ADV MATH SCI JOURNAL
Advances in Mathematics: Scientific Journal **9** (2020), no.6, 3567–3579
ISSN: 1857-8365 (printed); 1857-8438 (electronic)
https://doi.org/10.37418/amsj.9.6.35 Spec. Issue on RDESTM-2020

COMPARATIVE STUDY OF HYBRID SWARM OPTIMIZATION FOR FEATURE SELECTION AND CLASSIFICATION ACCURACY

MUKESH[1] AND JYOTI VASHISHTHA

ABSTRACT. Hybrid algorithms are the combination of two or more algorithms to use the advantages or eliminate the drawbacks of individual algorithms for improving overall search efficacy. In this regard various combinations of different algorithms at different levels have been proposed in the research literature of data mining. Artificial bee colony (ABC), genetic algorithms (GA), particle swarm optimization (PSO) etc. are various swarm optimization algorithms which have been hybridized in many recent research papers in the field of data mining. These hybrid algorithms play important role to improve the classification accuracy of the models. In this work, we compare the performance of different hybrid swarm optimization algorithms validated by most famous classifiers i.e. SVM (support vector machine) and KNN (k-nearest neighbour). This review provides a comparison table of the performance of various hybrid algorithms in terms of classification accuracy and feature selection. This work presents a review on hybrid algorithm based on data mining techniques which motivate us to work on hybridization approaches for classification purpose.

[1] *corresponding author*

2010 *Mathematics Subject Classification.* 90C26.

Key words and phrases. Data Mining, Genetic Algorithm, Particle Swarm Optimization, Hybrid Swarm Optimization, Artificial Bee Colony, Grey Wolf Optimization, Whale optimization algorithm.

1. Introduction

Swarm is a group of individuals, which communicate and co-ordinate with each other. Apparently the intelligent behaviour that surfaces from the collective behaviour of large number of these independent agents is termed as 'swarm intelligence'. This word was firstly introduced by G. Beni, Hackward and J. Wany in 1989 [1]. Swarm Intelligence, a branch of Artificial Intelligence, models the cooperative behavior of social swarms. It provides an idea to control and manage complex systems by interacting among the entities. The communication medium of agents can be of direct or indirect. In direct interaction the agents interact with each other by audio or visual contacts such as the waggling of honey bees and in indirect interaction, the communication is through environment. One of the agents changes the environment and other understands that change. For example, the pheromone traces of ants deposited on their way to hunt for food sources.

The evolving field swarm intelligence attracted many researchers and inspired them to model the collective behavior of social swarms in nature, such as ant colonies, honey bees, bird flocks etc. [2-4]. Inspired by natural swarm systems, a number of swarm optimization approaches like particle swarm optimization (PSO), cat swarm optimization (CSO), artificial bee colony optimization (ABC), ant colony optimization (ACO) are designed and developed and successfully applied in varied range of domains like function optimization, scheduling, structural optimization, finding optimal routes, machine learning, data mining, medical informatics, operations research, bioinformatics, image analysis, industrial problems, and even business. Data mining is the automated or semi-automated procedure of breaking down and demonstrating the huge information repository so as to concentrate interesting data [5]. It may be referred as an interdisciplinary field which involves the integration of various techniques and methods from multiple disciplines such as statistics, machine learning, neural networks, data visualization, image and signal processing, and data analysis. It may also be defined as the process of discovering meaningful information from stored data by using machine learning and data visualization techniques.

Data mining is often referred as a crucial step in knowledge discovery which deals with applying methods to mine interesting data patterns whereas knowledge discovery is the standardized process of finding unknown and interesting

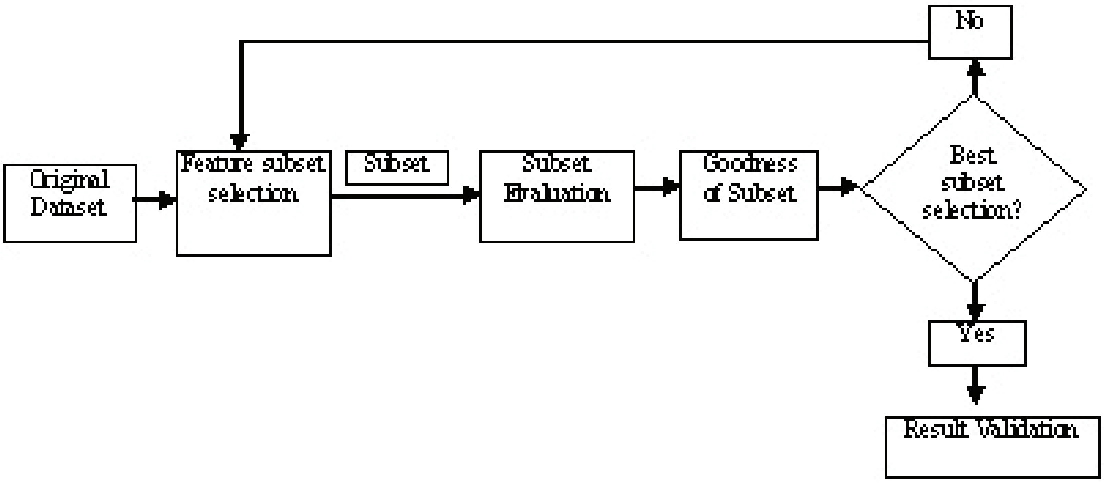

FIGURE 1. Feature Selection Process

information from raw data involving pre-processing and post-processing steps other than data mining [6]. Feature selection has become primary task of reducing the complexity, computational speed while improving the accuracy of a classification problem.

2. FEATURE SELECTION

Feature selection is very vital part of the classification process. In datasets, all the features are not important. The feature selection process selects only those attributes are which are actually needed and reduce the irrelevant and unnecessary features for the classification. With the reduction of the dimensions of data, the performance of classification process improves. Figure 1 shows the process of feature selection. Feature selection solves this problem by [7].

Recently, with the intensification of data dimensionality, a number of algorithms of feature selection confronts in terms of efficiency and effectiveness. Feature selection is done by using a classifier along with an optimization technique. There are many optimization techniques used in feature selection such as particle swarm optimization, genetic algorithms, ant colony optimization etc. For the classification purpose, support vector machine, decision trees techniques can be used.

2.1 Feature Selection Problem

Assume a dataset D with R number of records and N dimensions i.e. $D = R \times N$ matrix. Now the feature selection process finds n dimension where $n < N$. The two types of decision are taken in this process:

- Number of features
- Best subset of features

For all the n features, a pre-defined criterion function is found. One example of the criterions function is accuracy. Based on criterion function, the features that perform worst are discarded. Criterion function is the important decision of feature subset selection. There is no single criterion function which is best for all the data mining problems.

2.2 Importance of Feature Selection

Feature selection become more significant when the numbers of features are very huge [8]. There is no need to use all the features in the algorithm. You can only select those features that are important for your algorithm. I have myself witnessed feature subsets giving better results than the complete set of feature for the same algorithm. By feature selection not only training time reduced but also the evaluation time is reduced, you also have lesser things to worry about. Following are the top reasons to use feature selection:

(a) Provides better handling capabilities to machine learning techniques in lesser time.
(b) Reduces complexity and make model simpler to understand.
(c) Improves classification accuracy if relevant and less number of feature subset is selected.

2.3 Working of Feature Selection in Data Mining

Before training the model, feature selection is performed. Feature selection technique is 'by default built in' With a number of algorithms, so the inappropriate columns can be expelled and best features subset are repeatedly revealed. Each and every algorithm having feature selection techniques have their own way of applying the features intelligently. On the other hand, the parameters can be manually set to control feature selection behavior.

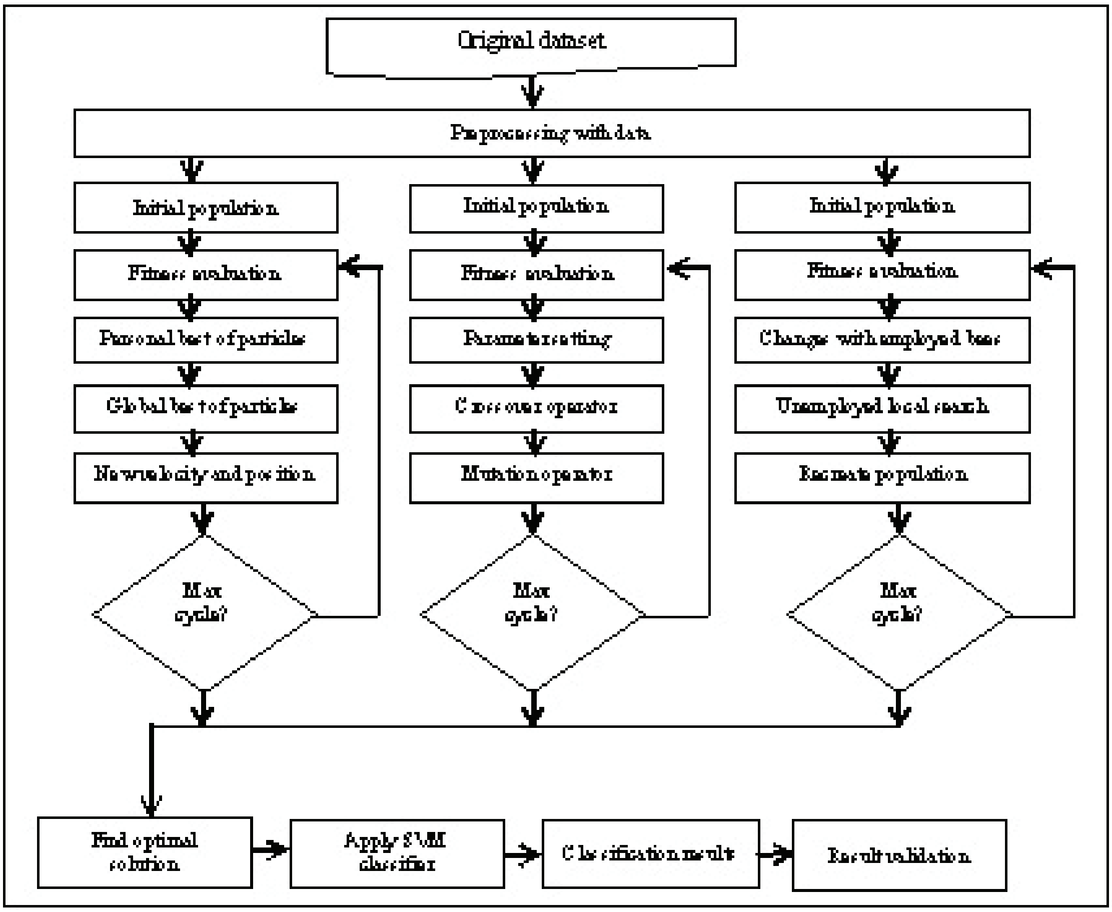

FIGURE 2. Classification Process based on Three Optimization Algorithms (PSO, GA, ABC)

3. SWARM INTELLIGENT ALGORITHMS

Various swarm intelligence algorithms have been widely used in the field of data mining, machine learning, parameter optimization and feature reduction due to their superior characteristics of parallel processing, fast optimization, and global optimization ability [9]. Figure 2 shows working process of three swarm optimization approaches named as Particle Swarm Optimization (PSO), Genetic Algorithm (GA) and Artificial Bee Colony algorithm (ABC). This figure proves that these approaches have different ways to provide the similar output.

3.1 Artificial Bee Colony Optimization Algorithm (ABC)

The Artificial Bee Colony (ABC) algorithm is a swarm based meta-heuristic algorithm. It is inspired by the intelligent searching behavior of honey bees. Work of honey bees is divided into four type i.e. foraging, storing, distributing and retrieving honey. Recruitment and information exchange are two main processes of ABC. Bees follow a different communication medium which is waggle dance. These bees can be separated by three groups named as employed, unemployed and scouts bees [10]. Employed bees are experienced bees, which move randomly from one flower to another for food source and when they achieved their goal they provide information to unemployed bees (onlooker) by waggle dance [11]. Scouts bees also search for food and after finding it, they also provide information to other by same process. Initially there were two algorithms named as ecological and bee system algorithm which are inspired by bees' collective behaviour [12].

3.2 Genetic Algorithm (GA)

Genetic algorithm is an intelligence and useful method that was introduced by Professor Holland in 1975 [13]. This technique was propelled by Darwin's organic advancement hypothesis, and it looks for ideal arrangements by reproducing the regular determination component of natural development in reality. In GA, the potential arrangement of the issue should be encoded as the chromosome that contains the parameters that should be improved (the arrangement vector), and a parameter in a chromosome is called as a quality [14]. The nature of chromosomes (potential arrangements) is determined through a wellness work. The chromosomes with higher wellness have higher likelihood to stay in the people to come. In the encoding procedure, the hyperspace is changed over into a pursuit space appropriate to the hereditary calculation, and an underlying populace (a subset of potential arrangements) is produced [15]. Consequently, the guardians' populace creates posterity (another age of arrangements) through hybrid and transformation tasks. In hybrid, the guardians' chromosomes trade a portion of their qualities to produce new age. After hybrid, the qualities of new age may infrequently be adjusted in transformation. The hybrid and transformation are the fundamental administrators of hereditary calculation, which can give progressively substitute chromosomes (arrangements) in progressive populaces. A choice task is utilized to hold the arrangement with the most noteworthy wellness [16].

3.3 Particle Swarm Optimization(PSO)

Kennedy and Eberhart in 1995 developed particle swarm optimization by studying social and cognitive behaviour of flock of birds [2]. They explore the word 'Boids' which was developed by Craig Reynolds in 1986. Boids are geometrical shapes i.e. birds fly in V-shaped pattern. Age, sex and body size play an important role for position of every bird in V-shaped flock of birds. Young birds always lead the group because juveniles usually slow down the entire group. V-shaped formation of birds help them to communicate and co-ordinate with each other and improve the capacity of group because when the leading bird tired, it falls to the rear of the V. V-shaped formation allow the birds to fly for long period of time without taking rest because of rotation of bird's position [17, 18].

3.4 Ant Colony Optimization Approach (ACO)

Ants are the best examples of indirect interaction of swarms. Ants are small tiny creatures which have limited intellectual ability. They use different type of pheromone for indication. Food trail pheromone is mostly considered for the swarm optimization algorithm which they leave on the way when they go to find the food [19]. Ant colony optimization was introduced by Macro Darigo in 1991 in his PhD thesis which is based on ants foraging behaviour. ACO solved the complex problem i.e. optimization problem, vehicle routing problem, scheduling problem, sequential ordering problem, assembly line balancing, multi-objective area etc. foraging behaviour generate an approach i.e. positive feedback process. Chemical substances always evaporate, so there is a chance to choose negative feedback which provides a new route [20, 21] to the ants.

4. Problem Formulation

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 the classification accuracy using feature selection methods. Swarm optimization algorithms have two distinct parameters: accuracy and computational time. Research work addresses the problem of improving both problems simultaneously. Hybrid swarm optimization approaches have solved these issues but there exist some space to improve the existing approaches by applying hybridization techniques. Moreover, hybridization of two technique of same swarm optimization is producing good results, for example, hybrid ant colony optimization improved by adding features of filter and wrapper approach [22]. An enhanced algorithm is proposed by using particle swarm optimization technique with Support Vector Machine for feature selection purpose. SVM is used as classifier for dividing the training dataset into two classes and PSO is used for optimizing the feature subset, by using PSO reduced numbers of features are selected [23, 24].

5. Comparative Study of Hybrid Swarm Optimization

Hybridization can be achieved by using four strategies.

- LLRH - Low Level Relay Hybridization method is used to improve an existing approach is used to improve an existing approach by improving a function of the same algorithm and it is single solution based approach.
- HLRH - In high level relay hybridization approach two different algorithm are used in pipelining way in a serial manner.
- LLTH - Low level teamwork hybridization is achieved by using global search strategy and achieved by population based metaheuristic.
- HLTH - Two algorithms are used in parallel manner in high level teamwork hybridization.

In 2009, a novel cancer classifier is introduced which is based on ACO and random forest approaches [25]. It provides promising results for feature selection and classification accuracy for micro array data. In 2010, a discrete PSO technique is proposed for binary classification [26]. This hybridization provides better results in term of feature selection and classification accuracy. In 2011, catfish binary particle swarm optimization algorithm is proposed for feature selection [27]. Most popular K-nearest neighbour classifier is used for cross validation. Comparison provide a conclusion that this catfishBPSO is superior in both the objectives i.e. feature selection and classification accuracy with respect to PSO alone. In 2012, another improvement is accomplished by introducing a novel approach named as ACO-TOFA (Trance oriented feature analysis) which provide better results in text classification and feature selection [28].

In 2015, a novel hybrid method is proposed by integrating two most famous swarm optimization algorithms named as HGAPSO (Hybrid Genetic algorithm and Particle Swarm Optimization). SVM classifier is used for cross-validation. This new approach has automatic most informative feature selection abilities. In 2016, Cat Swarm optimization algorithm is improved for text classification named as ICSO. This approach works more efficiently for big data and text classification. In 2016, another new hybrid approach named as HPSO-LS is introduced which is improved by local search strategy. This new hybrid method has been compared with various optimization techniques on 13 benchmark datasets which demonstrated that HPSO-LS is superior in both feature selection and classification accuracy. In 2017, an improved hybrid technique comes in nature by Y. Wan et al. which is based on ACO named as MBACO (modified binary coded ant colony optimization). This hybrid method is compared with PSO, GA and other meta-heuristic for feature selection. The results show that this method has significant and better performance with respect to other methods.

In 2017, another hybrid method name as AC-ABC is introduced which is performed by integrating ACO and ABC algorithms. 13 benchmark datasets have been used to find the significant results of proposed algorithm. Results indicate that this hybrid method have better performance in both objectives as feature selection and classification accuracy.

In 2019, GWO algorithm is hybridized in serial manner with WOA for feature selection and classification. Results of this hybridization proved the superiority of the HSGW approach. Another hybrid algorithm was proposed recently in 2020 which was hybridized by using two phase mutation operator with GWO algorithm. Proposed approach proved its capabilities and open the way of hybridization possibilities with new approach.

We have reviewed the research papers from year 2009 to 2019. The experimental areas of these techniques are quiet vast and it is not possible to display results for all the datasets. So, some of the best results are depicted in the tables. The performance, in terms of accuracy, of various hybrid swarm optimization algorithms and base techniques on a particular dataset is shown in the table 1. It is clearly evident from the comparison that there is an improvement in the classification accuracy when the base techniques are hybridized with other swarm optimization techniques. The table 2 shows the comparison of number of feature selected by the base techniques and the hybrid swarm optimization

TABLE 1. Comparison of classification accuracy of various hybrid swarm optimization algorithms

Year	Experimental Area	Base Technique	Accuracy of Base Technique	Hybridized Algorithm	Accuracy of Hybridized Algorithm
2009	9 data sets from UCI data repository	ABC	97.65	AC+GRASP	99.95
2009	13 datasets from UCI data repository	ACO	93.55	ACO+RF	96.77
2010	10 datasets from UCI data repository	PSO	79.80	DPSO	81.50
2011	10 datasets from UCI repository	PSO	94.47	Catfish + BPSO	96.92
2012	2 datasets from UCI repository	ACO	77.43	ACO+TOFA	90.12
2015	Indian Pines hyper-spectral data	GA	65.41	HGAPSO	73.39
2016	10 datasets from UCI data repository	CSO	81.50	ICSO	83.30
2016	13 datasets from UCI data repository	PSO	79.27	HPSO-LS	85.30
2017	10 datasets from UCI data repository	ACO	85.18	MBACO	95.10
2017	13 datasets from UCI data repository	ACO	85.30	ACO+ABC	93.06
2019	18 datasets from UCI data repository	GWO	97.60	GWO+WOA	98.60

algorithms. The comparative study and the results described in the tables establish that the hybridization of two or more swarm optimization approaches gives better and plays a prodigious role in feature selection and classification.

6. CONCLUSION

The hybrid algorithms are the logical blend of various existing techniques to boost the performance and provide better results. In this work, we put forward an explicit comparative analysis on various swarm optimization techniques and the hybrid algorithms in data mining. The performance of these algorithms is witnessed on multiple datasets. From the comparative study it is observed that hybrid optimization is better in terms of efficiency as compared to basic

TABLE 2. Comparison table for selected feature in various hybrid swarm optimization algorithms

Year	Base Technique	Hybridized Algorithm	Data set	Number of Features Selected by the	
				Base Technique	Hybridized Algorithm
2009	ABC	AC+GRASP	BCW1	8	5
2009	ACO	ACO+RF	Colon tumor	6	4
2010	PSO	DPSO	waveform	11	10
2011	PSO	Catfish + BPSO	Ionosphere	10.5	8
2012	ACO	ACO+TOFA	Brown	7	5
2016	CSO	ICSO	Wine	2.4	2.4
2016	PSO	HPSO-LS	Vowel	8	4
2017	ACO	MBACO	Handwritten	20	16
2017	ACO	ACO+ABC	Dermatology	26	24
2019	GWO	GWO+WOA	German	19	14

version of algorithms. They enhanced the accuracy of classification and reduced computational time of optimization.

REFERENCES

[1] G. BENI, J. WANG: *Swarm intelligence*, Proc. of the Seventh Annual Meeting of the, Tokyo, Japan, 1989, 425–428.

[2] J. KENNEDY, R. EBERHART: *Particle swarm optimization*, IEEE Int. Conf. on Neu. Net., **4** (1995), 1942–1948.

[3] T. SOUSA, A. SILVA, A. NEVES: *Particle Swarm based Data Mining Algorithms for classification tasks*, Parallel Computing, **30** (2004), 767–783.

[4] I. H. WITTEN, E. FRANK: *Data Mining: Practical machine learning tools and techniques*, 2nd Edition, Morgan Kaufmann, San Francisco, CA, 2005.

[5] S. C. CHU, T. SAI: *Computational intelligence based on the behavior of cats*, Int. J. Innov. Comput. Inf. Control., **3**(1) (2007), 163–173.

[6] K. C. LIN, Y. H. HUANG, J. C. HUNG: *Feature selection and parameter optimization of support vector machines based on modified artificial fish swarm algorithms*, Math. Probl. Eng., **15** (2015), 33–44.

[7] K. C. LIN, Y. H. HUANG, J. C. HUNG, Y. T. LIN: *Feature selection and parameter optimization of support vector machines based on modified cat swarm optimization*, Int. J. Distrib. Sens. Netw., **24** (2015), 116–127.

[8] G. SELVARAJ, S. JANAKIRAMAN: *Improved Feature Selection Based on Particle Swarm Optimization for Liver Disease Diagnosis*, Swarm, Evolutionary, and Memetic Computing: 4th International Conference, SEMCCO, Chennai, India, 2013, 214–225.

[9] M. J. VASHASITHA: *Hybrid Swarm Optimization Approaches: A Survey*, International Journal of Latest Trends in Engineering and Technology, **9**(2) (2017), 175–181.

[10] Y. YONEZAWA, T. KIKUCHI: *Ecological algorithm for optimal ordering used by collective honey bee behavior*, 7th International Symposium on Micro Machine and Human Science, 1996, 249–256.

[11] T. D. SEELEY, P. K. VISSCHER, K. M. PASSINO: *Group decision making in honey bee swarms*, American Scientist, **94**(3) (2006), 220–238.

[12] D. T. PHAM, A. GHANBARZADEH: *Multi-objective optimisation using the bees algorithm*, Innovative Production Machines and Systems IPROMS Conference, 2007.

[13] D. E. GOLDBERG, J. H. HOLLAND: *Genetic algorithms and machine learning*, Machine Learning, **3**(2) (2017), 95–99.

[14] K. DEEP, M. THAKUR: *A new mutation operator for real coded genetic algorithms*, Applied Mathematics and Computation, **193**(1) (2007), 211–230.

[15] J. GENLIN: *Survey on genetic algorithm*, Computer Applications and Software, **2** (2004), 69–73.

[16] J. J. GREFENSTETTE: *How genetic algorithms work: A critical look at implicit parallelism*, Proceedings of the 3rd International Joint Conference on Genetic Algorithms, 1989.

[17] A. CERVANTES, I. M. GALVAN, P. ISASI: *AMPSO: a new particle swarm method for nearest neighborhood classification*, IEEE Trans Syst Man Cybern Part B, 2009, 1082–1091.

[18] NEHA, J. VASHISHTHA: *Particle Swarm Optimization based Feature Selection*, Published in International Journal of Computer Applications, **146**(6) (2016), 11–17.

[19] S. RATNOO, A. PATHAK, J. AHUJA, J. VASHISHTHA: *Exception discovery using ant colony optimisation*, Published in International Journal of Computational Systems Engineering, **4**(1) (2018), 46–57.

[20] M. H. AGHDAM, N. G. AGHAEE, M. E. BASIRI: *Text feature selection using ant colony optimization*, Expert Systems with Appli-cations **36** (2009), 6843–6853.

[21] R. K. SIVAGAMINATHAN, S. RAMAKRISHNAN: *A hybrid approach for feature subset selection using neural networks and ant colony optimization*, Expert Systems with Applications, **33** (2007), 49–60.

[22] W. XIONG, C. WANG: *A Hybrid Improved Ant Colony Optimization and Random Forests Feature Selection Method for Microarray Data*, Fifth International Joint Conference on INC, IMS and IDC, (2009) 559–563.

[23] K. C. LIN, K. Y. ZHANG, J. C. HUNG: *Feature selection of support vector machine based on harmonious cat swarm optimization*, The 7th IEEE international conference on Ubi-Media computing (UME-DIA14), Mongolia, 2014.

[24] Y. MARINAKIS, M. MARINAKI, N. MATSATSINIS: *A hybrid discrete artificial bee colony grasp algorithm for clustering* , Proceedings of the International Conference on Computers and Industrial Engineering (CIE: 2009), **548** (2009), 1–3.

[25] W. XIONG, C. WANG: *A hybrid improved ant colony optimization and random forests feature selection method for microarray data*, Proc. International Conference on Networked Computing and Advanced Information Management (2009), 559–563.

[26] A. UNLER, A. MURAT:*A discrete particle swarm optimization method for feature selection in binary classification problems*, Eur. J. Oper. Res., **206**(3) (2010), 528–539.

[27] L. Y. CHUANG, S. W. TSAI, C. H. YANG:*Improved binary particle swarm optimization using catfish effect for feature selection*, Expert Syst. Appl., **38**(10) (2011), 12699–12707.

[28] H. S. ALGHAMDI, H. L. TANG, S. ALSHOMRANI: *Hybrid ACO and TOFA feature selection approach for text classification*, 2012 IEEE congress on evolutionary computation (CEC). IEEE, 2012, 1–6.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
GURU JAMBHESHWAR UNIVERSITY OF SCIENCE AND TECHNOLOGY
HISAR, HARYANA, INDIA
E-mail address: mukeshnimbiwal@gmail.com

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
GURU JAMBHESHWAR UNIVERSITY OF SCIENCE AND TECHNOLOGY
HISAR, HARYANA, INDIA
E-mail address: jyoti.vst@gmail.com