MRMR BA: A HYBRID GENE SELECTION ALGORITHM FOR CANCER CLASSIFICATION

[1]OSAMA AHMAD ALOMARI, [1]AHAMAD TAJUDIN KHADER, [2]MOHAMMED AZMI AL-BETAR, [1]LAITH MOHAMMAD ABUALIGAH

[1]School of Computer Science, Universiti Sains Malaysia, Penang, Malayisa
[2]Department of information technology, Al-Huson University College, Al-Balqa Applied University, Al-Huson, Irbid-Jordan
Email: [1]oasa14_com004@student.usm.my, [1]tajudin@cs.usm.my, [2]mohbetar@bau.edu.jo, [1]lmqa15_com072@student.usm.my

ABSTRACT

The microarray technology facilitates biologist in monitoring the activity of thousands of genes (features) in one experiment. This technology generates gene expression data, which are significantly applicable for cancer classification. However, gene expression data consider as high- dimensional data which consists of irrelevant, redundant, and noisy genes that are unnecessary from the classification point of view. Recently, researchers have tried to figure out the most informative genes that contribute to cancer classification using computational intelligence algorithms. In this paper, we propose a filter method (Minimum Redundancy Maximum Relevancy, MRMR) and a wrapper method (Bat algorithm, BA) for gene selection in microarray dataset. MRMR was used to find the most important genes from all genes in gene expression data, and BA was employed to find the most informative gene subset from the reduce set generated by MRMR that can contribute in identifying the cancers. The wrapper method using support vector machine (SVM) method with 10-fold cross-validation served as evaluator of the BA. In order to test the accuracy performance of the proposed method, extensive experiments were conducted. Three microarray datasets are used, which include: colon, Breast, and Ovarian. Same method procedure was performed to Genetic algorithm (GA) to conducts comparison with our proposed method (MRMR-BA). The results show that our proposed method is able to find the smallest gene subset with highest classification accuracy.

Key words: *Bat-inspired algorithm, Cancer Classification, Gene Selection, MRMR, SVM.*

1. INTRODUCTION

With the advent of DNA microarray technology, the biologist has enabled to analysis thousands of genes in one experiment. However, it is impossible to examine the expression of these huge numbers of genes through limited number of samples (high-dimensional data) [1]. As detecting and classifying of cancers are key issues in microarray technology, the existing of this huge number of genes form a challenge to classification algorithms. Gene expression dataset is such kind of huge dimensionality dataset extracted from DNA microarray technology. DNA microarray technology enables a new insight into the mechanisms of living systems by possibility of analyzing thousand of genes simultaneously and gets significant information about cell's function. This particular information can be utilized for diagnosis many diseases such as: Alzheimers disease [2], diabetes diseases [3], and cancer diseases [4].

Several studies have shown that most genes measured in a DNA microarray experiment are not relevant in the accurate classification of different classes of the problem [5]. To avoid the problem of the curse of dimensionality, gene selection (which is commonly known as feature selection) is process to find the most informative genes with respect to improve predictive accuracy of diseases [6]. Methods of gene selection are divided into 3 categories: the wrapper approach, the filter approach, and embedded approach. In the first category, the classifier is employed to assess the reliability of genes or genes subsets. In second category, a filter method does not involve the machine learning algorithm for removing irrelevant and redundant features, instead uses the principal characteristics of the training data to evaluate the significance of the genes subset or genes [7]. Wrapper methods tend to give better results but filter methods are usually computationally less expensive than wrappers [8].Embedded methods the last category of gene selection approaches, which perform gene selection in the process of training and are usually specific to given classifier algorithm. Basically, it is hybridization between filter and wrapper methods to integrating the advantages of both approaches that leads to finding informative genes with high classification accuracy [9].

The meta-heuristic techniques have been the mostly widely used in tackling gene selection problems, and their performance has been proved to be one of the better performing techniques that have been used for solving gene selection problems [10, 11]. Although many approaches have been proposed to tackle gene selection problem; most of them still suffer from problems of stagnation in local optima and high computational cost, and it cannot be guaranteed that the optimal subset of genes will be acquired, these problems as a result of huge search space [13][14]. Therefore, an efficient gene selection algorithm is needed.

Bat algorithm is a metaheuristic introduced by Xin-She Yang [15] is based on echolocation of behavior of bats. The advantage of bat algorithm is employ the idea of combining population based algorithm and local search. This combination result in providing the algorithm capability of global diverse exploration and local intensive exploitation which is the key point in metaheuristic. Furthermore, It has been successfully applied to a wide variety of optimization problems such as Continuous Optimization [15, 16], combinatorial optimization and scheduling [17, 18], Inverse problem and parameter estimation [16] and [19], Classifications [20], Clustering [21], and Image processing [22].

Due to it is strength search capabilities of BA algorithm; in this paper, we propose the application of BA to select the informative genes from microarray gene expression data. The new adaptation of BA involve change its continuous nature to binary nature. Moreover, when apply Bat algorithm to huge dimensional data, it will be encountered some challenging problem in computational efficiency, similar to other evolutionary algorithms. Toward avoiding these problems and further improvement to the performance of the Bat algorithm (BA), we adopted a filtering method, minimum redundancy maximum relevance (MRMR), as a preprocessing step to reduce the dimensionality of microarray datasets. Subsequently, Bat algorithm (BA) and Support Vector Machine (SVM) will form as a wrapper approach.

Bat algorithm produce a candidate gene subsets from elite pool of genes generated by MRMR, while SVM evaluate each candidate gene subsets, we called it (MRMR-BA). Experiments were carried out in order to evaluate the performance of the proposed algorithm using microarray benchmark datasets: Colon, Breast, and Ovarian. Results of MRMR-BA were compared with MRMR combined with GA (MRMR-GA) algorithm. When tested using all benchmark datasets, the MRMR-BA achieved better performance in terms of highest classification accuracy along with the lowest average number of selected genes. This proves that BA algorithm could be alternative approach for solving gene selection problem.

The paper is organized as follow: Section 2 introduce the related work. Section3 defines gene selection problem. Section 4 introduces a briefly description of MRMR, Bat Algorithm (BA), Binary Bat Algorithm (BBA) and SVM. Section 5 Illustrate the proposed method. The Experimental Setup and results are presented in Section 6. Finally, Section 7 conclude our paper and future work.

2. RELATED WORK

In the recent years, extensive research has been developed for gene selection problems. Various techniques for gene selection have been proposed. In the literature, there are several algorithms for gene selection and cancer classification using microarrays. Alshamlan et al. [12] proposed a new hybrid gene selection method namely Genetic Bee Colony (GBC) algorithm where the hybridization was done by combining Genetic Algorithm and Artificial Bee Colony (ABC) algorithm. The GBC algorithm proves to have a superior performance as it achieved the highest classification accuracy along with the lowest average number of genes. Shreem et al. [13] proposed a new method that hybridizes the Harmony Search Algorithm (HSA) and the Markov Blanket (MB), called HSA- MB for gene selection in classification problems. HSA utilizes naive Baye's classifier in its wrapper approach. During improvisation process, a new harmony solution is passed to MB(i.e. the filter approach) for more improvement. Experiments were carried out on ten microarray datasets. The HSA-MB performance revealed comparable results to the state-of-art approaches. El Akadi et al. [1] introduced a framework consisting of a two-stage algorithm for microarray data. In the first stage, Minimum Redundancy Maximum Relevance (MRMR) is employed to filter the genes. While in the second stage, a GA is used to generate the gene subsets and both classifiers Naïve Bayes (NB) and support vector machine (SVM) are utilized for the evaluation purpose.

3. GENE SELECTION PROBLEM

3.1 Problem Definition

Computationally, gene selection problems can be expressed as a combinatorial optimization problem in which the search space involves a set of all possible subsets [23, 24]. This problem is known to be an NP-hard problem [25], and is a highly combinatorial search in nature. The number of solutions in the search space exponentially increases when the number of genes increases, and there are [2N] possible subsets of genes, where N represents the number of genes.

3.2 Problem Formulation

The complete set of Genes is represented by a binary string of length N, where a bit in the string is set to 1 if it is to be kept, and set to 0 if it is to be discarded, and N is the original number of Genes. In context of optimization is called solution representation. The problem formulation is illustrated in Figure.1.

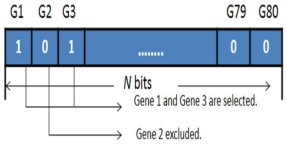

Figure 1: The problem formulation

4. RESEARCH BACKGROUND

4.1 MRMR

In this study, we employed minimum redundancy maximum relevancy (MRMR) [26] feature selection approach to address gene selection problem. The MRMR tries to find the most relevant features based on it correlation with class label. Furthermore, minimize the redundancy of the feature themselves. This filtering process reveals with the features that it has maximum relevancy and minimum redundancy. To quantify both relevancy and redundancy, mutual information (MI) is used to estimate the mutual dependency of two variables. MI is defined as in Eq.(1)) as follows:

$$I(X,Y) = \iint\, p(X,Y)\, log\, \frac{P(X,Y)}{P(X)P(Y)} \qquad (1)$$

Where X and Y are two features, P (X) and P (Y) are marginal probability functions, and P (X, Y) is the joint probability distribution. The mutual information value for two completely independently random variables is 0 [27].

Given *fi* , which represents the feature i, and the class label *c*, the Maximum-Relevance method selects the top m features in the descent order of I(fi; c), i.e. the best *m* individual features relevant to the class labels.

$$\max_S \frac{1}{|S|} \sum_{fi\in S} I(f_i; c) \qquad (2)$$

In order to eliminate the redundancy among features, a Minimum-Redundancy criterion is defined:

$$\min_S \frac{1}{|S|^2}\ \sum_{fi,fi\in S} I(\,f_i; f_j) \qquad (3)$$

A sequential incremental algorithm to solve the simultaneous optimizations of optimization criteria of Eqs. (2, and 3) is given as the following. Suppose set G represents the set of features and we already have Sm-1, the feature set with m-1 genes, then the task is to select the m-th feature from the set G-Sm -1. This feature is selected by maximizing the single-variable relevance minus redundancy function.

$$\text{Max}_{fi\in G-S_{m-1}}(I(f_i\ ; c) - \frac{1}{m-1} \sum_{f_i\in S_{\ m-1}}\ I(f_i\ ;\ f_j\ )) \qquad (4)$$

4.2 Bat Algorithm

Bat Algorithm (BA), introduce by Xin-She Yang in 2010 [15], emulates the echolocation behavior of bats. They are many kinds of bats in nature. All of them have quiet similar behaviors when navigating and hunting, whereas they are different in terms of size and weight. Among them, microbats extensively used echolocation feature. This feature assists microbats whiles it seeks for prey and/or avoids obstacles in a complete darkness. The behavior of microbats can be formulated as novel optimization technique. Furthermore, this behavior has been mathematically modeled as follows:

In the BA, the artificial bat has position vector, velocity vector, and frequency vector which are updated during the course of iterations. The BA can explore the search space through position and velocity vectors (or updated positions vectors). Each bat has a position Xi , frequency Fi , and Velocity Vi in a
d-dimensional search space.
The velocity, position, and frequency vectors is updated in Eq.((5), (6), (7).

$$V_i\,(t+1) = V_i(t) + (X_i(t) - Gbest)F_i \quad (5)$$

$$\mathrm{X_I}\,(\mathrm{t}+1) = \mathrm{X_i}(\mathrm{t}) + \mathrm{V_i}(\mathrm{t}+1) \quad (6)$$

Where Gbest is the best solution is obtained so far and Fi represent the frequency of the *i-th* bat which is update in each course of iteration as follows:

$$\mathrm{F_i} = \mathrm{F_{min+}}(\mathrm{F_{max}} - \mathrm{F_{min}})\beta \quad (7)$$

Where β is a random number of uniform distribution in [0,1]. The BA employed a random walk in order to improve its capability in exploitation as given below.

$$\mathrm{X_{new}} = \mathrm{X_{old}} + \varepsilon \mathrm{A^t} \quad (8)$$

Where *E* is a random number in [-1,1], and A is the loudness of emitted sound. At each iteration, the loudness and pulse emission I are updated as follows:

$$A_i(t+1) = \alpha A_i\,(t) \quad (9)$$

$$r_i(t+1) = r_i(0) + [1 - \exp(-\gamma t) \quad (10)$$

where α and γ are constant parameters lies between 0 and 1 used to update loudness rate A_i and pulse rate (r_i). The pseudo code of the algorithm is presented in Figure.2.

```
Initialize the bat population  $X_i$ (i=1,2,…..,n) and $V$i
Define pulse frequency $F_i$
Initialize pulse rate $r_i$ and the loudness $A_i$
While (t < Max number of iterations)
Generate new solutions by adjusting frequency,
Updating velocities and positions [equations (5) to
(7)]
If (rand > $r_i$ )
Select a solution among the best solutions randomly
Generate a local solution around the selected best
solution
End if
Generate a new solution by flying randomly
If( $rand < A_i$ & $f(x_i) < f(Gbest)$)
Accept the new solutions
Increase $r_i$ and reduce $A_i$
End if
Rank the bats and find the current Gbest
End while
```

Figure 2: Pseudocode of bat algorithm.

4.3 Binary bat algorithm

In continuous version of BA, the artificial bat can be moved around search space utilizing positions and velocity vectors (or updated position vectors) within continuous real domain. However, in dealing with binary search space where the position/ or solution is a series of 0's and 1's binary bits. As binary search space dealing with only two numbers ("0" and "1"), the updating positions process cannot be performed using Eq.6. Therefore, a transfer strategy should be found to reflect the velocity vector value in changing the elements of position vector from "0" to "1" or vice versa.

Nakamura et al. [28] have introduced a binary version of the Bat Algorithm restricting the new bat's position to only binary values using a sigmoid function as follows:

$$\mathrm{S}(\mathrm{v}_i^j) = \frac{1}{1+e^{-v_i^j}} \quad (11)$$

Therefore, Eq.11 can be replaced by:

$$\mathrm{x}_i^j = \begin{cases} 1, & \text{if } \mathrm{S}(\mathrm{v}_i^j) > \sigma, \\ 0, & \text{otherwise,} \end{cases} \quad (12)$$

In which σ ~ U(0,1). Therefore, Eq. (12) can provide only binary values for each bat's coordinates in the boolean lattice, which stand for the presence of absence of the features.

4.4 Support Vector Machine (SVM)

A support vector machine (SVMs) is a supervised learning algorithm method used for classification and regression [29]. SVM is a powerful classification algorithm; due to it efficient performance in pattern recognition domain. For example, SVM classifier has successfully applied to classify high-dimension data, such as microarray gene expression data. nonlinear data, we need to define a feature mapping function X → φ(x). This mechanism that defines feature mapping process is called kernel function. There are three common functions:

- Polynomial kernel

$$k(x_i, x_j) = (x_i \,.\, x_j + a)^b \,, \qquad (13)$$

- Radial basis kernel

$$k(x_i, x_j) = e^{-(\|x_i - x_j\|^2/2\sigma^2)}, \qquad (14)$$

- Sigmoidal kernel

$$k(\, x_i, x_j) = \tanh(\, \alpha x_i \,.\, x_j - b) \qquad (15)$$

Where a and b are parameters define the kernel's behavior.

5. PROPOSED METHOD FOR GENE SELECTION

Based on aforementioned characteristic of microarray dataset, it is impractical to adopt a evolutionary algorithm such as BA algorithm directly to such huge dimensionality data. A pre-process should take place to overcome this difficulty. Therefore, MRMR is employed to filter noisy and redundant genes. It utilizes a series of intuitive measures of relevance and redundancy to pick out promising features for both continuous and discrete datasets As shown in Eq.(1,2,3 and 4). The main role of MRMR is to select the genes that have minimum redundancy for input genes and maximum relevancy for cancer disease. To further explore reduced gene subset and identify a subset of informative genes, BA and SVM classifier combined to seek for the better gene subset in wrapper way, as shown in Figure.3.

BA has employed as a search technique to figure out the near optimal gene subset from the reduce set generated. Initially, as nature of gene selection problem is binary. Therefore, BA SVM constructs a hyperplanes or a set of hyperplanes in a high- dimensional space, which can be utilized for classification, regression, or other tasks.

SVM has the capability to deal with linear and nonlinear datasets. In linear data, SVM tries to find an optimal separating hyperplane that maximizes the margin between the training examples and the class boundary. In has changed to be applicable to binary problems, as mentioned in section 3.3.

Based on BA pseducode (as shown in Fig.1) BA algorithm starts to generate initial population or group of bats. Each bat consists a series of 0's and 1's bits, where bit value 1 represents that this gene is selected and bit value 0 represents that this gene is discarded. SVM has utilized as a wrapper approach to evaluate each candidate gene subset produced by BA. The evaluation function is combined classification accuracy and gene subset length, as shown in following equation:

$$\text{Fittness (R)} = \alpha\, \gamma \text{R(D)} + \beta\, \frac{|\text{C}|}{|\text{C}| - |\text{R}|} \qquad (16)$$

Where αγR(D) is the average of classification accuracy rate got by carrying out ten multiple cross-validation with SVM classifiers. In each round, SVM build a prediction model on the training dataset based on gene subset R and (class label)/Decision D, and a testing will perform on the prediction model on the testing dataset to obtain classification accuracy.

|R| is the length of selected gene subset. |C| is the total number of genes. $\boldsymbol{\alpha}$ and β are two parameters related to the significance of classification quality and subset length. $\boldsymbol{\alpha}$ϵ [0;1] and β= (1-$\boldsymbol{\alpha}$). The classification accuracy is more vital than subset length.

After initialization of bat population is performed and each candidate solution is evaluate, BA starts generate new solutions according to equations (5,6,7). With a probability range of pulse rate ri, each new bat location is updated using a local search strategy around the current selected best solutions. As already stated, the probability of accepting the new solution over the current solution depend on loudness Ai and if the new solution is better the current solution. It is obvious, that BA algorithm is controlled by two parameters: the pulse rate ri and the loudness rate Ai. Typically, the rate of pulse emission ri increases and the loudness Ai decreases when the population become near to the local optimum.

6. EXPERIMENTAL SETUP & RESULTS

The implementation of filter and wrapper approaches was programmed using two programming languages (i.e., java and matlab). In filter approach, MRMR was implemented using matlab whereas in wrapper approach, BA and SVM were implemented using java.

The SVM used in this approach is based on the one prepared in libsvm [30]. RBF kernel was assign for SVM classifier. Furthermore, grid search algorithm was running to tuning the parameter of SVM classifier.

In this study, we tested the proposed MRMR-BA method by comparing it with Genetic algorithm. MRMR was carried out to filter-out the genes. The highest 50 Genes based on MRMR scores form a search space to BA and GA. An SVM with 10-cross-validation is applied to validate and assess each candidate gene subset generated from both algorithms.

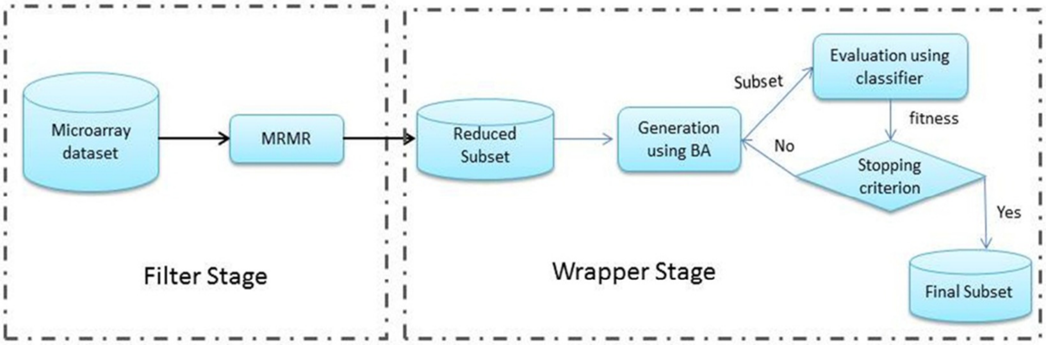

Figure 3: Framework of the Proposed Method For Gene Selection.

The 10-cross- validation method was performed over each dataset. The dataset is partitioned into (90%) for training set and (10%)for testing set. Thus, SVM build a model based on selected genes and test it on unseen data. This process is repeated 10 times for statistical validations. Both methods (BA and GA) are evaluated base on two measurements: the classification accuracy and the predictive gene subset length.

6.1 Dataset

To evaluate the proposed MRMR-BA approach, we carried out our experiments using 3 datasets of gene expression profile. These datasets can be freely downloaded from http://csse.szu.edu.cn/staff/zhuzx/Datasets.html. The datasets and their characteristics are summarized in Table 1.

Table 1: Dataset Information

Datasets	#Classes	# Samples	#Genes
Colon	2	2000	62
Breast	2	24481	97
Ovarian	2	15154	253

6.2 Parameter Setting

We examine the effect of BA on three microarray datasets with a different dimensionality i.e., a large dataset with 24481 genes (Breast), a medium dataset with 15154 genes (Ovarian) and a small dataset with 2000 genes (Colon). The parameter settings of BA for Gene selection problem is listed in Table 2. The parameters of the proposed algorithm were selected based on our preliminary experiments. They provide a good trade-off between solution quality and the computational time needed to reach good quality solutions.

Our preliminary tests show that increasing number of both iteration and population has impact of the algorithm performance, but the computational time is increased. We found the choosing of appropriate values for pulse rate and loudness lead to generate a good quality solution. Furthermore, the parameter values of GA are also determined as in [31]. Table 2 also shows the parameter values for the proposed algorithm and GA.

6.3 Results and Discussion

In this research, we re-implement mRMR with GA in order to conduct a fair comparison compare with mRMR-BA method. Moreover, we set the stopping criterion to be maximum number of iterations and we executed 30 independent runs. The average of both classification accuracy (ACC) and gene subset length (#G) were computed for all runs for both MRMR-BA and MRMR-GA. The results are presented in Table 3.

Algorithm	Parameter	Experimental Range	Selected Value
BA	Number of iterations	50-100	100
	Number of artificial bats	50-100	100
	Fmin	0-1	0.3
	Fmax	1-2	1
	A	0-1	0.5
	r	0-1	0.5
	α	0-1	0.9
	γ	0-1	0.9
GA	Population size	-	50
	Number of generations	-	50
	Crossover probability	-	0.6
	Mutation rate	-	0.5

Table 2: parameter setting of BA and GA.

From Table 3, the results show that the performance of MRMR-BA is superior in terms of average classification accuracy and average gene subset size for all dataset comparing with MRMR-GA. In respect to Breast dataset, MRMR-BA outperforms MRMR-GA with higher classification accuracy and lowest gene subset length. In Ovarian dataset, MRMR-BA has achieved 100 classification accuracy with only 3.83 averages of selected genes.

In contrast, the MRMR-GA obtained 100 classification accuracy with 19.35 average of selected genes. For Colon dataset, the result of MRMR-BA is quiet higher compared with MRMR-GA, as the former obtained 93.12 classification accuracy with only 8.13 average of selected genes. Whereas, MRMR-GA obtained 86.79 with 24.142 average of selected genes. This positive results return to search characteristics of BA algorithm in combining global diverse exploration and local intensive exploitation that result in boost it performance in selecting most informative genes with respect to high classification accuracy.

Datasets	M	MRMR-BA	MRMR-GA
Breast	|#G|	18.3	23.86
	ACC	88.6	86.606
Ovarian	|#G|	3.83	19.35
	ACC	100	100
Colon	|#G|	8.13	24.142
	ACC	93.12	86.79

Table 3: Results of comparison between MRMR-BA and MRMR-GA.

7. CONCLUSION AND FUTURE WORK

In this paper, a new approach proposed to solve gene selection problem which combined MRMR, BA, and SVM classifier. This approach is a hybrid filter-wrapper approach. MRMR filter approach run in the beginning to figure out the best genes. This process is to fine-tune the search space. Then, the reduced set of genes generated from MRMR represent solution dimension for BA. Each candidate gene subset is evaluated by SVM Classifier. Three cancer datasets were used to test the performance of the proposed approach. Furthermore, a comparison with MRMR-GA shows that our proposed approach achieved higher classification accuracy with less number of genes. The results exhibit that MRMR-BA is a promising approach for solving gene selection problem.

In the future work, experimental results on more real and benchmark datasets to verify and extend this proposed algorithm. Moreover, an enhancement to BA can be done by hybridize with existing local search algorithm to empower exploitation process that result in improve the performance of BA.

REFERENCES:

[1] A. El Akadi, A. Amine, A. El Ouardighi, D. Aboutajdine, A two-stage gene selection scheme utilizing mrmr filter and ga wrapper, Knowledge and Information Systems 26 (3) (2011) 487–500.

[2] P. P. Panigrahi, T. R. Singh, Computational studies on alzheimers disease associated pathways and regulatory patterns using microarray gene expression and network data: Revealed association with aging and other diseases, Journal of theoretical biology 334 (2013) 109–121.

[3] S. M. Yoo, J. H. Choi, S. Y. Lee, N. C. Yoo, J. Choi, N. Yoo, Applications of dna microarray in disease diagnostics.

[4] K.-H. Chen, K.-J. Wang, K.-M. Wang, M.-A. Angelia, Applying particle swarm optimization-based decision tree classifier for cancer classification on gene expression data, Applied Soft Computing 24 (2014) 773–780.

[5] T. R. Golub, D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R. Downing,

M. A. Caligiuri, et al., Molecular classification of cancer: class discovery and class prediction by gene expression monitoring, science 286 (5439) (1999) 531–537.

[6] A. Jain, D. Zongker, Feature selection: Evaluation, application, and small sample performance, Pattern Analysis and Machine Intelligence, IEEE Transactions on 19 (2) (1997) 153–158.

[7] R. Kohavi, G. H. John, Wrappers for feature subset selection, Artificial intelligence 97 (1) (1997) 273–324.

[8] B.-Q. Li, L.-L. Hu, L. Chen, K.-Y. Feng, Y.-D. Cai, K.-C. Chou, Prediction of protein domain with mrmr feature selection and analysis, PLoS One 7 (6) (2012) e39308.

[9] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, The Journal of Machine Learning Research 3 (2003) 1157–1182.

[10] P. Bermejo, J. A. G´amez, J. M. Puerta, A grasp algorithm for fast hybrid (filter-wrapper) feature subset selection in high-dimensional datasets, Pattern Recognition Letters 32 (5) (2011) 701–711.

[11] S. C. Yusta, Different metaheuristic strategies to solve the feature selection problem, Pattern Recognition Letters 30 (5) (2009) 525–534.

[12] H. M. Alshamlan, G. H. Badr, Y. A. Alohali, Genetic bee colony (gbc) algorithm: a new gene selection method for microarray cancer classification, Computational biology and chemistry 56 (2015) 49–60.

[13] S. S. Shreem, S. Abdullah, M. Z. A. Nazri, Hybridising harmony search with a markov blanket for gene selection problems, Information Sciences 258 (2014) 108–121.

[14] B. Xue, M. Zhang, W. N. Browne, Particle swarm optimization for feature selection in classification: A multi-objective approach, IEEE Transactions on Cybernetics 43 (6) (2013) 1656–1671. doi:10.1109/TSMCB.2012.2227469.

[15] X.-S. Yang, A new metaheuristic bat-inspired algorithm, in: Nature inspired cooperative strategies for optimization (NICSO 2010), Springer, 2010, pp. 65–74.

[16] X.-S. Yang, A. Hossein Gandomi, Bat algorithm: a novel approach for global engineering optimization, Engineering Computations 29 (5) (2012) 464–483.

[17] B. Ramesh, V. C. J. Mohan, V. V. Reddy, Application of bat algorithm for combined economic load and emission dispatch, Int. J. of Electricl Engineering and Telecommunications 2 (1) (2013) 1–9.

[18] P. Musikapun, P. Pongcharoen, Solving multi-stage multi-machine multi-product scheduling problem using bat algorithm, in: 2nd international conference on management and artificial intelligence, Vol. 35, IACSIT Press Singapore, 2012, pp. 98–102.

[19] J.-H. Lin, C.-W. Chou, C.-H. Yang, H.-L. Tsai, et al., A chaotic levy flight bat algorithm for parameter estimation in nonlinear dynamic biological systems, source: Journal of Computer and Information Technology 2 (2) (2012) 56–63.

[20] S. Mishra, K. Shaw, D. Mishra, A new meta-heuristic bat inspired classification approach for microarray data, Procedia Technology 4 (2012) 802–806.

[21] G. Komarasamy, A. Wahi, An optimized k-means clustering technique using bat algorithm, European Journal of Scientific Research 84 (2) (2012) 26–273.

[22] S. Akhtar, A. Ahmad, E. Abdel-Rahman, A metaheuristic bat-inspired algorithm for full body human pose estimation, in: Computer and Robot Vision (CRV), 2012 Ninth Conference on, IEEE, 2012, pp. 369–375.

[23] B. Duval, J.-K. Hao, J. C. Hernandez Hernandez, A memetic algorithm for gene selection and molecular classification of cancer, in: Proceedings of the 11th Annual conference on Genetic and evolutionary computation, ACM, 2009, pp. 201–208.

[24] M. Dash, H. Liu, Feature selection for classification, Intelligent data analysis 1 (3) (1997) 131–156.

[25] E. Amaldi, V. Kann, On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems, Theoretical Computer Science 209 (1) (1998) 237–260.

[26] Y. Cai, T. Huang, L. Hu, X. Shi, L. Xie, Y. Li, Prediction of lysine ubiquitination with mrmr feature selection and analysis, Amino acids 42 (4) (2012) 1387–1395.

[27] B. S¸en, M. Peker, A. C¸ avu¸so˘glu, F. V. C¸ elebi, A comparative study on classification of sleep stage based on eeg signals using feature selection and classification algorithms, Journal of medical systems 38 (3) (2014) 1–21.

[28] R. Y. Nakamura, L. A. Pereira, K. Costa, D. Rodrigues, J. P. Papa, X.-S. Yang, Bba: a binary bat algorithm for feature selection, in: Graphics, Patterns and Images (SIBGRAPI), 2012 25th SIBGRAPI Conference on, IEEE, 2012, pp. 291–297.

[29] V. N. Vapnik, An overview of statistical learning theory, Neural Networks, IEEE Transactions on 10 (5) (1999) 988–999. [30] C.-C. Chang, C.-J. Lin, Libsvm: a library for support vector machines, ACM Transactions on Intelligent Systems and Technology (TIST) 2 (3) (2011) 27.

[31] Z. Zhu, Y.-S. Ong, M. Dash, Markov blanket-embedded genetic algorithm for gene selection, Pattern Recognition 40 (11) (2007) 3236–3248.