IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received April 27, 2019, accepted May 12, 2019, date of publication May 20, 2019, date of current version June 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2918026

Predicting Intentions of Students for Master Programs Using a Chaos-Induced Sine Cosine-Based Fuzzy K-Nearest Neighbor Classifier

AIJU LIN[1], QUANQUAN WU[1], ALI ASGHAR HEIDARI[2,3], YUETING XU[1], HUILING CHEN[1], WUJUN GENG[4], YUPING LI[5], AND CHENGYE LI[5]

[1]College of Mathematics, Physics, and Electronic Information Engineering, Wenzhou University, Wenzhou 325035, China
[2]School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran 1439957131, Iran
[3]Department of Computer Science, School of Computing, National University of Singapore, Singapore
[4]Department of Anesthesiology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, China
[5]Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, China

Corresponding authors: Huiling Chen (chenhuiling.jlu@gmail.com), Wujun Geng (gengwujun@126.com), and Chengye Li (lichengye41@126.com)

This work was supported in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LY17F020012, in part by the Science and Technology Plan Project of Wenzhou, China, under Grant ZG2017019, in part by the Medical and Health Technology Projects of Zhejiang Province, China, under Grant 2019315504, in part by the Zhejiang 13th Five-Year Plan Teaching Reform Project under Grant jg20180188, in part by the National Education Information Technology Research Project under Grant 186130020, and in part by the Wenzhou Medical University Teaching Reform Project YBJG201819.

ABSTRACT One of the essential tasks for the planning and development of talents training programs in different colleges of universities is to find how we can reasonably guide students to pursue a master's degree concerning their comprehensive situations. The purpose of this study is to develop a modified fuzzy k-Nearest Neighbor (FKNN) framework to predict the college students' intentions for master programs in advance, that is, students choose to attend the postgraduate exam or find a job after graduation. The proposed integrated framework combines the random forest (RF), FKNN, and a new chaos-enhanced sine cosine-inspired algorithm (CESCA). In this model, RF is employed to evaluate the importance of features in the dataset, while the FKNN is utilized to establish the relationship framework between the features and the college students' decisions to earn a master's degree or not. The proposed CESCA is utilized to tune the key parameters of the FKNN automatically. All eight variants of SCA have been rigorously compared based on 13 benchmark problems to validate the effectiveness of the proposed CESCA. Then, the CESCA-based FKNN (CESCA-FKNN) has been further compared against the other three classical classifiers in terms of four common performance metrics. The experimental results indicate that the proposed CESCA-FKNN can obtain the best classification accuracy. The results indicate that the established adaptive FKNN framework can be served as a powerful tool for college students' intention before pursuing a master's degree.

INDEX TERMS Fuzzy k-nearest neighbor method, sine cosine algorithm, chaos theory, students' intentions for master programs, feature selection.

I. INTRODUCTION

In recent years, the number of postgraduate enrollments in Chinese universities has steadily increased. According to statistics from the Ministry of Education, several applicants for postgraduate entrance examination reached 2.38 million in 2018, an increase of 18.4% over 2017 million[1]. With the overall transformation and upgrading of China's national economy, the social demand for highly educated talents is increasing. The role of graduate and above literary talents is becoming more and more evident in the context of today's

The associate editor coordinating the review of this manuscript and approving it for publication was Huazhu Fu.

[1]Please visit: https://yz.chsi.com.cn/yzzt/fxbg2018, available at January 19, 2019

technological revolution and economic globalization. It can be predicted that the fever of the postgraduate entrance examination will last for a long time in the future.

At present, most students often hesitate to take part in the postgraduate entrance examination and employment. It is difficult for them to make rational choices in time. Therefore, it is necessary to make an in-depth analysis of the logical choice of postgraduate entrance examination and employment of college students to find an effective way for college students to choose the direction of graduation scientifically. At present, colleges and universities have accumulated a lot of data. Through mining and analyzing these data, an intelligent forecasting model is established to find out the factors that affect the graduate students' choice of postgraduate entrance examination or employment. Then, any potential correlation between the factors is further analyzed to guide the students for a better selection of postgraduate entrance examination or jobs.

So far, few pieces of literature used data mining techniques to mine the postgraduate students' relevant data. Koutina and Kermanidis [1] proposed to use the machine learning techniques for predicting the academic performance tendency of postgraduate students. The experimental results of their work demonstrated that Naive Bayes and 1 Nearest Neighbor achieved the best prediction results. Wang [2] proposed a residual modified grey model combined with an improved neural network to forecast the postgraduates' employment confidence index. Chi and Lin [3] proposed to utilize the genetic algorithms for optimizing the initial weights and thresholds of BP neural network in dealing with the prediction of the postgraduate entrance examination results. The results demonstrated that the proposed method was better than the BP neural network alone, and it was better than the traditional k-Nearest Neighbor (KNN) and Naive Bayes methods, as well. We can see that there is no study towards establishing an intelligent prediction model for predicting students' intentions of choosing the postgraduate entrance examination or employment. For the first time, this work develops an intelligent prediction model based on the fuzzy k-Nearest Neighbor (FKNN) method to predict the students' intention for master programs.

The KNN method is the most popular non-parametric pattern classification method due to its simplicity and ease of implementation. In KNN, a class is obtained by the most basic category in its k-nearest neighbors. FKNN [4], [5] classifier is a fuzzy version of traditional KNN, which combines the fuzzy theory with the KNN algorithm and it has been studied extensively since the establishment of the original work. Owing to its excellent characteristics, FKNN has been applied to many fields of sciences, for instance, bankruptcy prediction problems [6], protein identification and prediction problems [7], [8], slope collapse prediction problems [9], and medical diagnosis problems [10], [11]. The neighborhood size (k) and the fuzzy strength parameter (m) are fundamental determinants of the FKNN model. As it is obvious, the values of k and m affect the prediction results directly. It is difficult to determine the values of m and k because there is no theory and standard method that explains how to establish the fitting values of m and k. Hence, it is necessary to explore the full potential of FKNN by adjusting these parameters to exploit the maximum classification performance for students' intentions in the master programs. In this paper, the sine cosine algorithm (SCA) is employed and enhanced to tune the key parameters of FKNN automatically.

Mathematical models always need to a suitable technique to obtain the solution during a reasonable time [12]–[16]. The SCA is a new nature-inspired optimizer originally developed by Mirjalili [17]. Due to its simplicity and efficiency, this method has received extensive attention in different fields of research [18]–[22]. For example, the SCA was used to optimize parameters of an adaptive neuro-fuzzy inference system for forecasting the oil consumption in [23]. Moreover, the SCA has been applied to the field of power system economics to tackle the short-term hydrothermal scheduling cases [24]. Furthermore, it was used in a power system such as in [25] as an interactive process and applied it to specified important branches for improving the security of the power systems. However, like other metaheuristic algorithms [26]–[37], SCA has several drawbacks including easy to falling into local optima and slow convergence speed as well. To mitigate these deficiencies in dealing with the optimization problems, scholars tried to propose some modified variants of the original SCA. In 2018, Qu *et al.* [38] introduced a new, improved SCA by mixing two optimization mechanisms and a greedy Levy mutation strategy to avoid falling into the local optimum. Chegini *et al.* [39] combined the particle swarm optimizer (PSO), the updating equation in SCA and the levy flight to tackle the disadvantages of PSO such as falling into the local minimum. To obtain a better exploration of the search space and generate more accurate solutions, Elaziz *et al.* [40] used an improved version of SCA that considers opposition-based learning as an exploratory mechanism. Also, a combination of the SCA and the differential evolution (DE) algorithm was developed for tackling optimization problems and object tracking in [41].

Some researchers adopted the chaos mechanism and verified its constructive impact on improving the exploration and exploitation trends of a swarm-based optimizer. In order to further enhance the searching performance of SCA, eight different chaotic mapping strategies (CMS) were used in this paper to enrich the randomicity in the searching process of the standard SCA, thus, a series of SCA-based techniques were proposed, which are called 'LogisticSCA', 'IterativeSCA', 'CircleSCA', 'ChebyshevSCA', 'SingerSCA', 'SinusoidalSCA', 'SineSCA', and 'TentSCA'. Then, a series of comparative experiments based on 13 benchmark problems were conducted and the best method among all the proposed variants of basic SCA, which is called CESCA, was analyzed in the present study. The proposed CESCA was compared with several well-established optimizers using 13 classical benchmark functions to evaluate the performance of the proposed method, and the experimental results show that chaotic local search (CLS) strategy has indeed improved the efficacy of the basic SCA in a significant manner. Moreover, the random forest (RF) was employed to the feature evaluation procedure, and then, the CESCA was adopted to tune the key parameters of FKNN. It can be observed that the solution found by the proposed CESCA-FKNN is much better than that of the original SCA-FKNN. Also, the efficacy of the proposed CESCA-FKNN model was rigorously compared with other three well-regarded classifiers and three swarm intelligence algorithms-based FKNN approaches on the real-life dataset collected from Wenzhou University. The experimental results demonstrated that the proposed CESCA-FKNN had achieved better experimental results than other three methods in terms of the four standard performance metrics including Matthews Correlation Coefficients (MCC), classification accuracy (ACC), sensitivity and specificity.

The accuracy of predicting the college students' intentions for master programs was influenced to a great extent by the performance of the employed classifier. FKNN has been widely studied due to its excellent performance in dealing with some real-world problems. However, previous studies indicated that the performance of FKNN model was heavily influenced by two key parameters of the neighborhood size and the fuzzy strength parameter. Therefore, an enhanced SCA with chaotic mechanism was proposed to identify the two key parameters of FKNN to improve the prediction accuracy of college students' intentions for master programs. To the best of our knowledge, this is the first time that the FKNN has been employed to predict the college students' intentions for master programs.

The main contributions of this study are as follows:

a) First, to further make a delicate balance between the exploration and exploitation inclinations of SCA, we introduce several chaotic mechanisms for updating the key parameter of the primary method. This strategy helps conventional SCA to reveal a faster convergence rate and also achieves higher local optima avoidance.
b) Several variants of the chaos-based SCA are proposed with eight different chaotic maps, and the validation of the performance is based on 13 commonly used benchmark problems.
c) The improved CESCA strategy is successful in optimizing the two key parameters of FKNN. The resulting model, CESCA-FKNN, is applied to the prediction of students' intentions for master programs. Results of the proposed model show that the developed model outperforms other six effective machine learning methods.

The rest of this paper is organized as follows. Section 2 offers a brief description of the methodology including FKNN, SCA, and CLS. 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.

II. METHODS

This part will introduce the prediction framework of college students' intention of the postgraduate entrance examination, namely CESCA-FKNN. The main structure is shown in Figure 1. The whole process is divided into three parts. The first part is to normalize the data and evaluate the feature importance one by one by RF. The second part is to construct an optimized FKNN model with the chaos-enhanced SCA, and at the same time, the optimal feature set is obtained by the incremental evaluation of different feature sets. The main task of the third part is to use the optimal model constructed in the previous stage for predicting the new data samples.

It should be noted that the difference between the existing method and our proposed method is that we have proposed to use the chaotic mechanisms to update the parameter of r_3 in SCA for the first time, and its effectiveness was rigorously validated on a set of benchmarks including the unimodal and multimodal landscapes. Then, the proposed CESCA was used to tune the two key parameters of FKNN model in an adaptive manner. In the literature, there are few works which hybridize other swarm-intelligence methods for optimizing the parameters of FKNN, including PSO [42] and grey wolf optimizer (GWO) [43]. However, this is the first time that these improved SCA-based optimizers are proposed and assessed to train the FKNN model for dealing with our datasets and for the subject of this research.

A. PREDICTION ENGINE: FUZZY K-NEAREST NEIGHBOR (FKNN)

KNN is a generalization of the nearest neighbor method, that is, for the samples of unknown categories, the k-nearest neighbors in training samples are selected and count the number of samples in k-nearest neighbors. The category with the largest number of samples is regarded as the category of the unknown class samples. Let the number of samples be N, and the number of samples from w_i class is N_i. If the number of samples belonging to $\omega_1, \omega_2, \ldots, \omega_c$ among k-nearest neighbors are $k_1, k_2, \ldots, k_c$, respectively, then, the discriminant function can be defined as follows:

$$g_i(x) = k_i, \quad i = 1, 2, \ldots c \tag{1}$$

Here, c indicates the number of class. The decision rule is as follows:

$$\text{if } g_i(x) = max(k_i), \quad \text{then } x \in w_i \tag{2}$$

The KNN classification regards the weight of training samples as the same when considering training samples, and does not reflect the distance difference of training samples to classification decision, which is inconsistent with the actual situation. Therefore, considering introducing the idea of fuzzy classification, namely, the fuzzy KNN (FKNN), this problem is solved by adding the degree of membership function to training samples [4], [5]. The degree of the membership

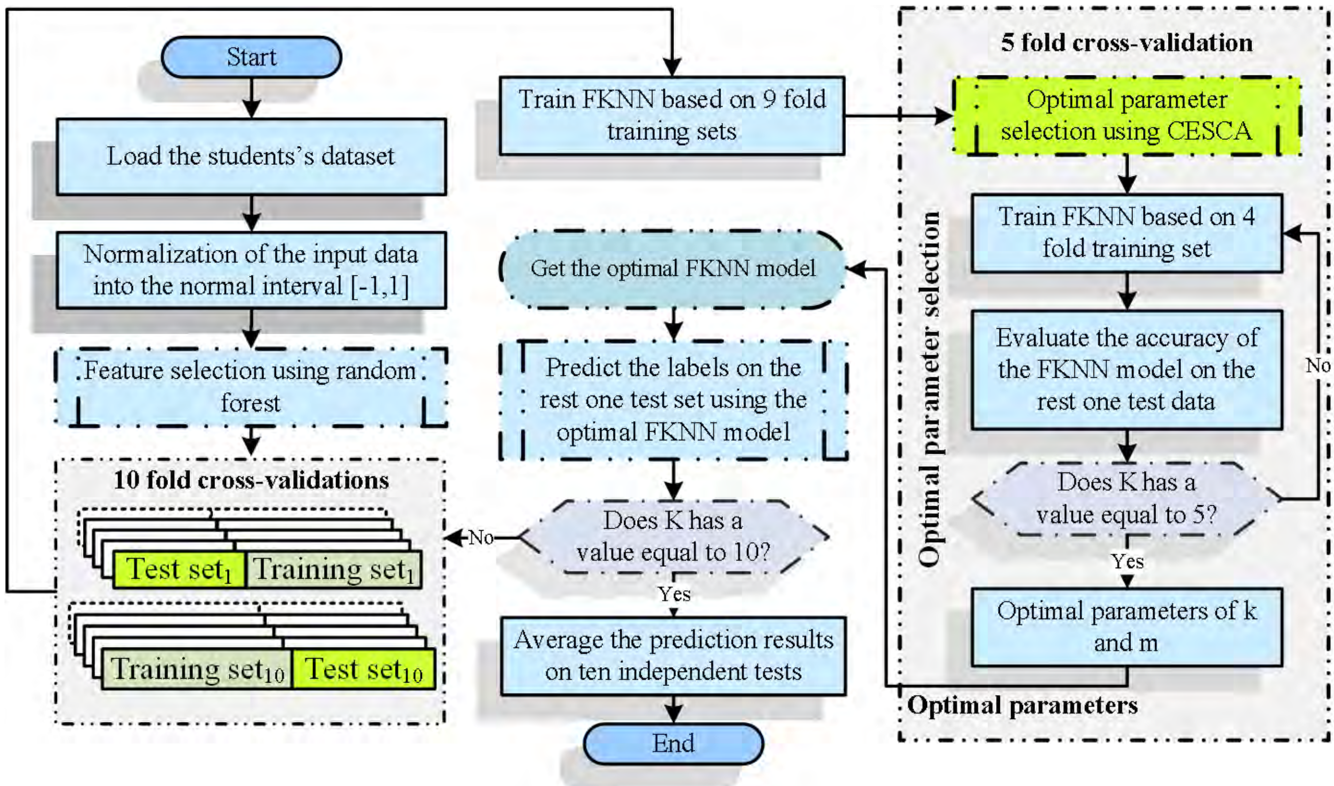

FIGURE 1. Flowchart of the proposed CESCA-FKNN framework.

function of the sample to be classified is defined as follows:

$$u_i(x) = \frac{\sum_{j=1}^{k} u_{ij}(1/\left\|x - x_j\right\|^{2/(m-1)})}{\sum_{j=1}^{k}(1/\left\|x - x_j\right\|^{2/(m-1)})} \tag{3}$$

where j $= 1, 2, \ldots, k$ and m is the fuzzy strength coefficient, representing the weight of the distance between each neighbor and the test sample; $\left\|x - x_j\right\|$ is indicates the distance between x and its jth nearest neighbor x_j. In this study, the Euclidean metric is used for $\left\|x - x_j\right\|$. u_{ij} is the membership degree of the pattern x_j belonging to the training sample of the ith category, among the k nearest neighbors of x. In this paper, the constrained fuzzy membership was adopted. Hence, the membership of x_k (the k-nearest neighbors of each training pattern) in each class is given as,

$$u_{ij}(x_k) = \begin{cases} 0.51 + \left(n_j/k\right) * 0.49, & j = i \\ \left(n_j/k\right) \times 0.49, & j \neq i \end{cases} \tag{4}$$

Here, the value n_i is the number of neighbors found which belong to the jth class. The brief steps to calculate the membership degree can be listed as the following steps:

a) Calculate the distance ($\left\|x - x_j\right\|$) between any two samples in the training sample set, and this paper uses the Euclidean distance.
b) For each sample, k samples which are nearest to it are selected, and their category information is counted, and the membership degree of training samples is calculated.
c) For the samples to be classified, calculate the distance between it and all the training samples, and also select the k samples closest to the distance. Rules in Eqs (1)-(4) obtain the membership degree of the samples to be classified.

B. THE HARMONIOUS SCA WITH CLS

Swarm-intelligence methods often inspire nature-based phenomena. However, it is possible to develop a mathematical model for an optimizer by inspiration from mathematical concepts. The SCA [17] is a new metaheuristic algorithm that starts the process with a set of random candidate solutions and updates the search agents outwards or towards the optimal solution. The formation of the movement is based on sine and cosine functions [29]. Random agents explore different regions of the search space while the core function returns a value less than -1 or greater than 1. Promising areas of the search space are exploited while the value returned by the core function is between -1 and 1. The updated agent locations ($X_i, i = 1...N$) used in the SCA are adjusted by Eq. (5) to make sure that solutions always update their positions around the optimal solution obtained so far.

$$\mathrm{X}_i^{t+1} = \mathrm{X}_i^t + r_1 \times \sin(r_2) \times \left|r_3\mathrm{P}_i^t - \mathrm{X}_i^t\right| \tag{5}$$

$$\mathrm{X}_i^{t+1} = \mathrm{X}_i^t + r_1 \times \cos(r_2) \times \left|r_3\mathrm{P}_i^t - \mathrm{X}_i^t\right| \tag{6}$$

TABLE 1. Eight chaotic maps.

NO	Name	Map (a= 0.5 and b=0.2, it can generate chaotic (0, 1))
Map_1	Circle map	$x_{k+1} = x_k + b - (a/2\pi)\sin(2\pi k) mod(1)$
Map_2	Chebyshev map	$x_{k+1} = \cos(k cos^{-1}(x_k))$
Map_3	Iterative map	$x_{k+1} = \sin(a\pi/x_k), a \in (0,1)$
Map_4	Logistic map	$x_{k+1} = a x_k(1 - x_k)$
Map_5	Sine map	$x_{k+1} = a/4\sin(\pi x_k), 0 < a \leq 4$
Map_6	Singer map	$x_{k+1} = \mu(7.86x_k - 23.31x_k^2 + 28.75x_k^3 - 13.302875x)$
Map_7	Sinusoidal map	$x_{k+1} = a x_k^2 \sin(\pi x_k)$
Map_8	Tent map	$x_{k+1} = \begin{cases} x_k/0.7, x_k < 0.7 \\ 10/3(1-x_k), x_k \geq 0.7 \end{cases}$

where X_i^t is the position of the current solution in the i-th dimension at t iteration, and P_i^t is the position of the target point in the i-th dimension at t iteration. r_1, r_2 and r_3 are random variables and || shows the absolute value. The range of sine and cosine is updated adaptively by using the parameter r_1 in Eq. (5) to balance exploration and exploitation abilities. Therefore, the random parameter r_1 defines the region of the next solution; this region may be either in the space between X_i^t and P_i^t or outside them.

$$r_1 = a - t \cdot \frac{a}{T} \quad (7)$$

where a is a constant, T means the maximum number of iterations. Then, Eq. (5) and Eq. (6) are combined by a random number $r_4 \in [0, 1]$ in SCA as follows,

$$X_i^{t+1} = \begin{cases} X_i^{t+1} = X_i^t + r_1 \times \sin(r_2) \times \left| r_3 P_i^t - X_i^t \right|, & r_4 < 0.5 \\ X_i^{t+1} = X_i^t + r_1 \times \cos(r_2) \times \left| r_3 P_i^t - X_i^t \right|, & r_4 \geq 0.5 \end{cases} \quad (8)$$

Due to the use of sine and cosine in the updating process in Eq. (8), this algorithm is named as SCA. The CLS mechanism was introduced into SCA to develop the basic SCA and to mitigate the possibility of the local optima stagnation, which is the core disadvantage of the basic SCA. Chaos is sensitive to its initial conditions. Also, it has two basic characteristics: randomness and ergodicity. The randomness explains that the variations of a chaotic system are random. Therefore, the optimization of the objective function can make full use of these characteristics. In this paper, eight CMS was used to produce chaotic sequences for updating the r_3 in the classical SCA. The eight maps are listed in Table 1.

The CLS mechanism, owing to its strong randomicity and ergodicity, can effectively help SCA to jump out of the local optima and avoid the premature convergence. 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 CESCA is shown in Figure 2.

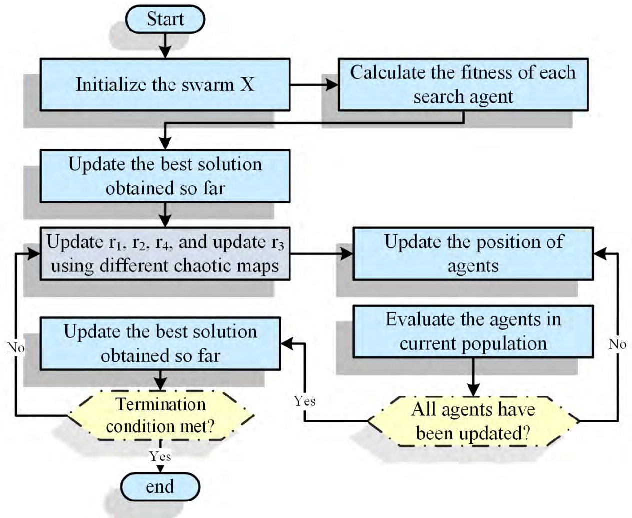

FIGURE 2. Flowchart of CESCA.

The computational complexity of the CESCA mainly depends on five producers: initialization, fitness evaluation, sequences of r_3 updating, the best solution searching and search agents updating. Hence, the overall computational complexity is O(CESCA) = O(*Initialization*) + O(*Calculating the fitness values of agents*) + O (*Producing the chaotic sequences*) + O(*Searching the best solutions of all search agents*) + O(*Updating the location vectors of all agents*). Considering that the calculation of time complexity is based on the specific optimization problem, the main focus of the complexity is on the other four producers. The computational complexity of the initialization process is $O(N)$. The time complexity of the chaotic sequences is $O(g)$, and the best location of all agents is $O(N)$, where g is the maximum number of iterations.The computational complexity of the updating mechanism is $O(g\times N\times D)+O\times(g\times N)$, which is composed of searching the best solution and updating the positions of all search agents, where D is the dimension of specific problems. Therefore, the final computational complexity of CESCA is $O(2N+g+N\times g(1+D))$.

The general procedure of CESCA is as follows:

III. EXPERIMENTAL DESIGNS

A. DATA COLLECTION AND DESCRIPTION

The data obtained in this study were from Wenzhou University. To achieve a balance between graduate employment samples and postgraduate entrance examination samples, 351 of samples received postgraduate degrees, and 351 graduate employment samples were randomly selected as experimental samples. So, there were a total of 702 students involved in this research. We analysis of the subjects' gender, type of standard school students, Grade Point Average (GPA), total credits, college English course, advanced mathematics course, programming language course, college physics

TABLE 2. Description of the 12 attributes.

Attributes	Name	Description
F1	Gender	Male and female students are represented by 1 and 2, respectively.
F2	Type of Normal School Students	It is divided into normal students and non-normal students, represented by 1 and 2, respectively.
F3	Grade Point Average (GPA)	GPA is a way for the school to assess students' learning quality. The score is within 0-4.
F4	Total Credits	It is a unit of measurement used to calculate students' learning volume. The more credits students receive, the more they learn.
F5	College English Course	Total Mark of the College English course. The score is between 0 and 100.
F6	Advanced Mathematics Course	It is an important basic course in science and engineering colleges. The score is within 0-100.
F7	Programming Language Course	It is a basic course for students of science and technology. The range of score is 0-100.
F8	College Physics	It is the fundamental courses of computer science and technology in universities. The score is also within 0-100.
F9	A score of Modern Chinese history	It is one of the fundamental courses in the new scheme of the Marxist theory curriculum. The score of the course for students is in the range of 0-100.
F10	Graduation Project or Thesis	It is the quality of graduation design paper, and it is divided into high level, middle level, low level and below standard level, represented by 1, 2, 3 and 4, respectively.
F11	College English Test Band Six (CET-6)	CET-6 is one of the tests to examine the English level of college students whose major is not English. The score of CET-6 for students is in the range of 0-710 intervals.
F12	College English Test Band Four (CET-4)	CET-4 is also an important standard to test students' English level. The score is within 0-710 intervals.

Procedure of CESCA

Begin
- Initialize a set of search agents (solutions)
- **Evaluate** each of the search agents by the objective function
- **Update** the best solution obtained so far ($P = X^*$)
- **Do until**
  - **Update** r_1, r_2, r_4
  - **Update** r_3 using eight different CMS
  - **Update** the positions of search agents using Eq. (8)
  - **Evaluate** each of the search agents by the objective function
- **End do**
- **Return** the best solution obtained so far as the global optimum

End

course, score of modern Chinese history course, graduation project or thesis, College English Test Band Six (CET-6), and College English Test Band Four (CET-4). This study intends to evaluate the importance and interrelationships of these twelve attributes to establish a predictive model for decision support. A detailed description of the twelve attributes is shown in Table 2.

To validate the effectiveness of the proposed method, firstly, the effectiveness of chaotic variants of SCA was compared on common unimodal and multimodal benchmark functions. The best one was selected to compare with other competitive metaheuristics including grasshopper optimization algorithm (GOA) [44], dragonfly algorithm (DA) [45], bat algorithm (BA) [46], moth-flame optimization (MFO) [47] and salp swarm algorithm (SSA). Secondly, we have applied the chaotic SCA to optimize the critical parameters of FKNN. The resultant CESCA-FKNN was compared against the SCA-FKNN and several popular machine learning methods such as RF, KELM, and SVM. The whole experiment was done in the MATLAB 2014b environment. We implemented CESCA-FKNN and SCA-FKNN from scratch. LIBSVM [48] was utilized for SVM implementation; the source code from http://www3.ntu.edu.sg/home/egbhuang was used for KELM implementation; the code from https://code.google.com/archive/p/randomforest-matlab was taken for RF implementation. 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. The k-fold cross-validation (CV) was used to evaluate the classification accuracy [49]. The range of two parameters in SVM and KELM, penalty factor C and kernel width γ are both set as $\{2^{-5}, 2^{-4}, \ldots, 2^4, 2^5\}$. The parameters in SCA, GOA, DA, BA, MFO, and SSA were set as those in their original papers. The initial value of all the chaotic maps was set as 0.7.

IV. EXPERIMENTAL RESULTS

A. CHAOTIC INITIALIZATION

In this part, the proposed chaotic SCA variants were verified on 13 benchmark functions which are listed in Tables 3 and 4. These two tables present seven unimodal (f_1-f_7) and

TABLE 3. Unimodal benchmark functions.

Function	Dim	Range	f_{min}
$f_1(x)=\sum_{i=1}^{n}x_i^2$	30	[-100, 100]	0
$f_2(x)=\sum_{i=1}^{n}\lvert x_i\rvert+\prod_{i=1}^{n}\lvert x_i\rvert$	30	[-10, 10]	0
$f_3(x)=\sum_{i=1}^{n}(\sum_{j-1}^{i}x_j)^2$	30	[-100, 100]	0
$f_4(x)=max_i\{\lvert x_i\rvert, 1\le i\le n\}$	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
$f_7(x)=\sum_{i=1}^{n}ix_i^4+random[0,1)$	30	[-1.28, 1.28]	0

TABLE 4. Multimodal benchmark functions.

Function	Dim	Range	f_{min}
$f_8(x)=\sum_{i=1}^{n}-x_i\sin(\sqrt{\lvert x_i\rvert})$	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
$f_{10}(x)=-20\exp\left\{-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i}\right\}-exp\{\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)\}+20+e$	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
$f_{12}(x)=\frac{\pi}{n}\{10\sin(ay_1)+\sum_{i=1}^{n-1}(y_i-1)^2[1+10sin^2(\pi y_{i+1})]+(y_n-1)^2+$ $0 \quad -a<x_i<a$ $k(-x_i-a)^m \quad x_i<-a$	30	[-50,50]	0
$f_{13}(x)=0.1\{sin^2(3\pi x_i)+\sum_{i=1}^{n}(x_i-1)^2[1+sin^2(3\pi x_i+1)]+(x_n-1)^2[1+sin^2(2\pi x_n)]+\sum_{i=1}^{n}\mu(x_i,5,100,4)$	30	[-50,50]	0

six multimodal (f_8-f_{13}) functions and include the following information for each test function: the function equation, the dimensionality of the specific optimization problem, the range of optimization variables and the optimal values.

Nine algorithms are compared in detail to evaluate the influence of eight CMS, the. Table 5 lists the average results (mean) and the standard deviation (std) of the best solution found by the nine algorithms over 30 independent runs. For a fair comparison purpose, all methods are implemented in the same testing environment on the same computing platform. The population size, dimensions and number of maximum iterations for the eight variants of SCA and basic SCA were set as 30, 30 and 500, respectively. As for the parameters of population number, iteration number and run times, we have tested by trial and error and determined the best parameters. Because too many iterations will consume a lot of CPU time. Therefore, we finally determined that 30 populations, 500 iterations, and 30 random runs were used in this paper.

Table 5 shows that all variants of SCA outperform the original algorithm. It indicates that the CMS leads the original SCA to move towards a better solution. The performance of the original SCA, LogisticSCA, IterativeSCA, CircleSCA, ChebyshevSCA, SingerSCA, SinusoidalSCA, SineSCA and TentSCA show a slight difference with each other. Inspecting the detailed results of algorithms on 13 problems in the table, the comparison results of the nine methods indicates that TentSCA has the smallest mean index on 5 out of 13 benchmarks problems. In the unimodal (f_2, f_4) and multimodal (f_9-f_{11}) functions, the TentSCA provides exact optimum results. For the rest of the variants, the results are competitive but TentSCA is far better. Apart from these, the Wilcoxon signed rank test [50], [51] has been used to evaluate the significant improvement of the TentSCA over the other seven variants of SCA and the original SCA. And the symbols of "+", "=" and "-" indicate that TentSCA is superior to, equal to, and inferior to other eight peers, respectively. In terms of the "+/=/-", we can observe that the proposed TentSCA is significantly better than the basic SCA, IterativeSCA, ChebyshevSCA, SingerSCA, SinusoidalSCA on 12,7, 7, 8 and 7 out of 13 functions, and inferior to them on 1,1,2,4 and 3 out of 13 ones, respectively. Moreover, the result of TentSCA is significantly better or equal than those obtained by the LogisticSCA, CircleSCA, and SineSCA in dealing with all functions. Furthermore, the Friedman test has also been employed to estimate the performance of all the variants of SCA for further statistical comparison, and the average ranking value (ARV) is reported in the results. It can be seen from Table 5, TentSCA obtains the lowest ARV for 13 functions, followed by the LogisticSCA, CircleSCA, SineSCA, IterativeSCA, ChebyshevSCA, SingerSCA, SinusoidalSCA and the basic SCA. In short, we can conclude that the proposed TentSCA achieves the best search performance and is well capable of escaping the local optimum value than all other competitors.

Also, to intuitively observe the difference of performance between eight different strategies, convergence curves of LogisticSCA, IterativeSCA, CircleSCA, ChebyshevSCA, SingerSCA, SinusoidalSCA, SineSCA, TentSCA, and original SCA on some typical benchmark functions are also provided in Figure 3. For f_1, f_4, and f_7, all the variants give competitive convergence trends, but TentSCA delivers the best result among all. For the multimodal benchmark function f_{11}, the TentSCA surpasses all other proposed versions. Compared with the other eight techniques, the TentSCA has more research value. It is also noteworthy that TentSCA renamed CESCA is selected as the best variant to compare with other methods in the following experiment.

The performance of the new method was compared with five recently proposed optimization algorithms on 13 classical benchmark tests to validate the proposed CESCA. The compared methods are GOA, DA, BA, MFO, and SSA. The same experimental condition of the competitors and

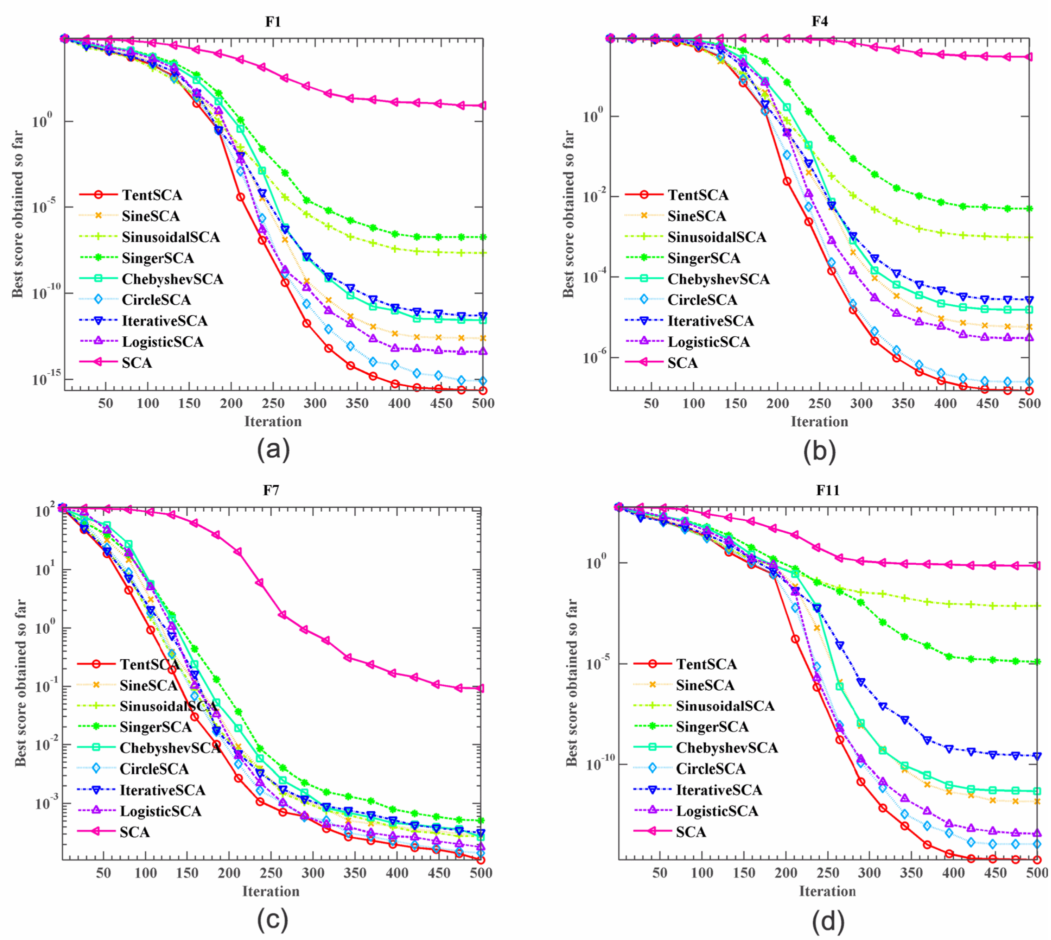

FIGURE 3. Convergence curves of some selected benchmark functions. (a) Convergence curve of F1. (b) Convergence curve of F4. (c) Convergence curve of F7. (d) Convergence curve of F11.

the proposed CESCA was adopted. As shown in Table 6, we can see that, on average, the presented CESCA obtains the lowest values for 12 problems, and it converges to the optimal solution on f_9. Followed by MFO, it provides the lowest mean value on one function. In terms of the statistic results of Wilcoxon's signed rank test, we can observe that the developed CESCA is significantly better than SSA, GOA, DA, and MFO on 12 out of 13 functions, and inferior to them on one case, respectively. Similarly, CESCA is superior to BA on 11 out of 13 problems, inferior to it on one case, and equal to them on one case. Moreover, the proposed algorithm achieves the best performance among all these competitors from the point of view. Therefore, we can say that CESCA produces the best results for these benchmark problems. The average results of the CESCA and all the involved algorithms were also compared by using the Friedman test. According to the ARV values of different methods, it can be observed that the proposed CESCA has the best performance for handling these benchmark problems, followed by GOA, BA, MFO, and SSA, while DA has the worst searching capability.

Furthermore, to clearly show the superiority of the CESCA, the evolution progress of SSA, GOA, DA, BA, MFO, and CESCA on some standard benchmarks are also provided in Figure 4. It can be detected that the proposed method has the ability of rapid convergence and it can be superior to all other competitors in realizing f_1, while SSA, GOA, DA, BA, and MFO cannot improve the quality of solutions in solving f_1 seen throughout more explorative steps. This trend also can be seen in dealing with the unimodal benchmark function (f_3) and multimodal problems (f_{10} and f_{11}). For f_5 and f_7, CESCA has converged so fast during few searching steps. To sum up, it can be indicated that the Tent mapping strategy is beneficial to improve the solution accuracy of SCA.

B. PREDICTION RESULTS OF STUDENTS' INTENTIONS FOR MASTER PROGRAMS

In this experiment, the RF was employed to evaluate the importance of the factors in the experimental dataset. The importance of every feature is plotted in Figure 5. The rank of the selected features based on the average values decrease in accuracy in subsets is F4, F11, F6, F5, F1, F8, F10, F12, F7, F3, F9, and F12. According to the ordered features, 12 feature subsets were incrementally constructed.

TABLE 5. Results of SCA variants with different chaotic maps.

Fun	Index	SCA	LogisticSCA	IterativeSCA	CircleSCA	ChebyshevSCA	SingerSCA	SinusoidalSCA	SineSCA	TentSCA
F1	mean	1.49E+01	6.94E-14	1.51E-11	**8.67E-16**	2.64E-12	1.67E-08	9.19E-08	1.56E-13	8.85E-16
	std	3.19E+01	2.26E-13	7.77E-11	2.85E-15	5.24E-12	3.26E-08	1.45E-07	2.83E-13	1.98E-15
F2	mean	1.46E-02	2.09E-09	1.47E-08	2.54E-10	9.02E-09	1.42E-06	4.71E-06	5.30E-09	**2.29E-10**
	std	1.87E-02	4.52E-09	2.38E-08	4.23E-10	2.23E-08	3.68E-06	5.48E-06	1.38E-08	4.09E-10
F3	mean	1.01E+04	6.65E-10	1.03E-06	**1.21E-11**	4.75E-07	1.64E-02	1.08E-03	1.64E-07	4.38E-11
	std	6.78E+03	1.49E-09	3.43E-06	2.91E-11	1.11E-06	3.57E-02	2.45E-03	4.05E-07	1.42E-10
F4	mean	3.76E+01	2.04E-06	5.23E-05	4.27E-07	7.62E-05	5.73E-03	1.68E-03	2.37E-05	**3.61E-07**
	std	1.55E+01	3.32E-06	9.26E-05	8.04E-07	2.22E-04	6.24E-03	1.78E-03	4.33E-05	5.63E-07
F5	mean	5.93E+04	2.84E+01	2.83E+01	2.85E+01	2.84E+01	2.83E+01	**2.83E+01**	2.85E+01	2.85E+01
	std	1.65E+05	4.08E-01	4.40E-01	3.49E-01	3.81E-01	2.94E-01	3.27E-01	3.94E-01	3.31E-01
F6	mean	2.66E+01	5.47E+00	5.18E+00	5.75E+00	5.24E+00	4.90E+00	**4.56E+00**	5.38E+00	5.42E+00
	std	4.05E+01	2.78E-01	1.92E-01	1.85E-01	3.07E-01	2.60E-01	2.16E-01	2.11E-01	2.64E-01
F7	mean	1.17E-01	3.28E-04	3.93E-04	**2.48E-04**	3.32E-04	6.08E-04	3.90E-04	3.41E-04	2.61E-04
	std	1.41E-01	2.64E-04	5.09E-04	2.03E-04	2.71E-04	5.27E-04	3.38E-04	2.71E-04	2.62E-04
F8	mean	-3.65E+03	-3.50E+03	-3.28E+03	-3.37E+03	**-3.68E+03**	-3.59E+03	-3.41E+03	-3.39E+03	-3.35E+03
	std	3.13E+02	3.55E+02	2.82E+02	3.55E+02	2.97E+02	2.46E+02	2.64E+02	2.95E+02	2.68E+02
F9	mean	3.71E+01	5.68E-15	4.03E-12	1.89E-15	2.35E-13	2.95E-06	1.34E+00	1.98E-12	**0.00E+00**
	std	3.57E+01	2.29E-14	1.61E-11	1.04E-14	4.93E-13	9.46E-06	7.12E+00	1.05E-11	0.00E+00
F10	mean	1.70E+01	1.82E-08	1.72E-07	2.78E-09	1.47E-07	1.77E-04	8.55E-01	1.15E-07	**2.08E-09**
	std	6.61E+00	2.38E-08	2.19E-07	4.27E-09	2.05E-07	7.78E-04	3.82E+00	2.37E-07	2.72E-09
F11	mean	1.10E+00	1.62E-14	2.42E-10	4.23E-14	4.70E-11	3.25E-07	7.33E-03	1.75E-12	**3.54E-15**
	std	5.18E-01	3.46E-14	9.83E-10	1.31E-13	1.07E-10	8.19E-07	1.56E-02	3.31E-12	9.95E-15
F12	mean	7.00E+05	6.82E-01	6.66E-01	8.14E-01	7.24E-01	5.74E-01	5.15E-01	**7.05E-01**	6.77E-01
	std	3.49E+06	5.07E-02	1.34E-01	1.30E-01	1.59E-01	1.09E-01	1.06E-01	1.33E-01	1.10E-01
F13	mean	1.21E+05	2.60E+00	2.56E+00	2.69E+00	2.59E+00	2.47E+00	2.39E+00	**2.56E+00**	2.60E+00
	std	2.80E+05	8.34E-02	8.88E-02	5.88E-02	8.56E-02	1.22E-01	1.13E-01	7.46E-02	1.21E-01
	+/=/-	12/0/1	5/8/0	7/5/1	3/10/0	7/4/2	8/1/4	7/3/3	7/6/0	-
	ARV	8.526	3.899	4.677	4.028	4.723	5.549	5.572	4.620	**3.406**

TABLE 6. Comparative results of CESCA with other popular algorithms.

Fun	Index	SSA	GOA	DA	BA	MFO	CESCA
F1	mean	8.87E+02	3.47E+01	2.75E+03	1.70E+01	3.45E+03	**3.16E-15**
	std	4.15E+02	1.66E+01	1.18E+03	2.47E+00	6.03E+03	5.32E-15
F2	mean	1.70E+01	1.86E+01	1.65E+01	5.40E+01	3.54E+01	**3.05E-10**
	std	3.18E+00	2.53E+01	6.97E+00	1.35E+02	2.03E+01	6.10E-10
F3	mean	5.06E+03	3.77E+03	1.39E+04	1.29E+02	2.06E+04	**2.69E-11**
	std	3.23E+03	2.22E+03	8.55E+03	4.09E+01	1.25E+04	5.37E-11
F4	mean	2.00E+01	1.53E+01	3.06E+01	3.81E+00	6.68E+01	**2.71E-07**
	std	3.50E+00	4.29E+00	7.98E+00	2.73E+00	9.53E+00	4.68E-07
F5	mean	1.44E+05	6.00E+03	2.96E+05	5.00E+03	2.69E+06	**2.85E+01**
	std	1.15E+05	8.65E+03	2.65E+05	1.60E+03	1.46E+07	3.66E-01
F6	mean	1.17E+03	3.61E+01	2.20E+03	1.62E+01	3.36E+03	**5.40E+00**
	std	6.78E+02	3.36E+01	1.17E+03	2.89E+00	5.48E+03	2.29E-01
F7	mean	3.79E-01	4.90E-02	6.36E-01	1.31E+01	4.51E+00	**1.80E-04**
	std	2.09E-01	1.76E-02	4.86E-01	7.17E+00	9.10E+00	1.60E-04
F8	mean	-6.13E+03	-7.56E+03	-5.56E+03	-7.24E+03	**-8.45E+03**	-3.31E+03
	std	8.59E+02	7.63E+02	5.94E+02	7.10E+02	7.23E+02	3.37E+02
F9	mean	9.48E+01	9.80E+01	1.60E+02	2.71E+02	1.71E+02	**0.00E+00**
	std	1.92E+01	4.17E+01	3.40E+01	2.48E+01	4.01E+01	0.00E+00
F10	mean	9.41E+00	5.19E+00	1.04E+01	5.59E+00	1.46E+01	**4.07E-09**
	std	1.40E+00	1.15E+00	1.84E+00	3.91E+00	7.39E+00	5.06E-09
F11	mean	1.02E+01	1.16E+00	1.79E+01	6.38E-01	1.30E+01	**9.61E-15**
	std	3.50E+00	1.35E-01	8.71E+00	5.54E-02	3.11E+01	2.02E-14
F12	mean	2.28E+01	9.43E+00	1.68E+04	1.48E+01	1.12E+04	**7.48E-01**
	std	1.18E+01	4.61E+00	4.87E+04	4.94E+00	6.15E+04	1.69E-01
F13	mean	1.82E+04	4.40E+01	4.79E+05	2.68E+00	2.08E+01	**2.60E+00**
	std	3.90E+04	3.00E+01	6.79E+05	4.00E-01	1.54E+01	8.62E-02
	+/=/-	12/0/1	12/0/1	12/0/1	11/1/1	12/0/1	
	ARV	4.323	3.062	5.015	3.313	3.846	**1.441**

Subset 1 is represented with the most crucial feature {F4}; Subset 2 denotes the top two features {F4, F11} and so on. The classification results of 12 feature subsets using the CESCA-FKNN are shown in Table 7. The results show that the Subset 5 {F4, F11, F6, F5, F1} is capable of achieving the best performance in four measurements: classification accuracy (ACC), Sensitivity, Specificity and MCC with values of 82.47%, 85.47%, 78.97% and 0.6565. In terms of the standard deviation, Table 7 also confirms the fact that Subset 5 is more stable than other feature subsets. Therefore, Subset 5 was selected as the best feature subset in the following experiment.

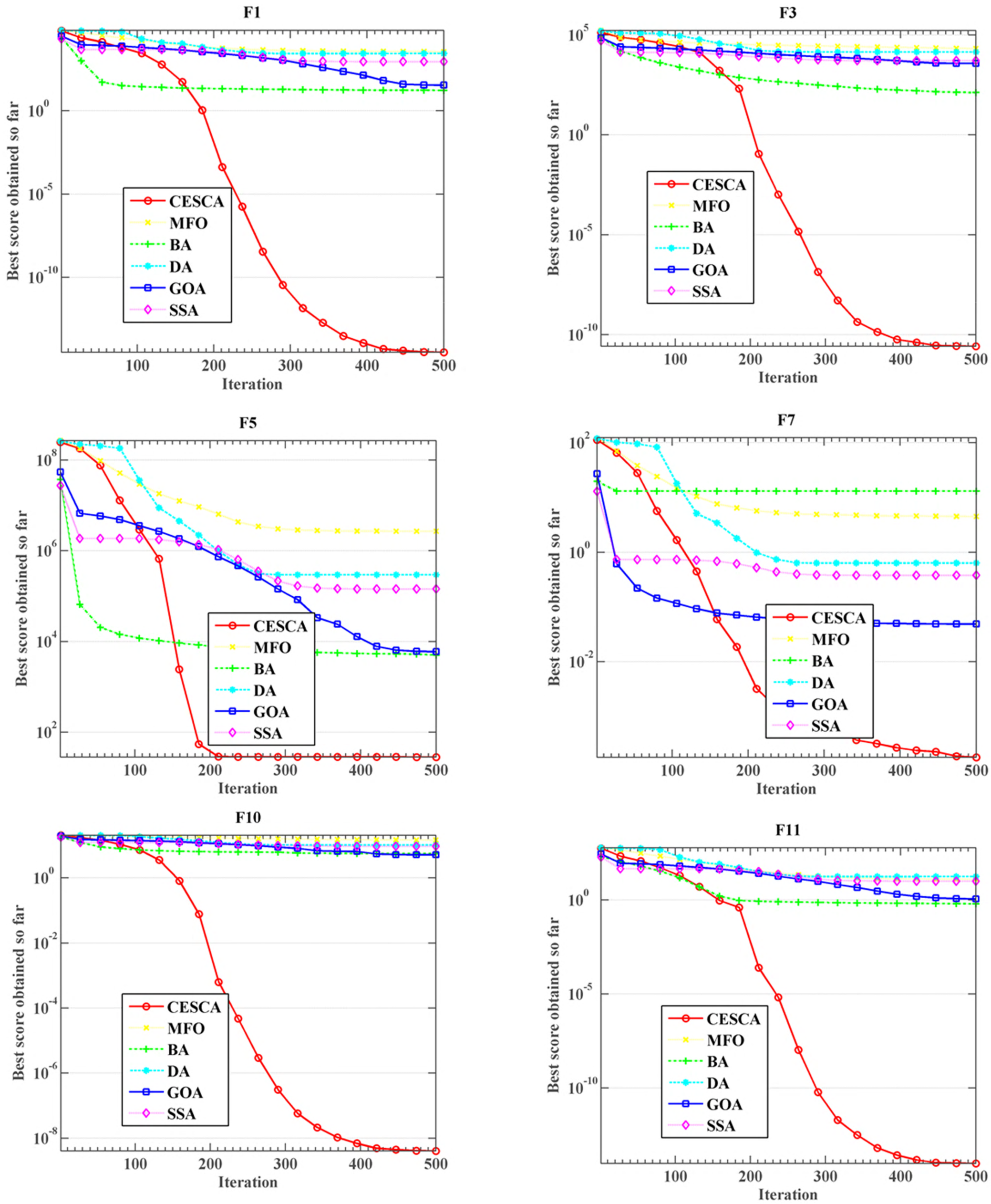

FIGURE 4. Convergence curves of some selected benchmark functions.

To verify the effectiveness of the proposed CESCA-FKNN method, we compared the proposed method with RF, kernel extreme learning machine (KELM) [52] and support vector machine (SVM) [53], [54]. Since RF, KELM, and SVM were the most commonly used and practical machine learning algorithms. Also, the original SCA-FKNN, the DA based FKNN approaches (DA-FKNN) and the MFO based FKNN model (MFO-FKNN) have also been employed to evaluate the performance of the developed CESCA-FKNN model. The experimental results of seven methods in four indicators are shown in Figure 6. As can be seen from the figure, CESCA-FKNN has better results than the other three methods in terms of the three indicators. Besides, CESCA-FKNN had the smallest standard deviation among the two indexes with ACC of 82.47%, sensitivity of 85.47%, the specificity of 78.97%, MCC of 0.6565.

According to the ACC metric, CESCA-FKNN achieves the best results which are slightly better than results yielded by DA-FKNN, whereas SCA-FKNN, MFO-FKNN, RF, KELM, and SVM produce inferior results. Regarding the sensitivity metric, RF obtains the best results, MFO-FKNN, CESCA-FKNN, SCA-FKNN, and DA-FKNN yielded similar results, whereas KELM and SVM yielded the worst results. As for the specificity metric, the developed

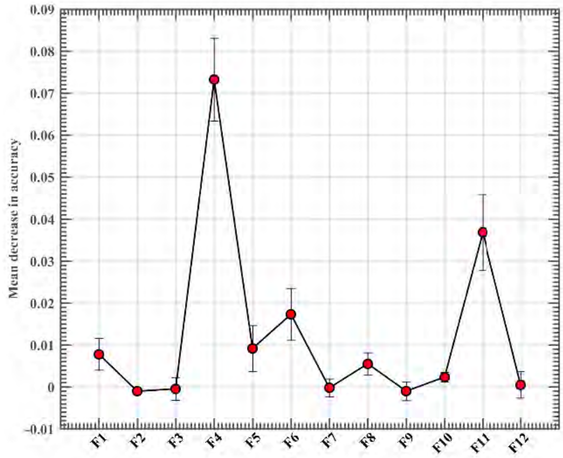

FIGURE 5. Mean decrease in accuracy for each feature.

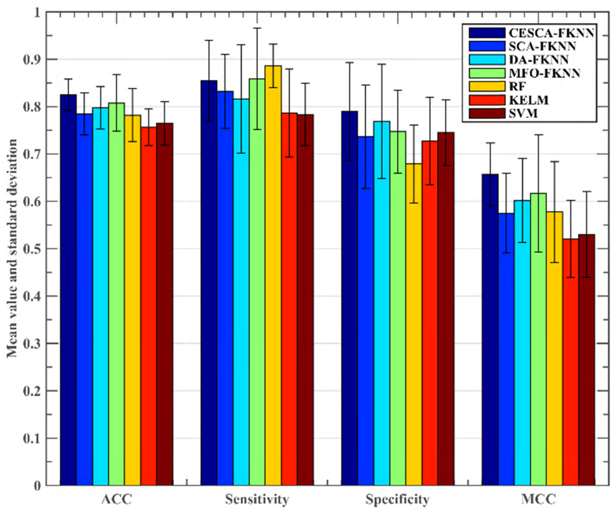

FIGURE 6. Comparison results among CESCA-FKNN, SCA-FKNN, DA-FKNN, MFO-FKNN, RF, KELM, and SVM.

CESCA-FKNN obtained the best results, followed successively by DA-FKNN, MFO-FKNN, SVM, SCA-FKNN, KELM, and RF. In terms of the MCC metric, CESCA-FKNN obtained the best results, whereas KELM again yields the worst results. The results indicate that the CESCA-FKNN model proposed in this paper is an effective method with strong potential to predict the students' intention of master programs.

The evolutionary curves of CESCA-FKNN, DA-FKNN, MFO-FKNN, and the original SCA-FKNN are depicted in Figure 7 to evaluate the performance of the proposed method, t. All methods were implemented in the same testing environment to ensure a fair comparison. The population size and number of maximum iterations were set as 20 and 50, respectively. As can be seen from the figure, the proposed CESCA-FKNN obtained the best convergence curve. It is also observed that solutions obtained by CESCA-FKNN are much better than the best solutions of

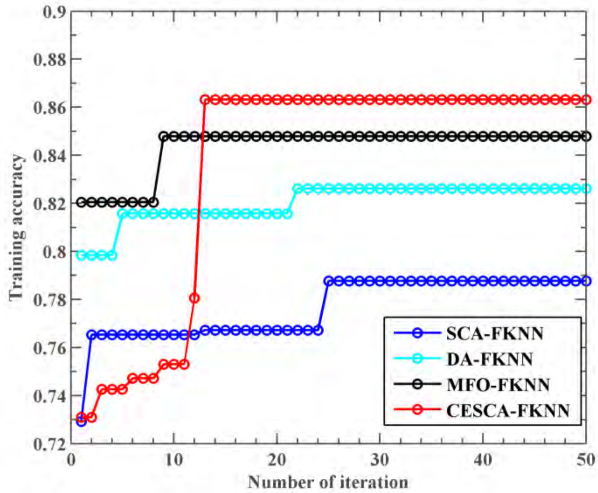

FIGURE 7. The convergence trend of four methods.

the original SCA-FKNN in the later period of optimization iterative. This fact can indicate that the exploitation tendency of the proposed CESCA in searching for optimal parameters has improved, considerably. Also, the original SCA-FKNN model may have poor ability to explore further during the iterative process, and thus, got the worse results.

V. DISCUSSION

Based on the above experimental results, it is observed that the most important features including total credits (F4), CET-6 (F11), Advanced Mathematics course (F6), College English course (F5), and gender (F1), the influence of these features on the choice of students' intentions for master programs is relatively prominent. On the whole, students with a higher level of language proficiency are more inclined to become a graduate student because English is an important subject for the postgraduate entrance examination, and English scores often become the lowest part of the barrel effect of the postgraduate entrance examination. Students with excellence score in Advanced Mathematics course have obvious opportunity to go to graduate school because math is an important basic course in science and engineering colleges. Gender differences also have a significant influence on students' intentions for master programs. The proportion of girls choosing to occupy a job is much higher than that of boys' maybe because boys are courageous, and girls prefer steady jobs. The academic achievement and the employment option have a significant influence on one's entire life. The academic performance mainly includes GPA, CET-6, Total Credits, scores of college English course and advanced mathematics course, and the academic achievement is generally divided into two categories: good and bad. Students with low academic scores do not have an obvious advantage to pursue further education, and they will actively think about how to improve their practical experience or job opportunities through various channels. Hence, the probability of searching

TABLE 7. Performance of CESCA-FKNN on different feature subsets.

Feature subsets	MCC	ACC	Sensitivity	Specificity
Subset 1	0.3388 (0.0910)	0.6680 (0.0493)	0.7409 (0.0486)	0.5959 (0.0848)
Subset 2	0.4370 (0.1217)	0.7178 (0.0660)	0.7452 (0.0937)	0.6824 (0.1736)
Subset 3	0.4490 (0.1038)	0.7192 (0.0520)	0.7781 (0.1186)	0.6455 (0.2196)
Subset 4	0.5879 (0.1193)	0.7919 (0.0606)	0.7970 (0.0880)	0.7890 (0.0903)
Subset 5	**0.6565 (0.0665)**	**0.8247 (0.0332)**	**0.8547 (0.0854)**	**0.7897 (0.1032)**
Subset 6	0.6455 (0.1170)	0.8192 (0.0579)	0.8752 (0.0779)	0.6455 (0.1170)
Subset 7	0.6053 (0.1083)	0.8035 (0.0528)	0.8179 (0.0937)	0.7809 (0.0881)
Subset 8	0.5775 (0.1057)	0.7833 (0.0587)	0.8620 (0.0484)	0.7072 (0.0932)
Subset 9	0.5689 (0.0767)	0.7821 (0.0415)	0.8150 (0.0630)	0.7502 (0.0989)
Subset 10	0.5511 (0.1038)	0.7722 (0.0549)	0.8278 (0.1011)	0.7083 (0.1200)
Subset 11	0.5441 (0.1154)	0.7623 (0.0598)	0.8435 (0.1168)	0.6724 (0.2112)
Subset 12	0.5958 (0.0946)	0.7935 (0.0507)	0.8346 (0.0678)	0.7536 (0.1165)

for a job is stronger than that of the students who got high total credits. On the contrary, students with better academic achievements prefer to study further to earn a master's degree.

It is noteworthy that the present study has several restrictions that can be addressed in the future. First, the research sample in this study was restricted. To obtain accurate results, more consecutive samples are needed to be collected for completing the unbiased learning model. Second, the research was accomplished only based on a relatively single university, which needs to add diversity, then, the model for decision support will be more reliable and practical. The third item is the limited involved attributes. More studies should be undertaken to investigate more attributes which may have any influence on students' intentions for Master Programs.

VI. CONCLUSIONS AND FUTURE WORKS

In this study, we developed an improved FKNN framework to predict the students' intentions for master programs. This paper's novelty lies in combing the RF algorithm with an enhanced SCA-based FKNN prediction engine. On the one hand, the RF algorithm aims at screening the key features in the data, and on the other hand, proposed enhanced SCA was employed to tune the two key parameters of FKNN to predict new samples. Simulation results have demonstrated that the proposed framework has better classification performance than three classical classifiers and four swarm intelligence algorithms-based FKNN approaches on the majority metrics of ACC, MCC, sensitivity, and specificity. Therefore, we can get a preliminary conclusion that an improved framework is a valuable tool for predicting students' employment intention. In future work, we plan to collect more consecutive samples to complete the unbiased learning model to improve the prediction performance. Also, we also plan to establish a set of the decision support system on the proposed model to assist the decision makers of university departments to predict the students' intentions of master programs.

ACKNOWLEDGEMENTS

Huiling Chen, Wujun Geng, and Chengye Li contributed equally to this work.

REFERENCES

[1] M. Koutina and K. L. Kermanidis, "Predicting postgraduate students' performance using machine learning techniques," in *Proc. IFIP Int. Conf. Artif. Intell. Appl. Innov.*, vol. 364, 2011, pp. 159–168.
[2] L. Wang, "Improved NN-GM(1,1) for postgraduates' employment confidence index forecasting," *Math. Problems Eng.*, vol. 2014, Aug. 2014, Art. no. 465208.
[3] L. Chi and L. Lin, "Application of BP neural network based on genetic algorithms optimization in prediction of postgraduate entrance examination," in *Proc. 3rd Int. Conf. Inf. Sci. Control Eng. (ICISCE)*, Jul. 2016, pp. 226–229.
[4] A. Jóźwik, "A learning scheme for a fuzzy *k*-NN rule," *Pattern Recognit. Lett.*, vol. 1, pp. 287–289, Jul. 1983.
[5] J. M. Keller, M. R. Gray, and J. A. Givens, "A fuzzy K-nearest neighbor algorithm," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, no. 4, pp. 580–585, Jul./Aug. 1985.
[6] H.-L. Chen *et al.*, "A novel bankruptcy prediction model based on an adaptive fuzzy *k*-nearest neighbor method," *Knowl.-Based Syst.*, vol. 24, pp. 1348–1359, Dec. 2011.
[7] J. Sim, S. Y. Kim, and J. Lee, "Prediction of protein solvent accessibility using fuzzy k-nearest neighbor method," *Bioinformatics*, vol. 21, no. 12, pp. 2844–2849, Jul. 2005.
[8] Y. Huang and Y. Li, "Prediction of protein subcellular locations using fuzzy k-NN method," *Bioinformatics*, vol. 20, no. 1, pp. 21–28, 2004.
[9] M.-Y. Cheng and N.-D. Hoang, "A swarm-optimized fuzzy instance-based learning approach for predicting slope collapses in mountain roads," *Knowl.-Based Syst.*, vol. 76, pp. 256–263, Mar. 2015.
[10] D.-Y. Liu, H.-L. Chen, B. Yang, X.-E. Lv, L.-N. Li, and J. Liu, "Design of an enhanced fuzzy k-nearest neighbor classifier based computer aided diagnostic system for thyroid disease," *J. Med. Syst.*, vol. 36, no. 5, pp. 3243–3254, 2012.
[11] H. L. Chen *et al.*, "An efficient diagnosis system for detection of Parkinson's disease using fuzzy *k*-nearest neighbor approach," *Expert Syst. Appl.*, vol. 40, no. 1, pp. 263–271, 2013.
[12] W. Gao, J. L. G. Guirao, B. Basavanagoud, and J. Wu, "Partial multi-dividing ontology learning algorithm," *Inf. Sci.*, vol. 467, pp. 35–58, Oct. 2018.
[13] W. Gao, W. Wang, D. Dimitrov, and Y. Wang, "Nano properties analysis via fourth multiplicative ABC indicator calculating," *Arabian J. Chem.*, vol. 11, no. 6, pp. 793–801, 2018.
[14] W. Gao, H. Wu, M. K. Siddiqui, and A. Q. Baig, "Study of biological networks using graph theory," *Saudi J. Biol. Sci.*, vol. 25, no. 6, pp. 1212–1219, 2018.

[15] W. Gao, J. L. G. Guirao, M. Abdel-Aty, and W. Xi, "An independent set degree condition for fractional critical deleted graphs," *Discrete, Continuous Dyn. Syst.*, vol. 12, nos. 4–5, pp. 877–886, 2019.

[16] G. Wei, D. Darko, and A. Hosam, "Tight independent set neighborhood union condition for fractional critical deleted graphs and ID deleted graphs," *Discrete Continuous Dyn. Syst., S*, vol. 12, pp. 711–721, Aug./Sep. 2019.

[17] S. Mirjalili, "SCA: A sine cosine algorithm for solving optimization problems," *Knowl.-Based Syst.*, vol. 96, pp. 120–133, Mar. 2016.

[18] I. Aljarah, M. Mafarja, A. A. Heidari, H. Faris, Y. Zhang, and S. Mirjalili, "Asynchronous accelerating multi-leader salp chains for feature selection," *Appl. Soft Comput.*, vol. 71, pp. 964–979, Oct. 2018.

[19] H. Faris *et al.*, "An intelligent system for spam detection and identification of the most relevant features based on evolutionary random weight networks," *Inf. Fusion*, vol. 48, pp. 67–83, Aug. 2019.

[20] H. Faris *et al.*, "An efficient binary salp swarm algorithm with crossover scheme for feature selection problems," *Knowl.-Based Syst.*, vol. 154, pp. 43–67, Aug. 2018.

[21] M. Mafarja *et al.*, "Binary dragonfly optimization for feature selection using time-varying transfer functions," *Knowl.-Based Syst.*, vol. 161, pp. 185–204, Dec. 2018.

[22] M. Mafarja *et al.*, "Evolutionary population dynamics and grasshopper optimization approaches for feature selection problems," *Knowl.-Based Syst.*, vol. 145, pp. 25–45, Apr. 2018.

[23] M. A. A. Al-Qaness, M. A. Elaziz, and A. A. Ewees, "Oil consumption forecasting using optimized adaptive neuro-fuzzy inference system based on sine cosine algorithm," *IEEE Access*, vol. 6, pp. 68394–68402, 2018.

[24] S. Das, A. Bhattacharya, and A. K. Chakraborty, "Solution of short-term hydrothermal scheduling using sine cosine algorithm," *Soft Comput.*, vol. 22, no. 19, pp. 6409–6427, 2018.

[25] B. Mahdad and K. Srairi, "A new interactive sine cosine algorithm for loading margin stability improvement under contingency," *Elect. Eng.*, vol. 100, no. 2, pp. 913–933, 2018.

[26] M. Wang *et al.*, "Toward an optimal kernel extreme learning machine using a chaotic moth-flame optimization strategy with applications in medical diagnoses," *Neurocomputing*, vol. 267, pp. 69–84, Dec. 2017.

[27] Y. Xu, H. Chen, J. Luo, Q. Zhang, S. Jiao, and X. Zhang, "Enhanced moth-flame optimizer with mutation strategy for global optimization," *Inf. Sci.*, vol. 492, pp. 181–203, Aug. 2019.

[28] Y. Xu *et al.*, "An efficient chaotic mutative moth-flame-inspired optimizer for global optimization tasks," *Expert Syst. Appl.*, vol. 129, pp. 135–155, Sep. 2019.

[29] Q. Zhang *et al.*, "Chaos-induced and mutation-driven schemes boosting salp chains-inspired optimizers," *IEEE Access*, vol. 7, pp. 31243–31261, 2019.

[30] X. Zhao *et al.*, "Chaos enhanced grey wolf optimization wrapped ELM for diagnosis of paraquat-poisoned patients," *Comput. Biol. Chem.*, vol. 78, pp. 481–490, Feb. 2019.

[31] J. Luo, H. Chen, A. A. Heidari, Y. Xu, Q. Zhang, and C. Li, "Multi-strategy boosted mutative whale-inspired optimization approaches," *Appl. Math. Model.*, vol. 73, pp. 109–123, Sep. 2019.

[32] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," *Future Gener. Comput. Syst.*, vol. 97, pp. 849–872, Aug. 2019.

[33] H. Chen, Y. Xu, M. Wang, and X. Zhao, "A balanced whale optimization algorithm for constrained engineering design problems," *Appl. Math. Model.*, vol. 71, pp. 45–59, Jul. 2019.

[34] J. Luo, H. Chen, Q. Zhang, Y. Xu, H. Huang, and X. Zhao, "An improved grasshopper optimization algorithm with application to financial stress prediction," *Appl. Math. Model.*, vol. 64, pp. 654–668, Dec. 2018.

[35] W. Deng, J. Xu, and H. Zhao, "An improved ant colony optimization algorithm based on hybrid strategies for scheduling problem," *IEEE Access*, vol. 7, pp. 20281–20292, 2019.

[36] W. Deng, H. Zhao, L. Zou, G. Li, X. Yang, and D. Wu, "A novel collaborative optimization algorithm in solving complex optimization problems," *Soft Comput.*, vol. 21, no. 15, pp. 4387–4398, Aug. 2017.

[37] A. A. Heidari, I. Aljarah, H. Faris, H. Chen, J. Luo, and S. Mirjalili, "An enhanced associative learning-based exploratory whale optimizer for global optimization," *Neural Comput. Appl.*, vol. 2019, no. 13. doi: 10.1007/s00521-019-04015-0.

[38] C. Qu, Z. Zeng, J. Dai, Z. Yi, and W. He, "A modified sine-cosine algorithm based on neighborhood search and greedy levy mutation," *Comput. Intell. Neurosci.*, vol. 2018, Jul. 2018, Art. no. 4231647.

[39] S. N. Chegini, A. Bagheri, and F. Najafi, "PSOSCALF: A new hybrid PSO based on sine cosine algorithm and Levy flight for solving optimization problems," *Appl. Soft Comput.*, vol. 73, pp. 697–726, Dec. 2018.

[40] M. A. Elaziz, D. Oliva, and S. Xiong, "An improved opposition-based sine cosine algorithm for global optimization," *Expert Sys. Appl.*, vol. 90, pp. 484–500, Dec. 2017.

[41] H. Nenavath and R. K. Jatoth, "Hybridizing sine cosine algorithm with differential evolution for global optimization and object tracking," *Appl. Soft Comput.*, vol. 62, pp. 1019–1043, Jan. 2018.

[42] J. Zhu, X. Zhao, H. Li, H. Chen, and G. Wu, "An effective machine learning approach for identifying the glyphosate poisoning status in rats using blood routine test," *IEEE Access*, vol. 6, pp. 15653–15662, 2018.

[43] J. Zhu *et al.*, "A new evolutionary machine learning approach for identifying pyrene induced hepatotoxicity and renal dysfunction in rats," *IEEE Access*, vol. 7, pp. 15320–15329, 2018. doi: 10.1109/ACCESS.2018.2889151.

[44] S. Saremi, S. Mirjalili, and A. Lewis, "Grasshopper optimisation algorithm: Theory and application," *Adv. Eng. Softw.*, vol. 105, pp. 30–47, Mar. 2017.

[45] S. Mirjalili, "Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems," *Neural Comput. Appl.*, vol. 27, no. 4, pp. 1053–1073, 2016.

[46] X. S. Yang, "A new metaheuristic bat-inspired algorithm," in *Studies in Computational Intelligence*, vol. 284, Apr. 2010, pp. 65–74.

[47] S. Mirjalili, "Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm," *Knowl.-Based Syst.*, vol. 89, pp. 228–249, Nov. 2015.

[48] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011.

[49] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," *IJCAI*, vol. 14, no. 2, pp. 1137–1145, 1995.

[50] S. García, A. Fernández, J. Luengo, and F. Herrera, "Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: Experimental analysis of power," *Inf. Sci.*, vol. 180, no. 10, pp. 2044–2064, 2010.

[51] S. García, D. Molina, M. Lozano, and F. Herrera, "A study on the use of non-parametric tests for analyzing the evolutionary algorithms' behaviour: A case study on the CEC'2005 special session on real parameter optimization," *J. Heuristics*, vol. 15, no. 6, pp. 617–644, 2009.

[52] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012.

[53] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. 5th Annu. Workshop Comput. Learn. Theory*, 1992, pp. 144–152.

[54] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer, 1995.

AIJU LIN is currently a Professor with the College of Mathematics, Physics, and Electronic Information Engineering, Wenzhou University. She is also the State Psychological Counselor (Level 2) and BCC Career Coach. She mainly studies the college student's social entrepreneurship education and the career development. She hosted and participated in researches of various national and provincial level, and published more than 10 academic papers in national core journals.

QUANQUAN WU received the M.S. degree from Wenzhou University, where he is currently an Ideological and Political Counselor with the College of Mathematics, Physics, and Electronic Information Engineering. He is also a Board Certified Coach for career (BCC). He is currently engaged in the management of student affairs and the guidance of graduates to take the postgraduate entrance examination. His current research interest is the application of big data technology in student affairs.

ALI ASGHAR HEIDARI is currently pursuing the Ph.D. degree with the University of Tehran. He is also the Ph.D. Research Intern with the School of Computing, National University of Singapore (NUS). He is awarded and funded by the Iran's National Elites Foundation (INEF). He has] published more than 20 papers in prestigious international journals, such as *Information Fusion*, *Information Sciences*, *Future Generation Computer Systems*, *Energy Conversion and Management*, *Applied Soft Computing*, *Knowledge-Based Systems*, and *Expert Systems with Applications*. His main research interests include advanced machine learning, evolutionary computation, meta-heuristics, prediction, information systems, and spatial modeling.

YUETING XU received the bachelor's degree from the Department of Computer Science and Technology, Ningbo Dahongying University, China. She is currently pursuing the degree in computer software and theory with Wenzhou University, China. She has published more than five papers in international journals, such as *Information Sciences*, IEEE Access and *Expert Systems with Applications*. Her research interests include data mining, machine learning, evolutionary computation, and their applications to medical diagnosis.

HUILING CHEN received the Ph.D. degree from the Department of Computer Science and Technology, Jilin University, China. He is currently an Associate Professor with the Department of Computer Science and Technology, Wenzhou University, China. He has published more than 100 papers in international journals and conference proceedings, including *Pattern Recognition*, *Information Sciences*, *Expert Systems with Applications*, *Knowledge-Based Systems*, *Soft Computing*, *Neurocomputing*, and PAKDD. His current research interests include machine learning and data mining, and their applications to medical diagnosis and bankruptcy prediction.

WUJUN GENG is currently a Tutor of master's degree, an Associate Professor with Wenzhou Medical University, and a Deputy Director of Science and Education Department, the First Affiliated Hospital of Wenzhou Medical University. He hosted and participated in researches of various national and provincial level, and published more than 10 academic papers in national core journals. His research direction is to improve students' teaching and scientific research ability.

YUPING LI received the M.D. degree from Wenzhou Medical University, China. She is currently a Professor/Senior Physician Director of the Department of Respiratory and Critical Care Medicine, the First Affiliated Hospital of Wenzhou Medical University. Her current research interests include lung cancer and pulmonary fungal infection.

CHENGYE LI received the M.D. degree from the Shanghai Jiaotong University School of Medicine, China. He is currently an Attending Physician with the Department of Pulmonary and Critical Care Medicine, the First Affiliated Hospital of Wenzhou Medical University, China. His research interests include data mining and machine learning, and their applications to medical diagnosis.