Research Article

Ensemble Classification Based on Feature Selection for Environmental Sound Recognition

Shuai Zhao ⓘ**, Yan Zhang** ⓘ**, Haifeng Xu** ⓘ**, and Te Han** ⓘ

College of Big Data and Intelligence Engineering, Southwest Forestry University, Kunming 650224, China

Correspondence should be addressed to Yan Zhang; zydyr@163.com

Environmental sound recognition has been a hot topic in the domain of audio recognition. How to select the optimal feature subsets and enhance the performance of classification precisely is an urgent problem to be solved. Ensemble learning, a new kind of method presented recently, has been an effective way to improve the accuracy of classification in feature selection. In this paper, experiments were performed on environmental sound dataset. An improved method based on constraint score and multimodels ensemble feature selection methods (MmEnFs) were exploited in the experiments. The experimental results show that when enough attributes are selected, the improved method can get a better performance compared to other feature selection methods. And the ensemble feature selection method, which combines other methods, can obtain the optimal performance in most cases.

1. Introduction

With the developing of artificial intelligence, Environmental Sound Recognition (ESR) has been the focus in the domain of speech recognition in recent years.

There are many environmental sounds in our daily life, such as birdcall, wind, thunder, and so on. By recognizing the sounds, a large amount of information can be obtained to understand our environment. Many applications can benefit from ESR, such as scene recognition, event detection, and surveillance systems [1]. By identifying the sounds happening in the scene, scene recognition is aimed at judging whether the scenes like the specific place such as home, square, field, or not [2]. Event detection is to identify the sounds occurring in our daily life such as talking, laughing, and police siren [3]. Surveillance systems are used to monitor the abnormal event by detecting danger sound during our daily life like gunshot, screams, alarm, and so on [4]. Traditionally, in the field of ESR, more attention has been paid to feature extraction, for example, Mel Frequency Cepstral Coefficients (MFCC), and Code Excited Linear Prediction (CELP). Chu *et al.* adopted a set of features with the Matching Pursuit (MP) technique; the main disadvantage of MP was the huge computational complexity [2]. Tsau *et al.* who used the different feature sets with the Bayesian network based on CELP achieved

a better performance for ESR [5]. Zhang *et al.* proposed the aggregation of multiple classifiers for environmental audio classification and improved the accuracy of prediction evidently [6]. It is difficult to find the optimal subset for ESR. Most researchers concentrate on the feature extraction and single feature selection method; few focus on the subspace for feature selection with ensemble learning. The idea of ensemble feature selection was proposed by Saeys *et al.* [7]. Ensemble feature selection [8] has many irreplaceable advantages; in the past, single feature selection method was used to select the optimal subsets and then to evaluate the performance. For the same datasets, different methods will generate different results. Combining the different outputs of several methods, ensemble feature selection may get a more accurate result [9]. In this work, in order to validate that whether the ensemble method is superior to other feature selection methods solely in the aspect of accuracy, experiments have been performed on environmental sound dataset by using several different approaches in Weka [10], i.e., Correlation [11], GainRatio [12], InfoGain [13], OneR [14], ReliefF [15], SymmetricalUncert [16], and two other methods, namely, csFs and MmEnFs, proposed by us.

This paper is made up of five sections. Introduction is in Section 1. Background is in Section 2. Section 3 briefly introduces the constraint score, csFs, and the process of

MmEnFs. Subsequently, the methods of experiments are described in detail and the quality of results are assessed in Section 4. Finally, we summarize the conclusion and the contribution of this work in Section 5.

2. Background

Feature selection is an important preprocessing step in many fields, such as data mining, pattern recognition, and machine learning [17, 18]. For a mass amount of datasets, it is a process to select a small amount of optimal feature subsets to enhance the performance of a specific model. Generally, one effective solution to reduce data dimensionality for feature selection is to remove irrelevant and redundant features [19, 20]. For the domain of supervised feature selection, it can be divided into three groups as follows: filters, wrappers, and embedded methods [17]. Filter methods evaluate the performance of features by using the training data, which are independent of any learning algorithm [21]. The evaluation criterion of features can be categorized into four types, namely, information, dependency, distance, and consistency [22]. Wrapper methods commonly use a specific learning algorithm to evaluate the features. And the embedded methods perform feature evaluation by using internal properties of the classification model [23]. With the development of feature selection techniques, many different learning algorithms are proposed to acquire a good performance in the domain of feature selection. In the process of feature selection, different algorithms will select different subsets and then yield different results.

In this paper, six different kinds of existed methods in Weka are mainly used to evaluate the performance of features. These methods are briefly introduced below. Correlation evaluates the value of an attribute by measuring the correlation between it and the class. And GainRatio evaluates the worth of an attribute by measuring the gain ratio with respect to the class. InfoGain assesses the value of an attribute by measuring the information gain with the class, while OneR uses the minimum error attribute to predict the performance [24]. ReliefF estimates the value of an attribute by repeatedly sampling an instance and considering the value of the given attribute for the nearest instance of the same and different class [25]. And SymmetricalUncert assesses the value of an attribute by measuring the symmetrical uncertainty with respect to the class.

Ensemble techniques are originally introduced to cope with the stability of feature selection techniques. In most cases, different feature subsets will generate different optimal results for high dimensional data with small sample size [23]. By using the ensemble feature selection, it can not only reduce the risk of selecting an unstable subset, but also avoid the problem of local optima. So the ensemble techniques are usually superior to the single model, when unstable models are combined [26, 27].

There are two essential points in ensemble learning: the first is the difference and the second is the diversity. To the best of our knowledge, in the aspect of feature selection, ensemble learning can be categorized into two types. One is to change the base learner to achieve ensemble; the other is to vary the samples of data to achieve ensemble [28, 29]. Both types of methods are all aims at getting a diverse set of feature selections. There are two key steps in constructing a feature selection ensemble. The first step is creating a set of different feature selectors, each providing an output, and then the second step aggregates the results of the single model [30].

3. Related Works

Constraint score [31] was first proposed by Zhang *et al.* in 2007. The prior knowledge of data can be obtained from two different approaches. The first is class labels, and the second is pairwise constraint score. In the past, in order to estimate the class of sample data precisely, we needed to know the details of class labels. Fortunately pairwise constraint score has changed this situation. It uses pairs of data samples to evaluate whether they belong to the same class (must-link) or the different classes (cannot-link) [32, 33].

The algorithm of constraint score makes use of some supervised information to select the most representative feature subsets. In this process, the key step is how to generate C and M. Given a dataset X, X = {x_1, x_2,...x_m}, two subsets C and M, M = {(x_i, x_j) | x_i and x_j belong to the same class}; C = {(x_i, x_j) | x_i and x_j belong to the different classes}, then using the supervised information in C and M to get the n features in initial dataset X. Equation (1) is used to evaluate the performance of r-th feature.

$$C_r^1 = \sum_{(x_i,x_j) \in M} \left(f_{ri} - f_{rj} \right)^2 - \lambda \sum_{(x_i,x_j) \in C} \left(f_{ri} - f_{rj} \right)^2 \quad (1)$$

Here, f_{ri} represents the r-th feature of the i-th sample x_i, i = 1,...,m; r = 1,...,n. In order to select features with the superior constraint preserving ability, in the same class, the distance between two samples in the same class is obviously shorter than that in the different classes($\lambda < 1$). So the regularization coefficient λ is set as 0.1 to coordinate the two classes.

3.1. The Proposed Method. The proposed method is an improvement of the constraint score for feature selection. Based on the constraint score, csFs method is presented to perform the experiments in this paper. At first, we select ten percent of columns in feature subsets as initial value. Then we employ the greedy method to add a column of attribute per time to compute its value in a forward sequence successively. Finally, from the perspective of columns of attribute set, we can get the range of peak value.

3.2. Ensemble Method. Different from the methods mentioned above, in this paper, we have tried another approach called multimodels ensemble feature selection. Firstly, several different kinds of feature selection methods are used to evaluate the performance of the training data, and then each generates a subset. Secondly, the single subset acquired before is used to train a model separately. Finally, we aggregate the results of the different models. Figure 1 illustrates the process of multimodels ensemble feature selection. According to the process of multimodels ensemble feature selection, the MmEnFs algorithm is shown below (Algorithm 1).

```
Algorithm: MmEnFs (multi-models ensemble feature selection)
Input: FSM = {FSMᵢ |  i = 1, 2, . . . p}; // p denotes the kinds of feature selection methods.
        £:learning algorithm method
        k:the ratio of selected features
        TrainD = {(xᵢ, yᵢ) | i = 1, 2, . . . m}; // m denotes the number of sample of traindata.
        F = {f₁, f₂,. . . fₙ}; // n denotes the numbers of features.
Output: H(x):Ensemble model
for t = 1, 2,. . .,p do
    Step 1: SubFsᵢ = Feature selector (TrainD, FSMᵢ, k)
    Step 2: SubDᵢ = map (TrainD, SubFsᵢ)
    Step 3: hᵢ(x) = Train (SubDᵢ, £)
end for
```

$$H(x) = \underset{y \in Y}{\operatorname{argmax}} \sum_{t=1}^{p} \mathbb{I}(h_t(x) = y)$$
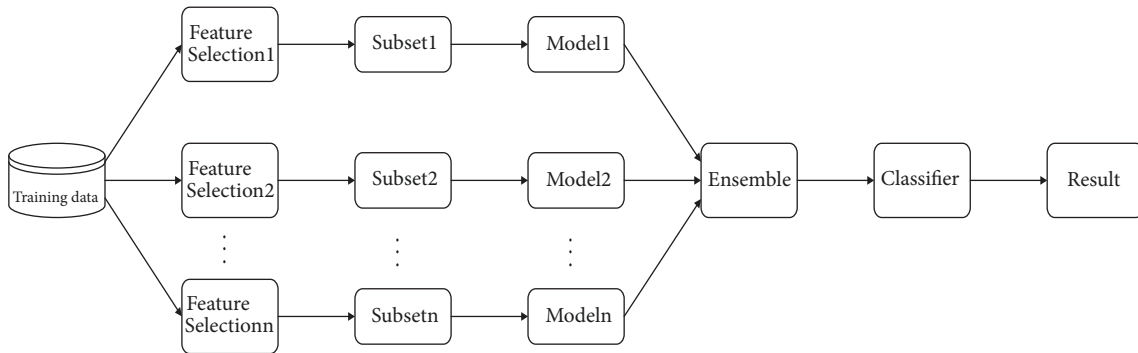
ALGORITHM 1: The proposed MmEnFs algorithm.

FIGURE 1: The process of multimodels ensemble feature selection.

4. Experiments

4.1. Data of Experiments. In this section, environmental sound dataset is selected as experimental data. It is made up of five classes and collected from the wild. It contains different sounds of birds, wind, rain, frogs, and thunder, totally 23 attributes. And 13 dimensionalities of features are involved in MFCC, while other 10 variables are composed of CELP. The details of dataset are shown in Table 1.

4.2. Methods of Experiments. In these experiments, seven feature selection methods tested on the data are derived from the platform of development Weka, open source data mining software, with the default settings. The decision tree is selected as the base classifier. And the whole process of experiments was run on Matlab. By using the bootstrap [34], the dataset was divided into 75% training data and 25% testing data at a ratio of 3:1 approximately. Then seven kinds of feature selection methods were chosen to select the different ratios of subsets to perform the experiment (Tables 2, 4, 5, and 6). If the number of selected attributes is not an integer, it will round down. Because of the instability in constraint score, we have performed the experiment ten times to get the average value. The MmEnFs1 is an ensemble method, which includes six existed methods, namely, Correlation, GainRatio, InfoGain, OneR, ReliefF, and SymmetricalUncert, while MmEnFs2 is

an aggregation of seven methods composed of Correlation, GainRatio, InfoGain, OneR, ReliefF, SymmetricalUncert, and csFs. Besides, as a comparison, we have built a decision tree classification model with the original data (Table 3). And the accuracy of decision tree is 0.686. Here, k denotes the ratio of selected features.

4.3. Result Analysis. According to Table 2 and Figure 2, much useful information can be obtained. Table 2 shows the accuracy of several methods on environmental sound dataset in different ratios. Figure 2 shows the variation tendencies of nine methods on this dataset. In terms of csFs, when 1/2 subsets are selected, it is inferior to most methods and lower than the model of decision tree. However, with the selected subsets increasing, it will get a good performance in an interval ranging from 2/3 features to 3/4 features, compared to other six existed methods in Weka. After that, csFs decreases slightly and is inferior to OneR and InfoGain. Then it reaches its peak value at the end. As for MmEnFs, the performances of MmEnFs1 and MmEnFs2 are both significantly better than that of other seven methods. When 13 features are selected, they both reach their peak values, respectively.

As for Tables 3, 4, 5, and 6, we can see the variation of accuracy on each class at different ratios. In terms of bird, csFs will reach its peak value when 16 features are selected. When k

TABLE 1: Information of dataset in the experiments.

Class	Total Frames	75%-Training	25%-Testing
Bird	800	600	200
Wind	800	600	200
Rain	800	600	200
Frog	800	600	200
Thunder	800	600	200
Total	4000	3000	1000

TABLE 2: Accuracy of different feature selection methods at k of subsets on environmental sound dataset.

Methods	Accuracy		
	k=1/2	k=2/3	k=3/4
Correlation	0.63	0.684	0.677
GainRatio	0.643	0.66	0.685
InfoGain	0.662	0.673	0.672
OneR	0.656	0.662	0.67
ReliefF	0.607	0.65	0.676
SymmetricalUncert	0.66	0.67	0.668
csFs	0.6374	0.6844	0.6899
MmEnFs1	0.734	0.73	0.706
MmEnFs2	**0.756**	**0.747**	**0.7215**

TABLE 3: Accuracy of decision tree at k = 1 of subsets on five classes.

Methods	Bird	Wind	Rain	Frog	Thunder
Decision tree	0.685	0.715	0.705	0.565	0.76

TABLE 4: Accuracy of different feature selection methods at k = 1/2 of subsets on five classes.

Methods	Bird	Wind	Rain	Frog	Thunder
Correlation	0.607	0.7028	0.5915	0.5808	0.6705
GainRatio	0.6124	0.7206	0.5313	**0.6646**	0.703
InfoGain	0.6169	0.7606	0.5936	0.6458	0.705
OneR	0.6418	0.6923	0.5769	0.6433	0.7236
ReliefF	0.5743	0.7321	0.5604	0.5758	0.587
SymmetricalUncert	0.6119	0.7634	0.5909	0.6458	0.7015
csFs	0.6419	0.7012	0.5676	0.5914	0.7005
MmEnFs1	0.765	0.79	0.695	0.655	0.765
MmEnFs2	**0.785**	**0.805**	**0.77**	0.635	**0.785**

TABLE 5: Accuracy of different feature selection methods at k = 2/3 of subsets on five classes.

Methods	Bird	Wind	Rain	Frog	Thunder
Correlation	0.6702	0.7568	0.6802	0.6184	0.702
GainRatio	0.6091	0.7461	0.6207	0.6223	0.7092
InfoGain	0.6382	0.7512	0.6321	0.6117	0.73
OneR	0.6019	0.7143	0.6419	0.6429	0.7172
ReliefF	0.6263	0.7716	0.6207	0.5924	0.6471
SymmetricalUncert	0.6443	0.7463	0.6394	0.5959	0.7206
csFs	0.6311	0.7719	0.657	0.6624	0.7534
MmEnFs1	0.74	0.775	**0.765**	0.62	0.75
MmEnFs2	**0.755**	**0.785**	0.76	**0.665**	**0.7715**

TABLE 6: Accuracy of different feature selection methods at k = 3/4 of subsets on five classes.

| Methods	Bird	Wind	Rain	Frog	Thunder
Correlation	0.6512	**0.7754**	0.6856	0.5829	0.6976
GainRatio	0.6215	0.7634	0.6618	0.64	0.75
InfoGain	0.6275	0.735	0.625	0.6064	**0.7708**
OneR	0.6436	0.7436	0.6186	0.6166	0.7333
ReliefF	0.6567	0.7143	0.648	0.6471	0.7143
SymmetricalUncert	0.63	0.7337	0.6233	0.5907	0.7668
csFs	0.6385	0.7563	0.6837	0.6493	0.7259
MmEnFs1	**0.735**	0.75	0.705	0.595	0.745
MmEnFs2	0.7325	0.7525	**0.7175**	**0.655**	0.75
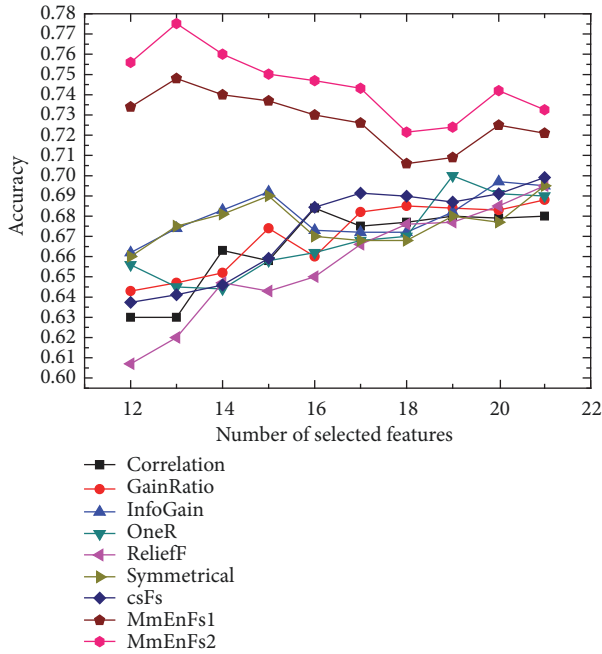
FIGURE 2: Accuracy versus different number of selected features on environmental sound dataset.

= 1/2, MmEnFs1 and MmEnFs2 will get the best performance. To the wind, when k = 1/2 and 2/3, ensemble methods are superior to other methods. When k = 2/3, csFs will have a better accuracy. While k = 3/4, MmEnFs1, MmEnFs2, and csFs are worse than Correlation and GainRatio. As for rain, ensemble methods are better than other methods all the time. However, at k = 2/3, MmEnFs2 is worse than EnFs1. When k = 3/4, csFs will get its perfect performance. For the frog, at k = 1/2, GainRatio will have the best performance. MmEnFs2 and csFs can reach their peak values at k =2/3. With ratio increasing, MmEnFs1 will get worse gradually. In respect of the thunder, ensemble methods can get better accuracy when k = 1/2 and 3/4. However, at k = 3/4, ensemble methods are worse than InfoGain and SymmetricalUncert. Compared to the decision tree with original data, MmEnFs2 has an evidently improvement all the time.

Obviously, the MmEnFs method is superior to the other single method. MmEnFs1 is not as good as MmEnFs2 in

terms of accuracy. It is noteworthy that csFs plays an important role in MmEnFs2. Combined with csFs, the accuracy of MmEnFs2 makes greater progress distinctly. Meanwhile, this result indicates that the aggregation of different feature selection methods will have a great impact on the performance. The MmEnFs achieves the aim which is selecting the less features to gain the better accuracy for the environmental sound dataset.

5. Conclusion

In this paper, we use the csFs based on constraint score to evaluate the performance of different feature subsets on environmental sound dataset. Compared to other six existed feature selection methods, csFs will get a good quality. And then we compare the results of single feature selection method and MmEnFs by varying the ratios of feature subsets. Experiments show that different methods will reach their perfect performance at different ratios or at least maintain the baseline results when enough subsets are selected for ensemble and other approaches. Obviously, in most cases, MmEnFs is superior to the single feature selection method in accuracy.

The advantage of csFs is that it selects the subspace randomly to achieve the assessment of feature subsets. But there still exists a problem that the result is not stable enough. As for the stability, three indices, Pearson's correlation coefficient [35], Spearman's rank correlation coefficient [36], and Tanimoto distance [37], are commonly used to assess the performance of stability thoroughly in recent years. In the future, further research will focus on the stability of csFs to obtain a better performance from the perspective of three indices. Combining the ensemble method, maybe we will make greater progress in the domain of environmental sound recognition.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] C. A. Ruiz-Martinez, M. T. Akhtar, Y. Washizawa, and E. Escamilla-Hernandez, "On investigating efficient methodology for environmental sound recognition," in *Proceedings of the 21st International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS '13)*, pp. 210–214, November 2013.

[2] S. Chu, S. Narayanan, and C.-C. J. Kuo, "Environmental sound recognition with timeFrequency audio features," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 17, no. 6, pp. 1142–1158, 2009.

[3] H. D. Tran and H. Li, "Sound event recognition with probabilistic distance SVMs," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 19, no. 6, pp. 1556–1568, 2011.

[4] B. Uzkent and B. D. Barkana, "Pitch-range based feature extraction for audio surveillance systems," in *Proceedings of the 8th International Conference on Information Technology: New Generations (ITNG '11)*, pp. 476–480, April 2011.

[5] E. Tsau, S. Kim, and C. J. Kuo, "Environmental sound recognition with CELP-based features," in *Proceedings of the 10th International Symposium on Signals, Circuits and Systems (ISSCS '11)*, pp. 1–4, Iasi, Romania, June 2011.

[6] Y. Zhang, D. J. Lv, and H. S. Wang, "The application of multiple classifier system for environmental audio classification," *Applied Mechanics and Materials*, vol. 462-463, pp. 225–229, 2014.

[7] Y. Saeys, T. Abeel, and Y. V. D. Peer, "Robust feature selection using ensemble feature selection techniques," in *Proceedings of the European Conference on Machine Learning & Knowledge Discovery in Databases*, 2008.

[8] B. Seijo-Pardo, I. Porto-Díaz, V. Bolón-Canedo, and A. Alonso-Betanzos, "Ensemble feature selection: Homogeneous and heterogeneous approaches," *Knowledge-Based Systems*, vol. 118, pp. 124–139, 2017.

[9] K. Dunne, P. Cunningham, and F. Azuaje, "Solutions to instability problems with sequential wrapper-based approaches to feature selection," *Journal of Machine Learning Research*, 2002.

[10] T. C. Smith and E. Frank, "Introducing machine learning concepts with WEKA," *Methods in Molecular Biology*, vol. 1418, pp. 353–378, 2016.

[11] S.-Y. Jiang and L.-X. Wang, "Efficient feature selection based on correlation measure between continuous and discrete features," *Information Processing Letters*, vol. 116, no. 2, pp. 203–215, 2016.

[12] M. Modinat, A. Abimbola, B. Abdullateef, and A. Opeyemi, "Gain ratio and decision tree classifier for intrusion detection," *International Journal of Computer Applications*, vol. 126, pp. 975–8887, 2015.

[13] G. Wu and J. Xu, "Optimized approach of feature selection based on information gain," *Computer Engineering & Applications*, vol. 47, pp. 157–161, 2011.

[14] H. Q. Wang, J. J. Chen, X. J. Hou, and K. Guo, "Research on an attribute section method for decision tree," *Journal of Taiyuan University of Technology*, 2011.

[15] M. Robnik-Šikonja and I. Kononenko, "Theoretical and empirical analysis of ReliefF and RReliefF," *Machine Learning*, vol. 53, no. 1-2, pp. 23–69, 2003.

[16] S. S. Kannan and N. Ramaraj, "A novel hybrid feature selection via Symmetrical Uncertainty ranking based local memetic search algorithm," *Knowledge-Based Systems*, vol. 23, no. 6, pp. 580–585, 2010.

[17] I. Iguyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.

[18] P. Zhou, X. Hu, P. Li, and X. Wu, "Online feature selection for high-dimensional class-imbalanced data," *Knowledge-Based Systems*, vol. 136, pp. 187–199, 2017.

[19] H. Liu and L. Yu, "Toward integrating feature selection algorithms for classification and clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 4, pp. 491–502, 2005.

[20] K. Benabdeslem and M. Hindawi, "Efficient semi-supervised feature selection: Constraint, relevance, and redundancy," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 5, pp. 1131–1143, 2014.

[21] C. Ding and H. Peng, "Minimum redundancy feature selection from microarray gene expression data," in *Proceedings of the 2nd International IEEE Computer Society Computational Systems Bioinformatics Conference (CSB '03)*, pp. 523–528, August 2003.

[22] M. Dash and H. Liu, "Feature selection for classification," *Intelligent Data Analysis*, vol. 1, no. 1–4, pp. 131–156, 1997.

[23] Y. Saeys, I. Inza, and P. Larrañaga, "A review of feature selection techniques in bioinformatics," *Bioinformatics*, vol. 23, no. 19, pp. 2507–2517, 2007.

[24] R. C. Holte, "Very simple classification rules perform well on most commonly used datasets," *Machine Learning*, vol. 11, no. 1, pp. 63–91, 1993.

[25] M. Robnik-Sikonja and I. Kononenko, "An adaptation of Relief for attribute estimation in regression," in *Proceedings of the Fourteenth International Conference on Machine Learning*, pp. 296–304, 1997.

[26] T. G. Dietterich, "Ensemble methods in machine learning," in *Proceedings of the 1st International Workshgp on Multiple Classifier Systems*, vol. 1857, pp. 1–15, 2000.

[27] M. Abdelwahab and C. Busso, "Ensemble feature selection for domain adaptation in speech emotion recognition," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '17)*, pp. 5000–5004, March 2017.

[28] A. B. Brahim and M. Limam, "Ensemble feature selection for high dimensional data: a new method and a comparative study," *Advances in Data Analysis and Classification*, vol. 12, no. 4, pp. 937–952, 2018.

[29] D. Guan, W. Yuan, Y.-K. Lee, K. Najeebullah, and M. K. Rasel, "A review of ensemble learning based feature selection," *IETE Technical Review*, vol. 31, no. 3, pp. 190–198, 2014.

[30] Y. Saeys, T. Abeel, and Y. Van De Peer, "Robust feature selection using ensemble feature selection techniques," *Lecture Notes in Computer Science*, vol. 5212, no. 2, pp. 313–325, 2008.

[31] D. Zhang, S. Chen, and Z.-H. Zhou, "Constraint score: a new filter method for feature selection with pairwise constraints," *Pattern Recognition*, vol. 41, no. 5, pp. 1440–1451, 2008.

[32] D. Zhang, Z. Zhou, and S. Chen, "Semi-Supervised Dimensionality Reduction," in *Proceedings of the SIAM International Conference on Data Mining*, pp. 11–393, Minneapolis, MN, USA, April 2007.

[33] A. Bar-Hillel, T. Hertz, N. Shental, and D. Weinshall, "Learning a Mahalanobis metric from equivalence constraints," *Journal of Machine Learning Research*, vol. 6, pp. 937–965, 2005.

[34] R. Duangsoithong and T. Windeatt, "Bootstrap feature selection for ensemble classifiers," *Industrial Conference on Advances in Data Mining: Applications & Theoretical Aspects Springer-Verlag*, 2010.

[35] S. Nogueira and G. Brown, *Measuring the Stability of Feature Selection*, Springer International Publishing, 2016.

[36] Q. Liu, C. Li, V. Wanga, and B. . Shepherd, "Covariate-adjusted Spearman's rank correlation with probability-scale residuals," *Biometrics*, vol. 74, no. 2, pp. 595–605, 2018.

[37] P. Yang, B. B. Zhou, J. Y. Yang, and A. Y. Zomaya, *Stability of Feature Selection Algorithms and Ensemble Feature Selection Methods in Bioinformatics*, John Wiley & Sons, Inc., 2014.