Journal of Network Intelligence
Taiwan Ubiquitous Information
©2024 ISSN 2414-8105 (Online)
Volume **9**, Number **2**, May **2024**

Binary Particle Swarm Optimization Predicting Students' Academic Performance in College English

Liya Yue
Fanli Business School
Nanyang Institute of Technology, Nanyang 473004, China
3172027@nyist.edu.cn

Pei Hu
School of Computer and Software
Nanyang Institute of Technology, Nanyang 473004, China
huxiaopei163@163.com

Jiulong Zhu*
Fanli Business School
Nanyang Institute of Technology, Nanyang 473004, China
zhujiulong7699@163.com

*Corresponding author: Jiulong Zhu

Received July 5, 2023, revised October 19, 2023, accepted March 20, 2024.

Abstract. *Educational data mining explores the features that are related to students' academic performance. To improve the prediction accuracy in college English, we propose a model based on feature selection and particle swarm optimization (PSO). Since feature selection is a typical binary optimization, to make the model solve the problem, an improved binary PSO (IBPSO) balances global and local search abilities through a new transfer function. A novel learning method develops the position update of PSO to facilitate the algorithm to search for solutions in more space, while advanced social cognitive factors accelerate convergence. IBPSO outperforms the original binary PSO, henry gas solubility optimization, and binary sine cosine algorithm on the benchmark datasets from UCI. In our proposed academic performance model, these algorithms predict the college English grades of sophomores at a university. IBPSO significantly improves classification accuracy, and the experiments illustrate that demographic features are the main factors affecting students' academic performance in college English.*
Keywords: Educational data mining; Academic performance; Particle swarm optimization; College English

---

1. **Introduction.** In the era of information technology, schools are evolving at a high rate to provide better education for students. Schools are committed to cultivating students with excellent academic performance, so they track students' actions in specific fields that require a lot of training [1, 2]. The performance depends on several factors such as personal, socio-economic and other environmental variables, and managing students is made easier by being aware of these aspects and how they affect students' actions. Due to the large amount of data available in education databases, it is a challenging task to predict the performance [3, 4]. If schools predict students' performance before examinations, they take additional measures to arrange appropriate assistance for students to help them develop their studies and achieve success. Additionally, predicting students'

academic performance assists specialists in creating association rules to make correct decisions [5, 6].

The prediction of students' academic performance has become an important research direction in higher education due to the growing usage of educational data mining (EDM) technology. This supports personalized education and also assists educators design timely interventions. Despite the development of numerous academic performance prediction systems for college students, the following issues still exist:

(1) Obtain sufficient personal data of students and integrate this data to acquire comprehensive views; (2) Explore the factors that influence students' academic performance and use the information to establish high-precision prediction models; and (3) Utilize the models to provide personalized services that enable students to optimize their learning.

It is a classification problem in machine learning in which academic data from students is utilized to train prediction models based on classifiers, and the models output the classes of students' performance [7, 8]. Unfortunately, students' academic data collected from digital systems often contains numerous features, some of which have no impact on the prediction. It is vital to choose the features that significantly affect the outcome of the prediction through feature selection.

In classification decisions, the prediction ability of a model is mainly determined by the characteristics of samples, including whether they are complete, redundant, and noisy. They affect the classification ability of the model, and even redundant and noisy data plays a negative effect [9, 10]. Feature selection is the process of selecting an optimal subset from original features and building a prediction model on this subset. An effective algorithm reduces data dimensionality by removing redundant information and noisy data. It also improves prediction accuracy and saves learning time.

Feature selection plays an important role in EDM, therefore, we predict students' academic performance in college English based on binary particle swarm optimization (PSO). The main contributions of this paper are summarized as follows:

(1) Establish a prediction model for students' performance of college English;

(2) Introduce a feature selection model;

(3) Propose an improved PSO algorithm to implement feature selection;

(4) The proposed algorithm is verified on benchmark datasets, and also successfully predicts the students' academic performance in college English at a university.

The structure of this paper is organized as follows. Section 2 introduces the recent research advances in feature selection and academic prediction based on feature selection. Section 3 presents the proposed prediction model, and Section 4 completes its experimental validation. In section 5, a brief summary is provided.

2. **Literature review.**

2.1. **Feature selection.**

According to different evaluation criteria, there are three main types of feature selection methods, filter, wrapper and embedded. Filter method has high computational efficiency and is suitable for large-scale datasets. Its process is independent from subsequent learners, and it has great generalization performance. However, the subset it obtained may not have the best classification performance. Wrapper method has the advantage of high prediction accuracy, but it needs to execute learning algorithm many times, which has high time complexity and weak generalization ability [11].

Classical feature selection methods have many limitations, such as the difficulty in setting key parameters and falling into local optimum. The feature selection based on evolutionary algorithm solves most of the defects of traditional methods mentioned above [12, 13, 14], and it achieves a satisfactory feature set without traversing all feasible regions.

Evolutionary algorithms solve continuous problems, and people usually adopt transfer functions to binarize where the common utilized is the Sigmoid function [15]. Mirjalili et al. first analyzed it in detail [16], and pointed out that it is not applicable to the feature selection of binary PSO. They evaluated new suggested transfer functions from two categories, S-type and V-type. Hu et al. analyzed the position update equations of grey wolf optimizer (GWO) in binary space [17], and proposed an improved update method and new transfer functions.

Based on Hu's analysis, [9, 18, 19, 20] claimed binary evolutionary algorithms for feature selection. Pan et al. suggested the first binary version of bamboo forest growth (BFGO) with long mutation by evaluating the search space of BFGO under binary conditions [9]. Wang et al. modified the step size and transfer function in grasshopper optimization algorithm (GOA) to improve the search ability of binary GOA and the quality of solutions [18]. To advance the performance of binary pigeon-inspired optimization (PIO), Pan et al. brought new transfer functions, velocity and position update equations [19]. Du et al. used 9 transfer functions to binarize symbiotic organism search (SOS) and analyzed the effect of each transfer function on binary SOS [20].

Although the above algorithms achieve the binarization of evolutionary algorithms through transfer functions, they do not take into account that during optimization processes, algorithms change from exploration to exploitation.

2.2. **Predicting students' academic performance based on feature selection.** Various methods have been adopted to predict students' academic performance, including conventional mathematical models and modern EDM technologies. In these methods, mathematical equations describe the quantitative relationships between output and input (prediction variables). The prediction is accurate if there is just a tiny difference between predicted and actual values.

Based on classification and clustering methods, Francis and Babu developed a new prediction algorithm for evaluating students' academic performance [21], and conducted a real-time test on mixed datasets of disciplines of higher education institutions in Kerala, India. The results demonstrate that the algorithm acquires superior results in achieving the prediction accuracy of students' academic performance. Liang et al. designed a hybrid framework of feature selection and feature fusion to identify important and relevant features from educational data to predict students' performance [22]. The main goals of the proposed algorithm are to improve prediction accuracy and to find the optimal features. Farissi et al. combined decision tree, KNN (K-Nearest Neighbor), naive Bayes, and random forest classifiers with genetic algorithm (GA) to enhance the prediction accuracy of students' academic performance [23]. Chen and Do predicted students' academic performance based on cuckoo search (CS)-hierarchical adaptive neuro-fuzzy inference system (ANFIS) [24]. ANFIS solves the curse of dimensionality, and CS optimizes the parameters of ANFIS.

Despite the advancements in predicting academic performance, the research has been constrained by data sources and limited samples, which hinders the establishment of general rules. Therefore, we predict the academic performance of students in college English from multi-source data.

3. **Methodology.** This section describes the proposed model and the improved binary PSO (IBPSO) to implement feature selection and realize prediction.

3.1. **The proposed model.** The primary purpose of this study is to develop a prediction model for students' academic performance in college English with feature selection, which

Table 1. The details of students' information.

Feature category	**Feature**	**Description**	**Data type**
Demographic features	Gender	Male & Female	Nominal
	PlaceOrigin	the region of student source	Nominal
Academic features	Major	Liberal Arts, science and engineering, arts, high fees, overseas classes	Nominal
	CET4/6	Whether passed CET4/6	Nominal
	Score	Previous English course grades	Nominal
Behavioral features	OnlineTime	The average online time through campus network or WiFi every day (minutes)	Numeric
	Cost	Average daily cost (RMB)	Numeric
	Character	Whether like communication/learning	Nominal
	LearningHabits	Study or review	Nominal
	Absence	Number of absences	Numeric
	Classroom	Classroom performance	Nominal
	StudyTime	The average study time through library or classroom (minutes)	Numeric
Examination features	Difficulty	Derived from students' scores	Nominal
	Matching	The degree of goal achievement	Nominal
Teacher features	Title	Professional ranks and titles	Nominal
	Education	Educational background	Nominal
	Age	The length of service as a teacher	Nominal
Family features	Income	Household income status	Nominal
	Importance	Level of parental attention	Nominal
Class	A & B & C & D & E	Students' academic performance	Nominal

mainly consists of data collection, data preprocessing, feature selection and performance evaluation.

1. Data collection

To establish the prediction model, Table 1 shows students' features and their descriptions. The data comes from a university in China, with demographic, academic, behavioral and additional features collected from Academic Affairs Office, Student Affairs Office, Network Information Center and Library based on student ID.

2. Data preprocessing

After collecting the data, the preprocessing method advances its quality. To facilitate the study, we map numeric data into five types of nominal data. There are about 5600 sophomores at the university. Through random sampling and removal of missing and invalid data, there are about 1200 data used in the investigation. One-pot method is utilized to convert nominal data.

3. Feature selection

In binary evolutionary algorithms, 0 and 1 represent unselected and selected features, respectively.

KNN is the most popular classification algorithm in data mining, machine learning and pattern recognition, where each sample is represented through its nearest K neighbors. KNN is classified based on the distance (such as Euclidean distance) between test data and training data.

K-cross validation randomly divides the original dataset into K parts. One part is utilized as test data, and the remaining $K-1$ parts are employed as training data. It repeats K times and finally obtains the average of K times.

In this paper, KNN (K=5) completes the modeling of features, and cross validation (10-fold) tests the model.

4. Performance evaluation

We adopt classification quality (evaluation criterion) as the objective function of evolutionary algorithms, which means that the more accurate of correct classification, the better.

$$fit = num_w/(num_r + num_w) \tag{1}$$

where num_w and num_r represent the numbers of wrong and correct classification samples.

3.2. **Improved PSO for feature selection.** PSO simulates the foraging activity of birds. It has the characteristics of few parameters and fast running speed. PSO randomly initializes a group of particles, and then iteratively searches for optimal solutions [25, 26]. In each iteration, particles update their velocity and position with following equations.

$$V_i^d(t+1) = wV_i^d(t) + c1r_1(pbest_i^d(t) - X_i^d(t)) + c2r_2(gbest^d(t) - X_i^d(t)) \tag{2}$$

$$X_i^d(t+1) = X_i^d(t) + V_i^d(t+1) \tag{3}$$

where X_i^d means the dth position of particle i, and V is the velocity. $pbest$ and $gbest$ denote personal and global solutions. $c1$ and $c2$ are two constants, and r_1 and r_2 are two random values between [0, 1]. w is a weight factor, which is calculated as follows:

$$w = (w_{max} - w_{min})(MAX\_IT - t)/MAX\_IT + w_{min} \tag{4}$$

where t and MAX_IT indicate current and max iterations, and w_{max} and w_{min} are the maximum and minimum values of w.

Binary PSO (BPSO) uses the Sigmoid function to map the velocity to the interval [0, 1], and then adopts Equation (5) to update the position of a particle.

$$X_i^d(t+1) = \begin{cases} 0 & if(rand \geq S(V_i^d(t+1))) \\ 1 & else \end{cases} \tag{5}$$

where $S(x) = 1/(1 + exp(-1 * x))$.

3.2.1. *New transfer function.* The transfer function controls the switching rate of 0 or 1. The transfer function's slope means that the switching of positions changes fast when the velocity is low. To balance exploration and exploitation, positions change rapidly when binary PSO is in the initial stage, while their switching gradually slows down when the algorithm is in the later stage. The new transfer function of IBPSO is described as follows:

$$S(x) = 1/(1 + exp(-\tau * x)) \tag{6}$$

$$\tau = \tau_{max} - (\tau_{max} - \tau_{min}) * it/MAX\_IT \tag{7}$$

where τ_{max} is 30 and τ_{min} is 0.1. τ decreases linearly from 30 to 0.1, and the slope of the transfer function gradually decreases, as shown in Figure 1. When τ is 30, the transfer function has a large slope and approaches a right angle; when τ is 0.1, it tends to a horizontal straight line.

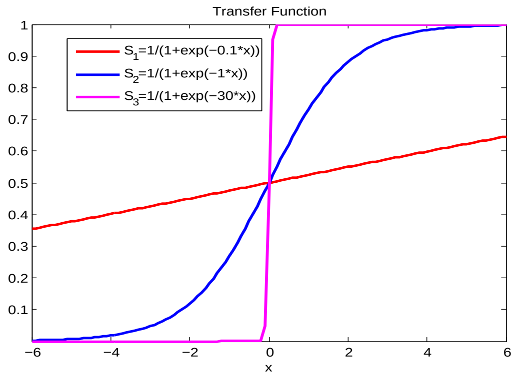

FIGURE 1. New transfer function.

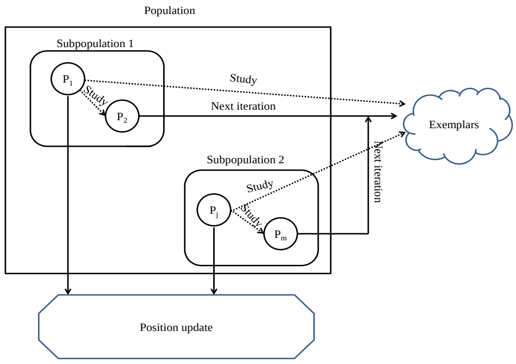

FIGURE 2. New learning exemplars.

3.2.2. *New learning exemplars.* Particles learn from *gbest* and *pbest*, and decide where they will move next time. Therefore, personal and historical optimal solutions influence the search direction of PSO. However, once they fall into local optimum or guide a wrong direction, the algorithm convergence stalls. New learning exemplars are introduced to improve the solution quality.

As shown in Figure 2, the population is randomly divided into multipopulations, and the particles in each subpopulation are randomly paired. The particles with great performance (winners) do not participate in position update, but directly enter to next iteration as exemplars. Loser particles learn from the exemplars and their winners, and the position update equation is described as follows.

$$V_i^d(t+1) = wV_i^d(t) + c1r_1(winner_i^d(t) - X_i^d(t)) + c2r_2(exemplar_i^d(t) - X_i^d(t)) \tag{8}$$

$$X_i^d(t+1) = \begin{cases} X_i^d(t) & if(rand \geq S(V_i^d(t+1))) \\ 1 - X_i^d(t) & else \end{cases} \tag{9}$$

$$X_i(t+1) = \begin{cases} X_i(t+1) & if(f(X_i(t+1)) \geq f(X_i(t))) \\ X_i(t) & else \end{cases} \tag{10}$$

To facilitate understanding, Figure 2 is used as an example for illustration. P_1 and P_2 in subpopulation 1 are randomly selected. Suppose the objective function value of P_1 is better than that of P_2. P_1 becomes an exemplar and P_2 learns from P_1 and a random one of exemplars. Similarly, P_m becomes an exemplar and does not participate in position update. P_j learns from P_m and exemplars.

Table 2. The main parameters setting of the compared algorithms.

Algorithm	Main parameters
IBPSO	wMax=0.9; wMin=0.4; Vmax=6;
BPSO	c1=2; c2=2; wMax=0.9; wMin=0.4; Vmax=6;
HGSO	K = 1; alpha = 1; beta = 1; L1 = 5E-3; L2 = 100; L3 = 1E-2;
BSCA	alpha = 2

Through the new learning approach, only half of particles participate in position update during each iteration.

IBPSO seeks a balance between exploration and exploitation to avoid the stagnation of local optimal solutions. In the early of IBPSO, the population is controlled by multiple solutions to bring more exploration, while in the later stage, a few optimal solutions guide population evolution to enhance exploitation. Therefore, the number of exemplars is determined as shown in Equation (11), and the presence of multiple exemplars promotes the adaptability and diversity of the population during search processes.

$$NL = 6 - ceil(5 * it/MAX\_IT) \tag{11}$$

It can accelerate the convergence of the algorithm, but may also lead to stagnation. So if a particle does not update within 5 iterations, it indicates that the particle may be caught in a local trap and needs to expand learners. The particle randomly selects two from the exemplars as its exemplar and winner to enlarge exploration ability.

3.2.3. *Advanced social cognitive factors.* In PSO, $c1$ and $c2$ are called social cognitive factors, and they are usually set as constants. The effect on the population remains unchanged regardless of the algorithm in any state and the positions of particles, leading to slow convergence and easy falling local optimum. The equations are suggested as follows to improve the performance of IBPSO.

$$\lambda(t) = (1 - \frac{t}{MAX\_IT})^{\frac{2t}{MAX\_IT}} \tag{12}$$

$$c1_i(t) = 2\lambda(t)(1 - \frac{1}{e^{r_3(winner_i(t) - X_i(t) + r_4\mu)}}) \tag{13}$$

$$c2_i(t) = 2\lambda(t)(1 - \frac{1}{e^{r_5(exemplar_i(t) - X_i(t) + r_6\mu)}}) \tag{14}$$

where r_3, r_4, r_5 and r_6 are four random numbers between [0,1], and μ a very small value.

It can be found by Equations (13) and (14) that if a particle is close to its exemplar and winner, $c1$ and $c2$ reduce their attraction, thus prompting the particle to fluctuate in less space and enhancing local search. If a particle is far from its exemplar and winner, $c1$ and $c2$ increase their guidance, promoting it to approach them quickly and improving convergence.

4. **Experimental results and analysis.** The benchmark datasets obtained from UCI are adopted to compare the performance of IBPSO with BPSO [27], HGSO [28] and BSCA [29]. Finally, they are validated in our collected data to predict students' academic performance. Each algorithm runs 20 times, with 100 iterations each time. The population size is 30. Since only half of the particles in IBPSO participate in position update during each circle, its iteration number is 200. Their parameters are shown in Table 2.

Table 3. UCI datasets.

Dataset	Attributes	Instances
Chess	36	3196
Divorce	54	170
HCC	49	165
Heart	13	270
Ionosphere	34	351
Sonar	60	208
Soybean	35	307
Spect	22	267
Waveform	40	5000
Zoo	17	101

4.1. **Benchmark datasets.** We select 10 datasets from UCI [30], and Table 3 presents a brief description of them.

Table 4 is the experimental statistical results, where *Error* means the obtained classification error, and *Len* represents the number of selected features. Table 4 implies that IBPSO performs best, and it outperforms BPSO, HGSO and BSCA on 6 datasets, Chess, Heart, Sonar, Soybean, Spect and Waveform. BPSO has the best results in HCC and ZOO, while HGSO and BSCA have excellent performance in Divorce and Ionosphere, respectively. IBPSO is superior to BPSO and other compared algorithms, which indicates that the proposed IBPSO improves the classification ability of feature selection.

The algorithms have small classification errors in Chess, Divorce, Ionosphere, and Zoo, which are no more than 1%, and they have large errors in Spect and Waveform. Although there are not many instances of Zoo, its features are simple to establish classification models with few features. The classification accuracy in Voting, Tic, and Waveform is relatively low. Voting and Tic have small features to construct classification models, and the data types of Waveform are complex to cause poor generalization and build an accurate model.

To verify the performance of the algorithms, two nonparametric verification methods, Wilcoxon rank sum and Friedman test, are utilized to confirm the effectiveness of the obtained experimental data. The last three rows of Table 4 are their results where ">", "=", and "<" respectively suggest significantly better, statistically similar, and significantly worse results.

It is found from Table 4 that IBPSO and BPSO have the same statistics in HCC, Heart, Spect and Zoo. In Spect and Ionosphere, BSCA cannot distinguish sample statistics with IBPSO and HGSO , respectively. Wilcoxon rank sum presents that BPSO, HGSO, BSCA and IBPSO perform well on 4, 2, 2 and 8 datasets. Their average ranks are 2.3, 3.4, 2.6, and 1.7, with P-Value less than 0.05. Wilcoxon rank sum and Friedman test confirm that IBPSO is superior to other algorithms.

Regarding the number of selected features, BSCA uses few feature ratios to complete classification, while BPSO and IBPSO have large ratios. The algorithms employ enormous features in Chess, Soybean and Waveform, and they adopt less than 30% features to complete the classification in Ionosphere. Feature selection improves classification ability by selecting appropriate features to complete data modeling, and more/fewer features are not conducive to improving the classification of models.

The experimental statistical results and the nonparametric verification illustrate that IBPSO improves the classification performance of feature selection.

TABLE 4. The classification errors and selected numbers of the compared algorithms.

Dataset	BPSO		HGSO		BSCA		IBPSO	
	Error	**Len**	**Error**	**Len**	**Error**	**Len**	**Error**	**Len**
Chess	0.0306	20.85	0.0594	22.15	0.0539	13.2	**0.0255**	20.8
Divorce	0.0168	24.5	**0.0103**	4.8	0.0118	13.85	0.0176	25.3
HCC	**0.0836**	24.7	0.1119	24.3	0.1037	13	0.084	22.9
Heart	0.1378	6.9	0.1598	4.7	0.1476	4.55	**0.1352**	6.45
Ionosphere	0.0890	8.65	0.0775	3.6	**0.0682**	4.25	0.0793	8.9
Sonar	0.1127	29.15	0.1310	20.85	0.1099	14.5	**0.1014**	27.1
Soybean	0.0921	23.55	0.1261	22.55	0.1350	15.4	**0.0806**	23.7
Spect	0.2434	8.55	0.2605	6.4	0.2423	5.25	**0.2369**	9.35
Waveform	0.1623	22.25	0.1775	29.5	0.174	15.2	**0.1572**	22.3
Zoo	**0.0297**	7.95	0.0446	8.65	0.0371	7.2	0.0307	8.35
¡=¿	2/2/6		1/1/8		1/1/8		6/2/2	
AVG	2.3		3.4		2.6		1.7	
P-value	0.029							

TABLE 5. The average running time of the compared algorithms (second).

Dataset	**BPSO**	**HGSO**	**BSCA**	**IBPSO**
Chess	656	554	**497**	643
Divorce	202	**199**	201	**199**
HCC	202	201	202	**199**
Heart	218	**184**	210	215
Ionosphere	208	**199**	211	207
Sonar	204	207	203	**201**
Soybean	224	239	226	**223**
Spect	213	**195**	212	209
Waveform	1382	**991**	1017	1347
Zoo	216	**203**	211	213

Table 5 is the average running time of the algorithms. It is noticed that HGSO has the highest efficiency on 5 datasets, and IBPSO and BSCA have excellent performance on 4 and 1 datasets. In feature selection, the performance of an algorithm is heavily influenced by classifier. The algorithms spend significantly more time on Chess and Waveform than on other datasets because they have the largest amounts of data.

4.2. **Students' academic performance in college English.** Figure 3 is the prediction results of students' academic performance with our proposed model in Section 3.1.

As shown in Figure 3, the classification errors of BPSO, HGSO, BSCA and IBPSO are 0.3043, 0.2756, 0.3033, 0.3011 and 0.1864, respectively. IBPSO performs the best, followed by HGSO, BSCA and BPSO. The numbers of selected features are 10.3, 2.6, 7.5 and 3.3. The algorithms are able to achieve classification with few features, especially IPBSO and HGSO, which predict students' performance with few features. The running time is 2171, 1877, 2080 and 1900 seconds, respectively. When the amount of data is large, the feature subset becomes one of the key factors affecting efficiency.

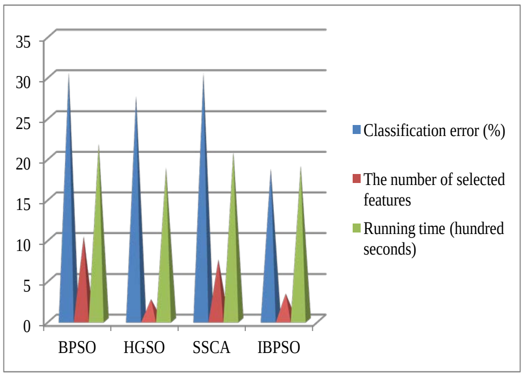

FIGURE 3. The prediction results of students' academic performance.

In IBPSO, the features that are commonly utilized are *PlaceOrigin*, *Gender*, *Importance*, *Absence* and *Difficulty*, while in HGSO, *PlaceOrigin* and *Gender* are frequently selected features. In SCA and BPSO, they use *Absence*, *PlaceOrigin*, *Importance*, *OnlineTime*, *LearningHabits*, *Classroom*, *StudyTime* to classify. The selected features of IBPSO and HGSO imply that demographic features affect students' performance in College English, while in SCA and BPSO, behavioral features are the main factors.

In China, the region of student source influences the quality of students' performance, and they generally have better grades in areas with developed education. The attention that families pay to students is also an important part that affects students' academic performance. This is probably because parental care and encouragement make students more motivated to study, so great home-school interaction can improve students' academic performance. Surprisingly, *gender* is also a more selected feature, which requires educators to research its impact on students' performance in the future.

Experiments on the real dataset show that the proposed algorithm is effective and can extract high-quality features that affect students' academic performance, providing classification models and teaching assistance for college English education.

5. **Conclusions.** With the development of higher education, it has become an urgent task that improves teaching management methods and teaching efficiency in traditional education processes. However, students' academic performance is a key point reflecting the quality of teaching and learning. It also illustrates students' mastery and proficiency of knowledge and reflects the effectiveness of teachers in teaching processes. To effectively extract features from educational data, this paper proposes a PSO-based model to predict students' academic performance in college English. By analyzing the characteristics of binary PSO, the transfer function is improved and a new learning scheme is proposed to avoid falling into local traps. In the benchmark datasets, IBPSO outperforms BPSO, HGSO and BSCA. In the students' academic performance dataset we obtained, its performance is also excellent. The algorithms expose that the region of student source and gender are the main factors affecting students' academic performance. In future research work, we can explore students' learning ability from multi-view multi-source data and feedback (English) teaching according to the acquired features.

Acknowledgments. This work is supported by the Support Program for Scientific and Technological Innovation Teams in Universities in Henan Province (24IRTSTHN028).

REFERENCES

[1] X. Huang, H. Xiong, J. Chen, and M. Yang, "Efficient revocable storage attribute-based encryption with arithmetic span programs in cloud-assisted internet of things," *IEEE Transactions on Cloud Computing*, vol. 11, no. 2, pp. 1273–1285, 2021.

[2] M.-E. Wu, H.-H. Tsai, W.-H. Chung, and C.-M. Chen, "Analysis of kelly betting on finite repeated games," *Applied Mathematics and Computation*, vol. 373, 125028, 2020.

[3] H. Xiong, T. Yao, H. Wang, J. Feng, and S. Yu, "A survey of public-key encryption with search functionality for cloud-assisted iot," *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 401–418, 2021.

[4] T.-L. Luo, M.-E. Wu, and C.-M. Chen, "A framework of deep reinforcement learning for stock evaluation functions," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 5, pp. 5639–5649, 2020.

[5] L. Ye, X. Meng, Y. Hang, and S. Tayeb, "Research on visual tracking method for students' browsing data in art literacy online education," *Journal of Network Intelligence*, vol. 5, pp. 218–225, 2020.

[6] H. Xiong, X. Huang, M. Yang, L. Wang, and S. Yu, "Unbounded and efficient revocable attribute-based encryption with adaptive security for cloud-assisted internet of things," *IEEE Internet of Things Journal*, vol. 9, no. 4, pp. 3097–3111, 2021.

[7] C.-M. Chen, Y. Gong, and J. M.-T. Wu, "Impact of technical indicators and leading indicators on stock trends on the internet of things," *Wireless Communications and Mobile Computing*, vol. 2022, 9916310, 2022.

[8] T.-Y. Wu, Q. Meng, Y.-C. Chen, S. Kumari, and C.-M. Chen, "Toward a secure smart-home iot access control scheme based on home registration approach," *Mathematics*, vol. 11, no. 9, 2123, 2023.

[9] J.-S. Pan, L. Yue, S.-C. Chu, P. Hu, B. Yan, and H. Yang, "Binary bamboo forest growth optimization algorithm for feature selection problem," *Entropy*, vol. 25, no. 2, 314, 2023.

[10] T.-Y. Wu, A. Shao, and J.-S. Pan, "Ctoa: toward a chaotic-based tumbleweed optimization algorithm," *Mathematics*, vol. 11, no. 10, 2339, 2023.

[11] T.-Y. Wu, H. Li, and S.-C. Chu, "Cppe: an improved phasmatodea population evolution algorithm with chaotic maps," *Mathematics*, vol. 11, no. 9, 1977, 2023.

[12] J. Xi, Y. Chen, X. Liu, and X. Chen, "Whale optimization algorithm based on nonlinear adjustment and random walk strategy," *Journal of Network Intelligence*, vol. 7, no. 2, pp. 306–318, 2022.

[13] J.-J. Yan, J.-S. Fang, J. S.-H. Tsai, C.-H. Huang, and S.-M. Guo, "Robust digital-redesign tracking control for uncertain systems: Pid sliding mode control and pso algorithm," *Journal of Network Intelligence*, vol. 6, no. 4, pp. 668–687, 2021.

[14] X. Cheng, Y. Jiang, D. Li, Z. Zhu, and N. Wu, "Optimal operation with parallel compact bee colony algorithm for cascade hydropower plants," *Journal of Network Intelligence*, vol. 6, no. 3, pp. 440–452, 2021.

[15] J.-S. Pan, P. Hu, V. Snášel, and S.-C. Chu, "A survey on binary metaheuristic algorithms and their engineering applications," *Artificial Intelligence Review*, pp. 1–67, 2022.

[16] S. Mirjalili and A. Lewis, "S-shaped versus v-shaped transfer functions for binary particle swarm optimization," *Swarm and Evolutionary Computation*, vol. 9, pp. 1–14, 2013.

[17] P. Hu, J.-S. Pan, and S.-C. Chu, "Improved binary grey wolf optimizer and its application for feature selection," *Knowledge-Based Systems*, vol. 195, 105746, 2020.

[18] G.-L. Wang, S.-C. Chu, A.-Q. Tian, T. Liu, and J.-S. Pan, "Improved binary grasshopper optimization algorithm for feature selection problem," *Entropy*, vol. 24, no. 6, 777, 2022.

[19] J.-S. Pan, A.-Q. Tian, S.-C. Chu, and J.-B. Li, "Improved binary pigeon-inspired optimization and its application for feature selection," *Applied Intelligence*, vol. 51, no. 12, pp. 8661–8679, 2021.

[20] Z.-G. Du, J.-S. Pan, S.-C. Chu, and Y.-J. Chiu, "Improved binary symbiotic organism search algorithm with transfer functions for feature selection," *IEEE Access*, vol. 8, pp. 225 730–225 744, 2020.

[21] B. K. Francis and S. S. Babu, "Predicting academic performance of students using a hybrid data mining approach," *Journal of Medical Systems*, vol. 43, pp. 1–15, 2019.

[22] Z. Liang, X. Di, Z. Liu, X. Liu, X. Zhang, and Z. Yu, "Predicting students' academic performance based on improved pso-xgboost: A campus behavior perspective," in *Algorithms and Architectures for Parallel Processing: 21st International Conference (ICA3PP 2021)*. Springer, 2022, pp. 402–421.

[23] A. Farissi and H. M. Dahlan, "Genetic algorithm based feature selection for predicting student's academic performance," in *Emerging Trends in Intelligent Computing and Informatics: Data Science, Intelligent Information Systems and Smart Computing 4 (ETCSIT 2020)*. Springer, 2020, pp. 110–117.

[24] J.-F. Chen and Q. H. Do, "A cooperative cuckoo search–hierarchical adaptive neuro-fuzzy inference system approach for predicting student academic performance," *Journal of Intelligent & Fuzzy Systems*, vol. 27, no. 5, pp. 2551–2561, 2014.

[25] P. Hu, J.-S. Pan, S.-C. Chu, and C. Sun, "Multi-surrogate assisted binary particle swarm optimization algorithm and its application for feature selection," *Applied Soft Computing*, vol. 121, 108736, 2022.

[26] S. Qin, C. Sun, G. Zhang, X. He, and Y. Tan, "A modified particle swarm optimization based on decomposition with different ideal points for many-objective optimization problems," *Complex & Intelligent Systems*, vol. 6, pp. 263–274, 2020.

[27] J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," in *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation (ICSMC 1997)*, vol. 5. IEEE, 1997, pp. 4104–4108.

[28] N. Neggaz, E. H. Houssein, and K. Hussain, "An efficient henry gas solubility optimization for feature selection," *Expert Systems with Applications*, vol. 152, 113364, 2020.

[29] K. S. Reddy, L. K. Panwar, B. Panigrahi, and R. Kumar, "A new binary variant of sine–cosine algorithm: development and application to solve profit-based unit commitment problem," *Arabian Journal for Science and Engineering*, vol. 43, pp. 4041–4056, 2018.

[30] M. Lichman, "(2013) uci machine learning repository," Available at: (http://archive.ics.uci.edu/ml. Irvine, CA: University of California, School of Information and Computer Science).