Binary Harmony Search Based Feature Selection and Data Classification Technique

A. Swatismita Das[1], B. Sudeep Padhi[2]
[1,2] *Swatismita Das, Gandhi Institute of Technology and Management*

***Abstract*- Nowadays data is rapidly growing at an exponential pace. To deal with such data explosion, we need effective data processing and analysis techniques. A most popular machine learning technique namely "feature selection" that plays the most vital role of machine learning technique where it selecting a subset of features from a dataset that still provides most of the useful information. In real-world applications misclassification costs of minority class could be extremely high. Therefore this is the most demanding query especially while the data are too high in dimensionality because of enhance in over-fitting and inferior representation interpretability. Recently Feature selection is one of the most popular way to deal with the trouble by figuring out the features that will be predict best to a minority class.**

***Index terms*- Binary Harmony Search, Feature Selection, Harmony Search Algorithm**

I.INTRODUCTION

We live in a world, where vast amount of data are collected daily. Nowadays data played a vital role in every economy, business function, organization, industry, medical, computer science, engineering, energy and individuals [1]. To analyze such huge data is most important need. Real-world data are generally stored on many other different platforms in distributed computing environments together such as individual systems, database, otherwise even on the internet system and it also heterogeneous, incomplete and noisy due to their typical huge size. So it is truly very complicated to handle all these different kind of information and extracting the useful documentation. There are many data mining issues such as noisy and incomplete data, performance, complex data, data visualization, distributed data, incorporation of background knowledge, pattern evaluation, data privacy and security and so on [2].So that more challenges get exposed since the actual data mining process begins and the success of data mining lies in overcoming all those majority issues or problems.

Optimization is defined as choosing the "best element" from some set of available alternatives. In short the term optimize is to make perfect. For compound optimization issues, evolutionary computation approaches are becoming most famous techniques to defeat over all those complications, for example Genetic Algorithm (GA) [3], Evolution Strategy (ES) [6], Harmony Search (HS) [10], Particle Swarm Optimization (PSO) [12], Ant Colony Optimization (ACO) [16], Tabu Search (TS) [18], Artificial Bee Colony (ABC) [19], Differential Evolution (DE) [20], Simulated Annealing (SA) [21], Firefly Algorithm (FA) [22] and so on.

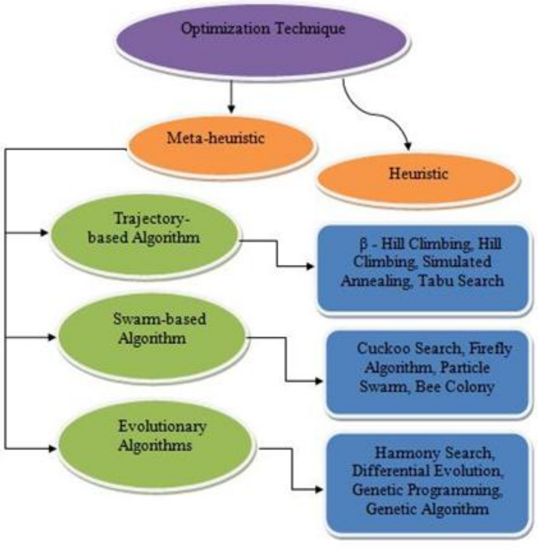


Fig.1.Optimization Algorithms

These techniques are generally motivated from biological judgment or in natural evolution mechanisms and that is very simple to implement and very effective to be used in various optimization issues.

In this research work optimization method which is logical in the real world is used. Harmony search

(HS) is a newly initiate population based meta-heuristic algorithm which imitates the improvisation procedure of musical performers [10]. HS algorithm have extremely fortunate in an extensive variations of manufacturing optimization problem and machine learning [9] tasks. It has demonstrated many advantages more than conventional optimization approaches. Harmony Search requires only some deficient arithmetical necessities and isn't sensitive to the primary rate settings. Even though it's a new population-based method, HS mechanism to produce a new vector which can be encodes a candidate result, later than that making an allowance for a collection of presented quality vectors.

II. FEATURE SELECTION

Feature selection (FS) is the term regularly used in data mining to explain the tools and strategies to be had for reducing the inputs to a convenient size of processing and to study. Lots of search procedures have been used for feature selection (FS) [23] tasks, within the attempt to recognize extra compressed and enhanced feature subsets. That determine the most favorable solution without departing through a complete search, such situation usually concern the exercise of nature-inspired heuristics, or greedy hill climbing (HC). Main motivation of Feature selection (FS) is to resolve a least feature subset as of the trouble area as retaining a properly high precision on behalf of the actual data. Feature selection has four essential steps in a typical feature selection process given below.

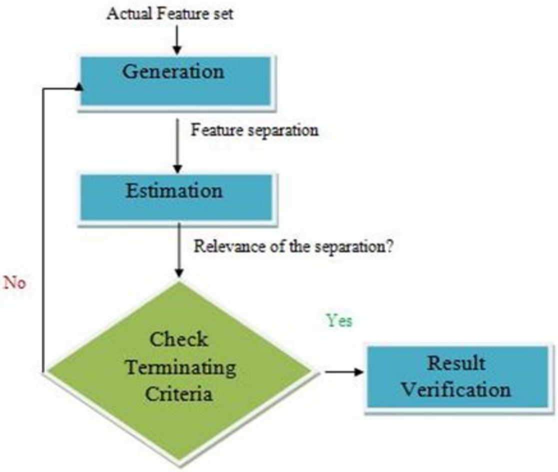
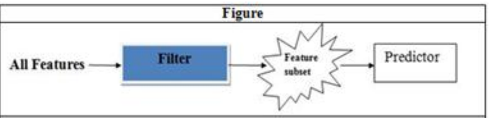
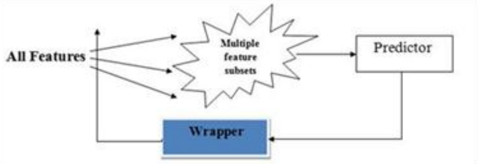
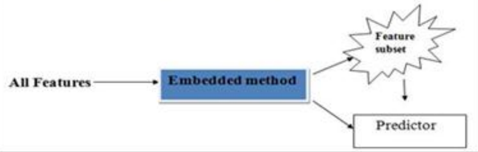


Fig.5. General Representation of feature selection process

This project work introduce a feature selection [24] method based on Harmony Search is to invent a least feature detachment from a problem field even as maintaining a properly excessive precision on behalf of the actual information [25]. Realistic problems that stand up while interpreting the information in real-world applications are frequently associated with the quantity of features, and the lack of aptitude to recognize and extract patterns or policy simply because of excessive interdependence between individual features. Individual assessment and successive pattern recognition are constrained while allowing for data sets that contain extremely huge amount of features. There are two main approaches for feature selection (FS), (a) Filter and (b) Wrapper and (c) Embedded methods.

Table 1 Different approaches of feature selection.

Method	Figure
Filter	
Wrapper	
Embedded	

Filter approaches are normally used as a preprocessing step. Here feature selection process is independent of any machine learning techniques. Instead of that features are also chosen on the basis of their scores in numerous statistical tests for their correlation with the outcome variable. Whereas Wrapper approaches, in which the features are chosen by filter methods and classifier, in which the selection of features is independent of the classifier used. Embedded approach, which attempt to mutually or concurrently guide both the classifier and a feature subset.

III. BACKGROUND OF FEATURE SELECTION METHOD

There are many evolutionary techniques that have been recommended in the literatures for feature selection in high dimensional dataset. These area are previously loaded along by the application of numerous evolutionary based techniques for feature selection models expansion similar to genetic algorithm (GA) [3], harmony search (HS) [10], particle swarm optimization (PSO) [12], differential evolution(DE) [20], artificial bee colony algorithm (ABC) [19], ant colony optimization (ACO) [16] etc.

Table 2 Variants of evolutionary framework applies in feature selection.

Sl. no	Method	Variation	Application areas & Reference
01	GA	Dynamic Genetic Algorithm	Vibration Rotor System [3]
02	GA	Hybrid Genetic Algorithm	Medical image of few cancer data [4]
03	GA	Enhanced Genetic Algorithm	Arabic text datasets [5]
04	ACO	Hybridised Ant Colony Optimization	Microarray Databases [17]
05	DE	Multi objective Differential Evolution	Benchmark Databases [31]
06	ABC	Artificial Bee Colony with Gene Recombination Operator	Microarray Data [33]
07	ABC	Hybridised Artificial Bee Colony	15 no. of Datasets from UCI Repository [32]
08	PSO	Improved Particle Swarm Optimization	Text Feature Selection [13]
09	PSO	Binary Particle Swarm Optimization	Benchmark Databases [34]
10	HS	Improved Harmony Search Algorithm	Pressure Vessel Design [8], Welded Beam Design [10]
11	HS	Adaptive Harmony Search	Microarray Databases [7]

In précis, only some of the evolutionary techniques along with their permutations are high-lighted in table 2. These approaches have been applying intend for designing an well-organized model for feature selection in various application fields.

IV. DESIGN AND IMPLEMENTATION

A. Harmony Search

Harmony search (HS) is a currently develop a population based meta-heuristic algorithm which imitate the improvisation procedure of musical performer. It imitates the improvisation process of players throughout which every performer plays a note for discover a best harmony simultaneously. Harmony search algorithm has marvelous piece in worldwide optimization issues, but it also very difficult in conducting the local search for numerical applications. Lots of studies have been conduct and various methods have been recommended to modify the basic HS algorithm to increase the performance of HS algorithm.

B. Basic Elements of Harmony Search Algorithm

Harmony: Harmony is the set of values of all the variables of the objective function.

Harmony Memory (HM): The spaces where harmonies are store.

Harmony Memory Size (HMS): The no. of spaces that harmony memories have.

Pitch Adjustment Rate (PAR): It is the rate of choosing neighboring value.

Number of Iteration (NI): It defines the stopping criterion.

C. Flow of Harmony Search Algorithm

Step 1: Initialization of the optimization problem definition and algorithm parameters.

Step 2: Initialization of the Harmony Memory. Generally Harmony Memory (HM) is contained with the random values as many as the Harmony Memory Size.

Step 3: Invention of a new harmony.

Step 4: Updating the Harmony Memory. If the new harmony is better than existing harmony, replace the worst one.

Step 5: Check the stopping criteria.

D. Harmony Search Algorithm

1. Initialization of Harmony Memory (HM)

Initial population Harmony Memory (HM) contain of Harmony Memory Size (HMS) vectors is produced randomly, where xi = xij , i = 1, …,HMS and j = 1, …, n. The HM matrix is filled with HMS vectors as follows:

$$HM = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{HMS1} & x_{HMS2} & x_{HMS3} & \cdots & x_{HMSn} \end{bmatrix} \quad (1)$$

2. Improvisation of New Harmony Vectors

2.1 Harmony Memory Consideration Rule

For this rule, a new random number r1 is generated within the range [0,1]. If r1 < HMCR, where HMCR is the harmony memory consideration rate, then the first decision variable in the new vector xij{new} is chosen randomly from the values in the current HM as follows:

$$x_{ij}^{new} = x_{ij}, x_{ij} \in \{x_{1j}, x_{2j}, x_{3j}, \dots\dots, x_{HMSj}\} \quad (2)$$

2.2 Pitch Adjustment Rule

The acquired selection variables from the harmony memory consideration rule is moreover investigated to discovered if it needs to pitch adjustment or not.

A latest random number r2 is produced within the range in between [0 1]. If ⟦rand⟧_2< PAR, where PAR is a pitch adjustment rate, then the pitch adjustment decision variable is calculated as follows:

$$x_{ij}^{new} = x_{ij}, \pm rand\,(0,1).BW \quad (3)$$

Where, BW is the bandwidth factor, which is used to control the local search just about the preferred decision variable in the new vector.

2.3 Randomization Rule

If the condition ⟦rand⟧_2< HMCR fails, then the new first decision variable in the new vector x ij{new} is produced arbitrarily as follows:

$$x_{ij}^{new} = l_{ij} + (u_{ij} - l_{ij}).rand\,(0,1) \quad (4)$$

Where, “l” is the lower and “u” is the upper bound for the given problem.

3. Harmony Memory Updation Rule

Once the harmony vector x {new} is generated, then it will swap the worst harmony vector x {worst} in the harmony memory if its objective function value is better than the worst harmony vector.

```
        if $(x^{new} < x^{worst})$ then
    Update the HM as $x^{worst} = x^{new}$
end if
```

V. PREVIOUS WORK ON HARMONY SEARCH

The main objective behind the execution of harmony search is, several application areas are there where harmony search is fruitfully applied. So, at this point we have emphasized some of them.

Hai-bin Ouyang a et al. In 2016 [8] proposed a new version of harmony search, namely LHS algorithm, for resolving numerical optimization problems. Briefly, LHS algorithm develops to be a better option to optimize several demanding and multidimensional real-world issues.

This is the recent paper of Rasmita Dash, (2018) [7] in this estimation high dimensional data classification problem, they have suggested a hybridised harmony search designed for feature selection. Here arithmetical investigation is supervised to demonstrate the superiority of proposed work through two other environment stimulated algorithms. Simulation product shows that the suggested hybridization provided excessive prospective in both sample classification and feature subset prediction potentiality for high dimensional databases.

VI. PROPOSED WORK

In our proposed work I have different datasets for high dimensional data. Here I have taken two types of datasets; first one is Tumors and Ap_Breast_Colon datasets. Then using min_max normalization, I have generated to the data within the range between [0-1]. Number of features is very high in high dimensional data. So, in first stage data reduction we are applying a ranking technique to extract the highly significant features. Where we are assume that higher value are the significant one. After first stage data reduction extend to some part, KNN classifier is applied. But still some insignificant values are left. So, to remove the insignificant value we have done the second stage reduction. Normally we need many samples or much data to make our model learn. Here I have taken 30% samples as testing data and 70% samples as traing data. If the model will be learned properly, then in future it can set any kind of input or it can recognized any kind of input. Here KNN classifier is applied because for high dimensional data classification KNN is a good classifier and it is very easy to implement.

A.Signal To Noise Ratio(STNR)

Here this process, rank of every gene be resolved, calculating the level of signal with level of noise. Subsequently the excessive STNR esteem speaks to the high need quality, or exceedingly noteworthy quality and some of these qualities are picked according to the determination strategy. The STNR methods are defined below:

$$STNR_{method1}(i) = \left|\frac{mean_D - mean_N}{std_D - std_N}\right| \quad (6)$$

$$STNR_{method2}(i) = \frac{(mean_D - mean_N)^2}{std_D{}^2 + std_N{}^2} \tag{7}$$

Where, mean_D and mean_N = mean of the sample of class1 for Healthy and class2 for Diseased. std_D and std_N = standard deviations of the samples of class1 for Healthy and class2 for Diseased. i = current genes under consideration.

B.Binary - Coding Harmony Search Algorithm

The HS with binary coding can be utilized to manage binary coded issues, where the pitch adjustment operator is lower in binary space, which damages the performance of the algorithm. To neutralize the issues of HS for binary-valued problems is analyzed and study and an adaptive binary HS calculation with a novel pitch alteration administrator are prescribed to enhance the advancement capacity.

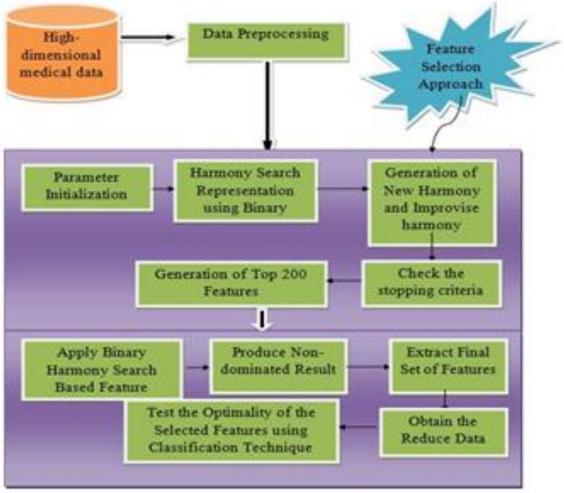


Fig:4: An arrangement of present feature selection and classification approach.

The details step wise description of binary harmony search algorithm is along these steps:

Step 1: Collection of high dimensional medical data or information.

Step 2: Accomplish data pre-processing by utilizing min-max normalization standardization strategy.

Step 3: Extend a preliminary feature selection technique by utilizing the binary harmony search algorithm to deal with pick best n features from the origin data.

Step 3.1: Initialization of algorithm parameters of binary harmony search like harmony memory (HM), pitch adjustment rate (PAR), harmony memory consideration rate (HMCR), bandwidth (BW) and number of iteration (NI).

Step 3.2: Generation of harmony memory (HM).

Step 3.3: Randomly initialize the vector within the range between [0-1] contained by the value of 1 to number of features in the dataset.

Step 3.4: locate the number of improvisation (NI) to zero.

Step 3.5: For every vector, discover the fitness function by using KNN classifier taking into consideration classification error as the fitness function.

Step 3.6: Generation of new solution vector is as follows

Step 3.6.1: Produce two random number $rand_1$ and $rand_2$ within the range between [0,1]

Step 3.6.2: If $rand_1 \leq$ HMCR, update the harmony vector with respect to equation-2.

Step 3.6.3: If $rand_2 \leq$ PAR, update the harmony vector with respect to equation-3.

Step 3.6.4: If not, than update the harmony vector with respect to equation-4.

Step 3.7: If fitness of new harmony is better than worst harmony, in that case upgrade the new harmony.

Step 3.8: Addition it to 1.

Step 3.9: Updation of HMCR, PAR and BW respecting the above equations.

Step 3.10: Imitate Steps 3.5 to 3.9 until NI is fulfilled.

Step 4: Generation of the best harmony with the highest fitness value and study the nominated index of n features to succeed the minimized feature subset.

Step 5**:** Check the stopping criteria.

VII. EXPERIMENTAL ANALYSIS

A. Datasets Used for Experimental Analysis

Complete research is supervised on the high dimensional medical data. Where number of features value is much higher than the number of samples. So, the illustrations of the databases are obtained in Table.3 which is signify for binary characterization issue.

Gene Expression Datasets

The complete experiment is supervised on two high dimensional gene expression data. Numbers of gene values are very expensive as compared to number of observation.

Tumors: It's an abnormal collection of tissue. Tumors are a classic sign of inflammation, and can be being or malignant (cancerous). The total number of genes to be tested is 7129 and number of samples is 218.

Ap_Breast_Colon: It contain 460 samples among them 240 are tumor biopsies which is labelled as negative and 220 are normal which is labelled as positive with total number of 10051 genes.

Table 3

Number of feature selected in the final stage.

Sl.no	Dataset	Genes	Samples	Class Level	
				Class1	Class0
01	Tumors	7129	218	3998	3131
02	Ap_Breast_Colon	10051	460	4064	5987

B. Data Normalization and Cross Vaslidation

Normally these data are accumulated from diverse sources._The mixture of feature cost misguides our goal while a training algorithm is being applied. Thus the feature values are necessary to be leveled within a particular altitude. This progression is called as "data normalization". In our experimental analysis we have used the min-max normalization where the feature values are leveled within the range between 0 to 10 as shown below.

$$Norm_{value} = \frac{Org_{value} - (org_{value} in\ series)}{(org_{value} in\ series) - (org\_value\ in\ series)} \quad (8)$$

There are many classifiers are there, but our experiment focuses on k-Nearest Neighbor (KNN). Where KNN classifier is supposed to be an easily implemented classifier where simply a single parameter (k value = no. of neighbor) is essential to optimized. Just as this approach, the entire test instance shift to the nearest k neighbor classes regarding the adjacent distance (we are calculating the Euclidian distance).

C. Matrices Evaluation

Here we discussed about the execution metrics examined for estimate of our case study. There are four possible combinations of outcomes in the ultimate period of classification that the actual outcome is compare with the target outcome.

1. True positive (TP): Disease instance perfectly classify
2. True negative (TN): Healthy instance perfectly identify
3. False positive (FP): Disease instance imperfectly classify
4. False negative (FN): Healthy instance imperfectly classify

Derived from those above consequences here we evaluate the performance criteria as shown below.

$$Accuracy = \frac{TP+TN}{R} \times 100 \quad (9)$$

$$Sensitivity = \left(\frac{TP}{TP+FP}\right) \times 100 \quad (10)$$

$$Specificity = \left(\frac{TN}{TN+FN}\right) \times 100 \quad (11)$$

Where, R contains the total number of samples.

D. Result Analysis

Table 4 Number of feature selected in the final stage.

Dataset	Tumor	Ap_Breast_Colon
Number of genes	12	8

Table 5 Performance of proposed model using KNN on Tumor dataset.

Method	Accuracy	Sensitivity	Specificity
Binary HS	90.00	80.95	94.87
HS using $STNR_1$	86.67	76.19	97.43
HS using $STNR_2$	91.66	95.23	94.87

Table 6 Performance of proposed model using KNN on Ap_Colon_Breast dataset.

Method	Accuracy	Sensitivity	Specificity
Binary HS	93.25	81.95	99.87
HS using $STNR_1$	83.97	72.18	98.69
HS using $STNR_2$	93.63	96.23	98.18

VIII. CONCLUSION

This research work has described flexible feature selection method based on binary-harmony search where harmony memory be initially packed by means of arbitrarily produce feature subsets. It focuses in new feature selection technique by join harmony search as well as binary harmony search for high dimensional data classification. Feature selection method used to produce top 200 features. That separation might be containing mutually significant features over and above only some insignificant features that further gene diminution is done during binary harmony search technique. Since the occurrence of relevant genes minimizes the classification fault, so it is used as fitness function for feature evaluation. Presentation of projected model is evaluated with esteem to KNN classifier over two different publicly available microarray data.

REFERENCES

[1] Manjarres, D., Landa-Torres, I., Gil-Lopez, S., Del Ser, J., Bilbao, M. N., Salcedo-Sanz, S., & Geem, Z. W. (2013). A survey on applications of the harmony search algorithm. Engineering Applications of Artificial Intelligence, 26(8), 1818-1831.

[2] Sharma, B. R., Kaur, D., & Manju, A. (2013). Review on Data Mining: Its Challenges, Issues and Applications. International Journal of Current Engineering and Technology, 3(2).

[3] Lu, L., Yan, J., & Meng, Y. (2016). Dynamic Genetic Algorithm-based Feature Selection Scheme for Machine Health Prognostics. Procedia CIRP, 56, 316-320.

[4] Nagarajan, G., Minu, R. I., Muthukumar, B., Vedanarayanan, V., & Sundarsingh, S. D. (2016). Hybrid Genetic Algorithm for Medical Image Feature Extraction and Selection. Procedia Computer Science, 85, 455-462.

[5] Ghareb, A. S., Bakar, A. A., & Hamdan, A. R. (2016). Hybrid feature selection based on enhanced genetic algorithm for text categorization. Expert Systems with Applications, 49, 31-47.

[6] Back, T., Hoffmeister, F., & Schwefel, H. P. (1991, July). A survey of evolution strategies. In Proceedings of the fourth international conference on genetic algorithms (Vol. 2, No. 9). Morgan Kaufmann Publishers San Mateo, CA.

[7] Dash, R. (2018). An Adaptive Harmony Search Approach for Gene Selection and Classification oh High Dimensional Medical Data. Journal of King Saud University-Computer and Information Sciences.

[8] Ouyang, H. B., Gao, L. Q., Li, S., Kong, X. Y., Wang, Q., & Zou, D. X. (2017). Improved Harmony Search Algorithm: LHS. Applied Soft Computing,53,133-167.

[9] El-Shafei, M., Ahmad, I., & Alfailakawi, M. G. (2016). Implementation of harmony search on embedded platform. Microprocessors and Microsystems, 45, 187-197.

[10] Geem ZW (2008) Harmony search for solving Sudoku, Knowledge-Based Intelligent Information and Engineering Systems, LNAI, 4692:371-378.

[11] Kumar, V., Chhabra, J. K., & Kumar, D. (2012). Effect of harmony search parameters' variation in clustering. Procedia technology, 6, 265-274.

[12] Kar, S., Sharma, K. D., & Maitra, M. (2015). Gene selection from microarray gene expression data for classification of cancer subgroups employing PSO and adaptive Knearest neighborhood technique. Expert Systems with Applications, 42(1), 612-627.

[13] Liu, Y., Wang, G., Chen, H., Dong, H., Zhu, X., & Wang, S. (2011). An improved particle swarm optimization for feature selection. Journal of Bionic Engineering, 8(2), 191-200.

[14] Gunasundari, S., Janakiraman, S., & Meenambal, S. (2016). Velocity bounded Boolean particle swarm optimization for improved feature selection in liver and kidney disease diagnosis. Expert Systems with Applications, 56, 28-47

[15] Lu, Y., Liang, M., Ye, Z., & Cao, L. (2015). Improved particle swarm optimization algorithm and its application in text feature selection. Applied Soft Computing, 35, 629-636

[16] Dadaneh, B. Z., Markid, H. Y., & Zakerolhosseini, A. (2016). Unsupervised probabilistic feature selection using ant colony optimization. Expert Systems with Applications, 53, 27-42.

[17] Sharbaf, F. V., Mosafer, S., & Moattar, M. H. (2016). A hybrid gene selection approach for microarray data classification using cellular learning automata and ant colony optimization. Genomics, 107(6), 231-238.

[18] Glover, F., & Laguna, M. (1997). Tabu search. Kluwer Academic Publishers.

[19] Shunmugapriya, P., & Kanmani, S. (2017). A hybrid algorithm using ant and bee colony optimization for feature selection and

classification (AC-ABC Hybrid). Swarm and Evolutionary Computation.

[20] Khushaba, R. N., Al-Ani, A., & Al-Jumaily, A. (2011). Feature subset selection using differential evolution and a statistical repair mechanism. Expert Systems with Applications, 38(9), 11515-11526.

[21] Laarhoven, P. J. M., & Aarts, E. H. L. (1988). Simulated annealing: Theory and applications. Kluwer Academic Publishers.

[22] Yang, X. S., & He, X. (2013). Firefly algorithm: recent advances and applications. International Journal of Swarm Intelligence, 1(1), 36-50.

[23] Boulesteix, A. L., Strobl, C., Augustin, T., & Daumer, M. (2008). Evaluating microarray-based classifiers: an overview. Cancer Informatics, 6

[24] E.P.Xing."Feature Selection in Microarray Analysis. A Practical Approach to Microarray Data Analysis", Kluwer Academic Publishers. 2003.

[25] B.S.Everitt. "An Introduction to Latent Variable Models, Monographs on Statistics and Applied Probability", Chapman & Hall, London. 1984.

[26] M.A. Carreira-Perpinan. "Continuous latent variable models for dimensionality reduction and sequential data reconstruction". PhDthesis, University of Sheffield, UK. 2001.

[27] Q. Shen and A. Chouchoulas. "A fuzzy-rough approach for generating classification Rules". Pattern Recognition, Vol. 35, No. 11, pp. 341–354. 2002.

[28] R. E. Bellman, "Dynamic Programming". Princeton, NJ: Princeton Univ. Press, 1957.

[29] M. Dash, K. Choi, P. Scheuermann and H. Liu. "Feature Selection for Clustering – A Filter Solution". In Proceedings of IEEE International Conference on Data Mining (ICDM), pp. 115–122. 2002.

[30] M. Shah, M. Marchand, and J. Corbeil, "Feature selection with conjunction of decision stumps and learning from microarray data," IEEE Trans.Pattern Anal. Mach. Intell., vol. 34, no.1, pp. 174–186, Jan. 2012

[31] Bhadra, T., & Bandyopadhyay, S. (2015). Unsupervised feature selection using an improved version of differential evolution. Expert Systems with Applications, 42(8), 4042-4053

[32] Zorarpacı, E., & Özel, S. A. (2016). A hybrid approach of differential evolution and artificial bee colony for feature selection. Expert Systems with Applications, 62, 91-103.

[33] Li, G., Cui, L., Fu, X., Wen, Z., Lu, N., & Lu, J. (2017). Artificial bee colony algorithm with gene recombination for numerical function optimization. Applied Soft Computing, 52,146-159.

[34] Chuang, L. Y., Yang, C. H., & Li, J. C. (2011). Chaotic maps based on binary particle swarm optimization for feature selection. Applied Soft Computing, 11(1), 239-248