

Research Article

Predict the Entrepreneurial Intention of Fresh Graduate Students Based on an Adaptive Support Vector Machine Framework

Jixia Tu,¹ Aiju Lin,² Huiling Chen ,² Yuping Li ,³ and Chengye Li ³

¹The Communist Youth League, Wenzhou University, Wenzhou, Zhejiang 325035, China

²College of Mathematics, Physics and Electronic Information, Wenzhou University, Wenzhou, Zhejiang 325035, China

³Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, China

Correspondence should be addressed to Huiling Chen; chenhuiling.jlu@gmail.com, Yuping Li; wzliyp@163.com, and Chengye Li; lichengye41@126.com

Received 21 August 2018; Revised 16 December 2018; Accepted 1 January 2019; Published 20 January 2019

Academic Editor: Erik Cuevas

Copyright © 2019 Jixia Tu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Under the background of “innovation and entrepreneurship,” how to scientifically and rationally choose employment or independent entrepreneurship according to their own comprehensive situation is of great significance to the planning and development of their own career and the social adaptation of university personnel training. This study aims to develop an adaptive support vector machine framework, called RF-CSCA-SVM, for predicting college students' entrepreneurial intention in advance; that is, students choose to start a business or find a job after graduation. RF-CSCA-SVM combines random forest (RF), support vector machine (SVM), sine cosine algorithm (SCA), and chaotic local search. In this framework, RF is used to select the most important factors; SVM is employed to establish the relationship model between the factors and the students' decision to choose to start their own business or look for jobs. SCA is used to tune the optimal parameters for SVM. Additionally, chaotic local search is utilized to enhance the search capability of SCA. A total of 300 students were collected to develop the predictive model. To validate the developed method, other four meta-heuristic based SVM methods were used for comparison in terms of classification accuracy, Matthews Correlation Coefficients (MCC), sensitivity, and specificity. The experimental results demonstrate that the proposed method can be regarded as a promising success with the excellent predictive performance. Promisingly, the established adaptive SVM framework might serve as a new candidate of powerful tools for entrepreneurial intention prediction.

1. Introduction

In recent years, with the enlargement of college enrollment in China, the number of college graduates is increasing every year and the employment situation of the society is becoming more and tenser. In order to solve the problem of employment, the government and universities are trying all kinds of solutions. For example, it is an important way to alleviate the employment problem of college students at present by encouraging some college students to carry out their own entrepreneurial activities and implementing the employment strategy of “employment driven by entrepreneurship.” However, statistics show that the proportion of college students

starting a business after graduation is still low. Less than 5% of the graduates will choose to start their own businesses. At the same time, the satisfaction rate of entrepreneurship after three years of graduation is not high. Although many students have received relevant training in entrepreneurship and innovation education at school, there will still be confusion in the choice of entrepreneurship and employment when they graduate. Therefore, it is necessary to make a deep analysis of the rational choice behavior of college students' employment and entrepreneurship, in order to find an effective way for students to choose the direction of graduation scientifically. At present, a large number of datasets have been accumulated in colleges and universities. We can make deep mining and

analysis of these data to establish an intelligent prediction model by which we can find out the factors that affect college students after graduation to choose entrepreneurship or employment in the data. And then we can further analyze the potential correlation between factors to guide students to better choose entrepreneurship or employment.

Up to now, data mining technology has been applied to establish models to analyze the issues related to the employment of college graduates. Decision tree (DT) is one of the commonly used models which are involved. Zang et al. [1] constructed DT model with ID3 algorithm in order to analyze and find out information that is helpful to the employment of graduates from huge amounts of data. To analyze the relationship between college students' employment and academic performance, reading status, and other factors, Liu et al. [2] proposed a novel method, Information Gain with Weight Based Decision Tree (IGWDT). The purpose of the information gain weight (IGW) defined in IGWDT is to improve the information gain method. Besides, genetic algorithm is used to get the value of IGW for reflecting the degree of influence of different factors. Finally, the most relevant factors are obtained. Certainly, experiments show that the method can achieve good results. Li et al. [3] applied DT model with C4.5 algorithm to trace useful information to improve the employment rate of college graduate students. The experimental result has shown that the proposed method can classify employment data more quickly and correctly, as well as providing valuable results for analysis and decision making. Wang [4] conducted a more in-depth study on the methods and patterns of college students' employment data mining and mainly studied the establishment of college students' employment data mining model based on uniform distribution of fuzzy decision tree (FDT), nonuniform distribution of FDT, increasing distribution of FDT, and decreasing distribution of FDT. Apart from DT models, there are many other models that are often mentioned in the literature. For example, Bayesian methods, neural networks (NN), sequential minimal optimization (SMO), and ensemble methods were utilized by Mishra et al. [5] to predict whether students will be at risk of losing their jobs and help schools take timely measures to train students for the improvement of their employment rate. In addition, Xu et al. [6] used Bayesian algorithm to establish a classification model of graduate employment choice. In this model, the inclusion degree is calculated by fuzzy mathematics to determine the categories of each student. The categories are set according to the degree of job satisfaction of graduates, which are satisfactory, general satisfactory, and unsatisfactory. And, proved by experiments, this method has achieved high accuracy. In [7], Rahman et al. tried to determine the model with the highest accuracy for predicting whether graduates would be employed or unemployed six months after graduation. Thakar et al. [8] proposed to construct a unified forecasting model by integrating clustering and classification techniques to identify whether students have unemployment risk. In this model, K-means kernel with chi-square analysis is applied to the data preprocessing stage and then the combination of k-star, random tree, simple cart, and the random forest was used to predict whether students would be at unemployment risk.

Tan et al. [9] used k-nearest neighbors, naïve Bayes (NB), DT, NN, logistic regression, and support vector machines (SVM) to predict and evaluate the attributes of student data sets to confirm that graduates with which qualifications are needed by enterprise. Lanka et al. [10] applied NB to predict whether students could be recognized by enterprises. Proven by experiments, KNN and NB had better prediction results. For the problem of unsatisfactory evaluation system of college students' entrepreneurship, literature [11] put forward an evaluation system of college students' entrepreneurship based on backpropagation NN theory to promote the development of college students' entrepreneurship education.

To sum up, there is a lot of work on predicting students' employment ability, but there is no report on predicting college students' entrepreneurial employment after graduation. This paper attempts to use a new machine learning framework to predict students' entrepreneurial intention after graduation, that is, to choose to start a business or not. The proposed framework consists of three main parts. The first part uses random forest (RF) method to select the key features in the data, the second part uses chaotic local search based sine cosine algorithm (CSCA) to optimize an optimal SVM model, and the third part uses CSCA-SVM obtained from the previous stage to predict the new sample. Sine cosine algorithm (SCA) is a new swarm intelligence method that was proposed recently by Mirjalili [12]. Since its introduction, SCA has successfully found its applications for many practical problems [13–18]. However, like other intelligent algorithms [19–25], the original SCA is easy to fall into the local optimum when solving the practical problems. In this study, we try to use the chaotic local search (CLS) strategy to enhance the local search capability of SCA, termed as CSCA. After that, the CSCA was used to determine the optimal parameters for SVM. As shown, the experimental results have shown that CLS strategy has indeed improved the solution quality and convergence speed for SCA. At the same time, it is also observed that the CSCA has also boosted the performance of SVM compared against many other nature inspired metaheuristic algorithms based SVM models. The efficacy of the proposed RF-CSCA-SVM framework was rigorously compared against four other SVM models based on different swarm intelligent algorithms including SCA, moth-flame optimization (MFO) [26] algorithm, bat algorithm (BA) [27], and grasshopper optimization algorithm (GOA) [28] on the real-life dataset collected from Wenzhou University. The classifiers were compared in terms of four common performance metrics including classification accuracy (ACC), sensitivity, specificity, and the Matthews Correlation Coefficients (MCC) criterion. The experimental results demonstrated that the proposed RF-CSCA-SVM approach achieved much better performance than other competitive counterparts.

The rest of this paper is structured as follows. Section 2 offers brief description on the methodology including random forest, support vector machine, sine cosine algorithm, and chaotic local search strategy. The experimental design is given in Section 3. Section 4 presents the detailed simulation results. The discussion on the experimental results is delivered in Section 5. Finally, Section 6 summarizes the conclusions and recommendations for future work.

2. Methods

In SVM model, Gaussian kernel function is needed to solve the calculation of high-dimensional vector inner product. Gamma value of Gaussian kernel function is mainly used to determine the kernel width of SVM. Meanwhile, soft margin support vector machine needs to introduce penalty factor to reduce noise interference. Values of gamma and penalty factor have a great impact on the classification results. Currently, grid search method and gradient descent method are commonly used for parameter optimization of SVM, but the main disadvantage of these methods is that they are easy to fall into local optimal solutions. SCA based on chaotic local search mechanism can find a good balance between local and global optimization ability, so searching these two key parameters of SVM with this method can better determine the optimal value of these two parameters.

This section will give a detailed introduction of the proposed framework for the entrepreneurial intention prediction of college students, named RF-CSCA-SVM. The main flow of the framework is shown in Figure 1. The whole process is divided into three parts. The first part is to normalize the data and evaluate each feature by random forest (RF) method. The second part is to construct an effective SVM model with optimal parameters based on the improved SCA; at the same time, this optimal model was used to evaluate different feature sets in an incremental way and obtain the best feature set. The main task of the third part is to use the optimal model constructed in the previous stage to predict the new data samples.

The main steps conducted by the RF-CSCA-SVM are described in detail as follows:

- (i) Step 1: Normalize the data to $[-1, 1]$.
- (ii) Step 2: Each feature of the data is evaluated using RF algorithm and the optimal subset is selected in an incremental manner based on the importance of each feature.
- (iii) Step 3: Initialize a population randomly based on the upper and lower bounds of the variables.
- (iv) Step 4: Evaluate the fitness of all search agents by SVM with agent as parameters and update the best solution obtained so far.
- (v) Step 5: Update the position of each search agent according to chaotic local search strategy.
- (vi) Step 6: Check if any search agent goes beyond the search space and amend it.
- (vii) Step 7: Evaluate the fitness of all search agents by SVM with agent as parameters and update the best solution obtained so far.
- (viii) Step 8: Update iteration t , $t=t+1$. If t is less than maximum number of iterations, go to step 5.
- (ix) Step 9: Return the best solution as the optimal SVM parameter pair (C, γ) .

The computational complexity of the CSCA-SVM method depends on the number of samples (S), the

scale of the problem (D), the population number (p), and the number of iterations (g). Therefore, the overall computational complexity is $O(\text{SVM, CSCA})=O(\text{Initialization})+g*(O(\text{Evaluate the fitness of all agents})+O(\text{Update the best solution obtained so far})+O(\text{SCA position updating})+O(\text{Calculate the CLS position of the best solution})+O(\text{Output}))$. As we all know, SVM's computational complexity on S samples is $O(S^3)$. Updating the positions of search agents is $O(p*D)$. Evaluating the fitness of all agents is $O(p*O(S^3))$. In addition, the computational complexity of CLS strategy for the best solution is $O(g*p)$. Therefore, the final computational complexity of the CSCA-SVM method is $O(\text{SVM, CSCA})=O(p*D)+g*(O(p*O(S^3))+O(N*\log N)+O(p*D)+O(p)) \approx O(p*D+g*(O(N*\log N+p(D+O(S^3)+1)))$.

2.1. Feature Selection Method: Random Forest (RF). Random forest (RF) [29] is an ensemble machine learning method that uses bootstrap and node random splitting technology to construct multiple decision trees and obtain final classification results by voting. RF has the ability to analyze the classification characteristics of complex interactions. It has good robustness for noisy data and data with missing values and has a faster learning speed. Its variable importance measure can be used as a feature selection tool [30]. In recent years, it has been widely used in various classification, feature selection, and outlier detection problems. This paper mainly uses the mean decrease accuracy to measure the importance of each feature. This method directly measures the impact of each feature on the accuracy of the model. The main idea is to disrupt the order of characteristic values of each feature and measure the impact of sequence changes on the accuracy of the model. Obviously, for unimportant variables, the disruption order will not affect the accuracy of the model much, but, for the important variables, this operation will reduce the accuracy of the model.

2.2. Prediction Engine: Support Vector Machine (SVM). Support vector machine (SVM) is an advanced artificial intelligence technology [31], which is mainly based on VC dimension theory and structural risk minimization principle, trying to find a compromise between minimizing training set error and maximizing classification interval, so as to obtain the best generalization ability. At present, SVM has been widely used to solve various practical problems [25, 32–42].

SVM is a learning algorithm mainly for small samples, which is very suitable for the prediction modeling of current cases. In this experiment, we will adopt the grid search algorithm to obtain the optimal parameters of the SVM model, which is used to construct the optimal classification function, as shown in the following:

$$f(s) = \text{sgn}(w^T \cdot x + b) = \text{sgn}\left(\sum_{i=1}^n a_i y_i K(x_i, x_j) + b\right) \quad (1)$$

In Eq. (1), $K(x_i, y_j) = \exp(-\gamma\|x_i - x_j\|)$, a_i is the Lagrange coefficient, b is the threshold value, x_i is the samples to be tested ($i=1\dots n$), and y_j indicates the label corresponding to the

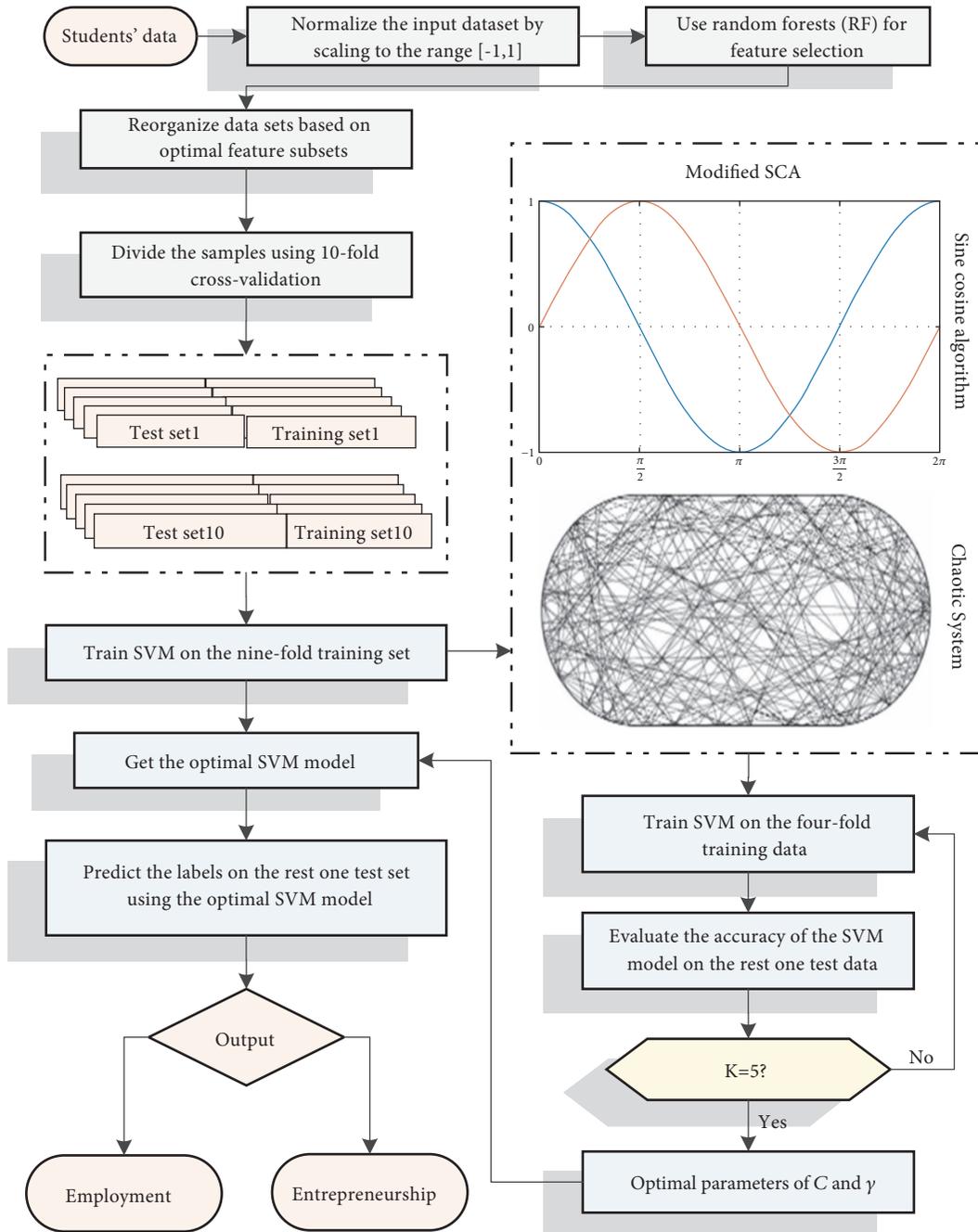


FIGURE 1: Flowchart of the proposed RF-CSCA-SVM framework.

training samples. y_j ($j=1\dots n$) takes the value of 1 or 2, where 1 and 2 represent the type of students who choose to start their own business and look for jobs, respectively, and T indicates target vector, $T = [t_1, t_2, \dots, t_n]$.

2.3. Chaotic Local Search Enhanced Sine Cosine Algorithm (CSCA). SCA was first put forward by Mirjalili *et al.* [12] which initializes many original stochastic positions and makes them move outwards or towards the best position. This movement is realized by a mathematical mean

whose core is sine and cosine formulas. While the core function returns a value less than -1 or greater than 1, this mechanism of movement enables different areas to be explored by random agents. While the value returned by the core function is between -1 and 1, the mechanism will drive the agents to exploit the most promising area. In order to make the exploration turn to development gradually, SCA algorithm adopts a very effective mechanism, which adds an adaptive range to the core function.

The mathematical formula of the update agent location used in the SCA algorithm is presented as follows:

$$X_i^{t+1} = \begin{cases} X_i^t + r_1 \times \sin(r_2) \times |r_3 P_i^t - X_i^t|, & r_4 < 0.5 \\ X_i^t + r_1 \times \cos(r_2) \times |r_3 P_i^t - X_i^t|, & r_4 \geq 0.5 \end{cases} \quad (2)$$

The general framework of SCA is as in Algorithm 1.

Chaos is a classic nonlinear natural phenomenon that has been of concern for a long time. It has many unique characteristics. Chaos is very sensitive to its initial conditions. And, also, it has the characteristics of randomness and ergodicity [43]. The explanation of randomness is that the variation of a chaotic system is random. And ergodicity means that, given enough time, chaotic systems can cover all regions. These characteristics, especially randomness and ergodicity, coincide with those required by a good search operator. Therefore, the optimization of objective function can make use of these characteristics. The time spent in chaos optimization is acceptable on a small scale. But time cost will be difficult to bear after the search space becomes large. In order to solve the above problem, chaotic local search (CLS) can be integrated into other intelligent optimization algorithms. For example, CLS has been introduced into GA, PSO, and BFO [23, 44–46]. Similar to other intelligent optimization algorithms, both the initial solution of SCA algorithm and the solution generated in the iterative process have strong randomness. This means that any solution of SCA algorithm has a great potential to be continuously optimized. The improved strategy based on this characteristic can be considered as an overall mutation to the population of SCA algorithm. If we repeatedly add the chaotic perturbation factor to the initial solution in the iterative process of SCA algorithm, many new solutions with higher fitness may be obtained. Consequently, we have reason to believe that the convergence speed of SCA algorithm will be greatly improved by continually replacing the original solution with the best solution obtained by chaotic local search.

The integration of CLS mechanism into SCA intelligent optimization algorithm can not only enhance its search ability, but also make it avoid falling into local optimum. In this paper, the famous logistic formula is used to generate chaotic factors and construct chaotic systems. The logistic formula [47] is very sensitive to original conditions. The chaotic system constructed by it will be integrated into the SCA algorithm, thus constructing a new hybrid algorithm. The following is the logistic formula:

$$x_{n+1} = \mu \cdot x_n (1 - x_n), \quad 0 \leq x_0 \leq 1, \quad (3)$$

where x is the variable quantity, $n = 0, 1, 2, \dots$, and μ is the adjustment factor. Judging from the expression of logistic formula, it is a deterministic formula. However, when we set μ to 4 and make x not equal to any one of 0, 0.25, 0.5, 0.75, and 1, it can generate a series of chaotic factors. This is the sensitivity to the original conditions mentioned above. Similarly, this is also an important feature of chaos. Any slight difference in the original conditions will result in a huge difference in the chaotic system that is subsequently generated. The traces of

chaotic factors generated by chaotic systems can be found throughout the whole area that needs searching. The traces of these chaotic factors contain important properties, including irregularity and randomness.

In this paper, chaotic local search is described as follows:

$$X_k^{(c)} = (1 - \lambda) X^* + \lambda (LB + \beta_k (UB - LB)) \quad (4)$$

where $X_k^{(c)}$ is the k th new position vector produced by chaotic local search. X^* is the position vector of the best solution obtained so far. β_k is k th chaotic value in chaotic sequence. LB and UB are the upper and lower bounds of the search space, respectively. λ is the shrinking scale given by the following:

$$\lambda = \frac{Max_iter - t + 1}{Max_iter} \quad (5)$$

where Max_iter is the maximum number of iterations and t is the current number of iterations.

For the SCA algorithm, the CLS strategy accelerates its convergence speed. This acceleration is achieved by continually generating new positions with chaotic factors in the iteration process and continually adopting greedy methods to preserve the optimal solution. In this way, the solutions of the whole population will be optimized, and the optimal solution will naturally move towards the global optimum. The general framework of CSCA is as shown in Figure 2.

3. Experimental Designs

3.1. Data Collection and Preparation. The data in this paper is mainly from the 2016 and 2017 graduates of Wenzhou University. From these graduates, 300 students were selected as research subjects, of which 136 were self-employed and 164 were employed. Through the analysis of the subjects' gender, political status, level of education, major, education year, type of normal school students, type of poor students, grade point average (GPA), and total credits, this study intends to investigate the importance and interrelationships of these nine attributes so as to establish a predictive model for decision support. Table 1 is a detailed description of the nine attributes.

3.2. Experimental Setup. To validate the proposed approach, we have conducted a comparative study between the proposed method and several SVM methods based on other nature inspired methods including PSO, GA, and SCA. LIBSVM [48] was utilized for SVM implementation. PSO, GA, SCA, and CSCA were implemented from scratch. The empirical experiment was conducted on a Windows Server 2008 R2 operating system with Intel (R) Xeon (R) CPU E5-2660 v3 (2.60 GHz) and 16GB of RAM.

Data was first scaled into the range $[-1, 1]$ before classification. In order to gain an unbiased estimate of the generalization accuracy, the k -fold cross-validation (CV) was used to evaluate the classification accuracy [49]. The two of SVM, penalty factor C and kernel width γ , are both set as $\{2^{-5}, 2^{-4}, \dots, 2^4, 2^5\}$. The parameters in PSO, GA, and SCA

```

Begin
  Initialize a set of search agents (solution) ( $X$ )
Do
  Evaluate each of the search agents by the objective function
  Update the best solution obtained so far ( $P=X^*$ )
  Update  $r_1, r_2, r_3,$  and  $r_4$ 
  Update the position of search agents using Eq. (4)
While ( $t <$  maximum number of generations)
  Return the best solution obtained so far as the global optimum
End
    
```

ALGORITHM 1: Procedure of SCA.

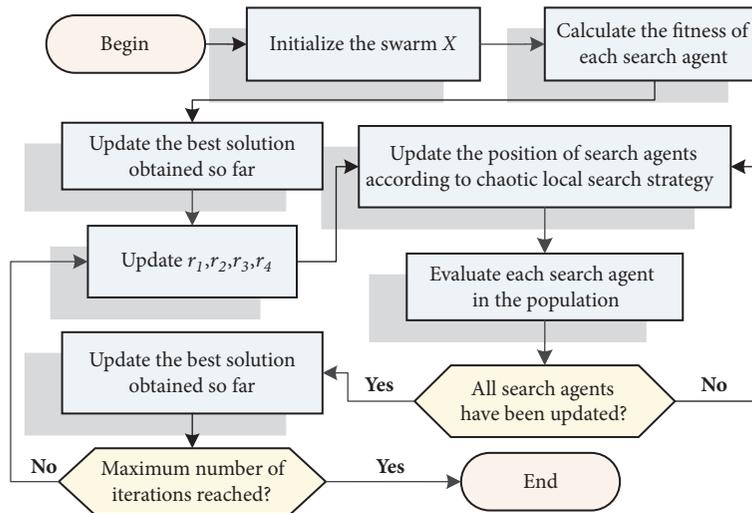


FIGURE 2: Flowchart of CSCA.

TABLE 1: Description of the Nine Attributes.

Attributes	Name	Description
F1	Gender	Male and female students are represented by 1 and 2, respectively.
F2	Political Status	It is divided into four categories: members of Communist Party of China, probationary members of the CPC, members of the Communist Youth League, and the masses, represented by 1, 2, 3, and 4, respectively.
F3	Level of Education	It is divided into master's degree, undergraduate degree, and college degree, represented by 1, 2, and 3, respectively.
F4	Major	It is divided into liberal arts and science, represented by 1 and 2, respectively.
F5	Education Year	It is divided into 2 years, 3 years, 4 years, and 5 years, represented by 1, 2, 3, and 4, respectively.
F6	Type of Normal School Students	It is divided into normal students and nonnormal students, represented by 1 and 2, respectively.
F7	Type of Poor Students	It is divided into four categories: nondifficult students, employment difficulties, family difficulties, and employment and family difficulties, represented by 1, 2, 3 and 4, respectively.
F8	Grade Point Average (GPA)	GPA is a way for school to assess students' learning quality. The score is within 0-4 intervals.
F9	Total Credits	It is a unit of measurement used to calculate students' learning volume. The more credits students receive, the more they learn.

TABLE 2: Unimodal benchmark functions.

Function	Dim	Range	f_{\min}
$f_1(x) = \sum_{i=1}^n x_i^2$	30	[-100, 100]	0
$f_2(x) = \sum_{i=1}^n x_i + \prod_{i=1}^n x_i $	30	[-10, 10]	0
$f_3(x) = \sum_{i=1}^n \left(\sum_{j=1}^i x_j \right)^2$	30	[-100, 100]	0
$f_4(x) = \max_i \{ x_i , 1 \leq i \leq n\}$	30	[-100, 100]	0
$f_5(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	30	[-30, 30]	0
$f_6(x) = \sum_{i=1}^n ([x_i + 0.5])^2$	30	[-100, 100]	0
$f_7(x) = \sum_{i=1}^n ix_i^4 + \text{random}[0, 1]$	30	[-1.28, 1.28]	0

TABLE 3: Multimodal benchmark functions.

Function	Dim	Range	f_{\min}
$f_8(x) = \sum_{i=1}^n -x_i \sin(\sqrt{ x_i })$	30	[-500, 500]	-418.9829×5
$f_9(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10]$	30	[-5.12, 5.12]	0
$f_{10}(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)\right) + 20 + e$	30	[-32, 32]	0
$f_{11}(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	[-600, 600]	0
$f_{12}(x) = \frac{\pi}{n} \left\{ 10 \sin(\pi y_1) + \sum_{i=1}^{n-1} (y_i - 1)^2 [1 + 10 \sin^2(\pi y_{i+1})] + (y_n - 1)^2 \right\} + \sum_{i=1}^n u(x_i, 10, 100, 4)$ $y_i = 1 + \frac{x_i + 1}{4}$	30	[-50, 50]	0
$u(x_i, a, k, m) = \begin{cases} k(x_i - a)^m & x_i > a \\ 0 & -a < x_i < a \\ k(-x_i - a)^m & x_i < -a \end{cases}$			
$f_{13}(x) = 0.1 \left\{ \sin^2(3\pi x_1) + \sum_{i=1}^n (x_i - 1)^2 [1 + \sin^2(3\pi x_i + 1)] + (x_n - 1)^2 [1 + \sin^2(2\pi x_n)] \right\} + \sum_{i=1}^n u(x_i, 5, 100, 4)$	30	[-50, 50]	0

were set as follows. The two constant factors c_1 and c_2 in PSO were set to 2 and the inertial weight w in PSO was set to 1. Crossover fraction and mutation probability in GA were set to 0.8 and 0.01, respectively. The constant a in SCA and CSCA is set to 2.

To evaluate the proposed method, commonly used evaluation criteria such as classification accuracy (ACC), sensitivity, specificity, and Matthews Correlation Coefficients (MCC) were analyzed. They are defined as follows:

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% \quad (6)$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \quad (7)$$

$$Specificity = \frac{TN}{FP + TN} \times 100\% \quad (8)$$

MCC

$$= \frac{TP * TN - FP * FN}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}} \times 100\% \quad (9)$$

4. Experimental Results and Discussion

4.1. Benchmark Function Validation. In order to evaluate the performance of the CSCA algorithm, a series of experiments on classical benchmark functions were conducted in this section. The benchmark functions can be divided into three

TABLE 4: Fixed-dimension multimodal benchmark functions.

Function	Dim	Range	f_{\min}
$f_{14}(x) = \left(\frac{1}{500} + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^2 (x_i - a_{ij})^6} \right)^{-1}$	2	[-65, 65]	1
$f_{15}(x) = \sum_{i=1}^{11} \left[a_i - \frac{x_1 (b_i^2 + b_i x_2)}{b_i^2 + b_i x_3 + x_4} \right]^2$	4	[-5, 5]	0.00030
$f_{16}(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$	2	[-5, 5]	-1.0316
$f_{17}(x) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6 \right)^2 + 10 \left(1 - \frac{1}{8\pi} \right) \cos x_1 + 10$	2	[-5, 5]	0.398
$f_{18}(x) = \left[1 + (x_1 + x_2 + 1)^2 (19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2) \right] \times \left[30 + (2x_1 - 3x_2)^2 \times (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2) \right]$	2	[-2, 2]	3
$f_{19}(x) = -\sum_{i=1}^4 c_i \exp \left(-\sum_{j=1}^3 a_{ij} (x_j - p_{ij})^2 \right)$	3	[1, 3]	-3.86
$f_{20}(x) = -\sum_{i=1}^4 c_i \exp \left(-\sum_{j=1}^6 a_{ij} (x_j - p_{ij})^2 \right)$	6	[0, 1]	-3.32
$f_{21}(x) = -\sum_{i=1}^5 \left[(X - a_i)(X - a_i)^T + c_i \right]^{-1}$	4	[0, 10]	-10.1532
$f_{22}(x) = -\sum_{i=1}^7 \left[(X - a_i)(X - a_i)^T + c_i \right]^{-1}$	4	[0, 10]	-10.4028
$f_{23}(x) = -\sum_{i=1}^{10} \left[(X - a_i)(X - a_i)^T + c_i \right]^{-1}$	4	[0, 10]	-10.5363

parts: unimodal (see Table 2), multimodal (see Table 3), and fixed-dimension multimodal (see Table 4). Tables 2–4 include the following aspects for each function: the mathematical equation, the dimension of solution space, the range of optimization variables, and the theoretical optimal value.

The performance of the CSCA is compared with MFO, BA, DA, FPA, GOA, SSA, and the original SCA. And the simulation results of CSCA on the benchmark functions are presented in Table 5. The table records the average value (mean) and the standard deviation (std.) of the best solution of each algorithm over 30 independent runs. For fair comparison, all algorithms are implemented in the same environment on the same computing platform. All algorithms can use the same global settings. The maximum number of iterations and the number of populations were set to 1000 and 30, respectively. In addition, the specific parameter values of the above algorithms, such as MFO, BA, and the original SCA, are set according to their original papers.

Inspecting the detailed results of algorithms on the 23 benchmark functions in Table 5, CSCA has the best results on 17 out of 23 test functions. For all seven unimodal f_1 - f_7 functions and six multimodal f_8 - f_{13} functions, the proposed CSCA algorithm can outperform all other algorithms. According to CSCA and SCA metrics, the CSCA performance is improved compared to the basic SCA. In addition, the ranking results also prove that CSCA provides the best solution among all in terms of the mean index. For ten fixed-dimension multimodal functions (f_{14} - f_{23}), CSCA has attained the exact optimal solutions for f_{15} , f_{21} , f_{22} , f_{23} . For other six functions (f_{14} , f_{16} , f_{17} , f_{18} , f_{19} , f_{20}), although

in dealing with some cases the improved CSCA is not better than other optimizers, it is observed that the optimization effect of proposed CSCA still obtains better results than the basic SCA in more than 90% of fixed-dimension multimodal cases. The results show that the utilized chaotic local search in CSCA has enhanced the efficacy of the SCA effectively. Moreover, based on rankings, it can be observed that the developed CSCA can achieve the best place. And the overall ranks show that the SSA, FPA, MFO, SCA, BA, and DA algorithms are in the next places, respectively.

Moreover, to visually show the performance of the CSCA, convergence curves of CSCA, MFO, BA, DA, FPA, GOA, SSA, and the original SCA on some typical benchmark functions are also provided in Figure 3. For f_1 , f_4 , and f_7 , the proposed CSCA algorithm has reached the best solution and it is seen that the worst results of CSCA are much better than the best values of the classical SCA and other six algorithms. It is clearly shown that the performance of the CSCA on unimodal cases has improved and the CSCA is much better than other algorithms. Regarding the convergence curves reported in Figure 3, it can be detected that the convergence speed of the CSCA is better than other optimizers in multimodal cases. In dealing with f_8 , the proposed technique has converged very quickly throughout early steps. For f_9 , it is seen that the fastest convergence also belongs to the CSCA algorithm, while MFO, BA, DA, FPA, GOA, SSA, and SCA cannot show a better trend. From f_{12} , the proposed CSCA shows the best function value in the early stages, while other optimizers have all fallen into local optima because of their weaker search capability. For f_{15} , the proposed CSCA algorithm has reached the best solution and it is clearly seen that the worst result of CSCA is

TABLE 5: Results of testing benchmark functions.

Function	Metric	CSCA	MFO	BA	DA	FPA	GOA	SSA	SCA
f_1	mean	1.85E - 09	2.12E - 01	1.59E + 01	1.57E + 03	5.04E + 02	3.76E + 01	1.60E - 08	2.89E - 02
	std	3.34E - 09	6.31E - 01	1.41E + 00	5.17E + 02	2.01E + 02	1.98E + 01	2.75E - 09	3.48E - 02
	rank	1	4	5	8	7	6	2	3
f_2	mean	3.04E - 08	3.61E + 01	2.71E + 01	2.32E + 01	1.57E + 01	1.62E + 01	2.30E + 00	3.94E - 05
	std	6.29E - 08	2.36E + 01	2.87E + 01	5.70E + 00	3.51E + 00	9.91E + 00	2.36E + 00	9.20E - 05
	rank	1	8	7	6	4	5	3	2
f_3	mean	7.94E - 07	2.61E + 04	9.55E + 01	1.78E + 04	4.73E + 02	2.67E + 03	8.15E + 02	4.44E + 03
	std	1.09E - 06	1.42E + 04	2.75E + 01	7.37E + 03	2.01E + 02	1.24E + 03	4.57E + 02	4.51E + 03
	rank	1	8	2	7	3	5	4	6
f_4	mean	1.99E - 06	7.87E + 01	2.63E + 00	2.71E + 01	1.85E + 01	1.40E + 01	1.30E + 01	2.63E + 01
	std	1.53E - 06	5.89E + 00	1.42E + 00	5.95E + 00	3.64E + 00	5.05E + 00	4.28E + 00	8.68E + 00
	rank	1	8	2	7	5	4	3	6
f_5	mean	7.32E - 10	9.69E + 03	4.73E + 03	1.96E + 05	3.45E + 04	4.59E + 03	1.97E + 02	3.10E + 02
	std	9.69E - 10	2.83E + 04	1.40E + 03	1.29E + 05	1.88E + 04	5.86E + 03	2.32E + 02	5.51E + 02
	rank	1	6	5	8	7	4	2	3
f_6	mean	9.97E - 12	9.90E + 02	1.53E + 01	1.54E + 03	5.14E + 02	3.63E + 01	1.64E - 08	5.99E + 00
	std	1.05E - 11	3.13E + 03	1.80E + 00	6.43E + 02	1.85E + 02	1.11E + 01	5.02E - 09	2.13E + 00
	rank	1	7	4	8	6	5	2	3
f_7	mean	3.21E - 06	5.34E + 00	3.23E + 01	6.37E - 01	3.10E - 01	1.21E + 00	1.51E - 01	2.45E - 02
	std	2.74E - 06	1.35E + 01	2.22E + 01	3.48E - 01	1.42E - 01	8.11E - 01	5.34E - 02	2.81E - 02
	rank	1	7	8	5	4	6	3	2
f_8	mean	-1.26E + 04	-7.83E + 03	-7.19E + 03	-5.39E + 03	-7.77E + 03	-7.63E + 03	-7.82E + 03	-3.92E + 03
	std	3.10E - 10	8.14E + 02	9.45E + 02	5.64E + 02	1.86E + 02	3.64E + 02	6.89E + 02	2.49E + 02
	rank	1	2	6	7	4	5	3	8
f_9	mean	1.31E - 10	1.78E + 02	2.87E + 02	1.84E + 02	9.84E + 01	1.06E + 02	7.65E + 01	1.84E + 01
	std	2.67E - 10	3.90E + 01	3.75E + 01	3.00E + 01	2.21E + 01	3.70E + 01	1.80E + 01	2.03E + 01
	rank	1	6	8	7	4	5	3	2
f_{10}	mean	3.78E - 06	1.73E + 01	1.04E + 01	9.77E + 00	9.19E + 00	6.92E + 00	2.50E + 00	1.15E + 01
	std	2.23E - 06	4.72E + 00	7.59E + 00	9.61E - 01	1.12E + 00	1.26E + 00	2.81E - 01	1.01E + 01
	rank	1	8	6	5	4	3	2	7
f_{11}	mean	2.23E - 09	3.65E + 01	6.41E - 01	1.54E + 01	5.03E + 00	1.08E + 00	8.61E - 03	3.41E - 01
	std	5.07E - 09	6.37E + 01	5.62E - 02	6.87E + 00	1.55E + 00	4.87E - 02	9.29E - 03	3.20E - 01
	rank	1	8	4	7	6	5	2	3
f_{12}	mean	3.26E - 12	6.32E + 03	1.40E + 01	1.16E + 03	7.33E + 00	9.51E + 00	8.54E + 00	4.84E + 01
	std	7.88E - 12	2.00E + 04	5.91E + 00	2.16E + 03	4.07E + 00	4.35E + 00	5.66E + 00	1.34E + 02
	rank	1	8	5	7	2	4	3	6
f_{13}	mean	1.38E - 11	7.28E + 00	2.42E + 00	1.58E + 05	5.93E + 03	3.64E + 01	1.40E + 01	8.21E + 03
	std	1.75E - 11	1.15E + 01	4.34E - 01	1.41E + 05	1.08E + 04	1.56E + 01	1.38E + 01	2.59E + 04
	rank	1	3	2	8	6	5	4	7
f_{14}	mean	9.98E - 01	3.73E + 00	8.30E + 00	9.98E - 01	9.98E - 01	5.39E + 00	9.98E - 01	2.20E + 00
	std	7.24E - 15	4.51E + 00	5.60E + 00	7.54E - 10	7.11E - 09	5.01E + 00	1.81E - 16	1.01E + 00
	rank	2	6	8	3	4	7	1	5
f_{15}	mean	3.49E - 04	3.03E - 03	8.76E - 03	7.11E - 03	4.90E - 04	1.98E - 02	7.57E - 04	1.04E - 03
	std	1.44E - 05	6.11E - 03	9.99E - 03	9.15E - 03	2.36E - 04	2.70E - 02	2.32E - 04	3.90E - 04
	rank	1	5	7	6	2	8	3	4
f_{16}	mean	-1.03E + 00							
	std	1.87E - 05	0.00E + 00	7.90E - 04	1.11E - 05	3.24E - 11	9.56E - 15	1.14E - 14	3.82E - 05
	rank	6	1	8	5	4	3	2	7

TABLE 5: Continued.

Function	Metric	CSCA	MFO	BA	DA	FPA	GOA	SSA	SCA
f_{17}	mean	3.99E - 01	3.98E - 01	3.99E - 01					
	std	7.87E - 04	0.00E + 00	5.13E - 04	1.36E - 06	6.71E - 15	1.41E - 14	7.41E - 15	2.00E - 03
	rank	7	1	6	5	2	4	3	8
f_{18}	mean	3.00E + 00	3.00E + 00	3.04E + 00	3.00E + 00	3.00E + 00	5.70E + 00	3.00E + 00	3.00E + 00
	std	2.31E - 05	1.75E - 15	3.09E - 02	4.13E - 05	8.27E - 13	8.54E + 00	1.29E - 13	4.81E - 05
	rank	4	1	7	5	3	8	2	6
f_{19}	mean	-3.85E + 00	-3.86E + 00	-3.83E + 00	-3.86E + 00	-3.86E + 00	-3.58E + 00	-3.86E + 00	-3.86E + 00
	std	2.12E - 03	2.49E - 03	1.25E - 02	1.37E - 04	5.98E - 12	5.48E - 01	5.84E - 14	3.13E - 03
	rank	6	4	7	3	2	8	1	5
f_{20}	mean	-3.08E + 00	-3.23E + 00	-2.86E + 00	-3.26E + 00	-3.32E + 00	-3.25E + 00	-3.22E + 00	-3.05E + 00
	std	5.03E - 01	4.24E - 01	5.40E - 02	1.29E - 01	1.22E - 01	4.20E - 03	1.74E - 01	3.53E - 01
	rank	6	4	8	2	1	3	5	7
f_{21}	mean	-1.02E + 01	-6.90E + 00	-3.24E + 00	-6.05E + 00	-1.02E + 01	-4.40E + 00	-6.63E + 00	-2.98E + 00
	std	2.21E - 10	3.55E + 00	1.64E + 00	2.86E + 00	2.00E - 04	3.12E + 00	3.17E + 00	2.53E + 00
	rank	1	3	7	5	2	6	4	8
f_{22}	mean	-1.04E + 01	-9.64E + 00	-6.87E + 00	-5.13E + 00	-1.04E + 01	-6.52E + 00	-8.58E + 00	-2.51E + 00
	std	2.71E - 10	2.42E + 00	3.02E + 00	3.03E + 00	9.23E - 04	3.48E + 00	3.00E + 00	1.78E + 00
	rank	1	3	5	7	2	6	4	8
f_{23}	mean	-1.05E + 01	-7.57E + 00	-5.04E + 00	-6.47E + 00	-1.05E + 01	-5.46E + 00	-7.65E + 00	-4.56E + 00
	std	1.41E - 10	3.84E + 00	2.85E + 00	3.55E + 00	2.38E - 02	3.67E + 00	3.80E + 00	1.49E + 00
	rank	1	4	7	5	2	6	3	8
Sum of rank		48	115	134	136	86	121	64	124
Average rank		2.087	5	5.826	5.913	3.739	5.261	2.783	5.391
Overall rank		1	4	7	8	3	5	2	6

TABLE 6: Performance of RF-CSCA-SVM on different sizes of feature subset.

Size of feature subset	ACC	Sensitivity	Specificity	MCC
1	0.6600(0.0625)	0.6357(0.0473)	0.7605 (0.1534)	0.3265(0.1420)
2	0.7300(0.0637)	0.7399(0.0889)	0.7323 (0.0719)	0.4600(0.1313)
3	0.7767 (0.0802)	0.7815 (0.1025)	0.7862 (0.0850)	0.5560 (0.1621)
4	0.7567 (0.0649)	0.7701 (0.0750)	0.7553 (0.1014)	0.5136 (0.1398)
5	0.8400 (0.0439)	0.8597 (0.0570)	0.8347 (0.0862)	0.6850 (0.0873)
6	0.7733 (0.0872)	0.8168 (0.1084)	0.7413 (0.1080)	0.5529 (0.1831)
7	0.7867 (0.0849)	0.8482 (0.1187)	0.7381 (0.0786)	0.5819 (0.1791)
8	0.7800 (0.0706)	0.8307 (0.0634)	0.7464 (0.1101)	0.5710 (0.1360)
9	0.6967 (0.1024)	0.7282 (0.1066)	0.6688 (0.1140)	0.3928 (0.2073)

much better than the best solutions of other methods. For f_{22} and f_{23} , CSCA has the best performance among all in terms of the std. index and is found to provide better results for these fixed dimensions cases.

To sum up, we can conclude that the proposed CSCA algorithm achieves better search performance and is well capable of escaping the local optimum values than all other competitors.

4.2. Prediction Results of Entrepreneurial Intention. In this experiment we first used the random forest (RF) to evaluate the importance of each feature of the data set. The results of the evaluation are shown in Figure 4. From the figure, we can find that the importance of each feature is very

different. Some features, like major (F4) and total credits (F9), are more important while some features, like political status (F3), education year (F5), and type of poor students (F7), are basically redundant features. Then, the optimal feature subset is formed based on the importance of these features in an incremental manner. Table 6 lists the results of the CSCA-SVM model on the incremental feature subset in the form of mean values with the standard deviation in the parentheses. From the table, it can be seen that the subset of features consisting of the five most important features can make the proposed model achieve the best performance. With these five top ranked features, CSCA-SVM has achieved the good prediction results with ACC of 84%, sensitivity of 85.97%, specificity of 83.47%, and MCC of 0.685. It is also

TABLE 7: Results obtained by the proposed RF-CSCA-SVM.

Fold No.	Confusion matrix		RF-CSCA-SVM			
			Accuracy	Sensitivity	Specificity	MCC
1	36 8	2 14	0.8333	0.9474	0.6364	0.6361
2	34 6	3 17	0.8500	0.9189	0.7391	0.6787
3	30 6	4 20	0.8333	0.8824	0.7692	0.6591
4	29 8	3 20	0.8167	0.9063	0.7143	0.6367
5	34 10	1 15	0.8167	0.9714	0.6000	0.6371
6	32 8	2 18	0.8333	0.9412	0.6923	0.6659
7	28 5	5 22	0.8333	0.8485	0.8148	0.6633
8	30 14	0 16	0.8500	0.9189	0.7391	0.6787
9	34 6	3 17	0.8500	0.9459	0.6957	0.6807
10	27 12	1 20	0.8333	0.8438	0.8214	0.6652
Avg.			0.8350	0.9125	0.7222	0.6602
Dev.			0.0123	0.0428	0.0708	0.0178

interesting to see that RF-CSCA-SVM has got the smallest standard deviation based on these five features. It validates the robustness and stableness of the proposed model. Therefore, these five best features were taken to build the predictive model for the subsequent experiment.

In order to verify the effectiveness of the proposed method, we conducted a comparative study between the RF-CSCA-SVM with other four SVM models based on different nature inspired metaheuristic algorithms including RF-SCA-SVM, RF-MFO-SVM, RF-GOA-SVM, and RF-BA-SVM. The detailed comparison of the five methods is shown in Figure 5. It is revealed that the RF-CSCA-SVM model is better than all other models in four evaluation indexes, and its standard deviation is also the smallest. It means that RF-CSCA-SVM model has better performance and stability in comparison with other models. In terms of the most important evaluation indicator ACC, RF-CSCA-SVM achieved the best result and the smallest standard deviation. RF-GOA-SVM achieved second place, followed by RF-MFO-SVM and RF-CSCA-SVM. The result obtained by RF-BA-SVM is the worst. In the other three evaluation indicators of sensitivity, specificity, and MCC, the results obtained by the RF-CSCA-SVM model are also much better than the other comparative counterparts. It is not difficult to find that the different results obtained by the model have obvious difference. The main reason lies in that there is a different parameter configuration obtained by each optimization algorithm for SVM. This also reveals that the parameters of SVM have a significant impact on the

classification performance of SVM. In sum, we can see that the proposed RF-CSCA-SVM can achieve better results or very competitive results than other involved counterparts in four performance metrics.

Figure 6 shows evolutionary curves of RF-CSCA-SVM and the other four models during the training stage. These curves are the average results of the 10 curves obtained by these models during the 10-fold cross-validation. As shown, it is apparent that the RF-CSCA-SVM model can quickly achieve a good performance during the model training process, which reveals that CSCA has a strong search capability and can find much more suitable parameters for SVM in an efficient manner. For RF-BA-SVM, it obtains the worst result compared with other models. The possible reason is that BA's global search ability is not strong enough to jump out of local optimum. Based on the analysis of the experimental results above, it is not difficult to find that the proposed CSCA strategy has improved a lot the performance of SVM compared to the original SCA. The main reason is owing to the fact that the embedded chaotic local search strategy has enhanced a lot the search capabilities of SCA.

To further evaluate the generalization capability of the proposed method, a hold-out way was conducted. The whole dataset was split into 80% and 20% for training and test, respectively. Owing to the randomness, the method has been run 10 times. The detailed results and the confusion matrix with different runs were recorded in Table 7. As

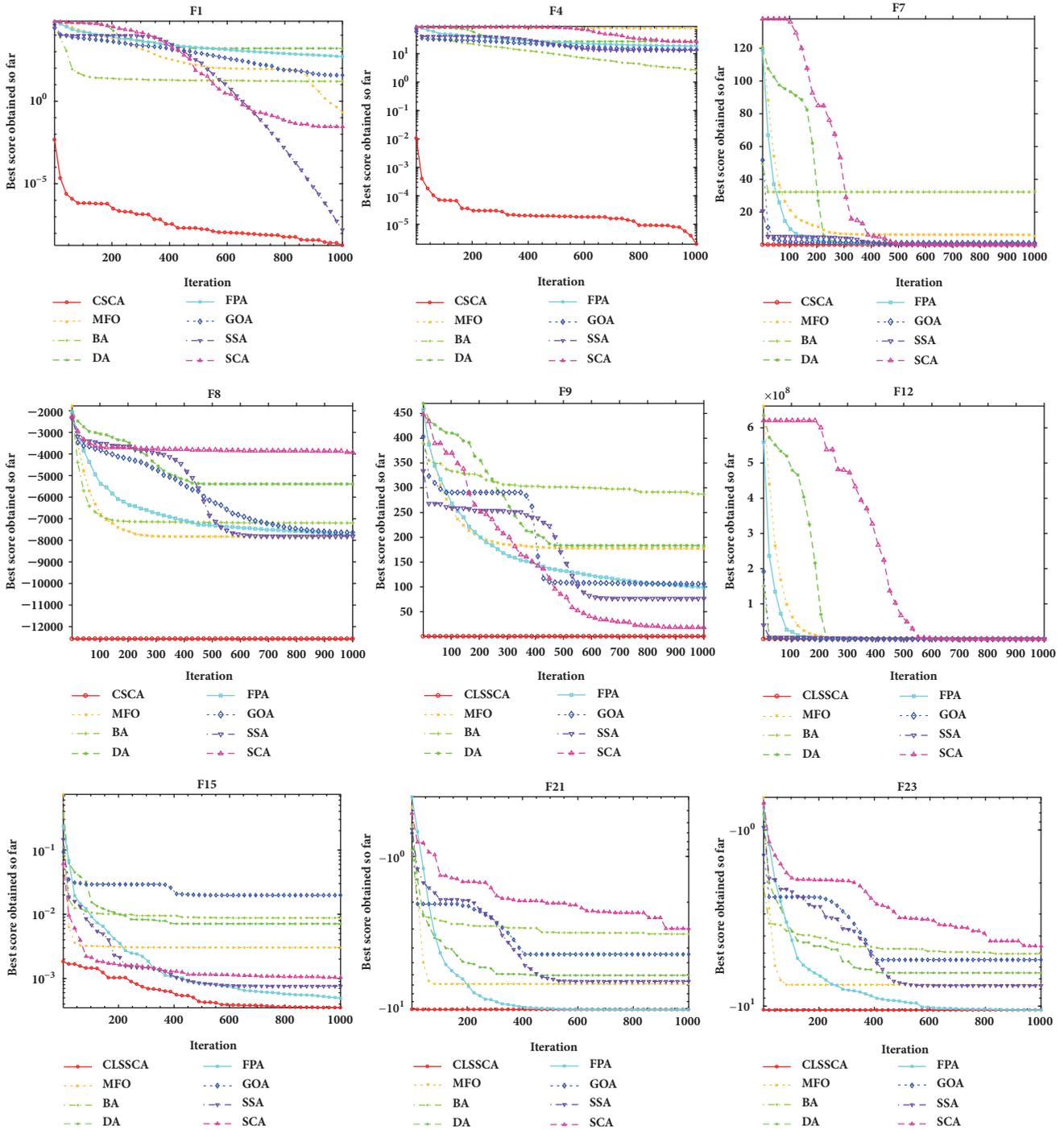


FIGURE 3: Convergence curve of CSCA and other algorithms.

shown, the test accuracy of 83.50% can be achieved, and the sensitivity, specificity, and MCC of the proposed RF-CSCA-SVM on the test set are 91.25%, 72.22%, and 0.6602, respectively. The results of the confusion matrix show that most entrepreneurship students can be accurately predicted as entrepreneurship. The recognition error rate mainly occurs when the students who are employed are misclassified as entrepreneurs.

5. Discussions

The study discovered some interesting results. From the experimental results, we can find that the most important features include major (F4), gender (F1), type of normal school students (F6), grade point average (F8), and total credits (F9); the influence of these features on the choice of entrepreneurial intention is relatively prominent. The

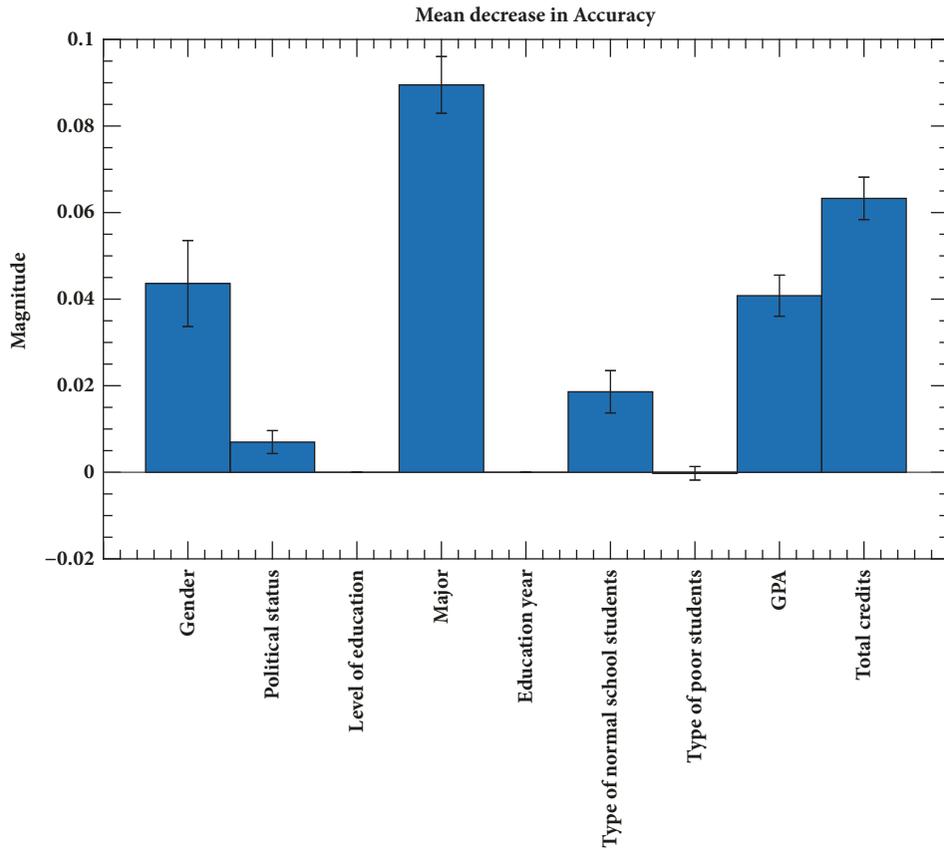


FIGURE 4: The importance of features evaluated by RF.

data show that different majors have obvious choice of entrepreneurial intention. On the whole, arts students have higher initiative intention than science students because they are more active in thinking, lower in employment, and more highly motivated to start their own businesses than students of science and engineering. Gender differences also have a significant impact on entrepreneurial intentions. The proportion of boys choosing to start a business is much higher than that of girls' maybe because boys are adventurous, and girls prefer stable jobs. Academic achievement and entrepreneurial choice have a clear impact. The academic achievement mainly includes the total average score point and the total credit score, and the academic achievement is generally divided into three grades: good, middle, and next. The students whose academic achievement is in the middle stream do not have obvious advantage in the employment choice, and they will actively think about how to improve their employment possibility through various channels. The intention of starting a business is obviously stronger than that of the students in the other streams. Students with better academic performance choose stable, high-paying, or prestigious careers. Students with poor academic performance do not have a clear plan. Family situation also has a significant impact on entrepreneurial willingness and entrepreneurial behavior of college students. Students with lower socioeconomic status have higher entrepreneurial willingness. They hope to change the social class status through

personal efforts. At the same time, the family situation in the aspects of financial experience, personal connection, and other aspects can provide the individual college students with the ability and convenience to start a business, thereby improving the success rate of entrepreneurship. Due to the influence of traditional ideas, normal school students' posts are considered more stable and decent than other professions, and the concept of entrepreneurship is weak, so the normal school students' entrepreneurial intention and choice are more indifferent than those of nonnormal school students.

It should be noted that the present study has several limitations that require further discussion. Firstly, the samples involved in this study were limited. To get more accurate results, a larger number of consecutive samples are required to be collected to take part in training the more unbiased learning model. Second, the study was accomplished in a single university. Confirmation of the model in multicenter studies would make the model more reliable for decision support. Furthermore, the involved attributes are limited; future studies should seek to investigate more attributes which may have impact on the students' entrepreneurial intention.

6. Conclusions and Future Work

In this study, we established an improved SVM framework to predict the employment intentions of college graduates.

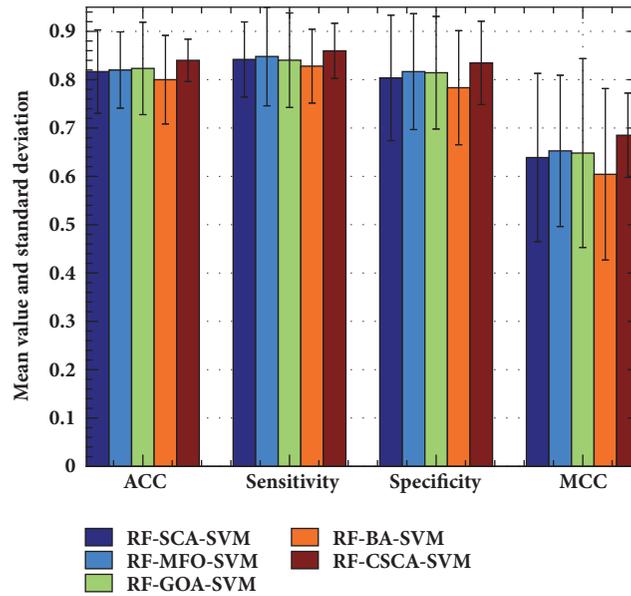


FIGURE 5: Classification performance obtained by the five methods in terms of ACC, MCC, sensitivity, and specificity.

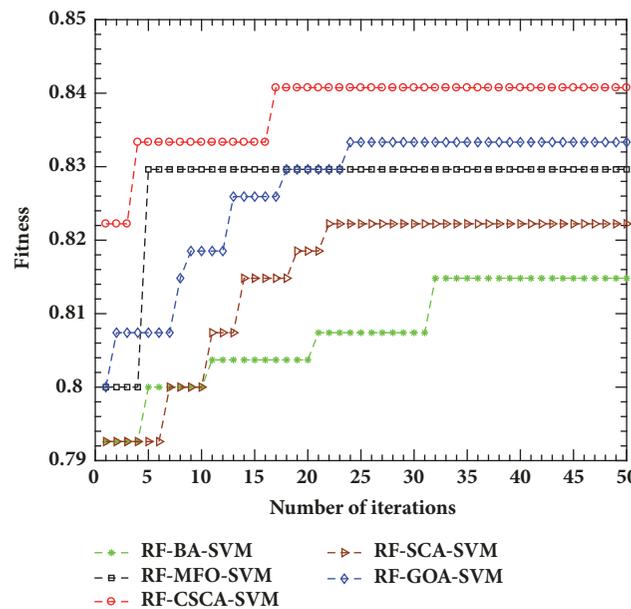


FIGURE 6: Comparison of convergence trends of several improved SVM methods based on swarm intelligence optimization.

In order to improve the accuracy of prediction, this paper firstly proposes to use RF to screen the key features in the data and further proposes an improved SCA strategy to tune the optimal parameters of SVM and finally uses the established CSCA-SVM model to predict new samples. Experimental results show that the proposed method has better classification performance than the SVM method based on other swarm intelligent optimization methods on the indicators of ACC, MCC, sensitivity, and specificity. Therefore, we can draw a preliminary conclusion that the proposed prediction framework can effectively predict students' employment

intention. In the future work, we plan to establish a set of decision support system based on the proposed framework to assist school departments to predict students' employment intention. In addition, we plan to collect more data samples in the future to improve the predictive performance of the proposed method.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This research is supported by the Zhejiang Provincial Natural Science Foundation of China (LY17F020012), Science and Technology Plan Project of Wenzhou, China (ZG2017019), and the Medical and Health Technology Projects of Zhejiang Province, China (2019315504).

References

- [1] W. Zang and R. Guo, "Application of data mining in employment instruction of undergraduate students," in *Proceedings of the 2010 IEEE International Conference on Intelligent Computing and Intelligent Systems, ICIS 2010*, pp. 772–775, China, October 2010.
- [2] Y. Liu, L. Hu, F. Yan, and B. Zhang, "Information gain with weight based decision tree for the employment forecasting of undergraduates," in *Proceedings of the 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, GreenCom-iThings-CPSCOM 2013*, pp. 2210–2213, China, August 2013.
- [3] L. Li, Y. Zheng, X. H. Sun, and F. S. Wang, "The application of decision tree algorithm in the employment management system," *Applied Mechanics and Materials*, vol. 543–547, pp. 1639–1642, 2014.
- [4] J. Wang, "Research of college students' employment data mining based on fuzzy decision tree (FDT)," *Agro FOOD Industry Hi Tech*, vol. 28, no. 3, pp. 2889–2893, 2017.
- [5] T. Mishra, D. Kumar, and S. Gupta, "Students' employability prediction model through data mining," *International Journal of Applied Engineering Research*, vol. 11, no. 4, pp. 2275–2282, 2016.
- [6] Y. Xu and B. Zhang, "Research on the employment option of college graduates based on bayesian algorithm of data mining and inclusion degree," *Advanced Science Letters*, vol. 11, no. 1, pp. 846–848, 2012.
- [7] N. A. B. A. Rahman, K. L. Tan, and C. K. Lim, "Supervised and unsupervised learning in data mining for employment prediction of fresh graduate students," *Journal of Telecommunication, Electronic and Computer Engineering*, vol. 9, no. 2–12, pp. 155–161, 2017.
- [8] P. Thakar, A. Mehta, and Manisha, "A unified model of clustering and classification to improve students' employability prediction," *International Journal of Intelligent Systems and Applications*, vol. 9, no. 9, pp. 10–18, 2017.
- [9] N. A. A. Rahman, K. L. Tan, and C. K. Lim, "Predictive analysis and data mining among the employment of fresh graduate students in HEI," in *Proceedings of the 2nd International Conference on Applied Science and Technology 2017, ICAST 2017*, Malaysia, April 2017.
- [10] S. Lanka and C. A. Pinto, "A research review of the student staging paradigm by Machine learning approach," *International Journal of Pure and Applied Mathematics*, vol. 119, no. 12, pp. 12849–12856, 2018.
- [11] K. Zhou and M. T. Jin, "The university students entrepreneurship evaluation system based on BP neural network," *Applied Mechanics and Materials*, vol. 556–562, pp. 6742–6745, 2014.
- [12] S. Mirjalili, "SCA: A Sine Cosine Algorithm for solving optimization problems," *Knowledge-Based Systems*, vol. 96, pp. 120–133, 2016.
- [13] S. Das, A. Bhattacharya, and A. K. Chakraborty, "Solution of short-term hydrothermal scheduling using sine cosine algorithm," *Soft Computing*, vol. 6, pp. 1–19, 2017.
- [14] A.-F. Attia, R. A. El Sehiemy, and H. M. Hasanien, "Optimal power flow solution in power systems using a novel Sine-Cosine algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 99, pp. 331–343, 2018.
- [15] B. Mahdad and K. Srairi, "A new interactive sine cosine algorithm for loading margin stability improvement under contingency," *Electrical Engineering*, vol. 100, no. 2, pp. 913–933, 2018.
- [16] D. R. Nayak, R. Dash, B. Majhi, and S. Wang, "Combining extreme learning machine with modified sine cosine algorithm for detection of pathological brain," *Computers & Electrical Engineering*, vol. 68, pp. 366–380, 2018.
- [17] H. Nenavath and R. K. Jatoth, "Hybridizing sine cosine algorithm with differential evolution for global optimization and object tracking," *Applied Soft Computing*, vol. 62, pp. 1019–1043, 2018.
- [18] J. Wang, W. Yang, P. Du, and T. Niu, "A novel hybrid forecasting system of wind speed based on a newly developed multi-objective sine cosine algorithm," *Energy Conversion and Management*, vol. 163, pp. 134–150, 2018.
- [19] W. Deng, R. Yao, H. Zhao, X. Yang, and G. Li, "A novel intelligent diagnosis method using optimal LS-SVM with improved PSO algorithm," *Soft Computing*, pp. 1–18, 2017.
- [20] W. Deng, H. Zhao, X. Yang, J. Xiong, M. Sun, and B. Li, "Study on an improved adaptive PSO algorithm for solving multi-objective gate assignment," *Applied Soft Computing*, vol. 59, pp. 288–302, 2017.
- [21] W. Deng, H. Zhao, L. Zou, G. Li, X. Yang, and D. Wu, "A novel collaborative optimization algorithm in solving complex optimization problems," *Soft Computing*, vol. 21, no. 15, pp. 4387–4398, 2017.
- [22] X. Zhao, X. Zhang, Z. Cai et al., "Chaos enhanced grey wolf optimization wrapped ELM for diagnosis of paraquat-poisoned patients," *Computational Biology and Chemistry*, 2018.
- [23] Q. Zhang, H. Chen, J. Luo, Y. Xu, C. Wu, and C. Li, "Chaos Enhanced Bacterial Foraging Optimization for Global Optimization," *IEEE Access*, vol. 6, pp. 64905–64919, 2018.
- [24] J. Luo, H. Chen, Q. zhang, Y. Xu, H. Huang, and X. Zhao, "An improved grasshopper optimization algorithm with application to financial stress prediction," *Applied Mathematical Modelling: Simulation and Computation for Engineering and Environmental Systems*, vol. 64, pp. 654–668, 2018.
- [25] L. Shen, H. Chen, Z. Yu et al., "Evolving support vector machines using fruit fly optimization for medical data classification," *Knowledge-Based Systems*, vol. 96, pp. 61–75, 2016.
- [26] S. Mirjalili, "Moth-flame optimization algorithm: a novel nature-inspired heuristic paradigm," *Knowledge-Based Systems*, vol. 89, pp. 228–249, 2015.
- [27] X.-S. Yang, "A new metaheuristic bat-inspired algorithm," in *Nature Inspired Cooperative Strategies for Optimization (NICSO 2010)*, J. R. Gonzalez, D. A. Pelta, C. Cruz, G. Terrazas, and N. Krasnogor, Eds., vol. 284 of *Studies in Computational Intelligence*, pp. 65–74, Springer, Berlin, Germany, 2010.

- [28] S. Saremi, S. Mirjalili, and A. Lewis, "Grasshopper optimisation algorithm: theory and application," *Advances in Engineering Software*, vol. 105, pp. 30–47, 2017.
- [29] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [30] X. Wang, Z. Wang, J. Weng, C. Wen, H. Chen, and X. Wang, "A New Effective Machine Learning Framework for Sepsis Diagnosis," *IEEE Access*, vol. 6, pp. 48300–48310, 2018.
- [31] B. Schölkopf and A. J. Smola, "Support Vector Machines," *Data Mining Knowledge Discovery Handbook*, vol. 1, no. 3, pp. 283–289, 2003.
- [32] H.-L. Chen, B. Yang, J. Liu, and D.-Y. Liu, "A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis," *Expert Systems with Applications*, vol. 38, no. 7, pp. 9014–9022, 2011.
- [33] H. L. Chen, B. Yang, S. J. Wang et al., "Towards an optimal support vector machine classifier using a parallel particle swarm optimization strategy," *Applied Mathematics and Computation*, vol. 239, pp. 180–197, 2014.
- [34] H.-L. Chen, D.-Y. Liu, B. Yang, J. Liu, and G. Wang, "A new hybrid method based on local fisher discriminant analysis and support vector machines for hepatitis disease diagnosis," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11796–11803, 2011.
- [35] M. Pontil and A. Verri, "Support vector machines for 3D object recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 6, pp. 637–646, 1998.
- [36] O. Chapelle, P. Haffner, and V. N. Vapnik, "Support vector machines for histogram-based image classification," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 10, no. 5, pp. 1055–1064, 1999.
- [37] H. Drucker, D. Wu, and V. N. Vapnik, "Support vector machines for spam categorization," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 10, no. 5, pp. 1048–1054, 1999.
- [38] S. Hua and Z. Sun, "A novel method of protein secondary structure prediction with high segment overlap measure: support vector machine approach," *Journal of Molecular Biology*, vol. 308, no. 2, pp. 397–407, 2001.
- [39] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Machine Learning*, vol. 46, no. 1–3, pp. 389–422, 2002.
- [40] C. Huang, L. S. Davis, and J. R. G. Townshend, "An assessment of support vector machines for land cover classification," *International Journal of Remote Sensing*, vol. 23, no. 4, pp. 725–749, 2002.
- [41] Z. Huang, H. Chen, C.-J. Hsu, W.-H. Chen, and S. Wu, "Credit rating analysis with support vector machines and neural networks: a market comparative study," *Decision Support Systems*, vol. 37, no. 4, pp. 543–558, 2004.
- [42] J. Mourão-Miranda, A. L. W. Bokde, C. Born, H. Hampel, and M. Stetter, "Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data," *NeuroImage*, vol. 28, no. 4, pp. 980–995, 2005.
- [43] T. Kapitaniak, "Continuous control and synchronization in chaotic systems," *Chaos, Solitons & Fractals*, vol. 6, no. C, pp. 237–244, 1995.
- [44] D. Jia, Y. Jiao, and J. Zhang, "Satisfactory design of IIR digital filter based on chaotic mutation particle swarm optimization," in *Proceedings of the 3rd International Conference on Genetic and Evolutionary Computing, WGEC 2009*, pp. 48–51, China, October 2009.
- [45] G.-C. Liao and T.-P. Tsao, "Application of a fuzzy neural network combined with a chaos genetic algorithm and simulated annealing to short-term load forecasting," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 3, pp. 330–340, 2006.
- [46] Q. Lü, G. Shen, and R. Yu, "A chaotic approach to maintain the population diversity of genetic algorithm in network training," *Computational Biology and Chemistry*, vol. 27, no. 3, pp. 363–371, 2003.
- [47] R. M. May, "Simple mathematical models with very complicated dynamics," *Nature*, vol. 261, no. 5560, pp. 459–467, 1976.
- [48] C. Chang and C. Lin, "LIBSVM: a Library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, article 27, 2011.
- [49] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proceedings of the IJCAI-95*, C. S. Mellish, Ed., pp. 1137–1143, Morgan Kaufmann, Los Altos, CA, USA, 1995.

