

## A Hybrid Algorithm Based on PSO and GA for Feature Selection

Yu Xue<sup>1,\*</sup>, Asma Aouari<sup>1</sup>, Romany F. Mansour<sup>2</sup> and Shoubao Su<sup>3</sup>

<sup>1</sup>Nanjing University of Information Science and Technology, Nanjing, 210044, China

<sup>2</sup>Department of Mathematics, Faculty of Science, New Valley University, El-Kharja, 72511, Egypt

<sup>3</sup>Jiangsu Key Laboratory of Data Science and Smart Software, Jinling Institute of Technology, Nanjing, 211169, China

\*Corresponding Author: Yu Xue. Email: xueyu@nuist.edu.cn

Received: 21 January 2021; Accepted: 14 July 2021

**Abstract:** One of the main problems of machine learning and data mining is to develop a basic model with a few features, to reduce the algorithms involved in classification's computational complexity. In this paper, the collection of features has an essential importance in the classification process to be able minimize computational time, which decreases data size and increases the precision and effectiveness of specific machine learning activities. Due to its superiority to conventional optimization methods, several metaheuristics have been used to resolve FS issues. This is why hybrid metaheuristics help increase the search and convergence rate of the critical algorithms. A modern hybrid selection algorithm combining the two algorithms; the genetic algorithm (GA) and the Particle Swarm Optimization (PSO) to enhance search capabilities is developed in this paper. The efficacy of our proposed method is illustrated in a series of simulation phases, using the UCI learning array as a benchmark dataset.

**Keywords:** Evolutionary computation; genetic algorithm; hybrid approach; meta-heuristic; feature selection; particle swarm optimization

### 1 Introduction

The selection of features is a complicated work because challenging functional interactions are possible. When interacting with other functions, an individually appropriate element can be redundant. Due to the vast search area, the feature selection process is complex. The search area is increasingly increasing in size with regard to the amount of data set features [1]. Consequently, a complete search is almost futile in most situations. Many methods for active filtering, such as Sequential Backward Selection (SBS) [2], have been applied in order to solve this problem. However, these methods to select features also face a range of challenges such as local optima inflation and high computational costs. An effective global search technique is essential in order to help address feature selection issues. Methods of Evolutionary Computing (EC) are famous for their global search capability. PSO is a relatively new EC technique focused on swarm intelligence [3,4].

The technology is relatively modern. PSO is computationally cheaper and can converge more rapidly than other EC algorithms like Gas. Reasons why PSO was brought into for this field, including a collection of features as an essential technique [5–7]. This technique was introduced by Eberhart et al. [8] as a population-based stochastic optimizer technique. PSO simulates a swarm of insects, an assembly of birds, or a fish school's social behavior. PSO shares certain parallels with evolutionary computer technologies such as genetic algorithms and artificial neural networks. It comes in-between with a



random population and upgrades generations to scan for optima. If one particle finds a decent way to eat, it is possible to immediately obey the rest of the colony, such as the Swarm of Ants.

In a multidimensional research area, the swarm behavior is modeled on particles with positions and velocity as features. In hyperspace, these particles feed by adapting their location and speed.

Inspired by PSO and GA, a new evolutionary hybrid algorithm (HEA) is developed by integrating the advantages of both algorithms to determine the most likely features in the chosen subset of features. The technique of the profit ratio is used to achieve this.

The swarm conduct is modeled on particles with locations and velocity in a multidimensional research environment. These particles feed through hyperspace by adjusting their own position and direction.

It is how the majority of the document is arranged. We provide history material in Chapter 2. The related works are discussed in Chapter 3. The new algorithm, the strategy bundle creation methods, is defined in Chapter 4 while Chapter 5 discusses the nature of the experiments, and Chapter 6 discusses the findings. The results and prospective research work are presented in Chapter 7.

## 2 Literature Survey

Numerous optimization algorithms to solve the complicated optimization problems have been recently proposed, as this vast number of algorithms have different behavior. We classify it into three major categories: evolutionary algorithm, dependent on Swarm and physics. The evolutionary algorithm is based on biological evolution that integrates Darwinian evolutionary principles. They are most famous (GA). The solutions are initially created by random means, and those are modified with each algorithm iteration. To effectively look for the ideal solution GA uses the crossover and mutation operators. As the perfect situation is more likely to take part in new offspring production, the new approach is predicted to be better [8]. Secondly, the Swarm is founded on the action of swarms, stoves, schools or animal flocks in nature. The search officers' activities are navigated through collective, social and intelligence simulation. The best-known swarm-based metaheuristic is PSO. It is inspired by birds flocking social behavior. It uses the best global approach in order to ensure research and the best local place for exploitation. ACO is another standard algorithm for this class. Thirdly, physics focused on physical law motivations. Annealing (SA) and HS algorithms simulation are the oldest algorithms concentrate on physics. SA is inspired by the metallurgical annealing process, in which metals are heated up to a certain degree and the impurity removal is accompanied by regulated cooling. The HS algorithm is influenced by concerts where musicians strive to produce the most delicate harmonic harmony. This phenomenon has been optimized for optimization in the HS algorithm. The GSA that is an imitation of Newton's gravitational law is another standard algorithm of this family. The search agents are called masses interacting with each other to find the optimum solution in accordance with this rule. Charged machine research (CSS), complies with the Electrostatics Law of Coulomb and the Mechanicals Law of Newton, aiming to find the best solutions effectively. Equilibrium Optimizer is the algorithm proposed recently, inspired by a well-mixed dynamic equilibrium on the volume of control. Mass balance equation is used to characterize non-reactive component concentration in the volume of control.

### 2.1 Feature Selection

The selection of features seeks to choose a range of features from the primary data collection, which can efficiently clarify input data while reducing the influence of noise and misfits.

1. Identification of a subset: the detection stage for candidate subsets' development. This stage is the stage of searching, which is possible to start with all variables or a subset of random variables without any function. The search methods at this point include traditional methodologies and EC techniques for producing subsets of features.

2. Subset assessment: an evaluation criterion used to calculate the desirability level of the subsets generated.

3. State of hold: the algorithm's execution starts with realizing the stop condition and ends until the stop situation has been realized. The separating condition may be specified based on an analysis function.

4. Validity: the accreditation of the chosen subset is the subject of this point. This accreditation is done using simulated databases (or real-world data sets), by experimenting and analyzing the effects of the proposed algorithm using former methods.

## **2.2 Particle Swarm Optimization (PSO)**

This approach of the Particle Swarm Optimization (PSO) [9], it simulates a swarm of flies, an assembly of birds, or a fish school's social behavior. PSO has many parallels with evolving calculation technologies, such as genetic algorithms and artificial neural networks. It is initialized by a random solution population and updates generations to look for optima. If one particle finds a decent way to eat, it is possible to immediately obey the rest of the colony, such as the Swarm of ants or even the fish school. In multidimensional research space [10], the swarming behavior is modeled on particles with characteristically location and speed. These particles drink their own position and velocity by changing the hyperspace.

Although the PSO has proven itself to be an efficient way to find optimum final sub-sets of functions, it has some deficiencies. One of the problems of PSO-based filtering processes is that the number of search features was abandoned [11] while only optimizing the precision of the classification. Related functions in the final function subset are more limited. Furthermore, the current PSO-based methods do not use features' correlation knowledge to direct the likelihood of the search process to be chosen in good functions that improve the classification accuracy.

## **2.3 Genetic Algorithm (GA)**

In 1960 Dr. Richenberg, whose thesis focused on evolutionary planning, first proposed evolutionary calculations. Later on, several scholars investigated his hypothesis in 1975, John Holland at the University of Michigan invented the genetic algorithm. This algorithm begins with an appropriate "k" response called the initial population, then calculates the value of each response using the fitness function, and two solutions are chosen from the available responses with the desired parent utility [11]; new children then come into being by mixing parent answers, added by the predominant population into the new population [11].

## **3 Related Work**

Existing feature selection algorithms may be categorized into two categories: approaches wrapper and filter methods. Wrapper approaches use an algorithm for learning/classification in the assessment process, while filter approaches do not. The solution of filters is argued to be less costly and generally computerized [12], while wrappers will usually produce improved results [13]. There have been several proposals for algorithms for selecting features over the last few years. This section reviews a typical function selection algorithm. The comparative output of the two networks demonstrates that the Wavelet network is statistically more critical than MLP. In identifying common mental disorders causes Ludermir et al. [14] used PSO techniques to identify them. Nasir et al. [15] present a hybrid algorithm that combines the Regrouping Particle Swarm Optimization (RegPSO) with the RFWN that is used to track, classify and characterize acoustic signals as a result of the surface drag operation. In order to deal with problems of nonlinear mixed-integer optimization in series, parallel series and bridge schemes, Shoo et al. [16] have created a hybrid approach. Naik et al. have suggested an effective classification system based on hybrid PSO and GA ANN [17] and are found in comparison to other alternatives to achieve relative improvement in efficiency.

	Population	F1	F2	F3	F4	F5	F6	F7	F8	Accuracy	WF1	WF2	WF3	WF4	WF5	WF6	WF7	WF8
GA	GA 1:	1	1	0	1	0	0	1	0	0.98	0.98	0.98	0	0.98	0	0	0.98	0
	GA 2:	1	1	0	0	0	1	1	0	0.93	0.93	0.93	0	0	0	0.93	0.93	0
	GA 3:	1	0	0	1	0	1	1	0	0.87	0.87	0	0	0.87	0	0.87	0.87	0
	GA 4:	1	1	0	1	0	1	1	1	0.75	0.75	0.75	0	0.75	0	0.75	0.75	0.75
	GA 5:	1	0	0	1	0	0	1	1	0.89	0.89	0	0	0.89	0	0	0.89	0.89
	GA 6:	0	1	0	1	0	1	1	0	0.98	0	0.98	0	0.98	0	0.98	0.98	0
	GA 7:	0	1	1	1	0	0	1	1	0.95	0	0.98	0.98	0.98	0	0	0.98	0.98
	GA 8:	1	1	0	1	0	1	1	0	0.79	0.79	0.79	0	0.79	0	0.79	0.79	0
	GA 9:	1	0	0	1	1	0	1	0	0.82	0.82	0	0	0.82	0.82	0	0.82	0
	GA 10:	0	1	0	1	0	0	1	1	0.91	0	0.91	0	0.91	0	0	0.91	0.91
PSO	PSO 1:	0	0	1	1	0	1	1	0	0.97	0	0	0.97	0.97	0	0.97	0.97	0
	PSO 2:	1	0	1	1	0	1	0	0	0.82	0.82	0	0.82	0.82	0	0.82	0	0
	PSO 3:	0	0	1	1	1	1	1	0	0.98	0	0	0.98	0.98	0.98	0.98	0.98	0
	PSO 4:	1	1	0	0	0	1	0	1	0.72	0.72	0.72	0	0	0	0.72	0	0.72
	PSO 5:	1	0	1	1	0	1	1	1	0.93	0.93	0	0.93	0.93	0	0.93	0.93	0.93
	PSO 6:	0	1	1	1	0	0	1	0	0.85	0	0.85	0.85	0.85	0	0	0.85	0
	PSO 7:	1	1	1	1	0	1	1	1	0.96	0.96	0.96	0.96	0.96	0	0.96	0.96	0.96
	PSO 8:	0	1	0	1	0	1	0	0	0.74	0	0.74	0	0.74	0	0.74	0	0
	PSO 9:	1	1	0	1	1	0	1	1	0.89	0.89	0.89	0	0.89	0.89	0	0.89	0.89
	PSO 10:	1	1	0	1	0	1	1	1	0.95	0.95	0.95	0	0.95	0	0.95	0.95	0.95
Feature importance											11.3	11.43	6.49	16.06	2.69	11.39	15.43	7.98

**Figure 1:** Measurement of the importance of a feature

PSO has been widely used in the field of FS. A sigmoid function was used to translate location values into probability to include features in binary PSO [18]. In certain instances, both continuous and discrete PSO were used in conjunction to optimize SVM and function parameters [19]. A distributed framework was used, with the server performing PSO calculations and the client performing SVM training and testing. PSO was changed to form geometric PSO in [20]. New agents were created using crossover operations on the current agents, the local best, and the best in the world, before mutating the agent created after crossover. For gene selection, the algorithm was used. The use of a rough set to perform FS was another slight change to PSO. To perform text clustering, Abualigah et al. proposed a hybrid algorithm combining PSO and genetic operators in [21]. By choosing a better collection of insightful features, the proposed hybrid model enhanced the efficiency of the k-means clustering algorithm. Kanan et al. [22] suggested another method for text document clustering that used adaptive PSO to find a more informative subset of features while also reducing the time required.

The best result of the two was taken in each iteration [23]. For protein function prediction, the hybrid algorithm used FS. A similar approach was taken, but this time the application domain was text classification. Using the SVM classifier, the researchers contrasted the use of ACO, GA, and PSO on siRNA data.

The fact that both GA and PSO outperformed ACO was a key finding. Ghosh et al. [12] proposed an embedded ACO called wrapper filter ACOFS, which uses a filter approach to test function subsets and reduces the overall model's time requirement. The writers have used memory to keep track of the best outcomes through generations. Abualigah et al. [21] proposed a novel text clustering approach based on Krill herd (KH).

MMKHA was developed for eight text datasets and used as a hybrid-enhanced KH algorithm. Chuang et al. [24] proposed approach for the combined objective and hybrid KH algorithms of the text record clustering. Abualigah et al. [21] an improved KH algorithm in the cluster of text papers.

#### 4 Proposed Method

Following the GA, our system is applied, and every step has pushed the population towards new, better answers. The cost function estimates the expense of a member of the population, thus retaining the speed of the population, and then changes the Gbest and Pbest of each member concerning new costs. Pbest refers to the best reaction each population has achieved up until now.

It must indeed be remembered that GA does not have the potential to exploit and the only root of GA manipulation is a mutation that causes very little chromosome disturbance. However, utilizing crossover operations, GA can do a remarkable exploration of the quest space. When considering PSO, it must be noted that PSO has strong local search capability but that it cannot achieve adequate exploration. PSO is always trapped in a local optima that obstructs the ability to explore. These complementary GA and PSO commercialization for exploitation and discovery encourage one to combine its findings to obtain efficient and usable results.

GA and PSO are two separate types of FS algorithms, namely evolutionary and swarm intelligence. GA is helpful for transferring useful functions from generation to generation. PSO has the benefit of thoroughly searching the search space with particles related to each other's function data. These advantages of PSO and GA are combined into exploiting and exploring in a balanced way.

This section will introduce our algorithm in detail as follows:

- Step 1 Determine how many features
- Step 2 Population Initialization.
- Step 3 Update the particle positions.
- Step 4 Cover ideological principles of the present method
- Step 5 Gbest mutation.
- Step 6 k-Nearest Neighbor (kNN)

At first, all the features of the dataset are provisional analyzed. The objective is to classify features more likely to exist in the chosen function subset. The technique of the profit ratio is used to achieve this. The algorithm aims to select features that make the most significant difference in each class as a split function for the decision tree branches. The profit ratio is a change in the knowledge gain, which decreases the difference between the high branch attributes. When selecting an attribute, the benefit ratio takes account of number and scale of branches.

---

**Algorithm 1.** The pseudo code of the proposed algorithm.

---

- 1) Calculate the GRI of each feature
  - 2) 75% feature selection with High GR
  - 3) Choose a start populace of individuals
  - 4) Fitness Evaluation
  - 5) Initialization of features and fitness evaluation
  - 6) Repeat
  - 7) Best individuals selection
  - 8) Crossover and mutation usage for new generation of individuals
  - 9) New individuals fitness evaluation
  - 10) Velocity and position updates
  - 11) Pbest and Gbest updates
  - 12) Fitness Evaluation of features
  - 13) Integrate the best values of population, velocity and positions
  - 14) The integrated population are sorted according to their fitness value
  - 15) Best part selection of the population
  - 16) Best Solutions Extraction in final population
  - 17) Classifier: take best values (kNN)
-

## 5 Experiment Design

### 5.1 Datasets and Parameter Setting

In our simulation experiments, seven data sets are followed to verify the efficiency of our suggested algorithm compared with the methods. These data sets are taken from the UCI learning device shown in Tab. 1. The data sets contain various variables, groups, and examples that prepare a comprehensive survey of the proposed approach [29] and the approaches used. One of the most common KNN algorithms was used to calculate the classification efficiency.  $k$  is calibrated for 5 in these experiments (5NN). The 7 data sets used in the trials selected from the UCI machine learning repository are seen above from Tab. 1. The function numbers of the PSO-based multi-target selection algorithms were chosen (from 13 to 166), groups (from 2 to 3), and instances (from 32 to 1 000).

In each sample of data, both cases are split randomly into two groups during the experiments: 70 percent as the testing set and 30 percent as the evaluation set. Each particle (single) constitutes a sub-set of functions during the training phase.

**Table 1:** Datasets description

Dataset	No. Features	No. Classes	No. Examples
Wine	13	2	178
German	24	3	1000
Lung	56	2	32
Hill-Valley	100	3	606
Musk 1	166	3	476
WBCD	30	3	564
Ionosphere	34	3	351

## 6 Experimental Results and Analysis

The maximum number of iterations and also the population play an essential role when talking about the multi-agent evolutionary algorithm in characterizing the actions of one agent from other agents' knowledge and the progressive progression of the agents in each other.

### 6.3 Comparison

The power of the proposed FS model will be determined by the application to different datasets. The model's efficiency is based on the performance of LFFS, LFS and our approach suggested in Tab. 2. This section provides the related material for the experiments.

**Table 2:** Accuracy rates of LFFS and LFS with our proposed method

Dataset	LFFS	LFS	Our proposed method
Wine			
Average-size	8	7	6.44
Average-acc	74.07	74.07	80.30
German			
Average-size	6	6	5.08
Ave-acc	69.33	69.33	78.17
WBCD			
Average-size	11	12	10.09

Ave-acc	91.18	92.35	93.48
Ionosphere			
Average-size	8	8	10.85
Ave-acc	90.48	90.58	91.37
Lung			
Average-size	5	6	30.25
Ave-acc	66.67	66.77	70.54
Hill-Valley			
Average-size	8	9	30.25
Ave-acc	53.85	55.49	78.53
Musk 1			
Average-size	13	13	90
Ave-acc	79.29	80.71	90.06

## 7 Conclusion

We also suggested a hybrid approach for FS in this paper based on a popular meta-heuristic called the algorithm PSO and GA. The proposed algorithm has been compared to ten state-of-the-art metaheuristic and hybrid metaheuristic FS methods using seven regular UCI datasets. The results obtained show that the algorithm proposed is superior to other approaches. We may then conclude that our way of solving FS problems is a competitive process. If we observe the findings, our algorithm manages to solve the PSO discovery and exploitation limitations. But in some situations, global optima may not be found according to the problem's need. We need to do a thorough experiment simultaneously to find the optimal value of the parameters used for various problems in this algorithm. The proposed algorithm can be applied to other common and interesting problems in research as a potential scope of work, such as facial recognition.

**Funding Statement:** This work was partially supported by the National Natural Science Foundation of China (61876089, 61876185, 61902281, 61375121), the Opening Project of Jiangsu Key Laboratory of Data Science and Smart Software (No. 2019DS301), the Engineering Research Center of Digital Forensics, Ministry of Education, the Key Research and Development Program of Jiangsu Province (BE2020633), and the Priority Academic Program Development of Jiangsu Higher Education Institutions.

**Conflicts of Interest:** The authors whose names are listed immediately below report the following details of affiliation or involvement in an organization or entity with a financial or non-financial interest in the subject matter or materials discussed in this manuscript.

## References

- [1] Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Resources*, vol. 3, pp. 1157–1182, 2003.
- [2] T. Marill and D. Green, "On the effectiveness of receptors in recognition systems," *IEEE Transaction on Information Theory*, vol. 9, no. 1, pp. 11–17, 1963.
- [3] L. Yu and H. Liu, "Feature selection for high-dimensional data: A fast correlation-based filter solution," in *Proc. of the Twentieth Int. Conf. (ICML 2003)*, Washington DC, USA, vol. 3, pp. 856–863, 2003.
- [4] A. Sadollah, A. Bahreinnejad and H. Eskandar, "Mine blast algorithm: A new population based algorithm for solving constrained engineering optimization problems," *Applied Soft Computing*, vol. 13, no. 5, pp. 2592–2612, 2013.

- [5] E. Njoku, W. Wilson, S. Yueh and Y. Rahmat-Samii, "A large-antenna microwave radiometer concept for ocean salinity and soil moisture sensing", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, pp. 2645–2655, 2000.
- [6] L. F. Chen, C. T. Su, K. H. Chen and P. C. Wang, "Particle swarm optimization for feature selection with application in obstructive sleep apnea diagnosis," *Neural Computing and Applications*, vol. 4, pp. 2087–2096, 2012.
- [7] P. C. Fourie and A. A. Groenwold, "The particle swarm optimization algorithm in size and shape optimization," *Structural and Multidisciplinary Optimization*, vol. 23, pp. 259–267, 2002.
- [8] R. C. Eberhart and J. Kennedy, "A new optimizer using particles swarm theory", in *Proc. Sixth Int. Sym. on Micro Machine and Human Science (Nagoya Japan)*, pp. 39–43, 1995.
- [9] W. D. Chang, "A modified particle swarm optimization with multiple subpopulations for multimodal function optimization problems," *Apple Software and Computer*, vol. 33, pp. 170–182, 2015.
- [10] M. H. Sabour, H. Eskandar and P. Salehi, "Imperialist competitive ant colony algorithm for truss structures," *World Applied Science Journal*, vol. 12, pp. 94–105, 2011.
- [11] Y. Zhang, S. H. Wang and G. L. Ji, "A comprehensive survey on particle swarm optimization algorithm and its applications," *Mathematical Problems in Engineering*, vol. 2015, pp. 38, 2015.
- [12] A. Ghosh, A. Datta and S. Ghosh, "Self-adaptive differential evolution for feature selection in hyperspectral image data," *Applied Soft Computing*, vol. 13, pp. 1969–1977, 2013.
- [13] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen and O. Shochet, "Novel type of phase transition in a system of self-driven particles," *Physical Review Letter*, vol. 75, no. 6, pp. 1226–1229, 1995.
- [14] T. B. Ludermir and A. T. Sergio, "PSO for reservoir computing optimization," in *Int. Conf. on Artificial Neural Networks*, pp. 685–692. Springer, Berlin, Heidelberg, 2012.
- [15] M. Nasir, S. Das, D. Maity, S. Sengupta, U. Halder *et al.*, "A dynamic neighborhood learning based particle swarm optimizer for global numerical optimization," *Information Sciences*, vol. 209, pp. 16–36, 2012.
- [16] L. P. Shoo, C. Storlie, Y. M. Williams and S. E. Williams, "Potential for mountaintop boulder fields to buffer species against extreme heat stress under climate change," *International Journal of Biometeorology*, vol. 54, pp. 475–478, 2010.
- [17] M. Gen, R. Cheng and L. Lin, *Network Models and Optimization: Multiobjective Genetic Algorithm Approach*, Springer, London, UK, 2008.
- [18] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom and J. Riedl, "GroupLens 1994: An open architecture for collaborative filtering of netnews," in *Proc. of the 1994 ACM Conf. on Computer Supported Cooperative Work*, pp. 175–186.
- [19] N. Dong, C. H. Wu, W. H. Ip, Z. Q. Chen, C. Y. Chan *et al.*, "An opposition-based chaotic GA/PSO hybrid algorithm and its application in circle detection," *Computer Mathematics with Applications*, vol. 64, no. 6, pp. 1886–1902, 2012.
- [20] H. Liu, H. Q. Tian and C. Chen, "An experimental investigation of two Wavelet-MLP hybrid frameworks for wind speed prediction using GA and PSO optimization," *Electronical Power and Energy System*, vol. 52, pp. 161–173, 2013.
- [21] L. M. Abualigah and A. T. Khader, "Unsupervised text feature selection technique based on hybrid particle swarm optimization algorithm with genetic operators for the text clustering," *The Journal of Supercomputing*, vol. 73, pp. 4773–4795, 2017.
- [22] H. R. Kanan and K. Faez, "An improved feature selection method based on ant colony optimization (ACO) evaluated on face recognition system," *Applied Mathematics and Computation*, vol. 205, no. 2, pp. 716–725, 2008.
- [23] Sahoo, L., A. Banerjee, A. KumarBhunia and S. Chattopadhyay, "An efficient GA–PSO approach for solving mixed-integer nonlinear programming problems in reliability optimization," *Swarm Evolutionary Computation*, vol. 19, pp. 43–51, 2014.
- [24] L. Y. Chuang, S. W. Tsai and C. H. Yang, "Improved binary particle swarm optimization using catfish effect for feature selection," *Expert Systems with Applications*, vol. 38, pp. 12699–12707, 2011.