RESEARCH   **Open Access**

Data feature selection based on Artificial Bee Colony algorithm

Mauricio Schiezaro and Helio Pedrini[*]

Abstract

Classification of data in large repositories requires efficient techniques for analysis since a large amount of features is created for better representation of such images. Optimization methods can be used in the process of feature selection to determine the most relevant subset of features from the data set while maintaining adequate accuracy rate represented by the original set of features. Several bioinspired algorithms, that is, based on the behavior of living beings of nature, have been proposed in the literature with the objective of solving optimization problems. This paper aims at investigating, implementing, and analyzing a feature selection method using the Artificial Bee Colony approach to classification of different data sets. Various UCI data sets have been used to demonstrate the effectiveness of the proposed method against other relevant approaches available in the literature.

Introduction

Data analysis aims at extracting and modeling information content to identify patterns within the data. As a manner of simplifying the amount of information to describe a large set of data, features are extracted from the data, serving as representative characteristics of its contents. In image analysis, for instance, examples of features include color, texture, edges, object shape, interest points, among others. These features usually are organized into an n-dimensional feature vector.

Feature selection is an important step used in several tasks, such as image classification, cluster analysis, data mining, pattern recognition, image retrieval, among others. It is a crucial preprocessing technique for effective data analysis, where only a subset from the original data features is chosen to eliminate noisy, irrelevant or redundant features. This task allows to reduce computational cost and improve accuracy of the data analysis process.

This paper proposes a feature selection method for data analysis based on *Artificial Bee Colony* (ABC) approach that can be used in several knowledge domains through *wrapper* and *forward* strategies. The ABC method has been widely used to solve optimization problems; however, there have been few works on feature selection. Our work proposes a binary version of the ABC algorithm,

where the number of new features to be analyzed in a neighborhood of a food source is determined through a perturbation parameter proposed by Karaboga and Akay [1]. The method is analyzed and compared to other relevant approaches available in the literature. Experimental results showed that a reduced number of features can achieve classification accuracy superior than that using the full set of features. The accuracy has significantly increased even though the number of selected features has drastically reduced. Furthermore, the proposed method presented better results for the majority of the tested data sets compared to other algorithms.

The paper is organized as follows: Initially, some relevant concepts and work related to feature selection are described. The proposed methodology for feature selection is then presented in detail. Experimental results obtained through the application of the proposed method to several data sets are described and discussed. Finally, the remaining section concludes the paper with final remarks and directions for future work.

Related concepts and work

The process of feature selection is responsible for electing a subset of features, which can be described as a search into a state space. One can perform a full search in which all the spaces are traversed; however, this approach is impractical for a large number of features. A heuris-

*Correspondence: helio@ic.unicamp.br
Institute of Computing, University of Campinas, Campinas, São Paulo
13083-852, Brazil

tic search considers the features, not yet selected at each iteration, for evaluation. A random search generates random subsets within the search space, such that several bioinspired and genetic algorithms use this approach [2].

Feature selection can be described as a search into a space of states, and according to the initialization and behavior during the search steps, we can divide the search into three different approaches [3]: *forward:* the feature subset is initialized empty and features are included in the subset during the feature selection; *backward:* the feature subset is initialized with a full set of features and the features are excluded from the subset during the feature selection process; *bidirectional:* features can be inserted or excluded during the feature selection process.

Feature selection methods can be classified into two main categories: *filter* approaches [4-9] and *wrapper* approaches [10-14]. In filter approaches, a filtering process is performed before the classification process; therefore, they are independent of the used classification algorithm [15]. A weight value is computed for each feature, such that those features with better weight values are selected to represent the original data set. On the other hand, wrapper approaches generate a set of candidate features by adding and removing features to compose a subset of features. Then, they employ accuracy to evaluate the resulting feature set. Wrapper methods usually achieve superior results than filter methods.

Many evolutionary algorithms have been used for feature selection, which include genetic algorithms and swarm algorithms [16]. Swarm algorithms include, in turn, Ant Colony Optimization (ACO) [5,17,18], Particle Swarm Optimization (PSO) [19], Bat Algorithm (BAT) [2], and Artificial Bee Colony [1,20-22].

The use of Swarm Intelligence for feature selection has increased in the last years. Suguna and Thanushkodi [23] proposed a rough set approach with ABC algorithm for dimensionality reduction using different medical data sets in the area of Dermatology for tests, whereas Shokouhifar and Sabet [24] employed the same algorithm (ABC) for feature selection using neural networks. Particle Swarm Optimization has been proposed for feature selection either as filter method [15] or as wrapper method [25-27]. Nakamura et al. [2] proposed a wrapper method using a BAT algorithm with OPF classifier. Among feature selection approaches to Ant Colony Optimization, we can highlight the ACO for image feature selection proposed by Chen et al. [28].

The Artificial Bee Colony is a Swarm Intelligent algorithm used to solve optimization problems in several research areas [29-33]. It was proposed by Karaboga [20] in 2005, based on forage for honeybees. Frisch [34], Frisch and Lindauer [35], and Seeley [36] have investigated the foraging behavior of bees, external information

(odor, location information in *waggle dance*, presence of other bees in the food source or between the hive and source), and internal information (source location and source odor). The process starts when bees leave the hive of a forage to search for a food source (nectar). After finding nectar, the bees store it in their stomach. After coming back to the hive, the bees unload the nectar and perform a waggle dance to share their information about the food source (nectar quantity, distance and direction from black the hive) and recruit new bees for exploring most rich food sources [37].

The minimum model of ABC to emerge a collective intelligence of bee swarm consists of three components: food sources, employed bees, and unemployed bees [38], which are described as follows:

- Food sources: each food source represents a probable solution to the problem.
- Employed bees: employed bees find a food source, store information about its quality, and share this information with other bees in the honeycomb. The number of food source and that of employed bees are the same.
- Unemployed bees: unemployed bees can be of two types: onlooker bees or scout bees.

  - Onlooker bees: onlooker bees receive information from employed bees about the quality of food sources and choose food sources with better quality to explore the neighborhood. At the moment that onlooker bees choose a food source to explore, they become employed bees.
  - Scout bees: employed bees become scout bees when a food source is exhausted. In other words, the employed bees explored a food source neighborhood *MAX LIMIT* times; however, they did not find any food source with better quality. Scout bees try to find new food sources.

A general pseudocode for the ABC optimization approach [22] is shown in Algorithm 1.

Algorithm 1 ABC optimization approach

1: Initialization Phase
2: **repeat**
3:     Employed Bee Phase
4:     Onlooker Bee Phase
5:     Scout Bee Phase
6:     Memorize the best solution achieved so far
7: **until** (Cycle = Maximum Cycle Number or a Maximum CPU time)

Initialization phase

The original algorithm [1] proposes a random creation of food sources, such that each one of them corresponds to a possible solution to the problem

$$x_{ij} = x_j^{\min} + \text{rand}(0,1)(x_j^{\max} - x_j^{\min}) \qquad (1)$$

where $i = 1, \ldots, N, j = 1, \ldots, D$, such that N is the number of food sources and D is the number of optimization parameters.

Employed bee phase

Each employed bee will explore the neighborhood of the food sources associated to them. The neighborhood exploration is defined as

$$v_{ij} = x_{ij} + \Phi_{ij}(x_{ij} - x_{kj}). \qquad (2)$$

For each food source, x_i, a food source v_i is determined through the modification of an optimization parameter j, that is, x_{ij} is modified. Indices j and k are random variables. The value of k is at the range $1, 2 \ldots, SN$ and must be different from i. Φ_{ij} is a real number between -1 and 1.

Once v_i is produced, the fitness value of the food source is obtained by

$$\text{fitness}_i = \begin{cases} \frac{1}{1+f_i}, & \text{if } f_i \geq 0 \\ 1 + \text{abs}(f_i), & \text{if } f_i < 0 \end{cases} \qquad (3)$$

where f_i is a cost function. For maximization problems, the cost function can be directly used as a fitness value.

After all employed bees have conducted their search, they share the information about the quality of the food source with the onlooker bees. The probability of an onlooker bee to choose a food source to be explored is associated to its fitness, that is,

$$p_i = \frac{\text{fitness}_i}{\sum\limits_{n=1}^{F} \text{fitness}_i}. \qquad (4)$$

Through the values of exploration probabilities, the food sources are selected by the onlooker bees.

Onlooker bee phase

The food sources with better probability to be explored are selected by the onlooker bees, which become the employed bees. The neighborhood of the selected food sources are explored as explained in the 'Employed bee phase' subsection.

Scout bee phase

The algorithm checks to see if there is any exhausted source to be abandoned. In order to decide if a source is to be abandoned, the LIMIT variable which has been updated during search is used. If the value of the LIMIT is greater than that of the MAX LIMIT, then the food source is assumed to be exhausted and is abandoned. The food source abandoned by its bee is replaced with a new food source discovered by the scout. The new food source associated with the scout bee is created randomly.

Artificial Bee Colony algorithm for feature selection

Unlike optimization problems, where the possible solutions to the problem can be represented by vectors with real values, the candidate solutions to the feature selection problem are represented by bit vectors.

Each food source is associated with a bit vector of size N, where N is the total number of features. The position in the vector corresponds to the number of features to be evaluated. If the value at the corresponding position is 1, this indicates that the feature is part of the subset to be evaluated. On the other hand, if the value is 0, it indicates that the feature is not part of the subset to be assessed. Additionally, each food source stores its quality (fitness), which is given by the accuracy of the classifier using the feature subset indicated by the bit vector.

The main steps of the proposed feature selection method are illustrated in Figure 1. Each step is described as follows:

1. Create initial food sources: for feature selection, it is desirable to search for the best accuracy using the lowest possible number of features. For this reason, the proposed method follows the forward search strategy. The algorithm is initialized with N food sources, where N is the total number of features. Each food source is initialized with a bit vector of size N, where only one feature will be presented in the feature subset, that is, only one position of the vector will be filled with 1.

2. Submit a feature subset of food sources to the classifier and use accuracy as fitness: the feature subset of each food source is submitted to the classifier, and accuracy is stored as the fitness of food source.

3. Determine neighbors of chosen food sources by employed bees using modification rate (MR) parameter: each employed bee visits a food source and explores its neighborhood. For feature selection, a neighbor is created from the bit vector of the original food source. In the basic version of ABC algorithm, the neighborhood is defined by performing a small perturbation in only an optimization parameter through Equation 2, which makes convergence slower. In the feature selection, the optimization parameters are represented by the bit vectors and their perturbation is performed by a
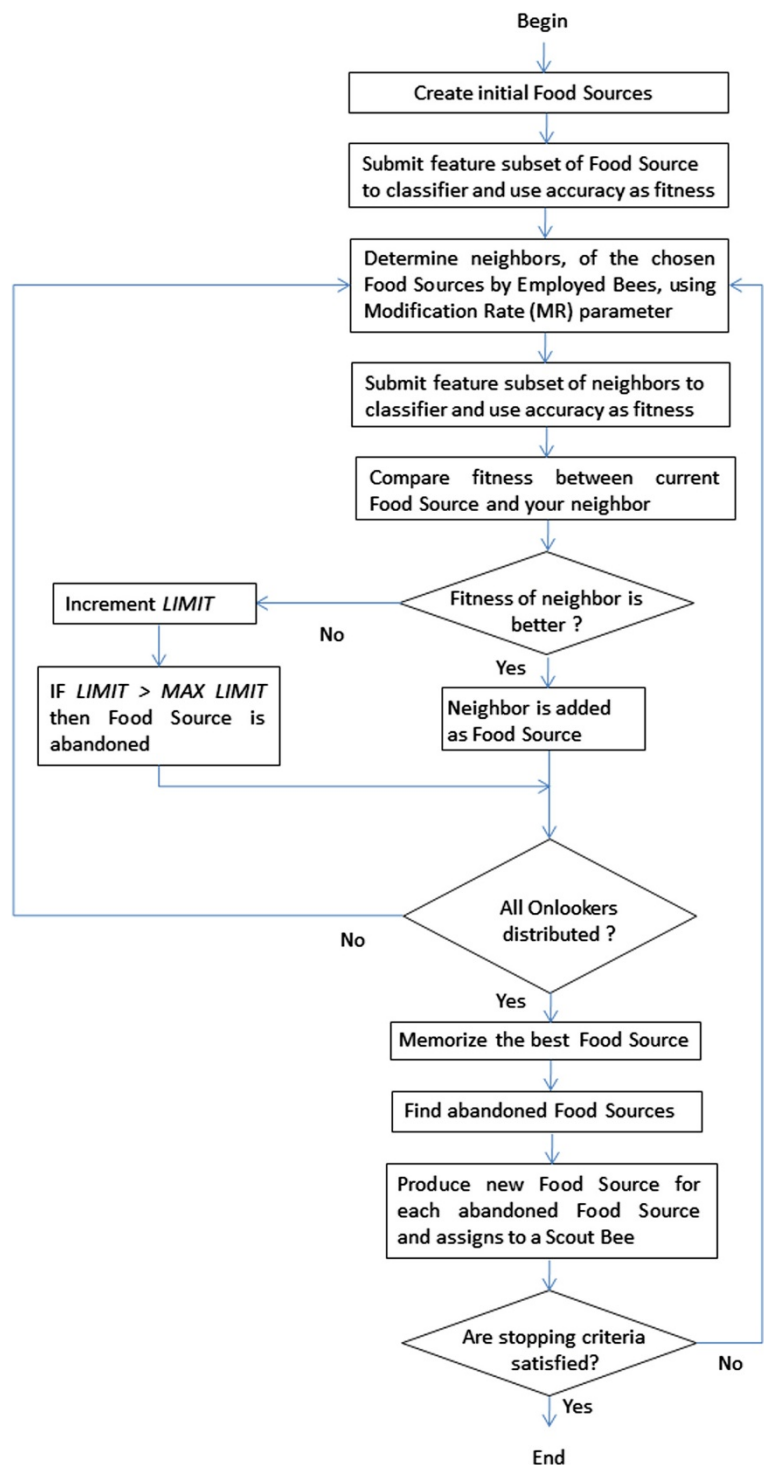
Figure 1 Steps of ABC feature selection. Diagram with the main steps of the proposed ABC feature selection method.

perturbation frequency or MR [1]. For each position of the bit vector or feature, a random and uniform number R_i is generated in the range between 0 and 1. If this value is lower than the perturbation parameter MR, the feature is inserted into the subset, that is, the vector value at that position is filled with 1. Otherwise, the value of the bit vector is not modified. This is expressed in Equation 5 :

$$x_i = \begin{cases} 1 & \text{if } R_i < \text{MR} \\ x_i & \text{otherwise} \end{cases} \qquad (5)$$

where x_i is the position i in the bit vector.

4. Submit a feature subset of neighbors to the classifier and use accuracy as fitness: the feature subset created for each neighbor is submitted to the classifier, and accuracy is stored as the neighbor's fitness.

5. Fitness of neighbor is better?: if the food source quality of the newly created neighbor is better than the food source under exploration, then the neighbor food source is considered as a new one and information about its quality will be shared with other bees. Otherwise, variable *LIMIT*, from the food source where the neighborhood is being explored, is incremented. If the value of LIMIT is greater than that of MAX LIMIT, then the food source is abandoned, that is, the food source is exhausted. In other words, the employed bees explored a food source neighborhood MAX LIMIT times; however, they did not find any food source with better quality, such that it is not worthwhile following a way where all food sources around it have worse quality than the current source. For each abandoned source, the method creates a scout bee to randomly search a new food source. The mechanism of search is illustrated in Figure 2.

6. All onlookers are distributed?: onlooker bees collect information about the fitness of food sources visited by employed bees and choose food sources with either better probability of exploration or better fitness. At the moment that onlooker bees choose the food source to be explored, they become employed bees and execute step 3.

7. Memorize the best food source: after all onlookers have been distributed, the food source with the best fitness is stored.

8. Find abandoned food sources and produce new scout bees: for each abandoned food source, a scout bee is created and a new food source is generated, where a bit vector with size N of features is randomly created and submitted to the classifier, and accuracy is stored. The new food source is assigned to scout bees, and then they become employed bees and execute step 3.

Experimental results

This section describes the data sets tested in our experiments, the computational resources used to implement and evaluate the proposed feature selection method, the strategies adopted in the data classification, the ABC parameters, as well as a discussion of the experimental results.

Data sets

The proposed method has been evaluated through ten data sets from different knowledge fields. The data sets are available from UCI Machine Learning Repository [39]. Table 1 presents a description of the tested data sets, including the number of instances, number of features, and number of classes for each data set.

UCI data sets have been widely used in the evaluation of data classification since they contain a varied number of features and classes, allowing the analysis of influence on accuracy and performance when features are selected (Table 2).

Comparison against other methods

The proposed method was compared to some relevant swarm approaches: ACO, PSO, and genetic algorithms (GAs) (Table 3).

Computational environment

All the experiments have been conducted on a computer with Intel Core I7-2600 3.4 GHz and 4-GB RAM. The Artificial Bee Colony feature selection algorithm
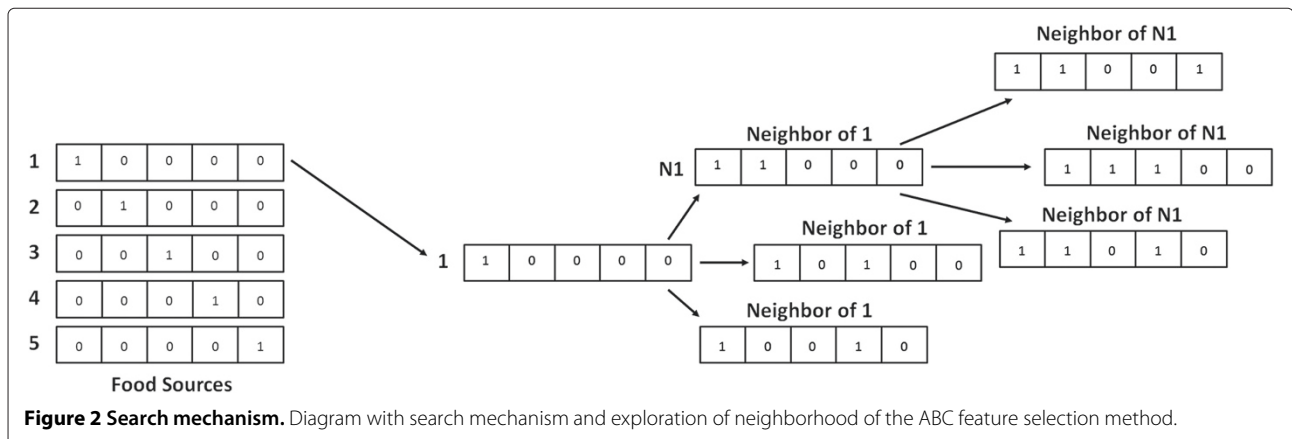

Figure 2 Search mechanism. Diagram with search mechanism and exploration of neighborhood of the ABC feature selection method.

Table 1 Summary of UCI data sets

Data set	Number of instances	Number of features	Number of classes
Image Segmentation	2,310	19	7
Auto	205	25	7
Breast Cancer	286	9	2
Diabetes	768	8	2
Glass	214	9	7
Heart-C	303	13	5
Heart-Statlog	270	13	2
Hepatic	155	19	2
Iris	150	4	2
Labor	57	16	2

was implemented using Java programming language with Weka [40] and LibSVM [41] libraries to execute the data classification.

Classification setup

To evaluate the accuracy and performance of the classification process with the original and selected feature sets, a ten fold cross-validation is used. In k-fold cross-validation, the data set is randomly partitioned into k equally sized folds (samples). One partition is retained as the test set, whereas the remaining $k - 1$ samples are used as the training set. This process is repeated k times, where one of the partitions becomes test data at each time. The average of k results produces an estimation of the accuracy. The accuracy measure employed for evaluating the results is the percentage of instances correctly classified, that is, for which a correct prediction was made.

In some tests, the feature vector has been normalized using z-score [42], that is, the features are normalized by

subtracting their mean value and dividing them by their standard deviation.

ABC parameters

The following parameters are used in the ABC algorithm:

- Food sources = N, where N is the total number of features
- MAX LIMIT = 3
- MR = 0.1
- Number of iterations = 100

PSO parameters

The following parameters are used in the PSO algorithm:

- Population size = 200
- Number of generations = 30
- C1 = 1
- C2 = 2
- Report frequency = 30

ACO parameters

The following parameters are used in the ACO algorithm:

- Population size = 10
- Number of generations = 10
- Alpha = 1
- Beta = 2
- Report frequency = 10

GA parameters

The following parameters are used in the GA:

- Population size = 200
- Number of generations = 20
- Probability of crossover = 0.6
- Probability of mutation = 0.033
- Report frequency = 20

Table 2 Results for UCI data sets

Data set	Original number of features	Selected number of features	Full set of feature average accuracy (%)	Average accuracy (%)	Normalization
Image Segmentation	19	12	65.37	91.13	None
Auto	25	9	32.68	82.93	None
Breast Cancer	9	4	73.08	75.87	z-score
Diabetes	8	1	65.10	71.48	None
Glass	9	6	68.69	71.50	None
Heart-C	13	7	54.79	83.17	None
Heart-Statlog	13	3	55.96	84.81	None
Hepatic	19	9	79.35	87.10	None
Iris	4	3	96.66	97.33	None
Labor	16	8	89.47	98.26	z-score

Table 3 Comparison of the selected features against the results for other algorithms

Data set	PSO	ACO	GA	ABC
Image Segmentation	16	16	17	12
Auto	8	9	9	9
Breast Cancer	8	9	8	4
Diabetes	8	8	8	1
Glass	8	8	8	6
Heart-C	8	7	7	7
Heart-Statlog	8	7	8	3
Hepatic	7	7	7	9
Iris	4	4	4	3
Labor	5	6	9	8

Discussion

Table 4 shows the results obtained by applying the proposed feature selection method for each data set. It is possible to observe that the selected feature set provides superior accuracy than the original feature set for all data sets, even though the number of selected features is much smaller than the original one for some data sets, such as Auto, Heart-Statlog, and Hepatic.

It can be observed that in terms of accuracy, the ABC algorithm obtained superior results (eight out ten tested data sets) when compared to other methods. Only for the Image Segmentation and Diabetes data sets, the accuracy of the proposed method was worse. For the Diabetes data set, although the other algorithms obtained a better accuracy, they did not reduce the set of features, that is, they used all the features. The proposed algorithm used only one feature; however, despite this fact, its accuracy was compatible to the other algorithms (75.65% against 71.48%). For the Image Segmentation data set, the proposed algorithm used 12 features against 16 and 17 of the

Table 4 Comparison of the accuracy using the features selected by the different algorithms

Data set	PSO (%)	ACO (%)	GA (%)	ABC (%)
Image Segmentation	94.42	94.42	94.26	91.13
Auto	68.78	72.20	69.27	82.93
Breast Cancer	73.08	73.08	73.08	75.87
Diabetes	75.65	75.65	75.65	71.48
Glass	71.03	71.03	71.03	71.50
Heart-C	83.17	80.86	80.20	83.17
Heart-Statlog	82.96	81.11	73.70	84.81
Hepatic	86.45	83.23	83.26	87.10
Iris	96.66	96.66	96.66	97.33
Labor	89.47	92.98	89.47	98.26

other algorithms; however, its accuracy was close to the other algorithms (94.26% against 91.13%).

Conclusions

This work presents a feature selection method based on ABC algorithm. The results show that a reduced number of features can achieve classification accuracy superior to that using the full set of features. For some data sets, the accuracy has significantly increased even though the number of selected features has drastically reduced. The proposed method presented better results for the majority of the tested data sets compared to other algorithms.

For future work, we plan to investigate alternative mechanisms to explore neighborhood of food sources, parallelize the exploration of employed bees in relation to the food sources, and create a filter approach combining ABC algorithm, entropy, and mutual information.

References
1. D Karaboga, B Akay, A modified artificial bee colony algorithm for real-parameter optimization. Inf. Sci. **192**, 120–142 (2012)
2. R Nakamura, L Pereira, K Costa, D Rodrigues, J Papa, in *SIBGRAPI - Conference on Graphics, Patterns and Images,* BBA: a binary bat algorithm for feature selection, (Ouro Preto, 22–25 Aug 2012)
3. H Liu, L Yu, Toward integrating feature selection algorithms for classification and clustering. IEEE Trans. Knowl. Data Eng. **17**(4), 491–502 (2005)
4. Y Jiang, J Ren, in *Machine Learning and Knowledge Discovery in Databases,* European Conference, ECML PKDD 2011, Athens, Greece, September 5–9 2011. Lecture Notes in Computer Science, ed. by D Gunopulos, T Hofmann, D Malerba, and M Vazirgiannis. Eigenvector sensitive feature selection for spectral clustering, vol. 6912 (Springer, Berlin, 2011), pp. 114–129
5. C Zhang, H Hu, in *18th Australian Joint Conference on Artificial Intelligence*, 18th Australian Joint Conference on Artificial Intelligence, Sydney, Australia, December 5–9 2005. Lecture Notes in Computer Science, ed. by S Zhang, R Jarvis. Ant colony optimization combining with mutual information for feature selection in support vector machines (Springer, Berlin, 2005), pp. 5–9
6. M Dash, K Choi, P Scheuermann, H Liu, in *2nd International Conference on Data Mining*. Feature selection for clustering - a filter solution (IEEE, Piscataway, 2002), pp. 115–122
7. M Hall, in *17th International Conference Machine Learning*. Correlation-based feature selection for discrete and numeric class machine learning (Morgan Kaufmann, San Francisco, 2000), pp. 359–366
8. H Liu, R Setiono, in *13th International Conference Machine Learning*. A probabilistic approach to feature selection - a filter solution (The International Machine Learning Society, Princeton, 1996), pp. 319–327
9. L Yu, H Liu, in *20th International Conference Machine Learning*. Feature selection for high-dimensional data: a fast correlation-based filter solution (The International Machine Learning Society, Princeton, 2003), pp. 856–863
10. E Hruschka, T Covoes, A feature selection for cluster analysis: an approach based on the simplified silhouette criterion. Int. Conf. Int. Agents, Web Tech. Internet Commerce **1**, 32–38 (2005)
11. R Caruana, D Freitag, in *11th International Conference Machine Learning*. Greedy attribute selection (The International Machine Learning Society, Princeton, 1994), pp. 28–36

12. J Dy, C Brodley, in *17th International Conference Machine Learning*. Feature subset selection and order identification for unsupervised learning (The International Machine Learning Society, Princeton, 2000), pp. 247–254

13. Y Kim, W Street, F Menczer, in *ACM SIGKDD International Conference in Knowledge Discovery and Data Mining*. Feature selection for unsupervised learning via evolutionary search (ACM, New York, 2000), pp. 365–369

14. R Kohavi, G John, Wrappers for feature subset selection. Artif. Intell. **97**(1–2), 273–324 (1997)

15. B Xue, L Cervante, L Shang, M Zhang, A particle swarm optimisation based multi-objective filter approach to feature selection for classification. Artif. Intell. Rev. **7458**, 673–685 (2012)

16. E Castro, M Tsuzuki, Swarm intelligence applied in synthesis of hunting strategies in a three-dimensional environment. Expert Syst. Appl. **34**(3), 1995–2003 (2008)

17. Z Yan, C Yuan, in *Biometric Authentication*, First International Conference, ICBA 2004, Hong Kong, China, July 15–17 2004. Lecture Notes in Computer Science, ed. by D Zhang, AK Jain. Ant colony optimization for feature selection in face recognition (Springer, Berlin, 2004), pp. 15–17

18. M Dorigo, C Blum, Ant colony optimization theory: a survey. Theor. Comput. Sci. **344**(2–3), 243–278 (2005)

19. J Kennedy, R Eberhart, in *IEEE International Conference on Neural Networks*, Particle swarm optimization, vol. 4 (IEEE, Piscataway, 1995), pp. 1942–1948

20. D Karaboga, An idea based on honey bee swarm for numerical optimization. Tech. Rep.-TR06. Computer Engineering Department, Engineering Faculty, Erciyes University,Turkey, (2005)

21. D Karaboga, B Basturk, On the performance of artificial bee colony (ABC) algorithm. Appl. Soft Comput. **8**, 687–697 (2008)

22. D Karaboga, B Gorkemli, C Ozturk, N Karaboga, A comprehensive survey: artificial bee colony (ABC) algorithm and applications. Artif. Int. Rev. (2012). doi:10.1007/s10462-012-9328-0

23. N Suguna, K Thanushkodi, An independent rough set approach hybrid with artificial bee colony algorithm for dimensionality reduction. Am. J. Appl. Sci. **8**(3), 261–266 (2011)

24. M Shokouhifar, S Sabet, in *3rd International Conference on Machine Vision*, Hybrid approach for effective feature selection using neural networks and artificial bee colony optimization (IEEE, Piscataway, 2010), pp. 502–506

25. J Zhao, C Han, B Wei, Q Zhao, P Xiao, K Zhang, in *15th International Conference Information Fusion (FUSION)*, Feature selection based on particle swarm optimal with multiple evolutionary strategies (IEEE, Piscataway, 2012), pp. 963–968

26. Y Liu, G Wang, H Chen, H Dong, X Zhu, S Wang, An improved particle swarm optimization for feature selection. J. Bionic Eng. **8**(2), 191–200 (2011)

27. A Unler, A Murat, A discrete particle swarm optimization method for feature selection in binary classification problems. Eur. J. Oper. Res. **206**(3), 528–539 (2010)

28. B Chen, L Chen, Y Chen, Efficient ant colony optimization for image feature selection. Signal Proc. **93**(6), 1566–1576 (2013)

29. D Karaboga, C Ozturk, Neural networks training by artificial bee colony algorithm on pattern classification. Neural Netw. World **19**(3), 279–292 (2009)

30. D Karaboga, C Ozturk, A novel clustering approach: Artificial Bee Colony (ABC) algorithm. Appl. Soft Comput. **11**, 652–657 (2011)

31. D Karaboga, C Ozturk, B Gorkemli, Probabilistic dynamic deployment of wireless sensor networks by artificial bee colony algorithm. Sensors **11**(6), 6056–6065 (2011)

32. D Karaboga, S Okdem, C Ozturk, Cluster based wireless sensor network routing using artificial bee colony algorithm. Wireless Netwo. **18**(7), 847–860 (2012)

33. D Karaboga, C Ozturk, N Karaboga, B Gorkemli, Artificial bee colony programming for symbolic regression. Inf. Sci. **209**, 1–15 (2012)

34. K Frisch, *The Dancing Bees: An Account of the Life and Senses of Honey Bee*. (Harcourt, Brace and World, New York, 1953)

35. K Frisch, M Lindauer, The "language" and orientation of the honey bee. Ann. Rev. Entomol. **1**, 45–58 (1956)

36. T Seeley, *Honeybee Ecology: A Study of Adaptation in Social Life*. (Princeton University Press, Princeton, 1985)

37. D Karaboga, B Akay, A survey: algorithms simulating bee swarm intelligence. Art. Int. Rev. **31**(1–4), 61–85 (2009)

38. L Bao, J Zeng, Comparison and analysis of the selection mechanism in the artificial bee colony algorithm. Hybrid Intell. Syst. **1**, 411–416 (2009)

39. UCI: UCI - Machine Learning Repository (2013). http://archive.ics.uci.edu/ml/. Accessed 28 Jan 2013

40. Weka: Weka (2013). http://www.cs.waikato.ac.nz/ml/weka/. Accessed 28 Jan 2013

41. libsvm: libsvm (2013). http://www.csie.ntu.edu.tw/~cjlin/libsvm/. Accessed 28 Jan 2013

42. S Aksoy, R Haralick, Feature normalization and likelihood-based similarity measures for image retrieval. Pattern Recognit. Lett. **5**(22), 563–582 (2001)