Hindawi
Mathematical Problems in Engineering
Volume 2022, Article ID 7132900, 11 pages
https://doi.org/10.1155/2022/7132900

Research Article

Evaluation Model of College English Teaching Effect Based on Particle Swarm Algorithm and Support Vector Machine

Chunyan Wei [1] **and Sang-Bing Tsai** [2]

[1] *School of Foreign Languages, Xuchang University, Xuchang, Henan 461000, China*
[2] *Regional Green Economy Development Research Center, School of Business, Wuyi University, Nanping, China*

Correspondence should be addressed to Chunyan Wei; villa22@163.com

Received 13 January 2022; Revised 14 February 2022; Accepted 25 March 2022; Published 27 April 2022

Academic Editor: Gengxin Sun

Based on the principle of particle swarm algorithm and support vector machine, this article aims to improve the classification performance of college English teaching effect and explores the best support vector machine parameter optimization algorithm to promote college English teaching for the theory and application research of data analysis. First, the advantages and disadvantages of common support vector machine parameter selection methods such as grid search algorithm, gradient descent method, and swarm intelligence algorithm are studied. Secondly, this article has a detailed analysis and comparison of various other algorithms. Finally, the study analyzed the advantages and disadvantages of the quantum particle swarm algorithm, introduced the dual-center idea into the quantum particle swarm algorithm, and proposed an improved quantum particle swarm algorithm. Through simulation experiments, it is proved that the improved quantum particle swarm algorithm is more superior in optimizing the parameters of support vector machine. In general, this paper uses the PSO algorithm to simultaneously solve the SVM feature selection and parameter optimization problems and has achieved good results. Within the scope of the literature that the author has, there is still a lack of work in this area. Compared with the existing algorithms, the algorithm proposed in this paper has stronger feature selection ability and higher efficiency.

1. Introduction

The theoretical basis of the support vector machine is Vapnick's structural risk minimization principle and VC dimension theory. It is difficult to find the optimal classification surface in the low-dimensional space, so the input space is mapped to the high-dimensional space through the transformation of the inner product function, so that the high-dimensional space becomes linearly separable to find the optimal classification surface [1]. It has great advantages in solving nonlinear, small-sample, and high-dimensional problems and has good generalization and generalization capabilities. It is widely used in pattern classification, regression analysis, and probability density function estimation. Classification is an important data mining and analysis method for college English teaching.

College English teaching data mining, which generally refers to searching for the hidden college English teaching data from a large amount of vague, irregular, and incomplete college English teaching data through certain algorithms. In the process of information and knowledge, there are potential uses [2–5]. As a decision support technology, college English teaching data mining mainly includes college English teaching data classification, regression analysis, college English teaching data clustering, feature changes, deviation analysis, college English teaching data visualization, association rule mining, and regression algorithms, etc. Algorithms belong to the category of particle swarm algorithm. Through training and learning, the machine is made intelligent, and then the existing college English teaching data is analyzed from different perspectives through the machine, and the rules are discovered [6–8].

In view of the importance of feature selection in the SVM classification problem, the discrete PSO algorithm is more suitable to deal with the combination optimization problem of feature selection than the continuous PSO algorithm; this article is in CPs. Based on the svM algorithm, a feature selection and SVM parameter synchronization optimization algorithm based on the discrete PSO algorithm (DPSO-SVI) is proposed to improve the feature selection ability of the CPSO-SVM algorithm. Aiming at the problem that it is easy for the particle swarm optimization algorithm to fall into the local optimal solution, this paper proposes an improved particle swarm optimization algorithm and applies the improved particle swarm optimization algorithm to the selective integration of classification SVM and proposes an improved particle swarm optimization algorithm. Using UCI College English Teaching Database to test APSOSEN, the experimental results show that this method can effectively solve the local optimal problem of PSOSEN, and it has improved accuracy, convergence, and integration scale for a more efficient implementation method of selective integration.

2. Related Work

Support vector machine has the advantages of simple structure, complete theory, strong adaptability, global optimization, short training time, and good generalization performance [9]. It has received extensive attention in the field of particle swarm optimization and has become the current international and domestic research hotspot. The two main parameters that affect the accuracy of support vector machine classification are the penalty coefficient C and the parameter of the kernel function. Previously, experimental and empirical methods were mainly used to select the parameters of the support vector machine, with very large, extremely low work efficiency, and more importantly, the results obtained are often not the global optimal parameters. The classification problem in support vector machines is to construct a support vector classifier [10, 11].

Xie et al. [12] pointed out that, in the PAC learning model, if there is a polynomial-level learning algorithm to identify a set of concepts, and the identification accuracy is high, then this set of concepts is strongly learnable; and if the learning algorithm identifies a set of concepts, the accuracy is only slightly better than random guessing, so this set of concepts is weakly learnable. If the two are equivalent, then when learning concepts, you only need to find a weak learning algorithm that is slightly better than random guessing, and it can be promoted to a strong learning algorithm, instead of directly looking for a strong learning that is difficult to obtain under normal circumstances algorithm. Rajamohana and Umamaheswari [13] used a constructive method to prove that any weak learning algorithm can be effectively transformed into a strong learning algorithm, and the proof process is boosting. Later Li et al. [14] proposed a more effective "boost-by-majority" algorithm. Both algorithms call the given weak learning algorithm multiple times, each time it is provided with a different distribution, and finally all the proposed hypotheses are merged into a single hypothesis. This intuitive idea is distributed in a way to increase the possibility of learning the "difficult to learn" parts, forcing learners to make new assumptions and make fewer mistakes in these parts.

But these two algorithms have a major flaw in solving practical problems; that is, they must know the lower limit of the learning accuracy of the weak learning algorithm in advance, which is difficult to achieve in practice. So Saputra et al. [15] proposed an algorithm. The accuracy of the final integrated hypothesis generated by the algorithm is based on the accuracy of the hypotheses generated by all weak learning algorithms, so the potential of weak learning algorithms can be more fully tapped. Moreover, because the AdaBoost algorithm solves the problem of "need to know the lower bound of the generalization ability of the weak learning algorithm in advance," it has the advantages of operability and simplicity. Juan and Hong Wei [16] proposed a technique similar to boosting. The researcher emphasized that the stability of the learning algorithm in the ensemble has a great influence on the final result. For unstable algorithms, such as neural networks and decision trees, the accuracy of prediction can be improved. However, the effect of stable learning algorithms is not obvious and sometimes even reduces the prediction accuracy [17–19]. The researchers used the simple average method to integrate the BP network with different numbers of hidden layer neurons and used it to replace the Gauss classifier in JARTOOL developed by NASA's Jet Propulsion Laboratory. The image is analyzed and it has reached the level of planetary geologists in volcano detection [20]. Researchers use boosting for text classification. They found through experiments that, on this issue, the integration effect of boosting is always better than or equivalent to Sleeping-experts, Rocchio, Naive-Bayes, and other commonly used technologies. In addition, integrated learning technology has also been successfully applied in many fields such as speech recognition, text filtering, and remote sensing information processing disease diagnosis [21, 22].

3. Algorithm Parameter Optimization

3.1. Cross Validation. The statistical learning theory provides a theoretical framework for the particle swarm algorithm to find the rules of the particle swarm algorithm when the number is small. The goal of the particle swarm algorithm is to find the relationship from the given college English teaching data in order to more accurately deal with the unknown results which are predicted. That is, it is necessary to find an optimal function in a function group based on n independent and identically distributed sample sequences to obtain the minimum empirical risk value [23, 24].

$$X \times \frac{\partial A[t]}{\partial C(c,t)} = \{c, t \in R | x(\alpha,1), x(\alpha,2), \ldots, x(\alpha,t)\}. \quad (1)$$

The VC dimension is an important concept in statistical learning. It is not only an important indicator for evaluating function learning ability, but also a measure of the complexity of learning machines. It can be defined as follows: If a set of functions can divide h samples into two categories according to all possible forms, it means that this function set can break up h samples, and the VC dimension of the function set uses this function set, the number of the largest sample set that can be broken, that is, the number of samples h. Table 1 shows the number of sample sets [25].

It can be seen from the above that, in the case of a small number of samples, the empirical risk and the actual risk are not equal, and it is unreasonable for the empirical risk to replace the actual risk. With the further in-depth research, researchers defined the sum of empirical risk and confidence range as structural risk and proposed the principle of structural risk minimization (SRM) as a condition for selecting predictive functions.

$$\frac{\nabla^2 \alpha}{\nabla^2 x - \nabla^2 y} - \frac{1}{a^2} \times \frac{\partial^2 \alpha}{\partial t^2} + \frac{\partial \rho(x,t)}{\partial \varepsilon} = 0. \quad (2)$$

It can be seen that, in the case of limited samples, the smallest empirical risk does not necessarily mean the smallest expected risk; the complexity of the learning machine not only is related to the system under study, but also needs to be adapted to the limited number of samples. Therefore, we need a theory that can guide us to learn and promote methods that are still effective in the case of small samples. This is the theory of statistics.

$$y(\alpha,\beta,n) = \int A \times X(n)x - \int t \times x(n-1)\mathrm{d}x. \quad (3)$$

It is a theory for the study of small-sample statistics and prediction. The core content includes statistical learning consistency conditions based on empirical risk minimization criteria; statistical learning method promotion-type circles; small-sample inductive reasoning rules established on the basis of the promotion circles to implement the new standards, etc.

3.2. Support Vector Machine Parameter Fitting. Because of the high application value of support vector machines, VC dimension theory, risk minimization theory, optimal classification hyperplane, kernel function, etc. are all theoretical research categories of support vector machines. The choice of penalty parameters and kernel function parameters will directly affect the performance of the support vector machine. Therefore, parameter selection is an important solution direction in terms of the performance optimization of the support vector machine algorithm. The penalty parameter is a compromise between the wrong sample and the classification interval in the determined college English teaching data space to ensure that the support vector machine not only satisfies the classification performance but also has a good promotion ability.

$$h(a,b,x) = \begin{cases} \sum \delta(i,n) \times x(t), & t > 1, \\ \sum \varepsilon(i,n) \times x(t), & 0 < t < 1. \end{cases} \quad (4)$$

The greater the value of the penalty parameter, the greater the punishment for experience errors, which reduces the risk of experience, which is called the "overlearning" situation. Similarly, the smaller the value of the penalty parameter, the smaller the penalty intensity, and the simpler the model supporting the vector machine, but the corresponding empirical risk also increases, which is called the "underlearning" situation.

$$\Omega[f(x+1) - f(x)] - \partial \frac{y(x,n)}{\|x(n)\|^2} * x(n) = 0. \quad (5)$$

The introduction of kernel functions is the reason why support vector machines can be widely used in linear inseparable problems. With limited training samples, the confidence range is proportional to the VC dimension. The higher the VC dimension, the larger the confidence range, which leads to the greater possible difference between the real risk and the experience risk, which leads to the phenomenon of overlearning.

$$\frac{\overline{x^2}(t,n) - |x(t,n)|^2 - \Delta x(t,n)}{\prod x(t,n) \times f(x)} - 1 = 0. \quad (6)$$

Support Vector Classifier (SVC) is proposed from the optimal classification hyperplane when two types of samples are linearly separable. The so-called optimal classification hyperplane requires the classification hyperplane not only to separate the two types of samples without error, but also to maximize the classification interval.

$$\oiiint_{A-X=t(c)} 2|\overline{x(t,n)}| * \cos \alpha \mathrm{d}x\mathrm{d}y\mathrm{d}z - \sqrt{\theta(t,z) + \theta(t,x)} = 0. \quad (7)$$

After the support vector machine classifier was proposed, its performance has been verified in the application of many practical problems, but the traditional SVC algorithm has some computational problems, including the slow training algorithm, the complexity of the algorithm, and the large amount of computation in the detection phase. The feature space is determined with the determination of the kernel function. However, the method based on the ant colony has relatively little information obtained due to the uncertainty of the initial test information, resulting in a very slow speed in the solution process. Assume that the standard voting principle is a majority decision. Based on the consensus voting method, the support set is updated during each iteration, and the "correct" elements of the support set can be found in some estimated support sets.

The two are corresponding. If the dimensionality of the feature space is high, the optimal hyperplane obtained may be very complicated; otherwise, it may be very simple. Whether the dimensionality is too high or too low will make the generalization ability of the support vector machine model worse. Therefore, the choice of penalty parameter and kernel parameter has a great influence on the promotion ability of support vector machine.

Table 1: Description of the number of sample sets.

Sample sets index	Function description	Weight indicator (%)
H-1	The empirical risk and the actual risk	11
H-2	The complexity of the learning machine	13
H-3	Confidence range as structural risk	24
H-4	The sum of empirical risk and confidence range	9
H-5	The principle of structural risk minimization	56
H-6	Statistical learning method promotion-type circles	21

3.3. Algorithm Network Search. The particle swarm algorithm takes the potential well as the basis and considers that each particle has quantum behavior. In classical mechanics, the two values of position and velocity can be used to express the state of a particle, and these states determine the trajectory of the particle when it moves.

$$\varphi(x, y, z, t) - \sum_{\cos\alpha=1}^{\cos\alpha=1} \frac{\beta(x, t=0) - \beta(x-1, t=1)}{\sum \alpha(x, y) = 1} = 1. \quad (8)$$

In Newton mechanics, particles move in a given trajectory at all times, but unlike classical mechanics, the quantum behavior of particles in quantum mechanics makes the trajectory of motion random and cannot be given. In the quantum world, due to the uncertainty principle, particles have no definite trajectory, and there is no way to determine the speed and position at the same time. Compared with particle swarm optimization, quantum particle swarm optimization requires fewer parameters and is easier to implement. When actually solving optimization problems, the performance of quantum particle swarm optimization has been proved to be better. Table 2 shows the adjustment parameters of the particle swarm algorithm.

This is a quadratic function optimization problem under inequality constraints, and there is a unique solution. It is easy to prove that only the part of the solution corresponding to the support vector is not zero. The optimal classification function obtained after solving the above problem is shown in the formula. The summation in the formula is actually only performed on the support vector. Mine is the classification threshold, which can be obtained by any support vector satisfying the equal sign in the formula or by taking the median of any pair of support vectors in the two categories.

$$\begin{cases} \dfrac{\partial^2 \alpha}{\partial x^2} + \dfrac{\partial^2 \alpha}{\partial y^2} + \dfrac{\partial^2 \alpha}{\partial z^2} - f(x, t) = 0, \\ \\ \dfrac{\partial^2 \beta}{\partial x^2} + \dfrac{\partial^2 \beta}{\partial y^2} + \dfrac{\partial^2 \beta}{\partial z^2} - f(x, t) = 0. \end{cases} \quad (9)$$

Therefore, in the process of particle swarm optimization, the empirical risk and VC dimension should be minimized in order to narrow the confidence range, so as to obtain a smaller actual risk and improve the sample promotion ability. It is assumed that, in the two-node support set estimation, the one with high probability of indexing the father is the common support set J.

But on the other hand, if the VC dimension of the function set is smaller, it is difficult to approximate the college English teaching data of the training sample, so this is a contradiction. When constructing the learning machine, two strategies can be adopted: the strategy adopted by the neural network and the strategy adopted by the support vector machine.

4. Construction of Teaching Effect Evaluation Model

4.1. Algorithm Space Iteration. For nonlinear problems, SVM first transforms the nonlinear problem in the original feature space into a linear problem in a high-dimensional space through nonlinear transformation and then finds the optimal classification surface in the high-dimensional space. Under normal circumstances, this transformation may be complicated and not easy to implement.

$$\frac{\alpha(x) - \alpha(x-1)}{\sum a(t) \times w(x)} - \frac{b(t) \times x - 1}{\sum w(x)} = 0. \quad (10)$$

However, as mentioned above, after converting the problem of finding the optimal classification surface into its dual problem, it can be found that both the optimization function and the classification function only involve the inner product operation between the training samples. In essence, the particle swarm algorithm is to summarize and refine an abstract cognitive model from a collection of real examples.

The model can be represented by a function from a mathematical point of view, such as fx. The instance set is usually the training set. The learning algorithm obtains the learner h from the training set. In fact, h is an approximation of the function $y = fx$. After getting the learner, given a new instance x, the learner outputs the corresponding result y. The output result y is the category of x for classification and the predicted value of x for regression.

$$\frac{\iint [\beta(x) - \beta(x-1)x] \mathrm{d}x \mathrm{d}y}{\Omega(x, y, t)} - 1 = 0. \quad (11)$$

Therefore, it is not necessary to calculate the image of the sample points in the original space in the high-dimensional space; that is to say, you do not need to know the form of the transformation, but only a function in the original space, which can be calculated based on the sample points in the original space.

The inner product of the image in the high-dimensional space is sufficient. This kind of function is called a kernel

Table 2: Particle swarm algorithm adjustment parameters.

	Parameter weight 1 (%)	Parameter weight 2 (%)	Parameter weight 3 (%)	Parameter weight 4 (%)	Parameter weight 5 (%)
Particle number 1	5	0	29	52	41
Particle number 2	43	2	4	36	0
Particle number 3	30	2	7	50	23
Particle number 4	3	0	48	47	55
Particle number 5	17	1	38	6	19

function. According to the related theory of functionals, as long as a kernel function $K(x_1,x-n)$ satisfies the Mercer condition, it corresponds to the inner product in a certain transformation space.

$$\frac{G(x(t-1\in\mu),k=1)}{\cap\overline{\overline{z(x-1)\times x(t)}}} = \frac{1}{1-\alpha(x)}\times\sum_{i=1}^{n}(1-\alpha(x))(1-\beta(x)). \tag{12}$$

The SRM principle has the following two aspects: first, we find the minimum empirical risk in each subset and then calculate the sum of the empirical risk R and the confidence range h corresponding to each corresponding minimum risk, and the minimum structural risk is the value of the minimum sum.

$$\begin{cases}\alpha(a,t)=(i,j,k)=i(1),i(2),\\ \beta(a,t)=j(1),j(2),\\ \varepsilon(a,t)=k(1),k(2),i,j,k\in Z.\end{cases} \tag{13}$$

The second is to calculate the minimum empirical risk value of each subset and then pass the appropriate subset to minimize the confidence range. At this time, the minimum structural risk is obtained. The first method is not feasible when the number of subsets is large; the second method has strong operability and at the same time lays the theoretical foundation of support vector machines.

4.2. Support Vector Machine Data Clustering. Feature selection for college English teaching can include two aspects: feature extraction and feature screening; feature extraction in a broad sense refers to a transformation, which converts samples in a high-dimensional space to a low-dimensional space through mapping or transformation to achieve dimensionality reduction. Feature screening refers to removing redundant or irrelevant features from a set of features to reduce dimensionality.

The two are often used in combination, such as mapping the high-dimensional feature space to the low-dimensional feature space through transformation and then removing redundant and irrelevant features to further reduce the dimensionality. Figure 1 shows the clustering distribution of support vector machine data.

The goal of the CPSO-S algorithm is to simultaneously optimize feature subsets and SVM parameters to improve the classification accuracy of SVM, while reducing the number of selected features as much as possible. Classification problem is the basic research problem of ensemble learning, which belongs to the category of concept learning. The classification problem is to classify a series of examples according to certain rules. In fact, it is to find a certain function $y=fx$, so that, for a given example x, the correct classification y can be obtained.

$$\sum_i|\langle\alpha(1),u(i)\rangle|^2+\sum_j|\langle\alpha(2),u(j)\rangle|^2=\|u(i,j)\|^2. \tag{14}$$

The solution idea in the particle swarm algorithm is to find a sufficiently good function in the hypothesis space to approximate it through a certain learning method. This approximate function is called a classifier. As the sampling rate increases, ASCE gradually decreases. When the sampling rate reaches a certain value, ASCE tends to 0 and finally equals 0; that is, there is no support set error.

Classification accuracy and the number of selected features are two criteria for designing fitness functions. A particle can make the classification accuracy produced by the classifier higher, and the fewer the number of features selected at the same time, the higher its fitness should be.

$$\left\{\prod_{\sigma(x,y)}z(x,y)-u(x,y)=\overline{u}(x,y)\prod_{\sigma(x,y)}z(x,y)-v(x,y)=\overline{v}(x,y),x,y\in S(u).\right. \tag{15}$$

There are two main reasons for the representation: one is that the assumed space is too small, and the other is that the assumed space is not closed. Since the hypothesis space is artificially specified, if the hypothesis space is too small, the actual target hypothesis in the application of particle swarm algorithm may not be in the hypothesis space. If the optimal value obtained by improper selection may be a local optimal value, there is a certain randomness.

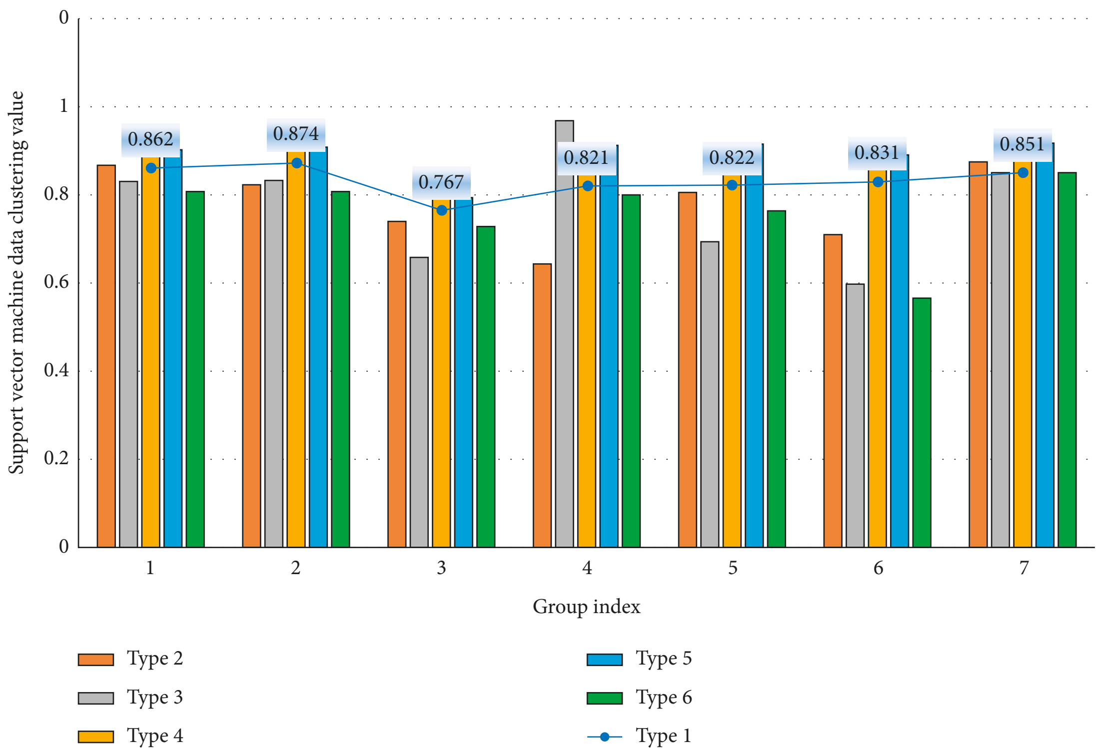

FIGURE 1: Data clustering distribution of support vector machine.

If the hypothesis space is not closed under a certain integrated operation, then by integrating a series of hypotheses in the hypothesis space, it is possible to express the target hypothesis that is not in the hypothesis space, causing judgment errors. This risk can be reduced by integrating multiple assumptions.

4.3. Periodic Analysis of Particle Swarm Optimization. The performance is evaluated, and the classification accuracy of the model is tested. In this dynamic change equation of particle state, the vector $y(f)$ is the state of the particle at time t, which is composed of its position and velocity at time t. It is the coefficient of the dynamic change equation. Its characteristics determine the dynamic behavior of the particle. P is the external input driving the particle to fly. At a specific location, the input matrix exerts an influence on the state of the particles through external input.

$$\frac{\partial u(\theta(t)-\theta(t-1))}{\partial(C=(\lambda-1)/(\lambda+1))}-\frac{\partial x\,\partial y}{\partial x(t)}=0. \tag{16}$$

Repeating the process described above K times is to ensure that each subset has a chance to be tested, and it is necessary to make sure that each subset reserved for testing is not repeated. The expected generalization error is estimated by the average of the test accuracy obtained after K predictions, and a set of optimal parameters is selected to build the model.

For example, if the number of groups K is 5, then the training sample needs to be divided into 5 equal parts, the last 4 parts are used as the training set, and the first part is used as the test set to train the classifier. The second part is used as the training set, and the second part is used as the test. By analogy, the standard for measuring the quality of the parameters is the average fitness value obtained in each test.

4.4. Evaluation Parameters of Teaching Effect. Teaching effect evaluation adopts cross validation technology to select model parameters to be combined with grid search method. In this paper, cross validation technology is used to optimize support vector machine parameters. The main steps are as follows: first set the model parameters (C, t) in $\lg C$. The choice of initial values for gradient descent-based methods is important and can have a large impact.

Within the range of $\in[-10,10]$, $\lg t \in [-10, 10]$, perform a grid search on the parameters c and mouth with a step length of 1 and then use the tenfold cross validation method to calculate the mean square error for the selected parameters (MSE). Finally, the selection of parameters is determined according to the global minimum of MSE as the criterion to evaluate the influence of the selected parameters on the generalization ability of the model. Figure 2 is the fitting distribution of the teaching effect evaluation parameters.

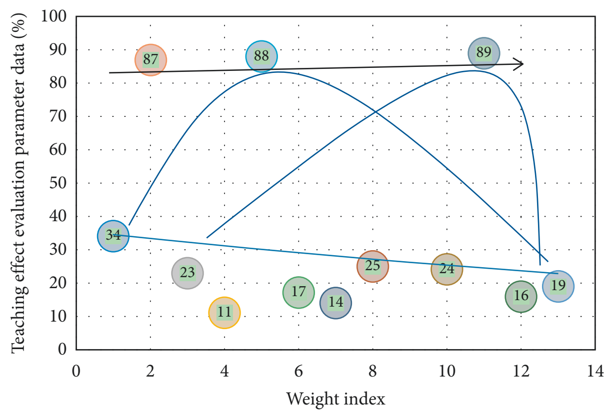

Figure 2: Fitting distribution of teaching effect evaluation parameters.

The number of support vectors and boundary support vectors of the two algorithms are different. For example, in the C-SVC algorithm, the support vector of the second classifier SVM2 is 10, the boundary support vector is 5, and the third classifier SVM support vector is 11, the boundary support vector is 5, the support vector of the fourth classifier SVM4 is 13, and the boundary support vector is 5; and in the y-SVC algorithm, the second classifier SVM2 support number of vectors is 22, the boundary support vector is 18, the support vector of the third classifier SVM3 is 12, the boundary support vector is 6, and the support vector of the fourth classifier SVM4 is 16, and the boundary support vector is 7. Comparing the simulation results of the two algorithms, the number of support vectors and boundary support vectors of the V-S ratio algorithm is more than that of the C-SVC algorithm.

5. Application and Analysis of Evaluation Model

5.1. Preprocessing of College English Teaching Data. The accuracy of particle swarm algorithm is higher than genetic algorithm, and the time is shorter than genetic algorithm. Particle swarm algorithm is also used a lot in optimization. During the operation of the algorithm, the operation of particle swarm algorithm is relatively simple and the space search speed is fast. The search efficiency is also high, but similar to the genetic algorithm; when the problem to be solved is more complicated, the particle swarm algorithm is also prone to premature phenomenon in the later stage of the algorithm; that is, the optimal parameter solution obtained is the local optimum. The solution is not the global optimal solution.

$$\begin{cases} \dfrac{\int [(\partial u/\partial x)\mathrm{d}x - (\partial v/\partial x)\mathrm{d}x]}{\varepsilon((\partial u/\partial x)\mathrm{d}u\mathrm{d}v)} = 0, \\ \bigcup\left[\dfrac{\partial u}{\partial x}\mathrm{d}x + \dfrac{\partial v}{\partial x}\mathrm{d}x\right] \subseteq \mathbb{N}. \end{cases} \tag{17}$$

The accuracy rate of quantum particle swarm algorithm is as high as that of particle swarm algorithm. Although the time is shorter than genetic algorithm, it is longer than particle swarm algorithm. Compared with the particle swarm algorithm, although the quantum particle swarm algorithm has improved global convergence, it is still easy to fall into local convergence, which occurs in the later stage of the algorithm.

The particle swarm algorithm of quantum behavior needs to consider few parameters and is easier to write. And because of the introduction of the average best position, the ability of the particles to work together is enhanced, and there is more room for improvement. Figure 3 is the data extraction and analysis of college English teaching.

The sampling rate of SamTrad directly affects the accuracy of SVM's classification of college English teaching data. The larger the SamTrad, the more the samples extracted; the SVM can get more sample information, and the probability of classification errors will be reduced a lot, but the training time of SVM will increase.

In the experiment, we set the number of learners to 20, and SamTrad takes (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 08, 0.9, 1.0) in turn. The gamma value of dna is 0.032, and C is 2. These two parameters are obtained in advance using an exhaustive method, which can achieve better results on rbf-SVM. Each algorithm has been tested 50 times on the college English teaching data set.

$$\oint\left(\frac{\partial u}{\partial x}\mathrm{d}x + \frac{\partial v}{\partial y}\mathrm{d}y + \frac{\partial w}{\partial z}\mathrm{d}z\right) = \begin{cases} 1, & \text{if } \dfrac{\partial u}{\partial x} > 0, \\ 0, & \text{if } \dfrac{\partial u}{\partial x} \leq 0. \end{cases} \tag{18}$$

The classification accuracy of the SVM algorithm has been greatly improved compared with the SVI algorithm. On each experimental college English teaching data set, the classification accuracy of the former is higher than that of the latter. For example, for the second-category college English teaching data set disease, the positive hit rate, negative hit

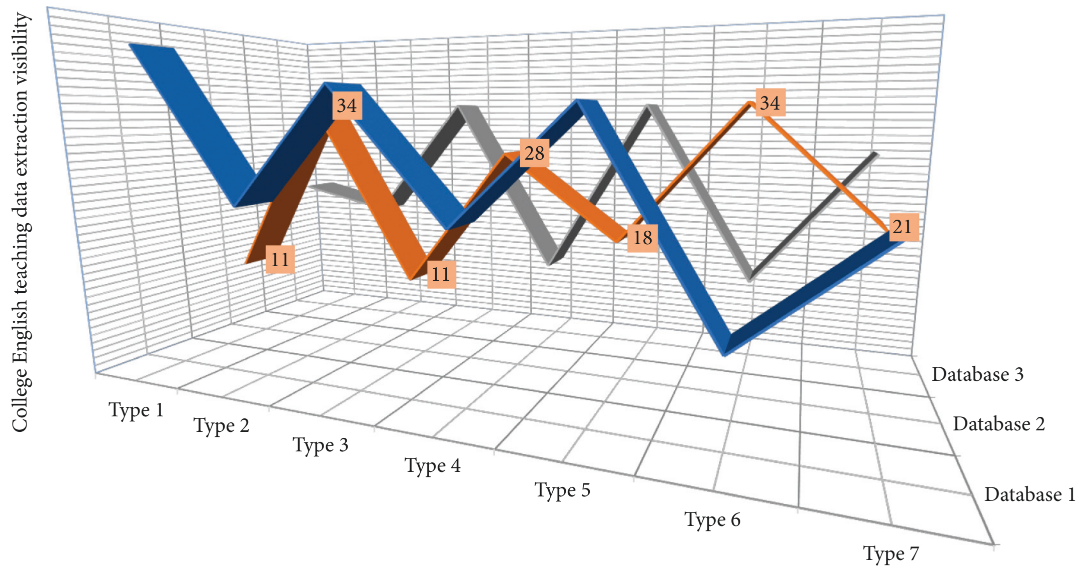

FIGURE 3: Extraction and analysis of college English teaching data.

rate, and overall hit rate of the SVM algorithm are only 0.8000 and 0.5, respectively, 8067 and 0.8037.

5.2. Teaching Model Simulation. The experiment uses cross validation method for evaluation. The college English teaching data set is randomly divided into k subsets, CPS0. The SVM algorithm is run k times on each college English teaching data set, each time a subset is taken as the test set, and the rest k. One subset is merged into the training set, and then the average of the results of k experiments is taken as the classification result of the college English teaching data set. In the experiment of this chapter, $k = 10$.

In the process of using AdaBoost to generate the subclassifier t-SVM, t-SVM will test the training set S and generate misclassified samples and samples that have not been misclassified and update the weight values of the samples according to this difference.

$$\begin{cases} \phi(i,1) = \coprod \dfrac{(\phi^{+}(i) + \phi^{-}(i))}{2} - \lim\limits_{x\longrightarrow\infty} \phi^{+}\dfrac{(i)}{2}, \\ \phi(i,2) = \coprod \dfrac{(\phi^{+}(i) - \phi^{-}(i))}{2} - \lim\limits_{x\longrightarrow\infty} \phi^{-}\dfrac{(i)}{2}. \end{cases} \tag{19}$$

In the training phase, the SVM algorithm is the fastest, because it does not require the use of optimization algorithms (PSO or GA) for feature selection and parameter optimization like the other two algorithms. However, the lack of feature selection in the training phase will reduce the efficiency of SVM in the use phase, and the large number of features will slow down the operation of SVI.

The use phase of a classifier is longer than the training phase, so in the long run, the efficiency of the SVM algorithm is the worst. The classifiers obtained by the SVM algorithm have better balance, which means that, for different classes, the hit rate is not much different. This is most obvious in Sonar, the second-class college English teaching data set.

5.3. Analysis of Teaching Effect Evaluation System. Since the number of rows in samples represents the number of features of the sample, it is necessary to know the number of features of the college English teaching data before the conversion of the college English teaching data. The process of converting the college English teaching data format of libSVM to the OSU-SVM college English teaching data format is shown below.

Step 1: Read in the S.txt file;

Step 2: Initialize two matrices Samples and Labels. The number of rows of Samples is FNum, the number of columns is R in S.txt, the number of Labels is 1, and the number of columns is R.

Step 3: Read the tth row of S.txt and put it into the temporary variable row_string.

Step 4: Find the first space position m in row_string, and store the number $[1, m-1]$ in the tth column of Labels.

Through the conversion of the above steps, the college English teaching data format of libSVM is converted to the OSU-SVM college English teaching data format, so that OSU-SVM can use the college English teaching data in libSVM. Figure 4 shows the weight simulation of updated samples at different points.

It can be seen from the result graph and result table of the simulation experiment that the simulation experiment is carried out on the SPECT Data Set college English teaching data set. When the cross-validated K value is 3, the correct

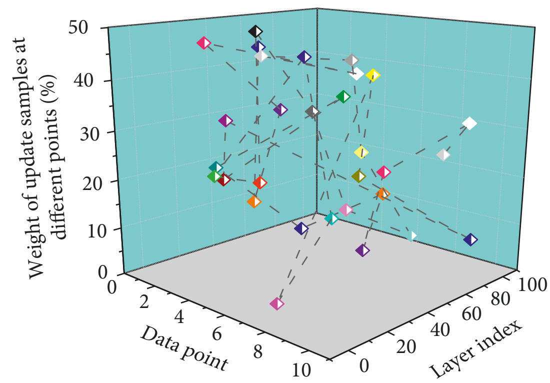

Figure 4: Weight simulation of updated samples at different points.

rate is 70.00%. The optimal parameter pair values are $C = 53.8961$ and $g = 0.0491$; when the value is 7, the classification accuracy rate is 72.50%, and the optimal parameter pair values are $C = 2.7040$ and $g = 0.0193$, respectively. A value of 5 has the highest classification accuracy rate, at this time the accuracy rate is 73.75%, and the corresponding optimal parameter pair is $C = 14.6752$, $g = 0.0239$. We present the results of applying DIAT to a 0-1 sparse signal with SMNR = 20 dB.

It can be seen from the simulation experiment that the best fitness value of the quantum particle swarm optimization algorithm has remained unchanged after the simulation experiment is carried out on the SPECT Heart Data Set college English teaching data set. The average fitness value rises from fast to slow and finally maintains a steady state. When the cross validation K value is 3, the correct rate is 77.50%. At this time, the optimal parameter pair values are $C = 49.5626$ and $g = 0.7550$.

5.4. Case Application and Analysis. In this paper, rbfSVM in OSU-SVM is used as the base classifier in the ensemble, and 20 SVM classifiers are trained using the Boostrap method on the Matlab platform. Then these 20 SVM classifiers use the selective ensemble methods PSOSEN and APSOSEN proposed in this paper to optimize the selection and select the optimal ensemble model. The validation set is randomly selected from the training set to select the optimal ensemble model. Finally, the selected integration model is tested on the test set. The experimental parameters are set as follows: population size $N = 40$, learning factor $c = 1.4962$, and maximum number of iterations DT = 100.

$$\sigma(x, y) = \frac{1}{E}\{\varepsilon(x, x-1) - t[\varepsilon(y-1, y)]\}, \quad \exists x, y, \in C(r = 0, t = 1). \tag{20}$$

When the number of classifiers is only 10, the improvement of Bagging_SVM to a single SVM is only about 4%. But when the number of classifiers reached 45, the improvement of Bagging_SVM to a single SVM reached about 17%. AdaBoost_SVM is also the same situation.

In the case of a relatively small number of classifiers, the improvement effect after integration is not as good as the case of a large number of classifiers. However, the increase in the number of classifiers comes at the cost of increased computing time. How to choose the appropriate number of classifiers also depends on the specific situation. Figure 5 shows the quantitative distribution of the classifier integration effect.

When the dimension is two, the penalty factor C is l, and the kernel function parameter d is 0.25. The parameter of the insensitive loss function is 0.000977, the number of support vectors is 10, the boundary support vector is 6, and the optimization error is 0.20596; when the dimension is three, the penalty factor C is 128. The kernel function parameter d is 0.00391, the parameter of the insensitive loss function accounts for 0.000977, the number of support vectors is 9, the boundary support vectors are 4, and the error is 0.147.

$$\delta(x, y) = \frac{1}{E} * \mu((x, x-1) - t(\mu(y-1, y))), \quad \exists x, y, \in C(r = 0, t = 1). \tag{21}$$

Since the search process of the particle swarm is a nonlinear and complex process, the linear adjustment method of the inertia weight 03 cannot correctly reflect the search process of the particle swarm. Therefore, the inertia weight needs to be adjusted nonlinearly. The general adjustment method is to use fuzzy rules for dynamic adjustment. This method formulates the corresponding membership function and fuzzy inference rules for the

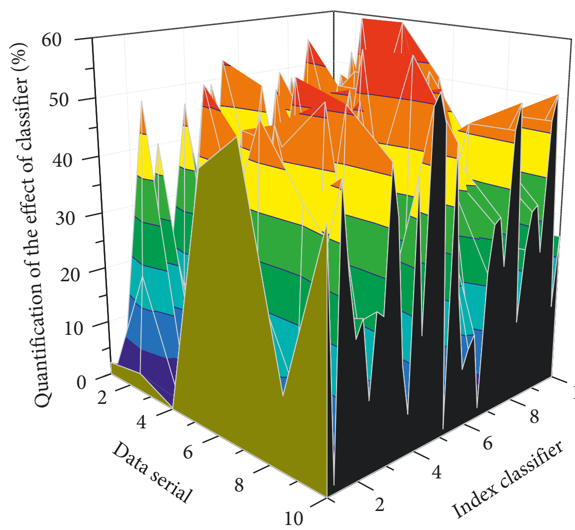

Figure 5: Quantitative distribution of classifier integration effect.

current best performance evaluation (CBPE) and the current inertia weight.

6. Conclusion

Aiming at the high complexity of selective integration, learning from the swarm intelligence method, this paper proposes a selective integration algorithm PSOSEN based on particle swarm optimization. Based on the particle swarm to select individual classifiers with large differences and high accuracy, establish the best integrated model of college English teaching effect evaluation. Secondly, starting with the theory of support vector machines, a brief analysis was carried out, and the related theories of parameter optimization were introduced, using grid search algorithm, genetic algorithm, particle swarm algorithm, and quantum particle swarm algorithm to penalize parameters of support vector machine. Then, in the teaching evaluation prediction model, a teaching evaluation prediction model based on the support vector machine regression algorithm and the time series algorithm is established. The selection of model parameters is determined according to the MSE global minimum as the criterion, and the particle swarm algorithm with convergence factor is used to optimize the penalty factor C of the support vector machine, the parameters of the kernel function, and the parameters of the insensitive loss function. Finally, the quantum particle swarm algorithm is improved, and the parameters of the support vector machine are optimized with the improved algorithm and experiments are carried out. According to the experimental results, the analysis verifies that the improved algorithm has higher performance when optimizing the parameters of the support vector machine.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] X. Li, S. Wu, and X. Li, "Particle swarm optimization-support vector machine model for machinery fault diagnoses in high-voltage circuit breakers," *Chinese Journal of Mechanical Engineering*, vol. 33, no. 1, pp. 8–10, 2020.

[2] A. Mustopa, E. B. Pratama, and A. Hendini, "Analysis of user reviews for the PeduliLindungi application on google play using the support vector machine and naive bayes algorithm based on particle swarm optimization," in *Proceedings of the 2020 Fifth International Conference on Informatics and Computing (ICIC)*, pp. 5–7, Gorontalo, Indonesia, November 2020.

[3] N. K. Wardhani, S. K. Rezkiani, H. Setiawan et al., "Sentiment analysis article news coordinator minister of maritime affairs using algorithm naive bayes and support vector machine with particle swarm optimization," *Journal of Theoretical and Applied Information Technology*, vol. 96, no. 24, pp. 8365–8378, 2018.

[4] Z. Zhong and T. R. Carr, "Application of a new hybrid particle swarm optimization-mixed kernels function-based support vector machine model for reservoir porosity prediction: a case study in Jacksonburg-Stringtown oil field, West Virginia, USA," *Interpretation*, vol. 7, no. 1, pp. T97–T112, West Virginia, USA, 2019.

[5] X. Liu, P. Hao, and A. Wang, "Non-destructive detection of highway hidden layer defects using a ground-penetrating radar and adaptive particle swarm support vector machine," *PeerJ Computer Science*, vol. 7, 2021.

[6] R. Ardianto, T. Rivanie, Y. Alkhalifi, F. S. Nugraha, and W. Gata, "Sentiment analysis on E-sports for education curriculum using naive bayes and support vector machine," *Jurnal Ilmu Komputer dan Informasi*, vol. 13, no. 2, pp. 109–122, 2020.

[7] R. Kaur, R. K. Sharma, and P. Kumar, "Speaker classification with support vector machine and crossover-based particle swarm optimization," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 34, no. 10, Article ID 2051010, 2020.

[8] W. Ali, "Hybrid intelligent android malware detection using evolving support vector machine based on genetic algorithm and particle swarm optimization," *Optimization*, vol. 19, no. 9, pp. 15–28, 2019.

[9] H. Cao, J. Zhang, F. Yang, Q. An, and Y. Wang, "Extending capture range for piston error in segmented primary mirror telescopes based on wavelet support vector machine with improved particle swarm optimization," *Ieee Access*, vol. 8, pp. 111585–111597, 2020.

[10] E. H. Houssein, A. A. Ewees, and M. A. ElAziz, "Improving twin support vector machine based on hybrid swarm optimizer for heartbeat classification," *Pattern Recognition and Image Analysis*, vol. 28, no. 2, pp. 243–253, 2018.

[11] H. Jing and X. Xiaoqiong, "Sports image detection based on FPGA hardware system and particle swarm algorithm," *Microprocessors and Microsystems*, vol. 80, Article ID 103348, 2021.

[12] T. Xie, J. Yao, and Z. Zhou, "DA-based parameter optimization of combined kernel support vector machine for cancer diagnosis," *Processes*, vol. 7, no. 5, p. 263, 2019.

[13] S. P. Rajamohana and K. Umamaheswari, "Hybrid approach of improved binary particle swarm optimization and shuffled frog leaping for feature selection," *Computers & Electrical Engineering*, vol. 67, pp. 497–508, 2018.
[14] X. Li, A. Luo, J. Li, and Y. Li, "Air pollutant concentration forecast based on support vector regression and quantum-behaved particle swarm optimization," *Environmental Modeling & Assessment*, vol. 24, no. 2, pp. 205–222, 2019.
[15] E. P. Saputra, S. A. Putri, and I. Indriyanti, "Prediction of successful e-learning based on activity logs with selection of support vector machine based on particle swarm optimization," *Indonesian Journal of Artificial Intelligence and Data Mining*, vol. 2, no. 1, pp. 10–17, 2019.
[16] D. Juan and Y. Hong Wei, "Particle swarm optimization neural network for research on artificial intelligence college English classroom teaching framework," *Journal of Intelligent and Fuzzy Systems*, vol. 40, pp. 8–13, 2021.
[17] R. Weizhi and S. Jian, "Application of support vector machine model in prediction of the settlement deformation of the hydropower station dam," *Jilin Water Resources*, vol. 21, p. 4, 2017.
[18] X. Wei, H. Huang, and W. Tang, "Parameters selection of twin support vector regression based on cloud particle swarm optimization," *Intelligent Computing Methodologies*, Springer, Berlin, Germany, pp. 388–399, 2020.
[19] H. Arora and G. M. Upadhyay, "A framework for the detection of suspicious discussion on online forums using integrated approach of support vector machine and particle swarm optimization," *International Journal of Advanced Research in Computer Science*, vol. 8, no. 5, pp. 134–137, 2017.
[20] H. Xue, Y. Bai, H. Hu, T. Xu, and H. Liang, "A novel hybrid model based on TVIW-PSO-GSA algorithm and support vector machine for classification problems," *IEEE Access*, vol. 7, pp. 27789–27801, 2019.
[21] N. Surantha, T. F. Lesmana, and S. M. Isa, "Sleep stage classification using extreme learning machine and particle swarm optimization for healthcare big data," *Journal of Big Data*, vol. 8, no. 1, pp. 15–17, 2021.
[22] J. S. Raj and J. V. Ananthi, "Recurrent neural networks and nonlinear prediction in support vector machines," *Journal of Soft Computing Paradigm (JSCP)*, vol. 1, no. 1, pp. 33–40, 2019.
[23] F. N. Hasan and M. Wahyudi, "Analisis sentimen artikel berita tokoh sepak bola dunia menggunakan algoritma support vector machine dan naive bayes berbasis particle swarm optimization," *Jurnal Akrab Juara*, vol. 3, no. 4, pp. 42–55, 2018.
[24] V. Sureshkumar, S. Balasubramaniam, and V. Ravi, "A hybrid optimization algorithm-based feature selection for thyroid disease classifier with rough type-2 fuzzy support vector machine," *Expert Systems*, vol. 39, Article ID e12811, 2021.
[25] I. O. Oyefolahan, S. Idris, and J. N. Ndunagu, "Hybridization of support vector machine with cat swarm algorithm for intrusion detection," *I-Manager's Journal on Computer Science*, vol. 8, no. 1, p. 13, 2020.