International Journal of Latest Trends in Engineering and Technology
Vol.(9)Issue(2), pp.175-181
DOI: http://dx.doi.org/10.21172/1.92.29
e-ISSN:2278-621X

HYBRID SWARM OPTIMIZATION APPROACHES: A SURVEY

Mukesh[1] & Jyoti Vashasitha[2]

Abstract- Data is continuously increasing due to millions of transactions done daily by the users. A plenty of classification techniques are available for discovering useful information out of this huge amount of data. Now-a-days swarm optimization approaches and the hybridization of these approaches are playing important role for classification in highly efficient manner. In this paper, we provide an overview and a detailed comparison of some of the swarm optimization approaches and hybrid swarm optimization approaches present in research literature.
Keywords – Swarm Optimization, PSO- Particle Swarm Optimization, ACO- Ant Colony Optimization, CSO- Cat Swarm Optimization.

1. INTRODUCTION

Classification, a task of data mining, generates rules from given training data and uses this rules (classification model) to classify test data. Comprehensibility, simplicity, and accuracy are the major parameters which determine the effectiveness of a model. A trade-off among these parameters is usually desired in practical situations. Swarm optimization algorithms, a class of evolutionary algorithms, make use of feature selection, rule discovery and exception discovery to ease the task of classification [1][2]. Particle Swarm Optimization (PSO) algorithm is one such swarm optimization and population based search algorithm which starts with an initial population of random solutions, termed as particles [3][4]. PSO combined with other swarm optimization algorithms, called Hybrid Swarm Optimization algorithms, has been widely used in past for multi-objective problems.

Cat Swarm Optimization (CSO) algorithm is another swarm optimization algorithm which imitates the natural behavior of cats [5]. Cats always remain alert and move very slowly - a behavior of cats which is termed as seeking mode in the context of CSO. Furthermore, on sensing a prey, the cats chase it with a very high speed - a behavior termed as tracing mode. These two modes have been mathematically modeled for solving optimization problems. Fitness value, position, and velocity of cat constitute an individual, called dimension, in CSO. Two different modes of a cat, seeking mode and tracing mode, are identified with the help of a flag.

Ant Colony Optimization (ACO) algorithm, another popular swarm optimization algorithm, models behavior of ants for discovering classification rules [6]. Pheromone concentration and evaporation help ants in discovering shortest path from source to destination. Ant drops pheromone as it moves along a path, which evaporates if the fellow ants do not follow it. Shorter paths will be traversed faster than their longer counterparts, leading to increased concentration of pheromones over shorter ones and eventually turning out to be the optimal choice for the ants. Longer paths are often rejected by the ants on the grounds of low pheromone concentration. In the context of classification rule set generation, ACO helps in selecting the rule which accounts for the majority of training data.

Hybridization of all such optimization algorithms can be achieved in two distinguished ways:

1. Merging features from two different approaches to yield a completely different approach.
2. Merging two or more different features of the underlying approach to facilitate better results.

In this paper, we have comparatively investigated the three major kinds of swarm optimization algorithms - PSO, CSO, and ACO and discussed their role in solving optimization problems pertaining to different problem domains.

2. CLASSIFICATION

The problems of classification such as the classification of criteria and database have been an important research topic in decision area. In the course of decision-making, a lot of uncertain factors and incomplete data have influenced the results of the decision. The methods including decision tree, BP neural network, rough set and Support Vector Machine (SVM) are used to solve the problem of classification [7]. SVM was based on the structural risk minimization (SRM) principle that seeks to minimize an upper bound of the generalization error consisting of the sum of the training error and a confidence interval. SVM classifies data with different class labels by determining a set of support Vectors that are the members of the set of training inputs that outline a hyper-plane in the feature space. Feature selection methods are categorized into three types: wrapper approach, filter approach, and embedded approach. Wrapper approach uses learning algorithm to select feature subset. It uses classifier accuracy as a fitness measure. A feature subset with the higher value is used to learn classifier. Most of Wrapper

[1] Department of Computer Science and Engineering, Guru Jambheshwar University of Science and Technology, Hisar, Haryana, India
[2] Department of Computer Science and Engineering, Guru Jambheshwar University of Science and Technology, Hisar, Haryana, India

Algorithms are categorized as follows: exact methods, greedy sequential subset selection method, partitioning methods, mathematical programming methods, and meta-heuristic methods. Two floating methods are Sequential Forward Floating Selection (SFFS) and Sequential Backward Floating Selection (SBFS). They can select or remove the features at different stages of the procedure until a suitable number of features is obtained. But these sequential floating may get stuck in local optimal solution. Wrapper approach based methods discussed above, encounter a variety of problems such as high computational cost and problem of getting trapped in local optima [8]. As opposed to wrapper method, Filter Methods do not use any classification algorithm [9]. It uses the different measures of information distance, dependency or consistency to select feature subset [10]. It is known that feature selection methods do not use any classifier, they provide a general view of feature space. Third, Embedded methods take advantages of both wrapper and filter methods for feature selection but they are rarely used because of their complex nature [10] Both SVM and feature selection methods are used in swarm optimization approaches and also for hybridization of swarm optimization approaches for improving classification accuracy.

3. SWARM OPTIMIZATION

Swarm optimization (SO) is a computational method that optimizes a problem by iteratively improving a candidate solution with regard to a given measure of quality. Various swarm optimization algorithms, e.g. Particle Swarm Optimization (PSO), Cat Swarm Optimization(CSO), Ant Colony Optimization(ACO), artificial bee colony algorithm (ABC) have been proposed in past.

3.1 Particle Swarm Optimization (PSO)–

PSO is an evolutionary technique inspired by bird flocking that uses the concept of fitness like other evolutionary approaches [4]. PSO is developed by simulating the social behaviour [11]. A new parameter called inertia weight (w) was added to the original PSO algorithm to balance global and local search [12] and Yuhui Shi [13] made efforts to improve the results produced by PSO. PSO was originally designed to solve real value problem. Original PSO was extended to Discrete/binary space to tackle the discrete problem in which velocity was squashed using the logistic function. Use of both binary and continuous representation of Particle Swarm Optimization (PSO) in classification was introduced which provide better classification results [14]. In the first phase of research three PSO variants have been compared with Genetic Algorithm (GA) and Tree Induction Algorithm (J48) named as Discrete Particle Swarm Optimizer (DPSO), Linear Decreasing Weight Particle Swarm Optimizer (LDWPSO) and Constricted Particle Swarm Optimizer (CPSO). The second phase of research improves PSO variant in term of attribute type support and temporal complexity. Experimental results show that PSO is competitive with evolutionary as well as tree induction algorithms [14].

3.2 Cat Swarm Optimization (CSO)–

CSO is another optimization algorithm [5], which mimics the behaviour of cats. Cats are excellent hunters but they also show high levels of alertness even at rest position. There exist two modes to describe their behaviour: seeking mode and tracing mode. Each cat has its own position and velocity (direction of movement).The positions are considered points in D-dimensions[15], and the velocities for each dimension change the values of these points. Deivaseelan et al.[16] have argued that search performance of CSO is better than that of PSO.

3.3 Ant Colony Optimization (ACO)–

Algorithm visualizes a decision rule as a combination of attribute-value pairs. Artificial ants are given entire training set for discovering rule set out of it. Using the given training set, ants begin to construct rules by incrementally adding terms (attribute-value pair) to partially constructed rule. There are two conditions for adding terms to a partially constructed rule, these are:

a) On adding the term, the rule must cover the minimum number of cases.
b) The attribute going to be added to a rule should not have been previously included in any other term otherwise rule is going to be an absurd one.

It is obvious that the Ant-Miner algorithm follows Michigan approach because each iteration of the algorithm yields one best rule and at the end, we have the list of best rules. The way Ant-Miner is different from decision tree algorithms is that decision tree algorithms calculate the entropy of an attribute whereas Ant-miner the entropy of an attribute-value pair. This entropy is incorporated in form of heuristic function. Another difference is that entropy measure is the only measure used for tree building whereas entropy measure along with pheromone updating is used in case of Ant-Miner.

4. CLASSIFICATION USING SWARM OPTIMIZATION

A vast amount of Research works exist on classification using swarm optimization techniques. Many evolutionary search techniques (PSO) comes under nature-inspired algorithm have been used to get optimal feature subset since these techniques have good global search ability. Meta-heuristic algorithms i.e. genetic algorithm (GA) and particle swarm optimization (PSO) follows search procedures to solve optimization problems. The GA solves optimization problems by simulating biological evolution[17]. In PSO, each solution is viewed as a particle and the algorithm searches for the best solution by considering the experience of all particles. Many other optimization approaches also have been proposed in last decades, such as ant colony optimization (ACO)[18], artificial bee colony algorithm (ABC)[19] and cat swarm optimization (CSO) [15]. The ACO algorithms solve the optimization problem by inspecting the intelligent behaviour of ant swarm for food-finding and the information exchange with the help of pheromone. The ABC algorithms were proposed by observing the food-finding behaviour of the bee swarm. Deivaseelan et al. have demonstrated that the search performance of CSO is better than that of PSO. However, CSO requires long computation times to identify the best solution[20]. ICSO (Improved Cat Swarm Optimization) has been adapted from an algorithm in [16]and performance of the overall ICSO algorithm in feature selection has been evaluated using SVM. A new hybrid ant colony optimization algorithm (ACOFS) proposed by Kabir et al. [21] combines advantages of wrapper and filter approaches.

4.1 Classification Using PSO–

Evolutionary search technique like PSO has been used to get optimal feature subset since these techniques have good global search ability. Pedrycz et al. [8] introduced PSO for feature selection using filter approach. In this paper, wrapper approach of feature selection has been used to evaluate the optimal feature subset using nearest neighbor classifier and minimization of classification error has been used as the fitness measure. Reduced feature set based on PSO enhances the classification performance. Size of Initial core feature set and training set plays an important role to get optimal feature subset. Multi-Swarm PSO (MSPSO) has been used by Liu et al. [9] to get optimal feature subset. Parameters of SVM are also optimized along with feature selection. MSPSO and support vector with F-measure have been used to improve the classification. To check the effectiveness of MSPSO, it has been compared with standard PSO, grid search and genetic algorithm (GA). MSPSO outperforms all these methods. Different sub-swarms in MSPSO have complicated communication rule. Due to the large population and complicated communication rules, computation cost of MSPSO is greater than all other three methods. Since imprecision, uncertainty can be easily handled by rough set theory, wang et al. [10] proposed feature selection method based on PSO and rough set theory. Experimental results show that PSO is a good approach for rough set reduction. Binary particle swarm optimization (PSO) with a rough set theory for dimension reduction was proposed by Liam et al. [22]. But there is a drawback of using rough set theory in feature selection problems, rough set consumes the most of the running time. BPSO Feature selection method was introduced again by Liam et al. [22] using filter approach. Two new algorithms were developed based on two information measures. Optimal feature subset selection is made on the basis of BPSO and information theory. BPSO-P uses mutual information of the pair of features as a measure. Fitness function uses both relevance and redundancy to select best feature subset. BPSO-G evaluates relevance and redundancy of selected feature subset based on the entropy of group of features. In this algorithm, fitness function of selected features is calculated as a whole rather than feature pair. Experimental results show that with suitable weight, both algorithms achieve better classification accuracy. However, these algorithms have not been compared with other related algorithms. To enhance the performance of BPSO, catfish effect was applied by Chuang et al. [23]. Those Particles having worst fitness in a number of consecutive iteration were replaced by new particles. Feature selection process is simplified by catfish-BPSO. Sahu et al. [24] proposed feature selection method using discrete PSO for classification. In this paper, clustering of the dataset is performed to obtain a feature subset. This feature subset is provided as an input to PSO to find out optimal feature subset. This approach provides the better result in comparison to previous approaches. Wrapper-based feature selection algorithm, based on modified binary particle swarm optimization (BPSO) and the linear regression model, and was introduced by unler et al. [25]. PSO is easier to implement as compared to genetic algorithm (GA) since mutation and crossover operator are not used in PSO. Social learning is introduced to update velocity in BPSO. An adaptive feature selection strategy has also been included to make feature selection more effective. In this algorithm, features are selected based on two parameters namely contribution of already selected feature subset and likelihood calculation by BPSO. Proposed algorithm achieves better performance in comparison to tabu search and scatters search algorithms. As multi-dimensional search space has higher complexity compared to one dimension searching space, Wang et al. proposed PSO based feature selection method in one dimension searching. In this paper, real-valued PSO is used rather than BPSO because BPSO can get trapped in local optima. Due to less complicated search space of one dimension, there are better chances of obtaining optimal feature subset. Experimental results show that proposed algorithm achieves remarkable results.

4.2 Classification using CSO –

Due to the complexity of two modes of CSO, there rarely exists much literature on classification using CSO. Classification accuracy was evaluated with proposed CSO+SVM [16]. Classification using this proposed technique gives better performance with high accuracy. A new improved CSO[26] introduced Common classifiers which include Neural Networks [27], Nave Bayes Classifiers [28], Decision Tree [29] and SVM [30]. For data which are linearly inseparable, SVM uses the concept of hyper-plane classification, which converts input vectors into a hyper-plane by applying a kernel function. In the hyper-plane, the SVM finds the maximum distance between different data clusters and divides them into two groups of information. This is achieved through the radial basis function (RBF) [31]. ICSO (Improved Cat Swarm Optimization) was adapted from an algorithm in [32] and performance of the overall ICSO algorithm in feature selection was evaluated using SVM. The process of ICSO follows two modes: seeking and tracing mode. Cat swarm optimization has been proposed for the improved seeking mode. Two methods have been applied on seeking mode to reduce the time required to find optimal solution and to change the position of cats.

4.3 Classification using ACO –

In year i.e. 2002 a modification to the Ant-Miner1 algorithm was suggested [33] as Ant-Miner2. The algorithm suggests a new way of calculating heuristic function and hence a new way of selection of term because the selection of term depends on the heuristic function. Another modification to above algorithm was suggested by Bo Liu et al.[34] in the year 2003, popularly known as Ant-miner 3. Ant-Miner 3 proposed two new things: (a) a new way of updating the pheromone values of terms used in rule construction. However, pheromone value of unused terms is still decreased by normalization. (b) It also suggested a new transition rule i.e. a new rule for selection of terms. A new version of Ant-Miner, proposed by Frietas et al. [35], for discovering unordered rule list, came in the year 2006. Earlier ant- miner algorithms discovered ordered rule list. New version discovered unordered rule set i.e. a set of rules which need not be applied to test data in the order in which they were discovered. It was possible with some modification in the high-level algorithm, heuristic function, and pheromone updating. Some parallelization in Ant-Miner algorithm was introduced by Parallel Ant Miner algorithm, proposed by Chen H. et al. [36] Instead of discovering consequents, later on, we fix the consequent and then discover antecedents corresponding to them. Note that the ants of the group excavate the classification rules in parallel with the corresponding consequent parts, thus introducing parallelism and hence the name parallel Ant-Miner. Modification to this parallel Ant-Miner was suggested by Omid Roozmand et al. in the year 2008 [37]. Ants now search all the space in parallel to discover classification rules and then communicate with the other ants of their own group and the best of the other groups to update the pheromone of terms. Parallelization and communication methods enable the algorithm to discover high-quality rules and avoid gathering irrelevant rules. The most important modification to original Ant-Miner was suggested by Frietas et al. in the year 2012 [38]. An ant creates a complete list of rules at each iteration of the algorithm instead of creating just a single rule and search is guided by the quality of a list of rules. Further improvement to the above algorithm has been suggested by Alex A. Freitas et al. in his paper [39]. This paper proposes an extension of the cAnt-MinerPB algorithm to create unordered rules. The main motivation is to improve the interpretation of individual rules. Results show that the predictions made by an unordered set of rules are potentially easier to be interpreted by a user, due to the nature of unordered rules (i.e., each rule has a modular meaning independent of the others) and there are fewer attribute-conditions involved in the predictions.

4.4 Classification using Hybrid swarm optimization approaches –

A hybrid ant colony optimization algorithm has been suggested by Md. M Kabir et al. [22] which uses the advantages of wrapper and filter approaches by selecting the subset of salient features of reduced size. The reason of distinctness of ACOFS from existing algorithms [40][41][42][43] is that it lies in following two aspects: First ACOFS not only the selection of a number of salient features but also the attainment of a reduced number of them. Second, ACOFS utilizes a hybrid search technique for selecting salient features that combine the advantages of the wrapper and filter approaches. Yannis et al. [44] proposed hybrid DABC-GRASP by combining the features of artificial bee's colony optimization and greedy random adaptive search procedure. Proposed algorithms have been compared with tabu search, GA, PSO, ACO and GRASP and the results advocates for better accuracy of the proposed algorithm. Ke Shang et al. [45] proposed hybrid GA-ACO approach which replaces the bed individuals of the GAs population with new individuals of ACO. The comparison results show that GA-ACO is competitive with existing approaches. WenXiong et al. [46] proposed hybrid ACO-RF algorithm also produced higher accuracy for classification. Alghamdi et al. [47] proposed algorithm combines the features of ACO and TOFA (Trace Oriented Feature Analysis) which can obtain better classification accuracy for a large amount of data by reducing the feature space to much smaller dimension.

5. COMPARATIVE STUDY

Most of the research work in swarm optimization addresses the problem of improving long execution time and classification accuracy. A lot of swarm optimization approaches have been proposed to improve classification accuracy using feature selection methods. Whereas very less work has been done for improvement of another parameter that is long execution time. A few Hybrid swarm optimization approaches are available which have worked on the improvement of both parameters. However, further improvement in this regard can be achieved by hybridization of two technique of same swarm optimization or two different swarm optimization techniques. Following table-1 describes the differences between various hybrid swarm optimization techniques for classification accuracy.

In the first paper, PSO hybridized with GA and SVM[48]. The proposed method was able to automatically select most important and informative features in term of classification accuracy with sustainable CPU processing time without requiring the number of features to be set initially by the user. Another hybridization approach invented by in 2015 named as IPSO-LDA (Improved particle swarm optimization and Linear Discriminate Analysis), which improved the image classification accuracy than other tested approaches[49]. In 2015, fang liv and Zhiguang Zhau [50]provide another hybrid approach, CPL-SVM (Cross-validation + Particle swarm optimization + Least Square + SVM) by hybridization of PSO, SVM (Support Vector Machine) and LS (Least Square) approaches. Proposed method have better learning performance, strong generalization ability, and classification accuracy. K.C. Lin et al. [51]in 2015 provided a new approach by combining the features of PSO, SVM, and ABC. Proposed hybrid approach reduced the number of features and improved accuracy of medical datasets. M.Celik et al.'s invention CoABCMiner (Co-operative artificial Bee Colony Miner)[52] performed better with non-parametric statistical data. K.C.Lin et al. [26] again in 2016, provide ICSO for big data classification. Experimental results concluded that proposed

ICSO+TF+IDE+SVM provide better results than TF+IDE+SVM for text data classification. Average accuracy rate with this hybridization is improved with less number of selected features. In last Paper, PSO is hybridized with Local Search strategy[53]. In this proposed method 12 data sets are taken from UCI data repository and all other techniques (GA, PSO, ACO, and SA) are applied to these data for classification. HPSO+LS hybridization outperform all techniques separately

Table -1 Observation with Hybridized Swarm Optimization Approaches

Year	Base Technique	Hybridized Technique	Experimental Area	Compared with	Observations
2015	GA	HGAPSO(Hybrid Genetic Algorithm And Particle Swarm Optimization)+ SVM(Support Vector Machine)	AVIRIS hyper-spectral data of Indian pines and Toronto Roads data set	• PSO+SVM • GA+SVM • HGAPSO+SVM	• Comparison observed in techniques. • HGAPSO+SVM perform better than the other approaches used for comparison in terms of performance metrics.
2015	PSO	IPSO(Improved Particle Swarm Optimization)+LDA(Linear Discriminant Analysis)	Indian face database	• PSO • PSO+LDA	• Average recognition rate of IPSO-LDA method is better than LDA and PSO-LDA • IPSO-LDA approach has higher classification accuracy rates than other tested approaches.
2015	PSO	CPL(chaotic particle swarm optimization +least square)+SVM	Iris flower Data from UCI data repository	• CVL-SVM(cross validation LS-SVM) • GL-SVM(GA-LS-SVM) • PL-SVM(PSO-LS-SVM) • APL-SVM (APSO-LS-SVM)	• CPL-SVM method gives better classification accuracy and has strong generalization ability and effective avoidance of isolated effects of sample in the active learning process.
			drug data (HIA, P-gp and TdP)for classification		• The proposed CPL-SVM algorithm can obtain the best sensitivity, specificity, classification precision, Matthews correlation coefficient and classification accuracy for HIA,P-gp and TdP
2015	PSO	Endocrine based PSO+SVM+ ABC(Artificial Bees Colony Optimization)	Medical data from UCI data repository, university of California, Irvine	• PSO-SVM • EPSO-SVM • ABC+SVM	• EPSO+SVM+ABC perform better than other hybridized techniques with higher accuracy using less number of features.
2016	ABC	CoABCMiner (Cooperative Rule Learning)	UCI data repository(32, 47-49)	• C4.5Rules • SIA • LDWPSO • CORE • ABCMiner • HIDER	• CoABCMiner provide the better results for classification rules
2016	TF-IDE+ SVM	ICSO(Improved Cat Swarm Optimization)+TF-IDE(Term Frequency-Inverse Document Frequency) + SVM	Food culture in Taiwan-food category from UCI data repository	• TF-IDF + SVM	• Average accuracy rate with ICSO improved with less number of selected features for text classification.
2016	PSO	HPSO+LS (Local search)	12 data set from UCI data repository i.e. wine, heart,	• Simulated Annealing (SA) • Genetic Algorithm (GA)	• Proposed method is compared for feature selection separately. • Experimental results

			cancer etc.	• Particle Swarm Optimization (PSO) • Ant Colony Optimization (ACO)	concluded that HPSO + Local Search provide better results with high accuracy in less time.

6. CONCLUSION

The research on swarm techniques for classification in several domains is on rise. Swarm optimization algorithms are optimization algorithms best suited for large datasets. The application areas include machine learning, data mining, artificial intelligence and pattern recognition. Hybrid swarm optimization approaches, when applied to above application areas, may result in more accurate and efficient solutions. After reviewing and comparing various hybrid swarm optimization algorithms for classification, it is observed that this area of research is still less explored and there exist a vast scope for the hybridization of swarm optimization approaches, in addition to existing hybridized swarm optimization approaches.

7. REFERENCES

[1] J. Vashishtha, D. Kumar, and S. Ratnoo, "An evolutionary approach to discover intra- and inter-class exceptions in databases," Int. J. Intell. Syst. Technol. Appl., vol. 12, no. 3/4, p. 283, 2013.

[2] J. Vashishtha, D. Kumar, S. Ratnoo, and K. Kundu, "Mining Comprehensible and Interesting Rules: A Genetic Algorithm Approach," Int. J. Comput. Appl., vol. 31, no. 1, pp. 39–47, 2011.

[3] A. Pathak and J. Vashistha, "Classification Rule and Exception Mining Using Nature Inspired Algorithms," vol. 6, no. 3, pp. 3023–3030, 2015.

[4] J. Kennedy and R. Eberhart, "Particle swarm optimization," Neural Networks, 1995. Proceedings., IEEE Int. Conf., vol. 4, pp. 1942–1948 vol.4, 1995.

[5] S. C. Chu and P. W. Tsai, "Computational intelligence based on the behavior of cats," Int. J. Innov. Comput. Inf. Control, vol. 3, no. 1, pp. 163–173, 2007.

[6] R. S. Parpinelli, H. S. Lopes, and a a Freitas, "Data mining with an ant colony optimization algorithm," IEEE Trans. Evol. Comput., vol. 6, no. 4, pp. 321–332, 2002.

[7] C. L. Huang and C. J. Wang, "A GA-based feature selection and parameters optimizationfor support vector machines," Expert Syst. Appl., vol. 31, no. 2, pp. 231–240, 2006.

[8] W. Pedrycz, B. J. Park, and N. J. Pizzi, "Identifying core sets of discriminatory features using particle swarm optimization," Expert Syst. Appl., vol. 36, no. 3 PART 1, pp. 4610–4616, 2009.

[9] Y. Liu, G. Wang, H. Chen, H. Dong, X. Zhu, and S. Wang, "An improved particle swarm optimization for feature selection," J. Bionic Eng., vol. 8, no. 2, pp. 191–200, 2011.

[10] X. Wang, J. Yang, X. Teng, W. Xia, and R. Jensen, "Feature selection based on rough sets and particle swarm optimization," Pattern Recognit. Lett., vol. 28, no. 4, pp. 459–471, 2007.

[11] R. C. Eberhart and Y. Shi, "Computational Intelligence: Concepts to Implementations," Knowl. Creat. Diffus. Util., no. March, p. 496, 2007.

[12] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," 1998 IEEE Int. Conf. Evol. Comput. Proceedings. IEEE World Congr. Comput. Intell. (Cat. No.98TH8360), no. June 1998, pp. 69–73, 2015.

[13] Y. Shi, "Particle swarm optimization," IEEE Connect., vol. 2, no. 1, pp. 8–13, 2004.

[14] T. Sousa, A. Silva, and A. Neves, "Particle Swarm based Data Mining Algorithms for classification tasks," Parallel Comput., vol. 30, no. 5–6, pp. 767–783, 2004.

[15] K.-C. L. K.-C. Lin and H.-Y. C. H.-Y. Chien, "CSO-based feature selection and parameter optimization for support vector machine," Pervasive Comput. (JCPC), 2009 Jt. Conf., pp. 783–788, 2009.

[16] Deivaseelan.A and P.Babu, "Modified Cat Swarm Optimization For Iir System Identification 1," Adv. Nat. Appl. Sci., vol. 6, no. 6, pp. 731–740, 2012.

[17] S. H. Cha and C. Tappert, "A genetic algorithm for constructing compact binary decision trees," J. Pattern Recognit. Res., vol. 4, no. 1, pp. 1–13, 2009.

[18] A. Colorni, M. Dorigo, and V. Maniezzo, "Distributed Optimization by Ant Colonies," Proc. first Eur. Conf. Artif. Life, vol. 142, no. or D, pp. 134–142, 1991.

[19] D. Karaboga, "An idea based on Honey Bee Swarm for Numerical Optimization," Tech. Rep. TR06, Erciyes Univ., no. TR06, p. 10, 2005.

[20] M. Orouskhani, Y. Orouskhani, M. Mansouri, and M. Teshnehlab, "A Novel Cat Swarm Optimization Algorithm for Unconstrained Optimization Problems," Int. J. Inf. Technol. Comput. Sci., vol. 5, no. 11, pp. 32–41, 2013.

[21] M. M. Kabir, M. Shahjahan, and K. Murase, "A new hybrid ant colony optimization algorithm for feature selection," Expert Syst. Appl., vol. 39, no. 3, pp. 3747–3763, 2012.

[22] L. Cervante, B. Xue, L. Shang, and M. Zhang, "A Dimension Reduction Approach to Classification Based on Particle Swarm Optimisation and Rough Set Theory," Springer-Verlag Berlin Heidelb., pp. 313–325, 2012.

[23] L. Y. Chuang, S. W. Tsai, and C. H. Yang, "Improved binary particle swarm optimization using catfish effect for feature selection," Expert Syst. Appl., vol. 38, no. 10, pp. 12699–12707, 2011.

[24] B. Sahu and D. Mishra, "A novel feature selection algorithm using particle swarm optimization for cancer microarray data," Procedia Eng., vol. 38, pp. 27–31, 2012.

[25] A. Unler and A. Murat, "A discrete particle swarm optimization method for feature selection in binary classification problems," Eur. J. Oper. Res., vol. 206, no. 3, pp. 528–539, 2010.

[26] K.-C. Lin, K.-Y. Zhang, Y.-H. Huang, J. C. Hung, and N. Yen, "Feature selection based on an improved cat swarm optimization algorithm for big data classification," J. Supercomput., vol. 72, no. 8, pp. 3210–3221, 2016.

[27] G. P. Zhang, "Neural networks for classification: a survey," IEEE Trans. Syst. Man Cybern. Part C (Applications Rev., vol. 30, no. 4, pp. 451–462, 2000.

[28] D. D. Lewis, "Naive (Bayes) at forty: The independence assumption in information retrieval," pp. 4–15, 1998.

[29] J. Ross, Q. Morgan, and K. Publishers, "Book Review : C4 . 5 : Programs for Machine Learning," Mach. Learn., vol. 240, pp. 235–240, 1994.

[30] C. Cortes and V. Vapnik, "Support-Vector Networks," Mach. Learn., vol. 20, no. 3, pp. 273–297, 1995.

[31] and C.-J. L. Chih-Wei Hsu, Chih-Chung Chang, "A Practical Guide to Support Vector Classification," BJU Int., vol. 101, no. 1, pp. 1396–400, 2008.

[32] K. C. Lin, K. Y. Zhang, and J. C. Hung, "Feature Selection of Support Vector Machine Based on Harmonious Cat Swarm Optimization," 2014 7th Int. Conf. Ubi-Media Comput. Work., pp. 205–208, 2014.

[33] B. Liu and B. Mckay, "Density-Based Heuristic for Rule Discovery with Ant Miner," 6th Australas. Jt. Work. Intell. Evol. Syst., pp. 1–6, 2002.

[34] B. Liu, H. A. Abbas, and B. McKay, "Classification rule discovery with ant colony optimization," 2003 IEEE/WIC Int. Conf. Intell. Agent Technol., pp. 83–88, 2003.

[35] J. Smaldon and A. A. Freitas, "A New Version of the Ant-Miner Algorithm Discovering Unordered Rule Sets," Proc. 8th Annu. Conf. Genet. Evol. Comput. GECCO 06, p. 43, 2006.

[36] Y. Chen and L. C. Member, "Parallel Ant Colony Algorithm for Mining Classification Rules," IEEE, pp. 85–90, 2006.

[37] O. Roozmand and K. Zamanifar, "Parallel ant miner 2," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 5097 LNAI, pp. 681–692, 2008.

[38] F. E. B. Otero, A. A. Freitas, and C. G. Johnson, "A new sequential covering strategy for inducing classification rules with ant colony algorithms," IEEE Trans. Evol. Comput., vol. 17, no. 1, pp. 64–76, 2013.

[39] F. E. B. Otero and A. A. Freitas, "Improving the interpretability of classification rules discovered by an ant colony algorithm," Proceeding fifteenth Annu. Conf. Genet. Evol. Comput. Conf. - GECCO '13, p. 73, 2013.

[40] A. Al-ani, "Feature Subset Selection Using Ant Colony Optimization," no. Feature Selection, p. 6, 2006.

[41] M. H. Aghdam, N. Ghasem-Aghaee, and M. E. Basiri, "Text feature selection using ant colony optimization," Expert Syst. Appl., vol. 36, no. 3 PART 2, pp. 6843–6853, 2009.

[42] R. K. Sivagaminathan and S. Ramakrishnan, "A hybrid approach for feature subset selection using neural networks and ant colony optimization," Expert Syst. Appl., vol. 33, no. 1, pp. 49–60, 2007.

[43] R. N. Khushaba, A. Alsukker, A. Al-ani, and A. Al-jumaily, "Enhanced Feature Selection Algorithm Using Ant Colony Optimization And Fuzzy Memberships," 6th IASTED Int. Conferance, pp. 34–39, 2008.

[44] Y. Marinakis, M. Marinaki, and N. Matsatsinis, "A hybrid discrete Artificial Bee Colony - GRASP algorithm for clustering," 2009 Int. Conf. Comput. Ind. Eng., pp. 548–553, 2009.

[45] K. Shang, S. Karungaru, Z. Feng, L. Ke, and K. Terada, "A GA-ACO hybrid algorithm for the multi-UAV mission planning problem," 14th Int. Symp. Commun. Inf. Technol. Isc. 2014, pp. 243–248, 2015.

[46] X. Wen and W. Cong, "A Hybrid Improved Ant Colony Optimization and Random Forests Feature Selection Method for Microarray Data," INC, IMS IDC, 2009. NCM '09. Fifth Int. Jt. Conf., pp. 559–563, 2009.

[47] H. S. Alghamdi, H. L. Tang, and S. Alshomrani, "Hybrid ACO and TOFA feature selection approach for text classification," 2012 IEEE Congr. Evol. Comput., pp. 1–6, 2012.

[48] P. Ghamisi and J. A. Benediktsson, "Feature Selection Based on Hybridization of Genetic Algorithm and Particle Swarm Optimization," Geosci. Remote Sens. Lett. IEEE, vol. 12, no. 2, pp. 309–313, 2015.

[49] S. Nema, "Improved Particle Swarm Optimization approach for Classification by using LDA," 9th Int. Conf. Intell. Syst. Control, no. 2, 2015.

[50] F. Liu and Z. Zhou, "A new data classification method based on chaotic particle swarm optimization and least square-support vector machine," Chemom. Intell. Lab. Syst., vol. 147, pp. 147–156, 2015.

[51] K. C. Lin and Y. H. Hsieh, "Classification of Medical Datasets Using SVMs with Hybrid Evolutionary Algorithms Based on Endocrine-Based Particle Swarm Optimization and Artificial Bee Colony Algorithms," J. Med. Syst., vol. 39, no. 10, 2015.

[52] M. Celik, F. Koylu, and D. Karaboga, CoABCMiner: An Algorithm for Cooperative Rule Classification System Based on Artificial Bee Colony, vol. 25, no. 1. 2016.

[53] P. Moradi and M. Gholampour, "A hybrid particle swarm optimization for feature subset selection by integrating a novel local search strategy," Appl. Soft Comput. J., vol. 43, pp. 117–130, 2016.