Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 72 (2015) 162 – 169

Procedia Computer Science

The Third Information Systems International Conference

Wrapper Feature Subset Selection for Dimension Reduction Based on Ensemble Learning Algorithm

Rattanawadee Panthong[a], Anongnart Srivihok[b]

[a]Computer Science Department, Kasetsart University, Bangkok and 10900, Thailand
[b]Computer Science Department, Kasetsart University, Bangkok and 10900, Thailand

Abstract

Feature selection is a technique to choose a subset of variables from the multidimensional data which can improve the classification accuracy in diversity datasets. In addition, the best feature subset selection method can reduce the cost of feature measurement. This work focuses on the use of wrapper feature selection. This study use methods of sequential forward selection (SFS), sequential backward selection (SBS) and optimize selection (evolutionary) based on ensemble algorithms namely Bagging and AdaBoost by subset evaluations which are performed using two classifiers; Decision Tree and Naïve Bayes. Thirteen datasets containing different numbers of attributes and dimensions are obtained from the UCI Machine Learning Repository. This study shows that the search technique using SFS based on the bagging algorithm using Decision Tree obtained better results in average accuracy (89.60%) than other methods. The benefits of the feature subset selection are an increased accuracy rate and a reduced run-time when searching multimedia data consisting of a large number of multidimensional datasets.

Peer-review under responsibility of organizing committee of Information Systems International Conference (ISICO2015)

Keywords Feature selection, Ensemble learning algorithm, dimension reduction, classification, data mining;

1. INTRODUCTION

Normally, multidimensional data may contain tens or hundreds of attributes, where several attributes in the dataset may be irrelevant to the pattern classification in machine learning [1]. The method to reduce a

* Corresponding author. Tel.: .+6-089-810-4181.
E-mail address: rattanawadee.p@ku.th

Peer-review under responsibility of organizing committee of Information Systems International Conference (ISICO2015)
doi:10.1016/j.procs.2015.12.117

large number of attributes or features by selecting only effective attributes from the original dataset in order to reduce irrelevant or redundant attributes is called feature selection [1-2]. The performance of classification is based on the attribute selected. Therefore, feature selection can serve as a pre-processing tool of great importance before solving the classification problems. Feature selection can reduce the number of irrelevant features from the input data which can efficiently describe the input data as well reduce effects from noise or irrelevant variables and still provide good prediction results [1]. The feature selection algorithms improve inductive learning, either in term of generalization capacity, learning speed, or reducing the complexity of the induced model and classification accuracy [1-2, 4].

In this paper, we present three search strategies based on the wrapper method to perform feature subset selection from a training set using sequential forward selection (SFS), sequential backward elimination selection (SBS) and optimize selection (evolutionary).

For ensemble learning algorithms which partition data into different segment, use Bagging and AdaBoost, two classifiers (Decision Tree and Naïve Bayes) are used as subset evaluators [1, 3]. Bagging (Bootstrap Aggregation) creates an ensemble by training individual classifiers on bootstrap samples of the different datasets [5]. Diversity of subset training data in bagging is obtained by using bootstrap subsampling in the training examples. That is, different training data subsets are randomly drawn with replacement from the original dataset. In AdaBoost, bootstrap training data samples are drawn from a distribution dataset that is iteratively updated such that subsequent classifiers focus on increasingly difficult instances. The effectiveness of bagging and boosting has been explained in terms of the bias-variance decomposition of the classification error [5].

This paper is organized as follows: Section 2 provides reviews of the literature on Wrapper Feature Selection and Ensemble learning algorithm. Section 3 presents our method. Section 4 describes the experiments and the datasets, and the measures used to evaluate the feature subset. Section 5 analyzes the experimental results. Lastly, Section 6 presents the conclusion and future work.

2. RELATED WORK

The feature selection is to select a subset of variables from the input data which can efficiently describe the input data while reducing effects from noise or relevant variables and still provide good prediction results[1-2]. Wrapper methods perform a search in the space of feature subsets such as classification performances on a cross-validation of the training set which provided better results than filter methods. But wrapper approaches increase the computational cost [12].

Donghai Guan, et al. [3] reviewed and compared two techniques of integrating feature selection and ensemble learning, (1) Feature selection for ensemble learning (ENfs) and (2) Ensemble learning for feature selection (FSen). This approach obtained predictive accuracy superior to conventional feature selection methods for supervised learning. Moreover, its most prominent advantage is the ability to handle stability issue that is usually poor in existing feature selection methods.

S´anchez-Maro˜no, et al. [10] proposed a new wrapper method, called Incremental ANOVA and Functional Networks-Feature Selection (IAFN-FS) for dealing with multiclass problems based in classical algorithms, such as C4.5 and Naïve Bayes. The multiple binary classifiers approach obtained better results in accuracy, although it has the drawback of selecting a higher number of features.

Akin Ozcift and Arif Gulten [11] used a rotation forest ensemble decision tree algorithm wrapped with best first search strategy. The wrapper uses forward selection to choose the optimum subset on the Erythemato-Squamous diseases dataset. The discrimination ability of selected features is evaluated using several machine learning algorithms and the diversity of the training data using the bagging algorithm.

Yvan Saeys, et al. [12] proposed the method of ensemble feature selection techniques for high dimension data which can be used to yield more robust feature selection techniques. As well Sangkyun Lee, et.al [13] presented a method of an extension to RapidMiner which delivers implementations of algorithms which is well suited for very high-dimensional data. These experiments were conducted on a microRNA-expression dataset.

3. PROPOSED METHOD

A wrapper approach is used to calculate attribute weights by using the classification model to measure the performance of attributes. Wrapper methods use the inductive algorithm as the evaluation function [6, 7]. This technique uses a classifier to evaluate subsets by their predictive accuracy (on test data) after statistical resampling or cross-validation of the dataset. Furthermore, the wrapper method achieves better recognition rates than a filter approach since the former is tuned to the specific interactions between the classifier and the dataset [1, 8]. Further, wrappers have a mechanism to avoid overfitting, since typically cross-validation measures of predictive accuracy are used [1, 9].

A sequential forward selection (SFS) is the simplest greedy search algorithm [14]. SFS starts with an empty selection of attributes and, in each round, it adds each unused attribute of the given example set. For each added attribute, the performance is estimated using the cross validation. Only the attribute giving the highest performance is added to the selection for the object function. Then a new round is started with the modified selection. Therefore, the SFS algorithm adds features which give a high value to the object function [1]; the forward greedy algorithm is shown in Figure 1(a).

Sequential backward selection (SBS) works in the opposite direction to SFS [14]. SBS starts with the full set of attributes and, in each round, it removes each remaining attribute of the given example set. For each iteration or attribute removal, the performance is estimated using the inner operators, such as a cross-validation. Only the attribute giving the least decreasing performance is finally removed from the selection. Then a new round is started with the modified selection [1]. This elimination process has two advantages: first, it can discard several features and second, it allows for backtracking, and so, when a subset of features worsens the results obtained by the previous one, some previously eliminated features can be included in the new subset for re-evaluation [10] as given in Figure 1(b) which shows the backward greedy algorithm.

Initialize feature set $F^k = \varnothing$ at $k = 0$
Iterate
- **find best feature j to add to F^k with most significant cost reduction**
- **$k++$ and $F^k = F^{k-1} \cup \{j\}$**

Initialize feature set $F^k = \{1, \ldots, d\}$ at $k = d$
Iterate
- **find best feature $j \in F^k$ to remove with least significant cost increase**
- **$F^{k-1} = F^k - \{j\}$ and $k--$**

Fig.1. (a) Forward Greedy Algorithm [13]; (b) Backward Greedy Algorithm [13]

Optimize selection (evolutionary) is a method which selects the most relevant attributes of the given example set. Genetic algorithm (GA) is a heuristic search algorithm that mimics the process of natural evolution [14]. Genetic algorithm (GA) parameters and operators can be modified within the general idea of an evolutionary algorithm to suit the data to obtain the best performance or the best search results [1]. This heuristic search is routinely used to generate useful solutions to optimization and search problems by using genetic algorithm [16].

We proposed 4 new wrapper approaches include OBDT, OBNB, OADT and OANB. Moreover, adapted methods from [11] include FBDT, BBDT, FBNB, BBNB, FADT, BADT, FANB and BANB.

Ensemble learning algorithms are generated by training individual classifiers on different datasets obtained by resampling a common training set such as Bagging and Boosting [17]. The component classifiers are built on different partitions of the training set obtained through 10-fold cross-validation [5]. Bagging uses component classifiers of the same type such as Decision Tree, and a simple combiner consisting of a majority vote across the ensemble. Applying the sampling with replacement procedure, each classifier is trained on the average of 63.2% of the unique training examples [5]. The bagging algorithm creates a diversity training set for feature subset selection to improve classification in terms of stability and classification accuracy. It also reduces variance and helps to avoid overfitting [5]. AdaBoost, (short for Adaptive Boosting) is a meta-algorithm which can be used in conjunction with many other learning algorithms to improve their performance. It is applicable for building ensembles that empirically improves generalization performance [5].

A Decision Tree (DT) classifier is typically a top-down greedy approach, which provides a rapid and effective method for classifying data instances [18]. Decision Trees are generated by recursive partitioning and the recursion stops when all the examples or instances are the same label value. Decision Tree classifier provides a rapid and useful solution for classifying instances in large datasets with a large number of variables classification using a Decision Tree has high accuracy but the performance usually depends on the characteristics of the dataset [18]. Naïve Bayes (NB) classifier is used for classification problems in data mining and machine learning. It is a simple probabilistic classifier based on applying Bayes' theorem from Bayesian statistics with strong independence assumptions [18].

The names of the wrapper feature subset selection algorithms and their abbreviations are given in Table 1.

Table 1 Search algorithms used in this study and their abbreviations

No.	Algorithms	Abbreviation
1	SFS+ Decision Tree	FDT
2	SBS+ Decision Tree	BDT
3	SFS+Bagging + Decision Tree	FBDT**
4	SBS+ Bagging + Decision Tree	BBDT**
5	Optimize selection (evolutionary) +Bagging +Decision Tree	OBDT*
6	SFS+ Naïve Bayes	FNB
7	SBS+ Naïve Bayes	BNB
8	SFS+ Bagging + Naïve Bayes	FBNB**
9	SBS+ Bagging + Naïve Bayes	BBNB**
10	Optimize selection (evolutionary) +Bagging + Naïve Bayes	OBNB*
11	SFS+AdaBoost+ Decision Tree	FADT**
12	SBS+AdaBoost+ Decision Tree	BADT**
13	SFS+AdaBoost+ Naïve Bayes	FANB**
14	SBS+AdaBoost+ Naïve Bayes	BANB**
15	Optimize selection (evolutionary) +AdaBoost+ Decision Tree	OADT*
16	Optimize selection (evolutionary) +AdaBoost+ Naïve Bayes	OANB*

* New proposed methods, ** Adapted methods

The proposed method for wrapper feature subset selection based on the ensemble learning algorithm is shown in Figure 2.

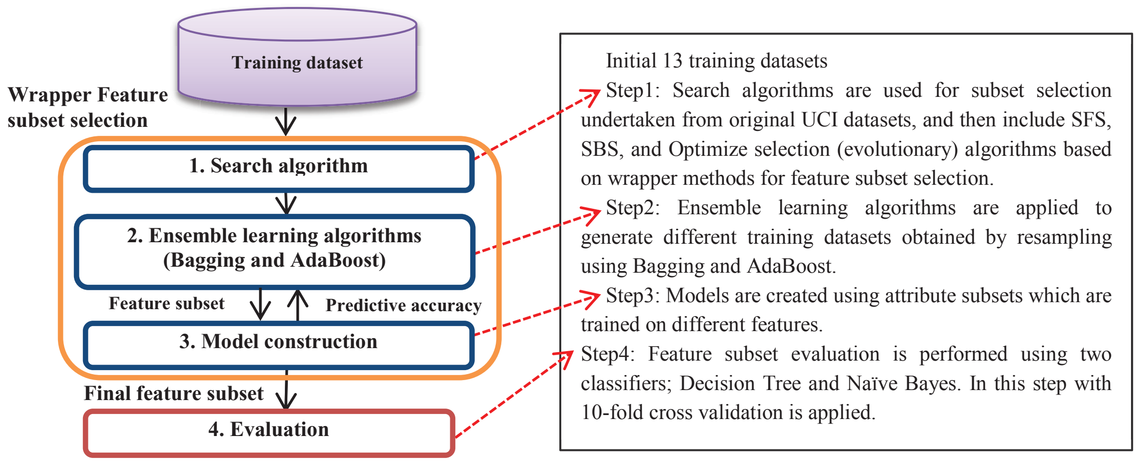

Fig.2. Wrapper feature subset selection based on ensemble learning algorithm

4. EXPERIMENTS

The data for this experiment uses 13 datasets from the UCI Machine Learning Repository [20]. The data are described in Table 2.

The performance of the different feature selector classifiers is calculated using 10-fold cross validation (CV). It is used to measure the classification evaluation on the datasets and to compare the accuracy of the classification models [18].

With Classification Performance Measurement, the calculation of classification accuracy is measured by Eq. (1)

$$\text{Accuracy} = (TP+TN) / (TP+FN+FP+TN) \quad (1)$$

where TP is the number of positive instances correctly classified. TN is the number of negative instances correctly classified. FN is the number of positive instances incorrectly classified as negative. FP is the number of negative instances incorrectly classified as positive [18].

This study uses RapidMiner studio version 6.4 Enterprise edition for model training and testing [19]. It is used as a tool for feature subset selection, training classifier and performance evaluation.

Table 2 Dataset description using in this study

Dataset	Number of attributes	Number of instances	Number of Class	Attribute characteristic
Iris	4	150	3	real
new-thyroid	5	215	3	Integer, real
glass	9	214	7	real
pageblocks	10	5473	5	Integer, real
wine	13	178	3	Integer, real
pendigits	16	7494	10	Integer
zoo	17	101	7	nominal, Integer
vehicle	18	846	4	Integer
segmentation	19	2310	7	Integer, real
dermatology	34	366	6	Integer
soybean	35	307	19	Integer
lung-cancer	56	32	3	Integer
movement	90	360	15	Integer

5. RESULTS AND DISCUSSIONS

In this section we present the results of feature selection using the wrapper approach based on the ensemble learning algorithm. The experimental results of proposed method using the classification accuracy on the training sets of the 13 benchmark datasets we depicted in Table 3, Table 4, and Table 5.

Table 3 shows that FBDT obtains a higher accuracy rate than other methods in 7 datasets (new-thyroid, glass, pageblocks, vehicle, segmentation, soybean, and movement) and also the highest average accuracy rate (89.60%) whereas, OBDT has the lowest accuracy in almost all datasets except for Iris dataset. For Table 4 illustrates that the FBNB obtains the highest average accuracy rate (85.56%) and also it obtains more accuracy rate than other methods in 4 datasets (pageblocks, dermatology, lung-cancer, and movement). In addition, OBNB receives the highest accuracy rates in 2 datasets (Iris and new-thyroid).

Table 5 demonstrates that FADT obtains accuracy rate superior to other methods in 4 datasets (segmentation, soybean, lung-cancer, and movement) and also the highest average accuracy rate (89.59%). Moreover, the OADT gains good accuracy rate in glass dataset and the FANB receives the best accuracy rate in wine dataset while, BANB obtains the lowest average accuracy rate (82.39%). The result in Table 6 indicates the maximum performance and the best subset of all datasets using wrapper feature selection with ensemble algorithms. The experiment results indicate that forward feature selection with the ensemble learning algorithm provides better classification accuracy values than single wrapper feature selection for the domain with a number of features is greater than 10. On the other hand, datasets having a low number of features, results were inferior to other methods.

The difference between previous work [10] and our study here is that we use three search algorithms based on wrapper approaches with ensemble learning algorithms by the two classifiers Decision Tree and Naive-Bayes with RapidMiner Studio. Previous work [10] implemented the wrapper methods for feature selection using forward and backward strategy with multiclass approaches based on the two classifiers C4.5 and Naive-Bayes using Weka and 5-fold cross-validation was used for evaluating each candidate feature subset. Also, feature selection is based on the sensitivity analysis and the ANOVA and Functional Networks (AFN) algorithm. Results of feature subset selection using wrapper approaches with ensemble algorithms are compared using the same datasets. The methods of FANB, FDT and FADT provide better accuracy rate in 3 datasets (wine, zoo, and segmentation) from all 6 datasets. On the other hand, this experiment provides lower accuracy rate in 3 datasets (Iris, soybean and vehicle). This method might be a good alternative for a domain where the number of attributes is greater than 10. Additionally, the effectiveness of adapted methods on classification performances obtained better results in some datasets. Due to Bagging provides the diversity of subsets which might be reduced the variance of datasets and AdaBoost reduced bias of multi-class datasets. Also, the accuracy rate is increased in balanced sample subsets.

Table 3 Performance of classification accuracies (%) for five feature subset selections: SFS, SBS and Optimize selection (evolutionary) with bagging by Decision Tree classifier as a subset evaluator.

Dataset	Used all features	FDT	BDT	FBDT	BBDT	OBDT
Iris	92.00	**95.33**	**95.33**	**95.33**	**95.33**	**95.33**
new-thyroid	93.46	94.44	93.98	**95.82**	93.96	93.05
glass	61.73	71.58	71.56	**73.79**	76.21	70.17
pageblocks	96.49	96.97	96.99	**97.13**	96.71	96.71
wine	92.71	95.56	94.38	96.67	**97.22**	93.27
pendigits	93.86	94.17	95.25	95.74	**96.29**	92.82
zoo	94.00	**98.00**	97.00	96.00	96.00	95.00
vehicle	62.89	71.99	73.88	**74.59**	74.00	70.10
segmentation	94.11	96.71	96.49	**97.62**	97.32	95.11
dermatology	85.76	96.99	97.27	97.53	**97.55**	94.27
soybean	61.10	89.57	76.69	**91.85**	91.81	84.66
lung-cancer	41.67	**81.67**	60.83	79.17	66.67	59.17
movement	21.67	67.50	35.28	**73.61**	59.17	35.28
Average	76.27	88.50	83.46	**89.60**	87.59	82.69

Table 4 Performance of classification accuracies (%) for five feature subset selections: SFS, SBS and Optimize selection (evolutionary) with bagging by Naïve Bayes classifier as a subset evaluator.

Dataset	Used all features	FNB	BNB	FBNB	BBNB	OBNB
Iris	95.33	96.00	96.67	96.00	96.00	96.67
new-thyroid	96.75	97.68	97.68	97.68	97.68	**98.12**
glass	45.39	**60.30**	47.75	59.87	46.82	56.08
pageblocks	90.04	94.63	94.63	**94.79**	94.63	94.54
wine	96.63	97.78	**98.86**	98.33	98.30	97.75
pendigits	87.84	86.87	**87.99**	86.88	87.98	86.83
zoo	95.00	95.09	**97.00**	96.00	96.00	94.09
vehicle	45.52	53.90	**58.76**	56.39	58.41	55.69
segmentation	79.65	87.36	**90.43**	87.36	89.35	85.15
dermatology	87.72	97.27	97.55	**98.36**	98.09	92.92
soybean	82.73	83.40	84.69	85.31	85.34	**86.00**
lung-cancer	44.17	80.00	55.83	**87.50**	65.83	52.50
movement	64.17	63.33	66.67	**67.78**	63.89	63.33

Average	77.76	84.12	82.65	**85.56**	82.95	81.51

Table 5 Performance of classification accuracies (%) for feature subset selection: SFS, SBS and Optimize selection (evolutionary) with AdaBoost by Decision Tree and Naïve Bayes as a subset evaluator.

Dataset	FADT	BADT	FANB	BANB	OADT	OANB
Iris	**96.00**	94.67	**96.00**	**96.00**	95.33	**96.00**
new-thyroid	95.82	95.82	95.82	**97.68**	95.82	**97.68**
glass	75.74	74.83	60.30	48.20	**76.67**	59.81
pageblocks	96.97	**96.99**	94.63	94.63	96.91	94.43
wine	95.56	96.11	**99.44**	96.00	97.19	96.67
pendigits	94.60	**96.86**	86.87	87.99	94.94	86.83
zoo	**97.00**	**97.00**	95.09	**97.00**	**97.00**	96.00
vehicle	71.99	**73.42**	56.39	60.42	72.58	54.02
segmentation	**98.31**	97.92	89.00	88.96	97.92	86.19
dermatology	97.54	97.27	96.70	**97.82**	97.27	93.71
soybean	**91.85**	73.77	85.31	84.69	88.89	86.67
lung-cancer	**81.67**	70.00	80.00	55.83	65.00	56.67
movement	**71.67**	65.83	67.78	65.83	63.61	67.22
verage	**89.59**	86.96	84.87	82.39	87.63	82.45

Table 6 Experimental results of the best feature subset selection

Dataset	No. attributes	No. instances	No. selected features	Maximum Performance	Methods
Iris	4	150	2	96.67	OBNB
new-thyroid	5	215	2	98.12	OBNB
glass	9	214	5	76.67	OADT
pageblocks	10	5,473	5	97.13	FBDT
wine	13	178	6	99.44	FANB
pendigits	16	7,494	14	96.86	BADT
zoo	17	101	6	98.00	FDT
vehicle	18	846	7	74.59	FBDT
segmentation	19	2,310	7	98.31	FADT
dermatology	34	366	9	98.36	FBNB
soybean	35	307	9,	91.85	FBDT,
soybean	35	307	10	91.85	FADT
lung-cancer	56	32	6	87.50	FBNB
movement	90	360	6	73.61	FBDT

6. CONCLUSION AND FUTURE WORK

In this paper, we present wrapper methods for feature subset selection based on the ensemble learning algorithm with two base learners (Decision Tree, Naïve Bayes). In multi-class datasets, we use 13 datasets from the UCI Machine Learning Repository. The output of this method is SFS with Bagging evaluated by Decision Tree (FBDT) is the best performance (mean accuracy is 89.60%) than other techniques. The FBDT provides better classification accuracy rate than single wrapper feature selection in 4 datasets (pageblocks, vehicle, soybean, and movement) from 13 datasets. Moreover, The FBNB is superior to other methods in 2 datasets (dermatology and lung-cancer). On the other hand, these methods (FBDT, FBNB) with a low number of attributes are worse than OBNB and OADT in 3 datasets (Iris, new-thyroid and glass).

We proposed 4 new wrapper approaches include OBDT, OBNB, OADT and OANB. The results of this method are that OBNB gains maximum performance in 2 datasets (Iris and new-thyroid). In addition, the OADT obtains maximum performance in glass dataset. The OBNB and OADT the algorithms provides better accuracy rate for the domain with a number of features is less than 10 and small multi-class datasets. Although feature selection based on ensemble algorithms (Bagging and AdaBoost) used more computation time, these methods obtained better average accuracy rate than single feature selection (without ensemble algorithms) on various attributes. The benefits of feature selection for this work are improvement in the performance of classification by increasing the accuracy rate and a reduction in the run-time for classification. Furthermore, we obtained good alternative for selecting subsets of features in predictive modelling.

In the future, feature subset selection might be applied in real world problems using a hybrid heuristic search and other methods with ensemble feature selection techniques.

Acknowledgements

This research was supported by the Department of Computer Science, Faculty of Science, Kasetsart University. The authors would also like to thank Delaware Corporation for supporting of the RapidMiner software usage under the Academic Program.

References

[1] Chandrashekar G, Sahin F. A survey on feature selection methods. *Computer and Electrical Engineering*; 2014, p.16-28.

[2] Kumar V, Minz S. Feature Selection: A literature Review. *Smart Computing Review*; 2014.

[3] Guana D, Yuana W, Leea YK, Najeebullaha K, Rasela MK. A Review of Ensemble Learning Based Feature Selection. *IETE Technical Review*; 2014.

[4] Janecek AGK, Wilfried N. Gansterer, Demel MA, Ecker GF. On the Relationship Between Feature Selection and Classification Accuracy. JMLR: Workshop and Conference Proceedings; 2008, p.90-105.

[5] Polikar R. Ensemble learning. Available from: http://www. scholarpedia.org/article/Ensemble_learning.2009.

[6] Pervez MS, Farid D. Md. Literature Review of Feature Selection for Mining Tasks. *International Journal of Computer Applications*; 2015, p.30-33.

[7] Beniwal S, Arora J. Classification and Feature Selection Techniques in Data Mining. *International Journal of Engineering Research & Technology (IJERT)*; 2012, p.1-6.

[8] Jantawan B, Tsai CF.A Comparison of Filter and Wrapper Approaches with Data Mining Techniques for Categorical Variables Selection. *International Journal of Innovative Research in Computer and Communication Engineering*; 2014.

[9] Naseriparsa M, Bidgoli AM, Varaee T. A Hybrid Feature Selection Method to Improve Performance of a Group of Classification Algorithms. *International Journal of Computer Applications;* 2013.

[10] S´anchez-Maro˜no N, Alonso-Betanzos A, and Calvo-Est´evez RM. A Wrapper Method for Feature Selection in Multiple Classes Datasets. J. Cabestany et al. (Eds.): IWANN; 2009, p. 456–463.

[11] Ozcift A, Gulten A. A Robust Multi-Class Feature Selection Strategy Based on Rotation Forest Ensemble Algorithm for Diagnosis. J Med Syst; 2012, p.941–949.

[12] Saeys Y, Abeel T, Peer YV. Robust Feature Selection Using Ensemble Feature Selection Techniques. W. Daelemans et al. (Eds.): ECML PKDD 2008, p. 313–325.

[13] Lee S, Schowe B, Sivakumar V. Feature Selection for High-Dimensional Data with RapidMiner. Technical Report of TU Dortmund University; 2011.

[14] Osuna R G. Pattern Analysis for Machine Olfaction: A Review. *IEEE SENSORS JOURNAL*; 2002, p.189-202.

[15] Zhang T. Adaptive Forward-Backward Greedy Algorithm for Sparse Learning with Linear Models. *IEEE Transactions on Information Theory*; 2011, p.1-19.

[16] Zhuo L, Zheng J, Wang F, Li X, Ai B, Qian, J. A Genetic Algorithm based Wrapper Feature Selection Method for Classification of Hyperspectral Images using Support Vector Machince. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*; 2008, p.397-402.

[17] Bryll R, Osuna RG, Queka F. Attribute bagging: improving accuracy of classifier ensembles by using random feature subsets. *The journal of pattern recognition society*; 2003, p.1291 – 1302.

[18] Farid D Md, Zhang L, Rahman CM, Hossain MA, Strachan R.Hybrid decision tree and Naïve Bayes classifiers for multi-class classification tasks. Expert Systems with Applications; 2014, p. 1937–1946.

[19] RapidMiner Studio Available from https://rapidminer.com/products/ studio.

[20] Frank A, Asuncion A.UCI Machine Learning Repository. Available from http://archive.ics.uci.edu/ml; 2010.