the other *K*-1 subsets as the training set. The *K* models can be obtained in this way, and the average of the final classification accuracy of the *K* models as the performance index of the *K*-CV classifier. Generally speaking, *K* is greater than or equal to 2. But in practice, *K* is taken from 3. *K* will try to take 2 only when the original data is very small.

C. Classification Forecasting Model Based on SVM

1) Data preprocessing: Raw data usually has a different order of magnitude, which directly affects the results of the data analysis. In order to eliminate the influence, data standardization is required to address the comparability between data indicators. The data is in the same order of magnitude after processing and it is suitable for comprehensive evaluation. Two commonly used normalized methods are as follows.

a) Normalized between [0,1]: The normalized mapping used in this method is as follows.

$$f: x \to y = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (4)$$

where $x, y \in R^*$; $x_{min} = \min(x)$; $x_{max} = \max(x)$. The result of the normalization is that the raw data is scaled to [0,1].

b) Normalized between [-1,1]: The normalized mapping used in this method is as follows.

$$f: x \to y = 2 * \frac{x - x_{min}}{x_{max} - x_{min}} - 1 \quad (5)$$

where $x, y \in R^*$; $x_{min} = \min(x)$; $x_{max} = \max(x)$. The result of the normalization is that the raw data is scaled to [-1, 1].

2) Classification forecasting model : The classification prediction model based on SVM is divided into two stages: parameter optimization and prediction classification. The first stage is parameter optimization. The training set and testing set are separated from the original data, then the normalized preprocessing method makes the data samples in the same order of magnitude with the same dimension. The parameters of SVM are optimized by random selection, GA, PSO and *K*-fold CV method. The optimal parameters are selected from

three angles: accuracy, MSE and r^2. The second stage is prediction classification. The *K*-CV-SVM model is established with the best parameters to find he support vector of the training sample and determine the optimal classification hyperplane, and then forecast the testing set. In order to further verify the superior performance of *K*-CV-SVM model, neural network and decision tree are designed as contrast model. The overall structure of the paper is shown in Fig. 4.

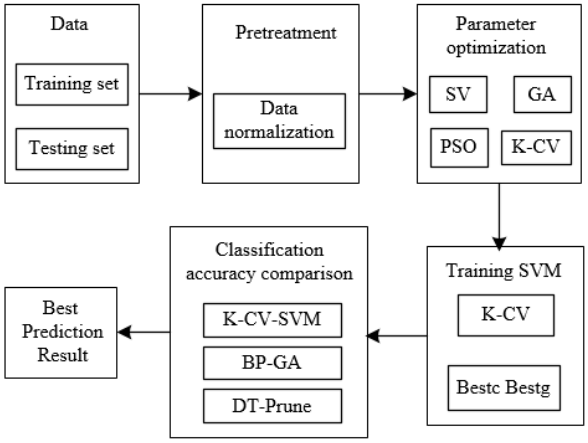


Fig. 4. Diagram of SVM Classification.

III. Experimental Discussion

A. Experimental Data Description

In order to verify the validity of the method proposed in this paper, we use the wine data to carry out the experimental. Wine data is from the UCI dataset and it can be download from the link: http://archive.ics.uci.edu/ml/datasets/Wine. These data are the results of a chemical analysis of wines grown in the same region in Italy but derived from three different cultivars. The analysis determines the composition of each wine. The data contains 178 samples and each contains 13 attributes and 1 label. The attributes which are the main chemical composition of wine donated by Riccardo Leardi are: 1) Alcohol; 2) Malic acid; 3) Ash; 4) Alcalinity of ash; 5) Magnesium; 6) Total phenols; 7) Flavanoids; 8) Nonflavanoid phenols; 9) Proanthocyanins; 10) Color intensity; 11) Hue; 12) OD280/OD315 of diluted wines; 13) Proline. Wine data is divided into two parts on average, one as a training set, and one as a testing set. The classification model is obtained by training the SVM, and then the model is used to predict the label of testing set. It is necessary to note that the selection of training set and testing set is fixed at the time of parameter optimization, but it is random selected when the classification prediction accuracy of SVM, BP and DT is analyzed.

B. Parameter Optimization of SVM

In order to test the effectiveness of different methods for SVM parameter optimization, four kinds of prediction models are designed to compare and analyze, such as classification model of random selection parameters of support vector machine (RS-SVM), classification model of genetic algorithm optimization parameters of support vector machine (GA-SVM), classification model of particle swarm algorithm optimization parameters of support vector machine (PSO-SVM), classification model of *K*-fold cross validation algorithm optimization parameters of support vector machine (*K*-CV-SVM).

1) RS-SVM: Random selection parameter refers to the absence of any restrictions, without any theoretical support, completely by the operator according to their own wishes to randomly specify the value of the parameter. If you do not manually set the parameters, SVM will use the default value to perform. In this paper, *c* is 13 and *g* is 17 by manual setting. The prediction accuracy of the model is 74.1573%, as shown in Table I. The classification prediction error of the model is shown in Fig. 9(a), and the classification chart is shown in Fig. 10(a).

2) GA-SVM: GA is a search heuristic algorithm used in solving the optimization problem in the field of artificial intelligence. The GA-SVM uses GA to find the best parameters for SVM. The parameters of GA are set as follows.

The number of iterations: 200.

The size of population: 20.

The probability of crossover: 0.4.

The probability of mutation: 0.2.

The result of model is that bestc is 3.2703, bestg is 3.6738, and classification accuracy is 97.7528%. They are showed in Table I. The best fitness curve is showed in Fig. 5. In the GA with the number of iterations of 200, it can be seen that the speed of parameter optimization converges faster. It is obvious that the average fitness of the population is stable between 85% and 93% after 10 generations of evolution, and tends to be stable. In this way, the classification prediction error of the model is shown in Fig. 9(b), and the classification chart is shown in Fig. 10(b).

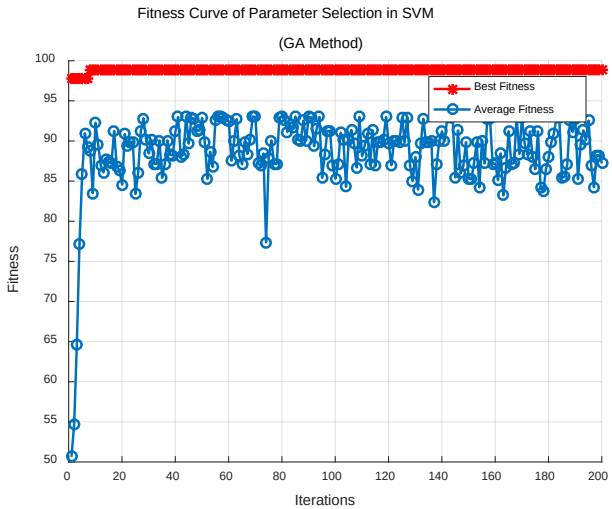


Fig. 5. Fitness Curve of GA-SVM.

3) PSO-SVM: PSO is an optimization algorithm based on group intelligence. Compared with GA, PSO has no selection, crossover and mutation operation, but it searches the result

through the particles in the solution space to follow the optimal way. The PSO-SVM uses PSO to find the best parameters for SVM. The parameters of PSO are set as follows.

The number of iterations: 200.

The size of population: 20.

The result of model is that bestc is 32.7842, bestg is 4.4621, and classification accuracy is 97.7528%. They are showed in Table I. The best fitness curve is showed in Fig. 6. Obviously, the stability of model fitness is poor and the variation range is large. The average fitness of the population fluctuates greatly between 50% and 80% after the population begins to evolve. It can be seen that there is still no steady trend with the number of iterations of 200. In this way, the classification prediction error of the model is shown in Fig. 9(c), and the classification chart is shown in Fig. 10(c).

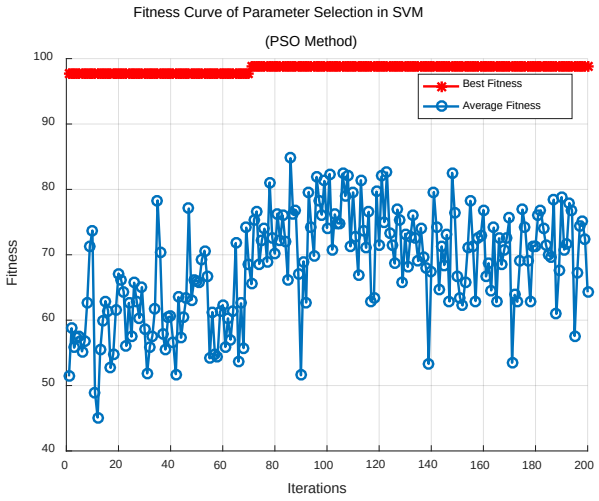


Fig. 6. Fitness Curve of PSO-SVM.

4) K-CV-SVM: The model is designed using the *K*-CV method to obtain the best parameters of SVM and divides into two phases. The first stage is a rough choice. Let us set the value of c and g in a certain range ($2^{-10} < c < 2^{10}$ and $2^{-10} < g < 2^{10}$). When the value of c and g is selected, the training set is taken as the original data and the classification accuracy is obtained by the *K*-CV method. Finally, a pair of c and g with the highest classification accuracy are selected as the optimal parameter. If there are multiple sets of c and g corresponding to the highest classification accuracy, then select the combination with the smallest c as the optimal parameters; if the corresponding c has multiple g, then select the first group of c and g as the best parameter. The reason is that the larger c will lead to the occurrence of the over learning, that is, the classification accuracy of training set is high, but the testing set is very low. So in all combinations that can achieve the highest classification accuracy, the smaller c is considered a better choice for the object. The rough selection of the parameters gives the contour map as shown in Fig. 7.

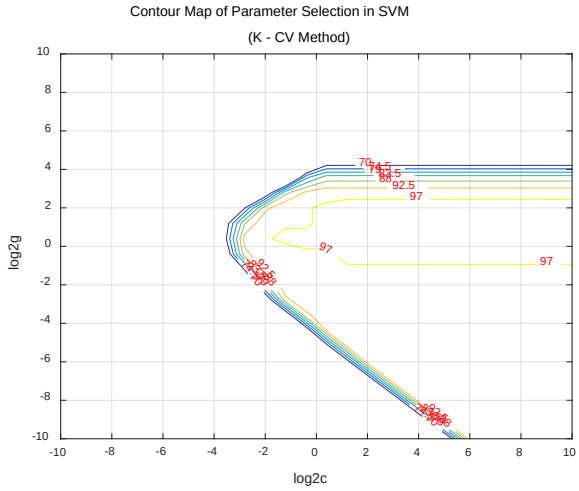


Fig. 7. Contour Map of Parameter Selected Roughly.

In the Fig. 7, the *x*-axis represents the value of c, which is the logarithm of 2, and the *y*-axis represents the value of g, which is the logarithm of 2, and the contour line indicates the classification accuracy with corresponding c and g which are gained by *K*-CV method. As can be seen from the Fig. 7, the range of c can be reduced to 2^{-2}~2^{4}, the range of g can be reduced to 2^{-2}~2^{4}, so that the fine selection can be made on the basis of rough selection. The fine selection of the parameters gives the contour map as shown in Fig. 8.

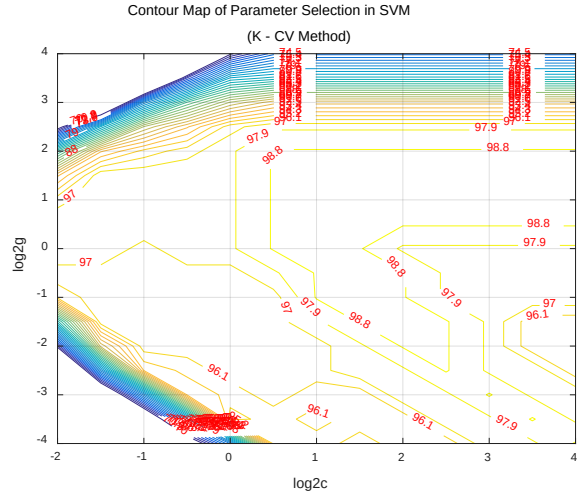
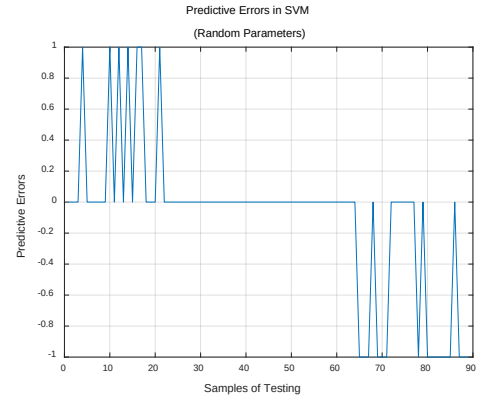


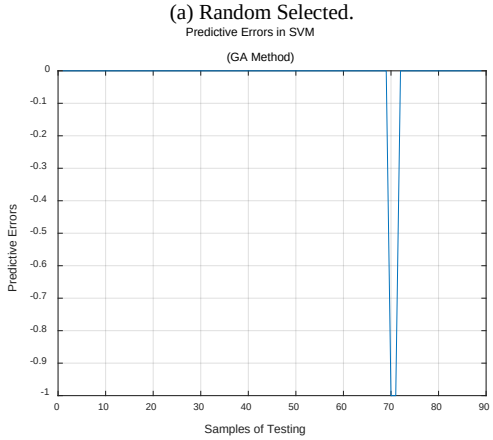
Fig. 8. Contour Map of Parameter Selected Finely.

The result of fine selection of parameters is that c is 1.4142 and g is 1. In this way, the classification accuracy of model is 98.8764 showed in Table I, the classification prediction error of the model is shown in Fig. 9(d), and the classification chart is shown in Fig. 10(d).

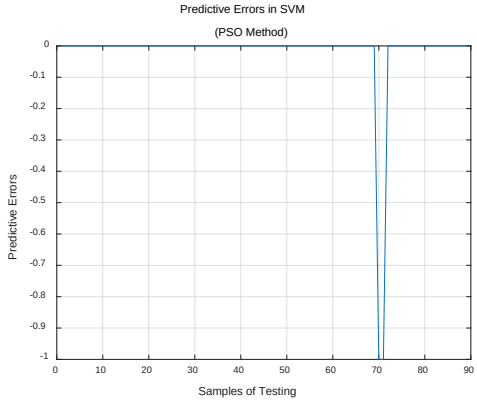
The prediction error curve showed in the Fig. 9 indicates that the prediction error of RS-SVM is the largest, while the prediction error of k-CV-SVM is the smallest. The prediction error of GA-SVM and PSO-SVM are equal. The classification accuracy showed in Table I also illustrates it.



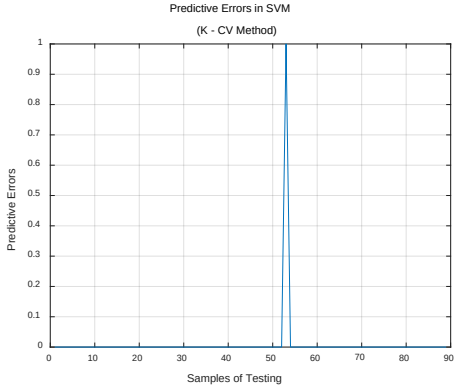
(a) Random Selected.



(b) GA Method.

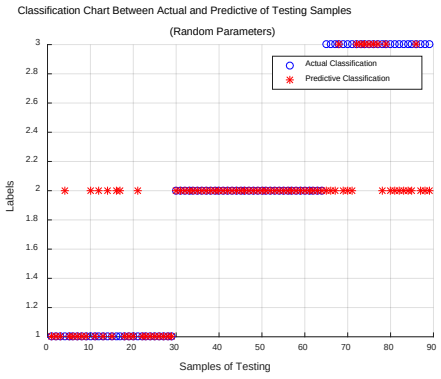


(c) PSO Method.

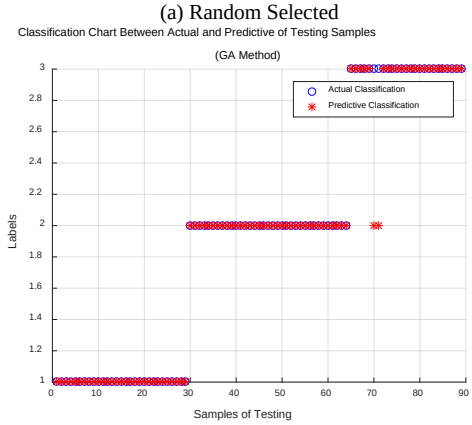


(d) K-CV Method.

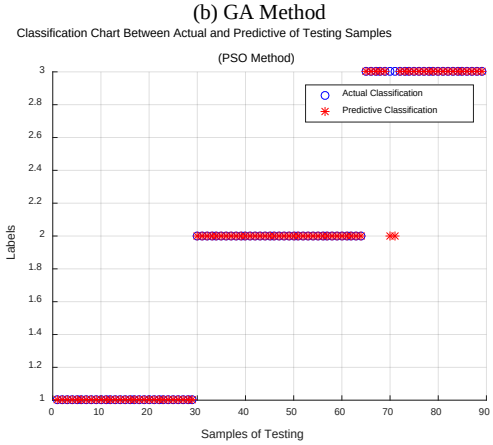
Fig. 9. Predictive Error Curve of Different Parameter Optimization Methods.



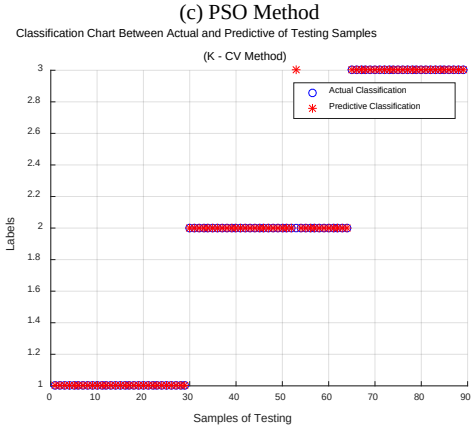
(a) Random Selected



(b) GA Method



(c) PSO Method



(d) K-CV Method

Fig. 10. Predictive and Practical Classification of different Parameter Optimization Methods.

Fig. 10 shows that the number of right classification is 66 and the number of wrong classification is 23 in the RS-SVM model. The number of misclassifications is the highest among the four models, that is, the prediction error of the model is the largest, which is consistent with Fig. 9(a). In the GA-SVM model, the number of right classification is 87 and the number of wrong classification is 2. The classification result of PSO-SVM model is equivalent to GA-SVM and the prediction error is similar, which is consistent with Fig. 9(b) and (c). In the K-CV-SVM model, the number of right classification is 88 and the number of wrong classification is 1. The number of misclassifications is the least among the four models, that is, the prediction error of the model is the least, which is consistent with Fig. 9(d). Overall, the K-CV-SVM model has the highest number of correct classification and the highest classification accuracy.

The other two parameters can also be obtained while the model is running. One is mean squared error (MSE), the other is squared correlation coefficient (r^2). Generally speaking, the accuracy is considered in the problem of classification and the MSE and r^2 are considered in regression issues. MSE is an indicator of the distance between the estimated and true values of the parameter. The smaller the value of MSE, the closer the results of the prediction model are to the real state and the model will have a better accuracy. r^2 is a statistical indicator used to describe the correlation and closeness between the variables. The larger the value of r^2, the closer is the relationship between the variables. The model optimizes the parameters of SVM with four different methods and carries out the prediction of classification. Table I shows that K-CV-SVM has the highest classification accuracy in four models. On the other hand, K-CV-SVM has the smallest MSE and maximum r^2, indicating that the prediction results of K-CV-SVM model are the closest to the real state. At the same time, compared with GA-SVM and PSO-SVM, the K-CV-SVM model consumes the least time when the classification accuracy is similar. Therefore, K-CV method is the best way to optimize the parameters of SVM, whether it is classification or regression.

C. The Application of Optimized SVM in Classification

In order to further verify the best effect of SVM by the K-CV method to optimize the parameter, the back propagation neural network (BP) and decision tree (DT) are used as contrast model based on the same data set. In order to eliminate the influence of the random probability on model, the BP and DT model are run 50 times, and the best and the worst state are selected to increase the contrast effect. The results of prediction are showed in Table II.

It can be seen from Table II that the highest classification accuracy is 100% and the worst case is only 75.28% in the BP prediction model, and the number of wrong samples reached 22 which reflects the natural defect that the BP is easy to fall into the local optimum. In order to overcome the defect, the GA is used to optimize the initial weights and thresholds of the BP (GA-BP), and then the training and prediction are carried out. The classification accuracy of the GA-BP is improved to 93.62%. Similarly, the performance of the DT is also uneven, the best accuracy of the classification is 97.75% and the worst is only 78.65% and the number of wrong samples is 19. The reason is that the original DT is flourishing, which contains a lot of noise and boundary nodes, which is prone to over-fitting. In general, pruning is performed on the DT to improve the classification accuracy. The result in Table II shows that the classification accuracy of the DT after pruning (DT-Prune) is improved to 96.63%, which further confirmed the effect of pruning. Compared with K-CV-SVM, GA-BP and DT-Prune found that the classification accuracy of K-CV-SVM model is 98.88%, and only one sample is wrong. The K-CV-SVM model has the best performance in three models.

TABLE I. PARAMETER OPTIMIZATION RESULT CONTRAST

	Random	**GA**	**PSO**	**K-CV**
c	13	3.2708	32.7842	1.4142
g	17	3.6738	4.4621	1
Accuracy (%)	74.1573	97.7528	97.7528	98.8764
MSE	0.2584	0.0225	0.0225	0.0112
r^2	0.5907	0.9637	0.9637	0.9820
Time(s)	0.9340	12.6566	16.7837	5.2956

TABLE II. COMPARISON WITH OVERALL PERFORMANCE OF THREE PREDICTION METHOD

	SVM		**BP**			**DT**	
	K-CV	Best	Worst	GA-BP	Best	Worst	DT-Prune
Right Number	88	89	67	83	87	70	86
Wrong Number	1	0	22	6	2	19	3
Accuracy (%)	98.88	100	75.28	93.26	97.75	78.65	96.63
Time(s)	5.36	1.22	2.76	0.96	1.56	1.55	2.65

IV. Conclusion

SVM classification prediction model solves the problem of nonlinear and small sample, but the effect of classification has a great relationship with the parameter setting. According to the theory of SVM, the penalty factor (c) and the kernel parameter (g) have the greatest influence on the classification result. Therefore, the paper analyzes the efficiency of the four methods to optimize the parameters of SVM. Experiments show that the optimal parameters of SVM can be obtained by K-CV and the classification accuracy is highest. In order to further illustrate the conclusion, the BP and DT are designed as the contrast model. The final results show that K-CV-SVM model has a good overall performance both in classification and regression, and it is an effective method for classification and prediction.

Parameter optimization is of great significance to ensure the accuracy of the prediction model. At the same time, the convergence speed of the algorithm and the running time of the program are also a factor that cannot be ignored, especially in the case of huge samples. The results in Table II show that the K-CV-SVM model has the highest classification accuracy, but the running time of system is also the most. Therefore, the next step is to study how to ensure the accuracy rate at the same time to speed up the convergence rate and reduce the time consumption.

References

[1] Liu J , Hu Y , Yang S . A SVM-Based Framework for Fault Detection in High-Speed Trains[J]. Measurement, 2020, 172:108779.

[2] Zhang H . Intelligent Detection of Small Faults Using a Support Vector Machine[J]. Energies, 2021, 14.

[3] Peng Q , Lin X , Shi H , et al. An SVM Classification based Signal Detection Method in UHF RFID Systems[J]. IEEE Transactions on Industrial Informatics, 2020, (99):1-1.

[4] Bajard J C , Martins P , Sousa L , et al. Improving the Efficiency of SVM Classification With FHE[J]. IEEE transactions on information forensics and security, 2020, 15:1709-1722.

[5] Wang W , Zhang B , Wang Z Q , et al. Garbage image recognition and classification based on hog feature and SVM-Boosting[J]. Journal of Physics Conference Series, 2021, 1966(1):012002.

[6] Kowalczyk A , Raskutti B . One class SVM for yeast regulation prediction[J]. ACM SIGKDD Explorations Newsletter, 2002, 4(2):99-100.

[7] Balachandran M , Shin T H , Gwang L . PVP-SVM: Sequence-Based Prediction of Phage Virion Proteins Using a Support Vector Machine[J]. Frontiers in Microbiology, 2018, 9:476.

[8] Mohamed, Goudjil, Mouloud, et al. A Novel Active Learning Method Using SVM for Text Classification[J]. International Journal of Automation and Computing, 2018, v.15(03):44-52.

[9] Rajeshwari R , Mandal S , Rajasekaran. Compressive strength prediction of SCC containing fly ash using SVM and PSO-SVM models[J]. Journal of Structural Engineering (Madras), 2021, 48(1):1-11.

[10] Yang P , Li Z , Yu Y , et al. Studies on fault diagnosis of dissolved oxygen sensor based on GA-SVM[J]. Mathematical Biosciences and Engineering, 2021, 18(1):386-399.

[11] Faris, H., Hassonah, M.A., Al-Zoubi, A.M., Mirjalili, S. and Aljarah, I. A multi-verse optimizer approach for feature selection and optimizing SVM parameters based on a robust system architecture. Neural Computing & Applications. 2017. pp.1-15.

[12] John R G , Ramachandran K I , Valavan K K , et al. Detection of Obstructive Sleep Apnea from ECG Signal using SVM based Grid Search[J]. International Journal of Electronics and Telecommunications, 2021, 67(1):5-12.

[13] Xia J , Wang Z , Yang D , et al. Performance optimization of support vector machine with oppositional grasshopper optimization for acute appendicitis diagnosis[J]. Computers in Biology and Medicine, 2022, 143:105206-.

[14] Liang B , Mao Z , Zhang K , et al. Analysis and Optimal Design of a WPT Coupler for Underwater Vehicles Using Non-Dominated Sorting Genetic Algorithm[J]. Applied Sciences, 2022, 12(4):2015.

[15] Wu Y , Li X , Liu Q , et al. The Analysis of Credit Risks in Agricultural Supply Chain Finance Assessment Model Based on Genetic Algorithm and Backpropagation Neural Network[J]. Computational Economics, 2021:1-24.

[16] Liu G , Zhu H . Displacement Estimation of Six-Pole Hybrid Magnetic Bearing Using Modified Particle Swarm Optimization Support Vector Machine[J]. Energies, 2022, 15.

[17] Alajmi M S , Almeshal A M . Least Squares Boosting Ensemble and Quantum-Behaved Particle Swarm Optimization for Predicting the Surface Roughness in Face Milling Process of Aluminum Material[J]. Applied Sciences, 2021, 11(5):2126.

[18] Rhoads T , Neale A C , Resch Z J , et al. Psychometric implications of failure on one performance validity test: a cross-validation study to inform criterion group definition[J]. Journal of Clinical and Experimental Neuropsychology, 2021(6):1-12.

[19] Sheppard L . Deployment, Calibration, and Cross-Validation of Low-Cost Electrochemical Sensors for Carbon Monoxide, Nitrogen Oxides, and Ozone for an Epidemiological Study[J]. Sensors, 2021, 21.