

Optimal and Novel Hybrid Feature Selection Framework for Effective Data Classification

Sivakumar Venkataraman and Rajalakshmi Selvaraj

Abstract Data mining methods are frequently applied in the framework of data classification. Under data mining methods, feature selection (FS) algorithms are essential for dealing with various dimensional data sets that may contain features in the range of small, medium, and large dimensions. Handling large number of features always raises the issues regarding the classifier accuracy and running time. A novel hybrid feature selection technique build on symmetrical uncertainty and genetic algorithm is proposed. The experiments' results on UCI datasets using this hybrid framework proved that proposed feature selector is efficient through minimizing the volume of initial features and accurate by providing better detection performance in the classification algorithms comparing with other feature selectors in the literature. It is evident from the earlier research work the prosed method promotes in optimizing and improves the performance. In summary, the proposed feature selection method has outperformed other methods in minimizing the selected features, classification performance and reduces the executing time.

Keywords Data mining · Feature selection · Classification
Symmetrical uncertainty and genetic algorithm

1 Introduction

Data analytics researchers need relevant and high-quality data from huge amount of stored data. Feature selection method helps in reducing the dimensionality of features by removing redundant, irrelevant, or noisy data through which improvisation in classification accuracy with minimum processing of data can be achieved.

S. Venkataraman (✉)

Faculty of Computing, Botho University, Gaborone, Botswana
e-mail: sivakumar.venkataraman@bothouniversity.ac.bw

S. Venkataraman · Rajalakshmi Selvaraj (✉)

Department of Information Systems, BIUST, Gaborone, Botswana
e-mail: selvarajr@biust.ac.bw

Due to increase of dimensionality of records and features in data repositories, there is shift in maintaining the records related to each and every individual. Formally, in recent times, data mining techniques are used to discover a novel and useful pattern from the historical data. Many research ideas are openly still needed to solve this. Mostly classification framework gives an effective result for classifying the datasets. Usually, large and high dimensional datasets contain complex information with errors, and in such a situation, classification algorithm plays a vital role.

By using the feature subset selection process, the relevant subset features can be determined from the original features. The process of ranking features according to their significance in improving the performance of classifiers is called as ranking-based feature selector. Under KDD process, feature selection is only one essential step that improves the detection performance of the classifiers, minimizes the time taken to build the data mining model, and reduces the number of initial features because there are no quality features and no quality results in the classifiers.

FS methods are classified as filter and wrapper method. Based on this key idea, several FS methods are introduced in machine learning paradigm. Wrapper method is used to select the features based on the accuracy estimate, and filter method is used to select the features not based on the accuracy estimate; rather, it uses the data characteristics with the relevancy or correlation measures. Filter-based approaches are not dependent on classifiers and usually faster and more scalable than wrapper-based methods. Moreover, they have low computational complexity too. Recently, numbers of hybrid approaches are also being proposed to achieve a good balance in the feature selection criteria by combing both filter and wrapper method.

2 Related Works

In [1], the author discussed the impact of the noise in the class labels by analyzing the traditional mutual information-based filter feature selection algorithm. This proposed idea brings the solution for the nearest neighbors-based entropy estimator to minimize the class label errors. In [2] the author presents an empirical study on many feature selection and classification algorithms to analyze their performance on diverse biological datasets. This study reveals combination of RFE with SVM and LR as best generalization model to perform feature selection and classification. In [3], the author has proposed a rule-based feature selection algorithm to improve the detection performance of multiclass support vector machine. In [4], the author has applied a hybrid algorithm combines the genetic algorithm with K-nearest neighbor for predicting the binding of protein–water from its X-ray crystallographic protein structure data. In [5], the author has proposed a redundancy demoting (RD) approach for making improvement in ranking by demoting the redundant features. For instance, for diagnosing the erythemato-squamous disease with the help of feature selection based on F -score measure has been applied [6]. In [7], the author proposed genetic algorithm (GA) as commonly used global searches for optimization. This method is used in feature selection process among various

applications and has exposed to be a good tool [8–12]. A recommended way of solving this issue is to group genetic algorithm and some memetic (search) operations [13, 14]. This helps in fine-tuning the search process and improves the quality of the results generated from genetic algorithm with relation to accuracy and efficiency. Recently, these types of evolutionary algorithms are called hybrid evolutionary algorithms (EAs), memetic algorithms (MAs), Lamarckian evolutionary algorithms, Baldwinian EAs, and local search or cultural genetic algorithms. These algorithms are not used for convergence to high-quality results only, also for further efficient search mechanism [13, 14]. Selecting only the minimum relevant features from the original subset is the main challenge in feature selection. This work aims to develop ‘symmetrical uncertainty and genetic algorithm (SU-GA)’-based feature selector named universal and novel feature selector. The goal of this study is to make accurate prediction with the help of least number of significant features. In this research work, the features undergo memetic (genetic) evolution such as ‘include’ and ‘remove’ to select the features. The examined results prove the proposed SU-GA-based classifier attains significant dimensionality reduction in various dimensional datasets from the UCI machine learning repository [15].

3 Proposed Method

3.1 Symmetrical Uncertainty

The symmetrical uncertainty (SU) between features and the target concept are used to obtain the best features for classification. Features with a larger SU value obtain greater weight. SU measures the relationship among X , Y variables based on the information theory. It can be calculated as follows.

$$SU(X, Y) = 2 \frac{I(X, Y)}{H(X) + H(Y)}$$

Considering $I(X, Y)$ as the MI among X , Y . $H(\cdot)$ as an entropy function for X , Y features. The SU admits the normalized range value [0,1] as correction factor value is 2. If SU value is 1, then the information of one feature is absolutely predictable. If SU value is 0, then X , Y are not associated.

3.2 Genetic Algorithm

The amount for genetic algorithm (GA) [16] is stated randomly for all individual chromosome that encrypt the feature subsets. Each chromosome are assembled with binary string, and the binary string encrypting describes that the value ‘1’ (‘0’)

shows the particular feature is selected (omitted). The Obj_Fun (Objective Function) for subset are obtained from the aptness of individual chromosome as,

$$\text{Fitness}(c) = \text{Max}(\text{Obj_Fun}(\text{SFc}))$$

$$\text{Obj_Fun}(\text{SFc}) = \alpha * (1/\tau) + (\text{RCF} * \text{Recall}) + (\text{PCF} * \text{Precision})$$

where

- τ No. of ones in the SFc
- α No. of minimum features selected
- RCF Recall credibility factor
- PCF Precision credibility factor

Considering SFc as the SF subset arranged with Chromosome c. Obj_Fun (SFc) estimates the feature subset contribution. If Chromosomes with similar aptness value obtained, then the first priority of surviving will be provided to the less number of SF.

3.3 Proposed SU-GA Feature Selector

The proposed hybrid feature selection algorithm including feature ranking and optimal feature selection (FR-OFS) method adopts the filter method with wrapper method to attain the optimal subset features. First phase, the proposed algorithm selects very few relevant features, by computing the SU values between features and the target concept. Second phase, GA is used to search the optimal subset of features with higher accuracy obtained in the first phase with symmetrical uncertainty. The number of GA is set based on features rank rendering from the examined results from SU. Features having higher SU value will have the higher possibility, where the feature can be selected, that states the resembling bits will have more possibility to be select in the chromosome. Each individual's fitness is evaluated using GA such as the value of the result by using the fitness function. The chromosomes might subjectively have transformed by the crossover and mutation functions. This impacts the aptness value is carried out by both crossover and mutation operation. The procedure is repeated until acceptable results are obtained. The features are chosen at the end of this space called as reliable features. The proposed SU-GA feature selector is shown in Figs. 1 and 2.

Fig. 1 Proposed SUGAFS method

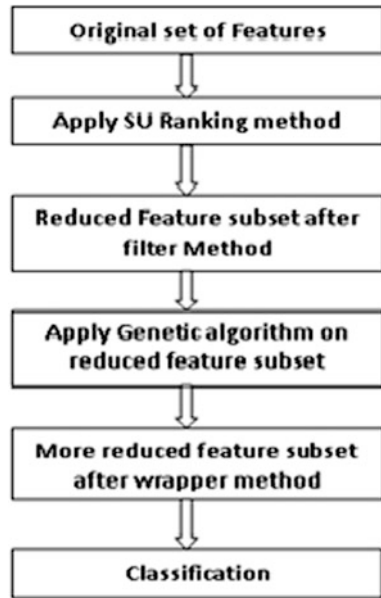


Fig. 2 Hybrid feature selection algorithm

Input: Full data set.

- Step 0. Start
- Step 1. Load Dataset
- Step 2. Compute SU value of each feature
- Step 3. Select the features with larger SU value.
- Step 4. Initialize the population P with selected features.
- Step 5. Evaluate the fitness value of the individuals.
- Step 6. While Termination criteria not Satisfied
 - {
 - Select parents for reproduction
 - Perform recombination and mutation
 - Evaluate population
 - }
- Step 7. Output the Selected feature subset.
- Step 8. Stop.

Output: Selected Feature subset.

4 System Implementation and Experimental Results

The implementation of proposed SU-GA Feature selector is on twofold. Firstly, the SU value of each feature is examined, and the features are prioritized based on the highest SU value, and those features having lesser SU are eliminated by considering

as an irrelevant and redundant feature. Both WEKA and MATLAB tool box are used to successfully implement the proposed hybrid feature selector SU-GA. In the second phase, genetic algorithm is applied on the selected feature subset by SU to find out the optimal feature with out compromising the classification accuracy. The final set of optimal features selected by the proposed hybrid feature selector SU-GA is tested with the help of various benchmark classification algorithms in the literature by using the classification accuracy, number of reduced features, and time taken to build a model by using 10-fold crossvalidation as a test method.

4.1 Experimental Setup

Totally, 10 benchmark dataset from UCI machine learning repository is chosen for carrying out the experiments on the existing and proposed feature selector in terms of number of feature selected by each method, the improvement in detection performance of classifiers by each method and time taken to build a model by each method are empirically evaluated and tested on various datasets including Soybean, Lung Cancer, Ionosphere, and Dermatology. The summary of this datasets such as the number of attributes and instances in each dataset is shown in Table 1.

Different classification algorithm, namely NB, J48, SMO, and JRIP, are applied on both original as well as number of features selected by SU, GA, and proposed SU-GA feature selector. The proposed method is implemented using WEKA and MATLAB tool box. The features having higher SU values are identified through ranker search algorithm available in WEKA. The number best-ranked features by SU was given as an input to the GA toolbox available in the MATLAB to select only the optimal feature set through eliminating both irrelevant and redundant features effectively. Finally, it varies supervised classifiers applied on both original and optimal subset of features selected by the proposed method. The classification accuracy, number of selected feature by each method, and time taken for classifiers are clearly proved that the proposed method is superior than other existing feature selector in the literature. The results are presented in Tables 3 and 4.

Table 1 Datasets for research

Dataset name	Attributes name	No. of instances
Ionosphere	35	351
Soybean	36	683
Diabetes	9	768
Segment challenge	20	1500
Vote	17	435
Dermatology	35	366
Lung cancer	57	32
Wine	14	178
Hepatitis	20	155
Vehicle	19	846

4.2 Results and Discussions

In this experiment, three widely used evaluation measures features selected; classification accuracy and processing time are adopted to evaluate the proposed method. This proposed SUGAFS algorithm and other two feature selection algorithms SU and GA are implemented in WEKA and MATLAB. These three algorithms are tested and compared on ten discrete UCI datasets. The results abstained from SU, genetic feature selection methods, and the proposed method have been tabulated. Tables 2, 3, 4, 5, 6, 7, and 8 exemplify the number of features selected, classification performance, and the time taken by the proposed method with SU and GA methods. This experiment shows the proposed method is more effective when compared with available feature selection methods.

4.3 Feature Selection

FS is a process of driving the subset features from the original feature space. The proposed method has been applied in all datasets to select the relevant features by removing the irrelevant one. Table 2 shows the features selected by using SU, GA, and SU-GA. The results indicate that the proposed method selects the least number of features than other two methods for all the ten datasets. Notably, this method selects around ten percentage of attributes for segment challenge and vote datasets. For lung cancer and vehicle data, it selects only around twenty percentage of features. If the selected features are decreased, the performance of classification algorithm does improve this advocates the necessity of feature selection. Therefore, effective feature selection may improve the accuracy and performance of learning algorithms (Fig. 3).

Table 2 Features selected by different FS methods

Datasets	All	SU	GA	SU-GA
Ionosphere	34	17	14	13
Soybean	35	18	22	16
Diabetes	8	4	4	3
Segment	19	10	8	4
Vote	16	8	4	2
Dermatology	35	18	22	11
Lung cancer	57	28	21	13
Wine	14	7	12	6
Hepatitis	20	10	11	8
Vehicle	19	10	11	4
Total	257	130	129	80

Table 3 Classification accuracy by different methods on various datasets

Dataset	NB			J48			SMO					
	All	SU	GA	SU-GA	All	SU	GA	SU-GA	All	SU	GA	SU-GA
Ionosphere	82.6	87.2	90.3	90.6	91.5	91.7	92.6	92.0	88.6	88.0	88.0	88.3
Soybean	93.0	90.6	92.1	89.3	91.5	90.6	90.8	90.5	93.7	93.4	93.5	93.4
Diabetes	76.3	75.4	77.5	76.4	73.8	74.3	74.9	74.6	77.3	76.0	76.8	75.9
Segment	81.1	76.9	82.9	82.8	95.7	94.9	95.4	94.9	91.9	90.2	89.4	82.1
Vote	90.1	91.3	96.1	95.6	96.3	95.2	96.1	95.6	96.1	95.6	95.6	95.6
Dermatology	97.3	91.8	98.1	92.9	94.0	88.0	94.3	90.4	95.4	84.1	97.3	98.2
Lung cancer	50.0	68.8	75.0	78.1	50.0	62.5	62.5	62.5	40.6	62.5	65.6	68.8
Wine	97.2	96.1	97.2	97.8	93.8	94.4	93.8	94.4	98.3	96.6	98.3	96.6
Hepatitis	84.5	84.5	84.5	85.8	83.9	82.6	83.2	83.9	85.2	83.2	83.2	83.2
Vehicle	44.8	43.4	48.5	44.1	72.5	69.5	68.3	66.1	74.3	66.7	58.0	46.7
	79.7	80.6	84.2	83.3	84.3	84.4	85.2	84.5	84.1	83.6	84.6	82.9

Table 4 Classification accuracy by different methods on various datasets

Dataset	NB			J48			SMO					
	All	SU	GA	SU-GA	All	SU	GA	SU-GA	All	SU	GA	SU-GA
Ionosphere	89.5	88.9	89.5	88.9	92.9	91.7	92.3	91.7	91.2	92.0	92.6	92.9
Soybean	84.3	82.3	80.8	81.0	92.1	90.3	91.7	88.6	93.4	93.0	94.0	93.3
Diabetes	71.2	72.1	73.6	73.4	73.8	72.0	74.0	72.0	75.4	77.2	75.5	76.4
Segment	87.4	87.1	88.2	87.1	96.9	96.4	97.9	96.9	96.7	96.3	94.5	91.5
Vote	94.9	95.6	95.6	95.6	95.6	94.3	95.2	95.6	94.7	95.2	95.9	95.4
Dermatology	86.7	86.4	87.4	89.3	94.8	91.5	94.2	90.4	96.2	92.1	96.4	93.7
Lung cancer	62.5	65.6	62.5	62.5	56.2	62.5	65.6	65.6	37.5	53.1	65.6	71.9
Wine	88.8	88.8	88.8	88.8	97.2	96.6	96.6	97.2	97.2	98.3	99.4	98.3
Hepatitis	76.1	76.8	77.4	79.4	80.0	81.9	83.9	80.0	80.0	83.9	83.9	82.6
Vehicle	65.7	64.9	66.0	65.7	77.0	73.8	72.8	64.7	81.7	70.1	71.1	64.2
	80.7	80.8	81.0	81.2	85.7	85.1	86.4	84.3	84.4	85.1	86.9	86.0

Table 5 Average performance of different classification algorithms

Classification algorithms	All	SU	GA	SU-GA
NB	79.7	80.6	84.2	83.3
J48	84.3	84.4	85.2	84.5
SMO	84.1	83.6	84.6	82.9
JRIP	81.6	82.5	81.6	83.0
DT	80.7	80.8	81.0	81.2
Rand Frst	85.7	85.1	86.4	84.3
Multi Perptn	84.4	85.1	86.9	86.0
Kstar	81.9	83.5	84.3	83.7
Average	82.8	83.2	84.3	83.7

4.4 Classification Performance

In order to evaluate how well original features and each selected feature by different feature selector including SU, GA and SU-GA will able to improve the detection performance of various classifiers is empirical evaluated. The various classification algorithms are used in numerous sets in predicting the results, and the value are observed. It is observed that the selected features by SU-GA feature selector will able to improve the detection performance of all the classifiers. However, this method selects more number of features as compared with the proposed method which selects only 30% of features. It is worth to note that the difference of classification performance between GA and the proposed method is very small with better running time and number of features selected. And also, the proposed method is effective and efficient when compared with other available feature selection methods.

Table 5 shows the average learning accuracy and observed the proposed SU-GA feature selector improves the detection performance of JRIP and decision table algorithms and is comparatively improved than the other methods. This experimentation concludes that the attributes reduced through SUGAFS, and classification accuracy may increase or remain comparatively identical. The exploratory results illustrate the classification accuracy from the chose feature subset indicates prevalent results than all other existing strategies (Fig. 4).

4.5 Processing Time

In the third phase, several tests are carried out to assess the running time of the proposed method across all the datasets. It is also conducted that the same experiments for SU and GA feature selection methods. After obtaining the results, the performance of the three feature selection methods are compared with the original dataset in terms of running time. The detailed results are provided in Tables 6, 7 and 8. From the experimental results, it can be noted that the proposed feature selection method drastically reduces the running time of a learning algorithms.

Table 6 Computational time of different methods on various datasets

Dataset	NB				J48				SMO			
	All	SU	GA	SU-GA	All	SU	GA	SU-GA	All	SU	GA	SU-GA
Ionosphere	0.01	0.00	0.00	0.00	0.08	0.01	0.01	0.01	0.00	0.08	0.08	0.01
Soybean	0.00	0.00	0.00	0.00	0.03	0.01	0.01	0.01	1.49	1.24	1.12	1.08
Diabetes	0.01	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.09	0.07	0.06	0.01
Segment	0.02	0.01	0.00	0.00	0.04	0.03	0.02	0.02	0.4	0.13	0.28	0.11
Vote	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.01	0.01	0.01
Dermatology	0.01	0.00	0.01	0.00	0.03	0.02	0.03	0.00	0.28	0.26	0.21	0.08
Lung cancer	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.01	0.02	0.01
Wine	0.00	0.00	0.01	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02
Hepatitis	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.01	0.05	0.01	0.01	0.01
Vehicle	0.01	0.00	0.00	0.00	0.04	0.02	0.02	0.00	0.07	0.05	0.05	0.05

Table 7 Computational time of different methods on various datasets

Dataset	NB			J48			SMO					
	All	SU	GA	SU-GA	All	SU	GA	SU-GA	All	SU	GA	SU-GA
Ionosphere	0.05	0.14	0.13	0.02	0.04	0.11	0.1	0.02	2.27	0.65	0.54	0.43
Soybean	0.24	0.19	0.23	0.08	0.19	0.22	0.29	0.07	31.78	15.67	15.94	13.47
Diabetes	0.17	0.05	0.04	0.01	0.16	0.12	0.13	0.03	0.66	0.45	0.46	0.32
Segment	0.16	0.08	0.07	0.02	0.13	0.08	0.07	0.05	5.31	3.00	2.53	1.88
Vote	0.15	0.01	0.00	0.00	0.15	0.02	0.01	0.00	0.84	0.42	0.2	0.14
Dermatology	0.07	0.12	0.17	0.02	0.14	0.11	0.1	0.01	23.66	7.39	10.68	4.77
Lung cancer	0.00	0.01	0.00	0.00	0.01	0.01	0.01	0.00	5.75	1.43	0.78	0.36
Wine	0.1	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.29	0.15	0.23	0.13
Hepatitis	0.01	0.01	0.01	0.01	0.09	0.01	0.01	0.02	0.36	0.15	0.15	0.13
Vehicle	0.06	0.03	0.03	0.01	0.07	0.06	0.06	0.05	2.09	0.95	1.13	0.63

Table 8 Processing time average by different FS methods

Classification algorithms	All	SU	GA	SU-GA
NB	0.006	0.002	0.003	0.00
J48	0.029	0.01	0.011	0.006
SMO	0.251	0.188	0.186	0.139
JRIP	0.065	0.055	0.058	0.022
DT	0.101	0.065	0.069	0.018
Rand Frst	0.099	0.075	0.079	0.026
Multi Perptn	7.301	3.026	3.264	2.226
Kstar	0.00	0.00	0.00	0.00
Average	0.98	0.43	0.46	0.30

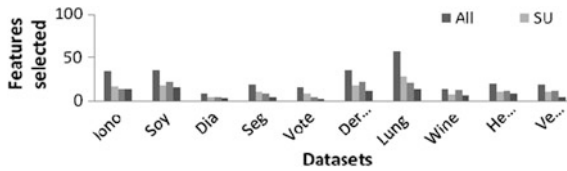


Fig. 3 Number features selected on UCI datasets

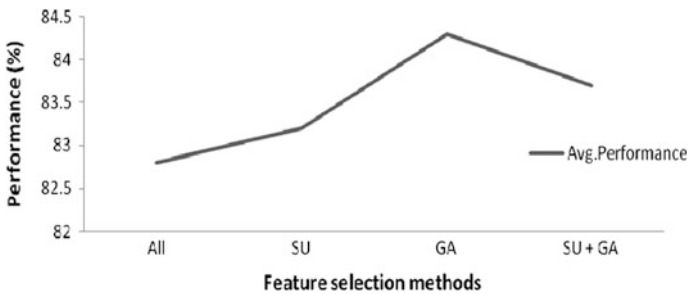


Fig. 4 Performance of FS methods

It is found that the average running time of the proposed model significantly improves the processing time than any other FS methods.

From the experiment, it has observed the following points are in favor of the proposed method:

1. Running time for all classification algorithms is lesser than other methods.
2. GA feature selection method takes much time than other two feature selection methods because of its global search nature.
3. The average processing time of the proposed method is considerably lesser than other two feature selection methods (Figs. 5 and 6).

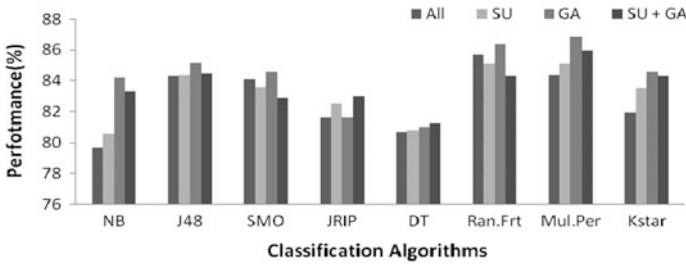
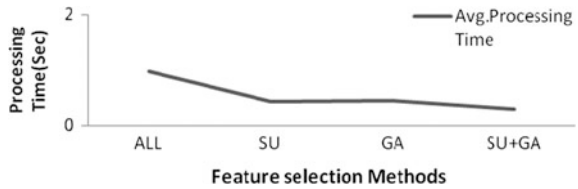


Fig. 5 Processing time average by different FS methods

Fig. 6 Processing time average by different FS methods



5 Conclusion

In this work, both SU and GA are combined named as hybrid feature selector for the purpose of eliminating both irrelevant and redundant features and to select only the most relevant features for improving data classification. The performance of the proposed feature selector was evaluated in terms of three quality measures such as number of selected features, detection performance of classifiers, and time taken to build the model with the dataset from University of California Irvine dataset. The proposed SU and GA should probably become part of the standard tool box of feature selection method for effective data classification. The proposed method can more clearly be stated as follows:

1. SU and GA are combined named as SU-GA hybrid feature selector for selecting only most relevant features for supervised. The system is aimed at making improvements over the existing work in three perspectives such as reduction in feature set, improvement in classification accuracy, and finally, minimizing the running time of achieving the goal.
2. The proposed method significantly reduces processing time than any other feature selection methods with minimal number of features. The result of SU-GA imparts higher classification accuracy rate for some dataset with minimum selected features and minimum running time.
3. The proposed features and learning paradigm SU-GA are promising strategies to be applied on any data classification problems.

References

1. Frenay, B., Doquire, G., Verleysen, M.: Estimating mutual information for feature selection in the presence of label noise. *Comput. Stat. Data Anal.* **71**(1), 832–848 (2014)
2. Hemphill, E., Lindsay, J., Lee, C., Mandoiu, I., Nelson, C.E.: Feature selection and classifier performance on diverse bio-logical datasets. *BMC Bioinf.* **15**(13) (2014)
3. Ganapathy, S., Kulothungan, K., Muthurajkumar, S., Vijayalakshmi, M., Yogesh, P., Kannan, A.: Intelligent feature selection and classification techniques for intrusion detection in networks: a survey. *EURASIP J. Wirel. Commun. Netw.* (2013)
4. Raymer, M.L., Doom, T.E., Kuhn, L.A., Punch, W.F.: Knowledge discovery in medical and biological datasets using a hybrid bayes classifier/evolutionary algorithm. *IEEE Trans. Syst. Man Cybern.* **33**(5), 802–810 (2003)
5. Osl, M., Dreiseitl, S., Cerqueira, F., Netzer, M., Pfeifer, B., Baumgartner, C.: Demoting redundant features to improve the discriminatory ability in cancer data. *J. Biomed. Inform.* **42**(4), 721–725 (2009)
6. Xie, J., Wang, C.: Using support vector machines with a novel hybrid feature selection method for diagnosis of erythemato-squamous diseases. *Expert Syst. Appl.* **38**, 5809–5815 (2010)
7. Holland, J.H.: *Adaptation in Natural Artificial Systems*, 2nd edn. MIT Press (1992)
8. Deutsch, J.M.: Evolutionary algorithms for finding optimal gene sets in microarray prediction. *Bioinformatics* **19**(1), 45–52 (2003)
9. Jirapech-Umpai, T., Aitken, S.: Feature selection and classification for microarray data analysis: evolutionary methods for identifying predictive genes. *BMC Bioinf.* **6**, 148 (2005)
10. Li, L., Weinberg, C.R., Darden, T.A., Pedersen, L.G.: Gene selection for sample classification based on gene expression data: study of sensitivity to choice of parameters of the GA/KNN method. *Bioinformatics* **17**(12), 1131–1142 (2001)
11. Li, L., Pedersen, L.G., Darden, T.A., Weinberg, C.R.: Computational analysis of leukemia microarray expression data using GA/KNN method. In: *Proceeding of the 1st Conference on Critical Assessment of Microarray Data Analysis, CAMDA* (2000)
12. Ooi, C.H., Tan, P.: Genetic algorithm applied to multi-class prediction for the analysis of gene expression data. *Bioinformatics* **19**(1), 37–44 (2003)
13. Moscato, P.: *On evolution, search, optimization, genetic algorithms and martial arts: toward memetic algorithms*. Technical Report Caltech Concurrent Computation Program, Rep. 826, California Institute of Technology, Pasadena, CA (1989)
14. Zhu, Z., Ong, Y.S., Dash, M.: Wrapper-Filter feature selection algorithm using a memetic framework. *IEEE Trans. Syst. Man Cybern. Part B* **10**(4), 392–404 (2006)
15. Hettich, S., Blake, C., Merz, C.: UCI repository of machine learning databases. <http://www.ics.uci.edu/mllearn/MLRepository.html> (1998)
16. Moretti, S., van Leeuwen, D., Gmuender, H., Bonassi, S., Van Delft, J., Kleinjans, J., Patrone, F., Merlo, D.F.: Combining Shapley value and statistics to the analysis of gene expression data in children exposed to air pollution. *BMC Bioinf.* **9**(361), 1–21 (2008)
17. Aitkenhead, M.J.A.: Co-evolving decision tree classification method. *Expert Syst. Appl.* **34**(1), 18–25 (2006)
18. Baker, J.E.: Adaptive selection methods for genetic algorithms. In: *Proceedings of International Conference in Genetic Algorithm and Their Applications*, pp. 101–111 (1985)
19. Hualonga, B., Jingb, X.: Hybrid feature selection mechanism based high dimensional data sets reduction. *Energy Procedia* **11**, 4973–4978 (2011)
20. Tan, F., Fu, X., Zhang, Y., Bourgeois, A.G.: A genetic algorithm—based method for feature subset selection. *Soft Comput.* **11**, 111–120 (2008)
21. Jinyan, L., Huiqing, L.: Kentridge bio-medical data set repository. <http://datam.i2r.a-star.edu.sg/datasets/krbd> (2001)
22. Keinan, A., Sandbank, B., Hilgetag, C.C., Ellison, I., Ruppim, E.: Fair attribution of functional contribution in artificial and biological networks. *Neural Comput.* **16**(9), 1887–1915 (2004)

23. Qi, Z., Tian, Y., Shi, Y.: Robust twin support vector machine for pattern classification. *J. Pattern Recognit.* **46**(1), 305–316 (2013)
24. Senthamarai Kannan, S., Ramaraj, N.: A novel hybrid feature selection via Symmetrical Uncertainty ranking based local memetic search algorithm. *Knowl. Based Syst.* **23**, 580–585 (2010)
25. Shao, Y.H., Chen, W.J., Zhang, J.J. et al.: An efficient weighted Lagrangian twin support vector machine for imbalanced data classification. *J. Pattern Recognit.* **47**(9), 3158–3167 (2014)
26. Weka.: Machine Learning Software in Java. The University of Waikato software documentation. http://www.cs.waikato.ac.nz/_ml/wek
27. Eswa, J., Yang, J.H., Honavar, V.: Feature selection using a genetic algorithm. *IEEE Intell. Syst.* **13**(2), 44–49 (1998)
28. Yildirim, P.: Filter based feature selection methods for prediction of risks in hepatitis disease. *Int. J. Mach. Learn. Comput.* **5**(4), 258–263 (2015)