ICIC Express Letters
Part B: Applications
Volume **8**, Number **7**, July **2017**

ICIC International ©2017 ISSN 2185-2766
pp. 1015–1021

AN IMPROVED FRUIT FLY OPTIMIZATION ENHANCED KERNEL EXTREME LEARNING MACHINE WITH APPLICATION TO SECOND MAJOR PREDICTION

Yan Wei[1], Huiling Chen[2,*], Jie Luo[2] and Qiang Li[2]

[1]Institute of Information Technology
Wenzhou Vocational College of Science and Technology
No. 1000, Liuhongqiao Road, Wenzhou 325006, P. R. China
weiyan@wzvcst.edu.cn

[2]College of Physics and Electronic Information Engineering
Wenzhou University
Wenzhou University Town, Wenzhou 325035, P. R. China
*Corresponding author: chenhuiling.jlu@gmail.com; { 767340601; 1727849802 }@qq.com

Received February 2017; accepted April 2017

Abstract. *There are significant potential benefits for building a scientific and major assistant selection system that can help students scientifically choose majors. This study proposes an effective prediction model for choosing the second major based on an improved fruit fly optimization algorithm (IFOA) enhanced kernel extreme learning machine (KELM). The IFOA strategy was adopted to adaptively determine the optimal parameters in KELM. The resulting prediction model, IFOA-KELM, was compared against the other four competitive methods, including support vector machines and several KELM models optimized by original FOA, particle swarm optimization and a grid search technique on a real-life dataset via a 10-fold cross validation scheme. The obtained results clearly confirm the superiority of our proposed model in classification accuracy, area under the receiver operating characteristic curve (AUC), sensitivity, and specificity. Promisingly, the proposed method has the potential to serve as an excellent new candidate for powerful major selection assistant systems for undergraduate students.*

Keywords: Kernel extreme learning machine, Fruit fly optimization algorithm, Second major selection

1. **Introduction.** Data mining techniques have drawn increasing attention from the higher education domain, and have been introduced in this field for course selection, scores prediction, and career guidance [1]. Kardan et al. [2] discussed course selection prediction using neural network techniques in the context of e-learning. This study focuses on identifying the potential factors that affect student satisfaction concerning their online courses. Guo [3] proposed using statistical analysis and neural networks to establish dynamic models for analyzing and predicting students' course satisfaction. The results showed that the three-layer multilayer perceptron models outperformed linear regressions in predicting student course satisfaction; the best outcome was achieved by combining the number of students and high distinction as inputs in the networks. Campagni et al. [4] proposed a clustering and series pattern-based method for analyzing the careers of college graduates, which was helpful in graduate career planning. Huang and Xu [5] analyzed the evaluation theories of statistics and psychology, and developed a second major selection system based on a rough set-based association rule mining algorithm. The simulation results demonstrated that their method was accurate for software engineering, networking, and programming majors. Lee [6] performed logical regression analysis and found that a large number of computer courses learned at the middle school level had a significant impact on STEM subject selection in the U.S. colleges.

To improve the performance of second major selection, this study proposes and examines a new machine learning method, kernel extreme learning machine (KELM) [7]. To the best of the authors' knowledge, KELM has not yet been used for selecting second majors. Several studies have shown that parameters such as the penalty parameter C and kernel width γ in KELM have a significant impact on performance. Therefore, we explored a swarm intelligence technique to overcome parameter optimization problems in KELM just as done in support vector machines (SVM). Like other global optimization algorithms (such as genetic algorithms, PSO), the FOA [8] is easy to fall into local optimum, resulting in slower convergence speed at late time and the convergence precision is gradually reduced, especially for high-dimensional multi-polar complex optimization problems. So we try to extend the two-dimensional coordinates to three-dimensional coordinates where fruit flies are more likely to shift freely. An improved fruit fly optimization algorithm (FOA) strategy was considered to tune KELM parameters due to its simple implementation and good optimization capability. The efficacy of the resultant method, the IFOA-KELM-based prediction system, was rigorously compared against SVM and KELM models optimized by original FOA (FOA-KELM), particle swarm optimization (PSO-KELM) [9] and a grid search technique (Grid-KELM) [10] on the real-life dataset collected from Wenzhou Vocational College of Science and Technology.

The rest of this paper is organized as follows. Section 2 presents the detailed implementation of the proposed method. Section 3 describes the experimental design. Section 4 presents the experimental results and discusses the proposed approach. Finally, Section 5 summarizes the conclusions and recommendations for future work.

2. Proposed Prediction Model.

In this section, we briefly describe the proposed system for second major selection. The proposed system is mainly constructed based on KELM, where the input data is comprised of a series of factors that influence the choice of the specific major. The data in the input space was mapped into the hidden-layer feature space by using the (radial basis function) RBF kernel. The optimal penalty parameter C and kernel width γ were dynamically specified by the IFOA strategy. When the optimal parameter pair was obtained through the training phase, it was fed into the KELM model to perform the prediction task.

The pseudo-code of the whole procedure is given below.

3. Experimental Design.

3.1. Data description.

The data used for this study was acquired from Wenzhou Vocational College of Science and Technology. The set contained 402 students that majored in Digital Media Technology, which includes Graphic Design and Video Production. At the end of the second term, 195 of the students decided on the Graphic Design major, and 207 students selected the Video Production. The 12 factors include gender, type of college entrance applications, whether they came from Zhejiang, whether they came from Wenzhou, whether they were science students, whether they volunteered to major in digital media, basic course scores, participation in after-class graphic design activities, participation in after-class video production activities, whether they enrolled in self-study undergraduate courses, the scores of the basic course relevant to graphic design and the scores of the basic course relevant to video production.

3.2. Experimental setup.

IFOA, FOA and PSO were implemented from scratch. For SVM, LIBSVM implementation developed by Chang and Lin [11] was utilized. The implementation code that was used to construct the KELM models is available at http://www3.ntu.edu.sg/home/egbhuang. Data was scaled into the range $[-1, 1]$ before classification. The empirical experiment was conducted on an AMD Athlon 64 X2 Dual Core Processor 5000+ (2.6 GHz) with 4 GB of RAM, running Windows 7.

```
Begin
    For i = 1 to sizepop
        Set the KELM parameters with the initialized distance reciprocal S(i,1) and
        S(i,2);
        Calculate the initial fitness;
        Train the KELM model with the distance reciprocal, and record test results
        into the Smell array;
    End
    [bestSmell, bestindex] = max(Smell);
    past position = current position;
    bestCV = bestSmell;
    bestC = S(bestIndex, 1);
    bestg = S(bestIndex, 2);
    For j = 1: maxgen
        For i = 1 to sizepop
            X(i,:) = X_axis + ax * rand() - bx;
            Y(i,:) = Y_axis + ax * rand() - bx;
            Z(i,:) = Z_axis + ax * rand() - bx;
            D(i,1) = sqrt(X(i,1)^2 + Y(i,1)^2 + Z(i,1)^2);
            D(i,2) = sqrt(X(i,2)^2 + Y(i,2)^2 + Z(i,2)^2);
            S(i,1) = 1/D(i,1);
            S(i,2) = 1/D(i,2);
            Set the KELM parameters with S(i,1) and S(i,2);
            Calculate the initial fitness;
           Train the KELM model with the distance reciprocal, and record test
           results into the Smell array;
        End
        [bestSmell, bestIndex] = max(Smell);
        If (bestSmell > bestCV)
            past position = current position;
            bestC = S(bestIndex, 1);
            bestg = S(bestIndex, 2);
            bestCV = bestSmell;
        End If
    End
    Return bestC, bestg;
End
```

The 10-fold cross validation (CV) was used to evaluate classification performance to guarantee unbiased results. The parameter settings for other involved algorithms were as follows. The numbers of the maximum iterations and swarm size were set to 100 and 25, respectively. The common parameter searching range was set as $C = \{2^{-10}, 2^{-8}, \ldots, 2^{10}\}$ and $\gamma = \{2^{-10}, 2^{-8}, \ldots, 2^{10}\}$ for all the methods. For PSO, the maximum velocity was set to about 60% of the dynamic range of the variable on each dimension for the continuous type of dimensions. The two acceleration coefficients c_1 and c_2 were set as 2.05, the inertia weight was set to 1. For FOA, the parameters of *ax*, *bx*, *ay*, and *by* in the distance equations were set as 20, 10, 20, and 10, respectively.

To evaluate the performance of the second major selection by the IFOA-KELM approach, we mainly examined four metrics: classification accuracy (ACC), the area under the Receiver Operating Characteristic (ROC) curve (AUC), sensitivity, and specificity.

4. **Experimental Results and Discussion.** In this experiment, we firstly evaluated the effectiveness of the proposed IFOA approach on six multidimensional benchmark functions as shown in Table 1. For comparison purpose, the original FOA was tested on the same functions. We tested each function for 10 times, and calculated the average (Ave) and standard deviation (Std) respectively.

Table 2 lists the average and standard deviation of the six benchmark functions in the dimension 2, 10, 50 and 100, respectively. From the table we can see that the proposed IFOA has achieved better test testing results than that of FOA at almost all the functions with smaller average values. Figure 1 records the convergence curve of each function in the dimension 2 for the two methods. As shown, the IFOA algorithm proposed in this

Table 1. Benchmark functions

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

Table 2. Results of testing benchmark functions

Benchmark function	Dimension	FOA		IFOA	
		Ave	Std	Ave	Std
f_1	2	1.4415e-04	2.1948e-05	9.0553e-05	1.3691e-05
	10	0.0012	1.6483e-04	0.0007	2.1339e-04
	50	0.0103	0.0015	0.0046	0.0017
	100	0.0244	0.0021	0.0119	0.0026
f_2	2	0.1697	0.0168	0.1213	0.0056
	10	0.9922	0.1229	0.5774	0
	50	6.5693	0.2733	2.8868	4.6811e-16
	100	13.6200	0.9842	5.7735	9.3622e-16
f_3	2	3.1733e-04	3.6984e-05	2.2451e-04	4.2667e-05
	10	0.0401	0.0078	0.0242	0.0063
	50	7.3781	0.3497	3.4982	1.1930
	100	66.7212	4.6032	26.8823	10.9707
f_4	2	0.0091	6.7821e-04	0.0069	7.8237e-04
	10	0.0138	0.0015	0.0105	0.0014
	50	0.0270	0.0040	0.0184	0.0012
	100	0.0394	0.0041	0.0224	0.0018
f_5	2	0.6613	0.0500	0.8494	0.0767
	10	9.2716	0.2123	9.0534	0.1021
	50	54.8054	1.4254	49.8958	1.2753
	100	113.4997	1.7723	101.0448	2.9794
f_6	2	0.2992	0.0222	0.1545	0.0270
	10	5.1638	0.0406	2.3059	0.0285
	50	118.7720	0.0208	52.8057	0.0216
	100	470.3355	0.0150	209.0674	0.0274

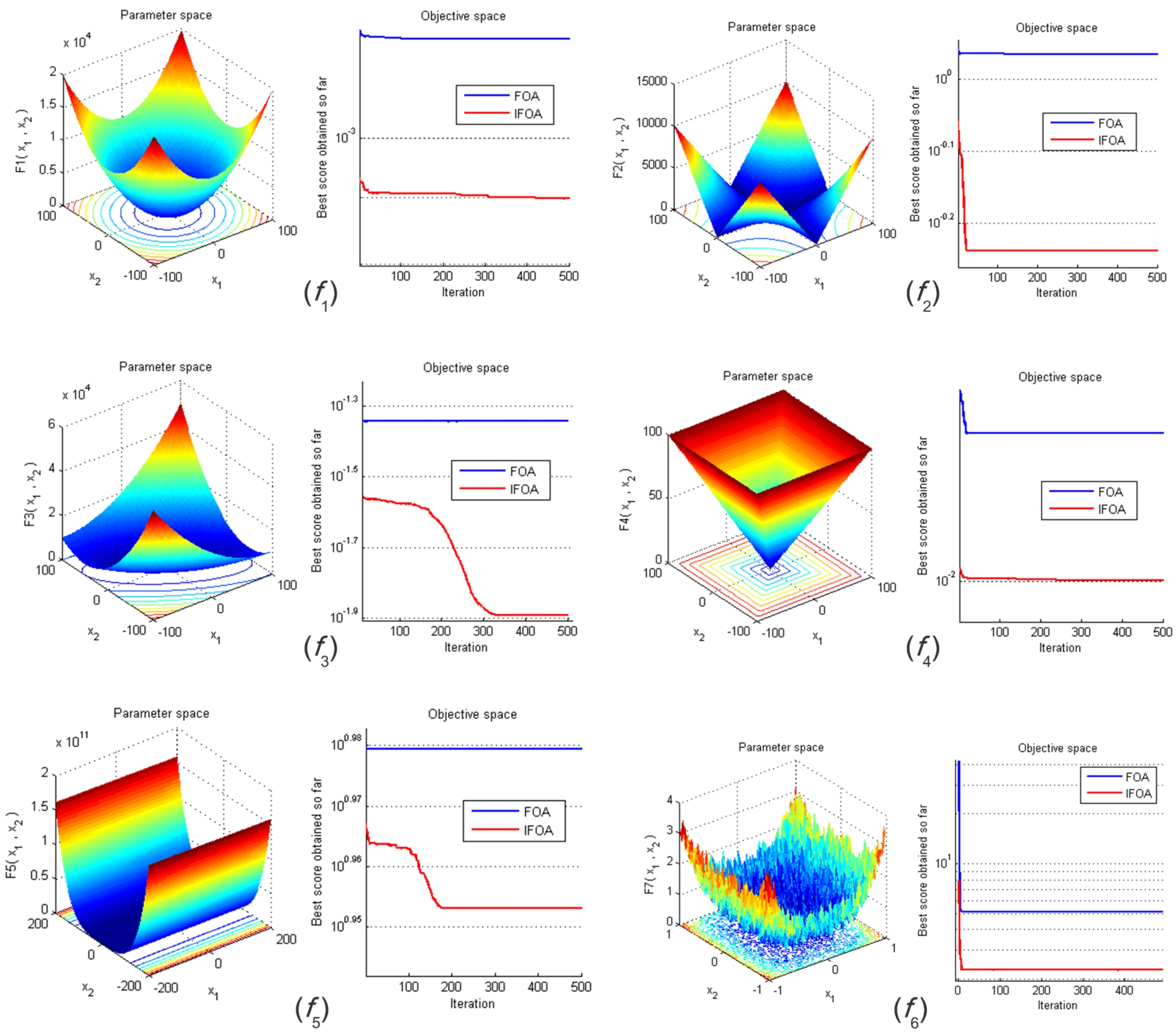

FIGURE 1. Convergence curves of six functions for IFOA and FOA when dimension = 2

paper has much higher convergence speed than FOA at all the functions. It indicates that the IFOA has better capability to find the better solutions in a much wider space, and thus accelerating the fruit flies to overstep the local extremum.

In the following part, we evaluated the effectiveness of the IFOA-KELM model on the original feature space. Table 3 shows the detailed results achieved by IFOA-KELM. As illustrated in the table, IFOA-KELM achieved the average results of 81.55% ACC, 80.63% AUC, 83.17% sensitivity, and 81.90% specificity. In addition, it can be observed that the values of C and γ can be specified adaptively for each data fold. The explanation lies in the fact that the two parameters can be adaptively determined by the IFOA strategy according to the specific distribution of the training data.

The comparative results of the five methods are recorded in Figure 2. As shown, we can see that IFOA-KELM surpassed FOA-KELM, PSO-KELM, SVM and Grid-KELM in terms of ACC, AUC, and sensitivity. The IFOA-KELM specificity was also comparable to that of the other four methods. FOA-KELM and PSO-KELM have achieved almost the same ACC and AUC, while the FOA-KELM can give more stable results than PSO-KELM with smaller standard deviation. SVM produced the worst result among the five methods, while the standard deviation produced for ACC and AUC was much smaller than Grid-KELM and PSO-KELM. IFOA-KELM's better performance may be due to the fact that the IFOA aided the KELM classifier in achieving the maximum classification performance by automatically detecting the optimal parameter pair. Additionally, it can

Table 3. The detailed results obtained by IFOA-KELM

Fold	C	γ	ACC	AUC	Sensitivity	Specificity
#1	218.932	771.89	0.8000	0.8667	0.7600	0.8133
#2	1024	987.022	0.7500	0.8000	0.7200	0.7600
#3	229.711	234.826	0.8049	0.8000	0.8095	0.8048
#4	863.118	820.916	0.8500	0.8095	0.8947	0.8521
#5	343.889	515.911	0.8000	0.7727	0.8333	0.8030
#6	295.04	462.727	0.8500	0.8095	0.8947	0.8521
#7	925.947	804.954	0.8000	0.7619	0.8421	0.8020
#8	198.499	708.903	0.8250	0.8333	0.8182	0.8258
#9	611.139	788.709	0.8250	0.8000	0.8500	0.8250
#10	923.641	934.93	0.8500	0.8095	0.8947	0.8521
Avg.	563.392	703.079	0.8155	0.8063	0.8317	0.8190
Dev.	341.376	231.614	0.0314	0.0291	0.0584	0.0291

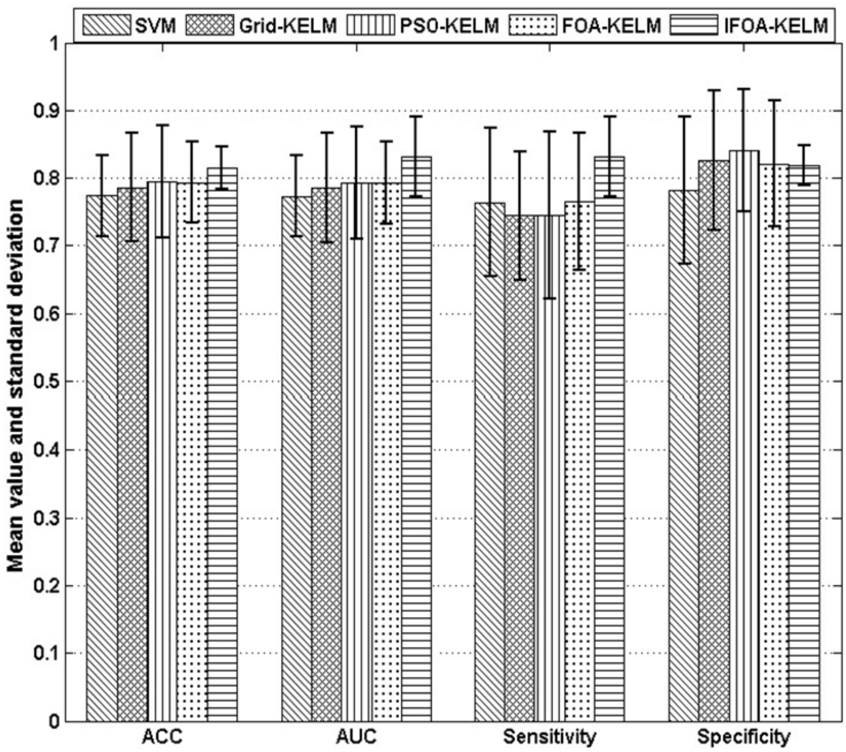

Figure 2. The classification performance obtained by IFOA-KELM, FOA-KELM, PSO-KELM, Grid-KELM, and SVM in terms of ACC, AUC, sensitivity, and specificity

be seen from the figure that the standard deviation of the IFOA-KELM was much smaller than the other four competitors in terms of the four evaluation metrics. The comparison results indicate that IFOA-KELM is the most stable and robust method for second major selection.

5. **Conclusions and Future Work.** In this study, we explored the feasibility of applying the KELM classifier to effectively assisting students in selecting a major. To exploit the maximum potential of KELM, an improved FOA strategy, IFOA, was employed to search for the optimal parameters. The experimental results demonstrated that the developed model performed better than other four advanced machine learning models in terms of the ACC, AUC, sensitivity, and specificity on the real life-life dataset. Therefore, it can be

safely concluded that the developed intelligent system can serve as a promising alternative decision support system for students' second major selection.

Acknowledgment. This study was financially supported by the National Natural Science Foundation of China (61303113), Zhejiang Provincial Natural Science Foundation of China (LY17F020012), the Science and Technology Plan Project of Wenzhou, China (G20140048, R20160035), Scientific Research Fund of Zhejiang Provincial Education Department (Y201533884).

REFERENCES

[1] S. Natek and M. Zwilling, Student data mining solution – Knowledge management system related to higher education institutions, *Expert Systems with Applications*, vol.41, no.14, pp.6400-6407, 2014.

[2] A. A. Kardan et al., Prediction of student course selection in online higher education institutes using neural network, *Computers and Education*, vol.65, pp.1-11, 2013.

[3] W. W. Guo, Incorporating statistical and neural network approaches for student course satisfaction analysis and prediction, *Expert Systems with Applications*, vol.37, no.4, pp.3358-3365, 2010.

[4] R. Campagni et al., Data mining models for student careers, *Expert Systems with Applications*, vol.42, no.13, pp.5508-5521, 2015.

[5] C. Huang and J. Xu, Second major selection system based on data mining, *Journal of Information and Computational Science*, vol.6, no.6, pp.2333-2340, 2009.

[6] A. Lee, Determining the effects of computer science education at the secondary level on STEM major choices in postsecondary institutions in the United States, *Computers & Education*, vol.88, pp.241-255, 2015.

[7] G. B. Huang et al., Extreme learning machine for regression and multiclass classification, *IEEE Trans. Systems Man and Cybernetics Part B – Cybernetics*, vol.42, no.2, pp.513-529, 2012.

[8] W.-T. Pan, A new fruit fly optimization algorithm: Taking the financial distress model as an example, *Knowledge-Based Systems*, vol.26, pp.69-74, 2012.

[9] X. C. Yang, X. R. Yan and C. F. Song, Pressure prediction of coal slurry transportation pipeline based on particle swarm optimization kernel function extreme learning machine, *Mathematical Problems in Engineering*, vol.2015, no.8, pp.1-7, 2015.

[10] D. Zhao et al., An effective computational model for bankruptcy prediction using kernel extreme learning machine approach, *Computational Economics*, vol.49, no.2, pp.325-341, 2017.

[11] C.-C. Chang and C.-J. Lin, LIBSVM: A library for support vector machines, *ACM Trans. Intelligent Systems and Technology (TIST)*, vol.2, no.3, p.27, 2011.