

Research Article

Modified Hybrid Discriminant Analysis Methods and Their Applications in Machine Learning

Liwen Huang ^{1,2}

¹College of Mathematics and Computer Science, Quanzhou Normal University, Quanzhou 362000, Fujian, China

²Key Laboratory of Financial Mathematics (Putian University), Fujian Province University, Putian 351100, Fujian, China

Correspondence should be addressed to Liwen Huang; livern@126.com

Received 17 May 2019; Revised 20 December 2019; Accepted 30 December 2019; Published 24 February 2020

Academic Editor: Seenith Sivasundaram

Copyright © 2020 Liwen Huang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a new hybrid discriminant analysis method, and this method combines the ideas of linearity and nonlinearity to establish a two-layer discriminant model. The first layer is a linear discriminant model, which is mainly used to determine the distinguishable samples and subsample; the second layer is a nonlinear discriminant model, which is used to determine the subsample type. Numerical experiments on real data sets show that this method performs well compared to other classification algorithms, and its stability is better than the common discriminant models.

1. Introduction

Multigroup classification or discrimination is an important problem with application in many fields [1], and Fisher's linear discriminant analysis (FLDA) [2] is a statistical method for multiclass discrimination. Despite being almost 80 years old, it remains to be one of the most widely used methods in light of its simplicity and effectiveness.

At present, there are many alternative methods for classification [3, 4], such as FLDA, Quadratic discriminant analysis, genetic algorithm, artificial neural network approaches, decision trees, and SVM. In fact, using a high number of data sets, it is probable that each classifier will work well in some data sets and not so well in others [5], so no method will be universally superior to other methods [6]. However, as the number of classes increases or the number of variables increases, data sets will become more complex, and there is also a corresponding need for a discriminant model with good stability and applicability. Therefore, it is a good idea to use multiple methods to solve the classification problem in discriminant analysis.

In general, when constructing discriminant methods, various assumptions are often made, which may not be appropriate because of the complexity of the actual data. FLDA has no requirement on the distribution of data sets, which attracts the

interest of many scholars. Hence, FLDA has developed numerous variations for different purposes since its first publication 80 years ago [7–10]. Recently, various methods, including [11–14], have been developed to combine multiple methods to efficiently process complex data or high-dimensional data [15–18]. However, when the data set is more complex, the characteristics of the data may have different relationships in different parts of the data. To solve this problem, Giles et al. [19] used an iterative denoising method to discover the structure and relationship of the data set. Using the linear and nonlinear ideas, Huang and Su [20] proposed a hierarchical discriminant analysis method (HDAM), but this method was not very effective in finding the characteristics of the data. One-hidden-layer neural network is also a linear and nonlinear classification method, but the classification effect is related to the selection of nonlinear functions. SVM is a good classifier, and its effect is related to the choice of kernel function. Generally, some data sets are sensitive to kernel functions, while others are not. Therefore, in order to solve the discriminant problem of more complex data and improve the stability of the model, this paper tries to establish hybrid discriminant analysis based on HDAM and uses an adaptive method to identify the features of the data set.

In the following sections, the paper will discuss the new hybrid discriminant analysis methods, their discriminant criterion, numerical experiments conducted, and its conclusion.

2. Modified Hybrid Discriminant Analysis Methods (MHDA)

2.1. Modified FLDA Criterion. Suppose there are k classes (G_1, G_2, \dots, G_k) , $k > 0$, where $G_i: x_{(1)}^{(i)}, x_{(2)}^{(i)}, \dots, x_{(n_i)}^{(i)}$. Here, n_i is the sample size of the class G_i , $i = 1, 2, \dots, k$. In this paper, assuming that x is an arbitrary given sample with $x = (x_{11}, x_{12}, \dots, x_{1m})'$.

According to the idea of the large between-class difference and the small within-class difference, the goal of FLDA is to find an optimal discriminant vector u by maximizing the following criterion [21, 22]:

$$J(u) = \frac{u' S_b u}{u' S_w u}. \quad (1)$$

Here, $S_b = \sum_{i=1}^k n_i (\bar{x}^{(i)} - \bar{x})(\bar{x}^{(i)} - \bar{x})'$, $S_w = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_j^{(i)} - \bar{x}^{(i)})(x_j^{(i)} - \bar{x}^{(i)})'$, $\bar{x}^{(i)}$ is the mean of G_i , and \bar{x} is the mean of all classes.

Suppose u is the optimal discriminant vectors obtained by the training samples, then the following linear discriminant function can classify the sample type of each class as much as possible:

$$z = u' x. \quad (2)$$

Let C_i be the discriminative threshold between the class G_i and the class G_{i+1} and $C_1 < C_2 < \dots < C_{k-1}$, then the Fisher criterion can be described as follows.

For any given sample x , the value of z can be calculated by formula (2), then

- (1) If $z < C_1$, then $x \in G_1$
- (2) If $C_{i-1} < z < C_i$, then $x \in G_i$, $i = 2, 3, \dots, k-1$
- (3) If $z > C_{k-1}$, then $x \in G_k$

However, depending on the projection direction of u , there may be overlaps among the samples of different classes, which will lead to misclassification of the samples. In order to improve the classification performance, a two-layer discriminant model is established by using the combination of linear and nonlinear discriminant methods. Its main idea can be described as follows: for the first layer, a modified FLDA is constructed to separate the distinguishable samples from each class, and the rest samples are treated as subsamples; for the second layer, a modified nonlinear discriminant method is established to classify the subsample type.

Using formula (2), the projection values of the samples in each class can be calculated. Let $a_i = \min\{z = u' x, x \in G_i\}$ and $b_i = \max\{z = u' x, x \in G_i\}$, then interval (a_i, b_i) is the discriminant range of the class G_i , $i = 1, 2, \dots, k$. However, for the discriminant ranges of different classes, if there is an intersection, then there will be misclassification samples. Therefore, it is necessary to adjust the discriminant ranges of each class so that these discriminant ranges do not intersect each other.

Suppose the new discriminant range of the class G_i is denoted by (a'_i, b'_i) , then for any given two classes, denoted by G_i and G_j , their discriminant ranges will satisfy the following condition:

$$(a'_i, b'_i) \cap (a'_j, b'_j) = \phi, \quad 1 \leq i < j \leq k. \quad (3)$$

Thus, the new linear discriminant criterion can be described by the following form.

For any given sample x , let $z = u' x$, then,

$$x \in \begin{cases} G_i, & z \in (a'_i, b'_i), \\ \text{subsample,} & \text{otherwise.} \end{cases} \quad (4)$$

2.2. Modified Nonlinear Discriminant Criterion. Suppose subsamples still have k classes $(G'_1, G'_2, \dots, G'_k)$, $\bar{x}'^{(i)}$ is the mean of G'_i and n'_i is the number of G'_i , $i = 1, 2, \dots, k$.

For a given sample x , if $x \in G'_i$, then the distance between the sample x and G'_i is defined by the following form:

$$d(x, G'_i) = \sqrt{(x - \bar{x}'^{(i)})' (x - \bar{x}'^{(i)})}. \quad (5)$$

Let $d^{(i)}$ be the maximum distance between the samples of G'_i and $\bar{x}'^{(i)}$, then $d^{(i)} = \max_{1 \leq j \leq n'_i} \left\{ \sqrt{(x_j^{(i)} - \bar{x}'^{(i)})' (x_j^{(i)} - \bar{x}'^{(i)})} \right\}$. Especially, if G'_i is a spherical-shaped class, $d^{(i)}$ can be regard as the radius of G'_i .

To better distinguish the sample type between two spherical-shaped classes, some results were obtained from the perspective of the inclusion or disjoint of the classes, which can be described as follows [20, 23].

Let G'_p be spherical-shaped class or spherical shell-shaped class, and let G'_q be spherical shell-shaped class and $d^{(i)}$ is the radius of G'_i , $i = p, q$.

- (1) Suppose $G'_p \subset G'_q$ and $G'_p \cap G'_q = \phi$, namely, the samples of G'_p are surrounded by the samples of G'_q and there are no cross samples between G'_p and G'_q , which indicates that the relationship between the two classes is inclusive. For any given sample x ,
 - (i) If $x \in G'_p$, then $(d(x, G'_p)/d^{(p)}) < 1$ and $(d(x, G'_q)/d^{(q)}) < 1$
 - (ii) If $x \in G'_q$ and $x \notin G'_p$, then $(d(x, G'_p)/d^{(p)}) < 1$ and $(d(x, G'_q)/d^{(q)}) < 1$
 - (iii) If $x \notin G'_p$ and $x \notin G'_q$, then $(d(x, G'_p)/d^{(p)}) < 1$ and $(d(x, G'_q)/d^{(q)}) < 1$
- (2) Suppose $G'_p \not\subset G'_q$, $G'_q \not\subset G'_p$, and $G'_p \cap G'_q = \phi$, which indicates that the relationship between the two classes is disjoint. For any given sample x ,
 - (i) If $x \in G'_p$, then $(d(x, G'_p)/d^{(p)}) < 1$ and $(d(x, G'_q)/d^{(q)}) > 1$
 - (ii) If $x \in G'_q$, then $(d(x, G'_p)/d^{(p)}) < 1$ and $(d(x, G'_q)/d^{(q)}) < 1$
 - (iii) If $x \notin G'_p$ and $x \notin G'_q$, then $(d(x, G'_p)/d^{(p)}) > 1$ and $(d(x, G'_q)/d^{(q)}) > 1$

The abovementioned two results can be generalized to the case of more than two classes, but this method can only solve the discriminant problem when the relationships among classes are inclusive or disjoint.

However, in practical application, many classes may not be spherical classes or spherical shell-shaped classes, and their relationships are not necessarily inclusive or disjoint, which limit the application of this method. A feasible idea is to divide nonspherical class into several spherical-shaped classes [24–27], and the purpose is to find the feature of each class and improve the classification performance of discriminant problem.

Suppose the class G_p is the original class corresponding to G'_p . In order to better find the feature of G'_p , G_p is divided into k_p spherical-shaped classes, denoted by $B_1^{(p)}, B_2^{(p)}, \dots, B_{k_p}^{(p)}$, $1 \leq k_p < n_p$, and $c_i^{(p)}$ and $d_i^{(p)}$ are the center point and the radius of $B_i^{(p)}$, where $B_i^{(p)} \cap B_j^{(p)} = \phi$, $1 \leq i \neq j \leq k_p$.

Let $Q(x, G_p)$ be $d(x, G_p)/d^{(p)}$, for any given sample x , then

$$Q(x, G_p) = \min_{1 \leq i \leq k_p} \left\{ \frac{d(x, B_i^{(p)})}{d_i^{(p)}} \right\} = \begin{cases} 0 \sim 1, & \text{if } x \in \bigcup_{1 \leq i \leq k_p} B_i^{(p)}, \\ >1, & \text{otherwise.} \end{cases} \quad (6)$$

Thus, the samples of G'_p can be classified by the features of G_p , and using the principle of neighbors, the nonlinear discriminant criterion can be established as follows.

For any given sample x , if $Q(x, G'_i) = \min_{1 \leq p \leq k} (Q(x, G_p))$, then $x \in G'_i$.

2.3. Modified Hybrid Discriminant Analysis Method (MHDA). From Section 2.1 and Section 2.2, MHDA can be briefly described in the following five steps:

- (1) Find the optimal discriminant vector u according to training data
- (2) Determine the discriminant range of each class, denoted by (a'_i, b'_i) , $i = 1, 2, \dots, k$
- (3) Classify the sample type of each class from modified FLDA criterion (Section 2.1) and determine the subsamples of each class
- (4) Decompose each class into several spherical-shaped classes, denoted by $B_1^{(p)}, B_2^{(p)}, \dots, B_{k_p}^{(p)}$, $1 \leq k_p < n_p$, and $c_i^{(p)}$ and $d_i^{(p)}$ are the center point and the radius of $B_i^{(p)}$, where $B_i^{(p)} \cap B_j^{(p)} = \phi$, $1 \leq i \neq j \leq k_p$
- (5) For any given sample x in the subclass, classify its type according to the modified nonlinear discriminant criterion (Section 2.2)

HDAM is an improved discriminant method based on FLDA. This method has advantages of FLDA, and it can deal with the discriminant problem of one class surrounded by the other class. Its computational complexity is $O(k * n * m) \sim O(k * n \log n * m)$. For more extensive application, MHDA is proposed on the basis of HDAM. This method can identify the features of the data set well, but its algorithm runs slower than HDAM. In general, the computational complexity of this method is related to the complexity of the data. If the number of features in the data set is represented as t_n , then its complexity is $O(k * n * m) \sim O(k * n \log n * m * t_n)$.

3. Numerical Examples

To demonstrate the improvements that can be achieved by this method of the paper, nine data sets are derived from UCI Machine Learning Repository [28]; these data sets are Abalone Data set, Balance Scale Data Set, Banknote Authentication Data Set, Breast Tissue Data Set, Cryotherapy Data Set, Iris Data Set, Vehicle Silhouettes Data Set, Vertebral2c Data Set, and Vertebral3c Data Set, respectively, and their basic information is shown in Table 1.

Then, for the nine data sets above, the discriminant model is established by FLDA [29], SDAM [23], Bayes Discriminant analysis method (BDAM) [29], HDAM [20], SVM-Kernel (SVM-Kernel method is derived from SVM-KM-matlab, and learning parameters are set as follows: $c = 1000$, $\lambda = 1e-7$, $\text{kerneloption} = 2$, and $\text{kernel} = \text{"poly"}$, $\text{verbose} = 1$), Ensembles (Ensembles method is derived from Matlab 18, and parameters are set to AdaboostM1 or AdaboostM2) and MHDA, respectively, and their results are given in Table 2.

Generally, FLDA is suitable for the problem of the large between-class difference and the small within-class difference. SDAM and HDAM can achieve good effect for the discriminant problem of spherical-shaped classes or classes with inclusive relations. BDAM has good classification performance on the discriminant problem of multivariate normality and equal class covariance matrix. Since the real-world data are usually more complex, these four methods do not always work well. Results given in Table 2 also illustrate this point and indicate the robustness of these four methods which needs to be improved. Furthermore, from Table 2, although SVM-Kernel and Ensembles are better than MHDA on some data sets, their stability and overall effect are inferior to MHDA.

MHDA is an extension of HDAM, and from Table 2, its classification accuracy ratio is superior to HDAM. However, the results in Table 3 show that the run time of MHDA is generally longer than that of HDAM.

Numerical examples indicate, in some data sets, MHDA does not achieve the best results, but it still has a high accuracy ratio. However, it can be seen from Tables 1 and 2, when the number of classes in multiple data sets is equal, such as Abalone, Balance Scale, and Vertebral3c, the accuracy ratio tends to decrease as the number of variables increases; when the number of variables in multiple data sets is equal, for example, Vertebral2c and Vertebral3c, or Banknote Authentication and Iris, the accuracy ratio tends to decrease as the number of classes increases; when the number of classes and the number of variables in multiple data sets are all equal, such as Balance Scale and Iris, or Cryotherapy and Vertebral2c, the accuracy ratio is usually related to the complexity of the data set. The run time of MHDA is not necessarily related to the number of variables and classes, but it is related to the size of samples, the number of variables, the number of classes, and the complexity of the data set. From the results of the data sets of Abalone, Breast Tissue, and Vehicle Silhouettes, as the number of variables increases or the number of classes

TABLE 1: Information of different data sets.

UCI data set	Number	Variable	Class
Abalone	4177	8	3
Balance scale	625	4	3
Banknote authentication	1372	4	2
Breast tissue	106	9	6
Cryotherapy	90	6	2
Iris	150	4	3
Vehicle silhouettes	846	18	4
Vertebral2c	310	6	2
Vertebral3c	310	6	3

TABLE 2: Accuracy ratio of different methods.

UCI data set	Accuracy ratio (%)						
	FLDA	SDAM	BDAM	HDAM	SVM-Kernel	Ensembles	MHDA
Abalone	52.89	49.03	54.32	48.86	N	53.82	79.79
Balance scale	68.64	52.00	69.28	79.84	99.36	88.16	93.12
Banknote	97.67	75.44	97.67	99.20	100.00	100.00	99.78
Breast tissue	52.83	46.23	72.64	66.98	33.96	57.55	78.30
Cryotherapy	90.00	65.56	90.00	67.78	56.67	100.00	96.67
Iris	98.67	93.33	98.00	98.00	99.30	96.67	98.67
Vehicle silhouettes	63.71	31.00	79.67	30.73	25.65	53.31	77.31
Vertebral2c	80.65	76.77	80.65	76.77	67.74	91.94	94.52
Vertebral3c	68.39	76.45	80.97	73.55	48.39	78.06	93.55

Note. In Table 2, N denotes the exception of the corresponding method.

TABLE 3: Run time of HDAM and MHDA.

UCI data set	Run time (s)	
	HDAM	MHDA
Abalone	1.6853	44.2816
Balance scale	0.0384	0.6347
Banknote	0.0380	0.4328
Breast tissue	0.0121	0.0839
Cryotherapy	0.0264	0.0525
Iris	0.0199	0.0435
Vehicle silhouettes	0.1583	0.7580
Vertebral2c	0.0155	0.2686
Vertebral3c	0.0344	0.1664

increases, the corresponding data sets tends to become more complex, and the performance of MHDA tends to decline. Thus, for the data set of higher dimensions or more classes, the next goal is to improve the classification performance through the variable selection or dimensionality reduction.

4. Conclusions

In this paper, MHDA is proposed based on HDAM, and it combines the ideas of linearity and nonlinearity to establish a two-layer discriminant model. HDAM can better solve the discriminant problem that the relationships among classes are inclusive or disjoint, but the real-world data are usually more complex, which limits its wide application. MHDA overcomes these shortcomings and improves the accuracy ratio and robustness.

Numerical experiments show that MHDA works well for the real data sets, and its accuracy ratio and stability are better

than some common discriminant methods. However, for the data set of higher dimensions or more classes, the effect achieved by MHDA needs to be further improved. In the future, the method will need to be further modified so that it can be better suit for the discrimination problem of such data sets.

Data Availability

The data sets used to support the findings of this study have been deposited in UCI machine learning repository, <http://archive.ics.uci.edu/ml/datasets>, and the Vehicle Silhouettes data used to support the findings of this study are included within the supplementary information file. Abalone: <http://archive.ics.uci.edu/ml/datasets/Abalone>. Balance scale: <http://archive.ics.uci.edu/ml/datasets/Balance+Scale>. Banknote authentication: <http://archive.ics.uci.edu/ml/datasets/banknote+authentication>. Breast Tissue: <http://archive.ics.uci.edu/ml/datasets/Breast+Tissue>. Cryotherapy: <http://archive.ics.uci.edu/ml/datasets/Cryotherapy+Dataset+>. Iris: <http://archive.ics.uci.edu/ml/datasets/Iris>. Vehicle Silhouettes: [http://archive.ics.uci.edu/ml/datasets/Statlog+\(Vehicle+Silhouettes\)](http://archive.ics.uci.edu/ml/datasets/Statlog+(Vehicle+Silhouettes)). Vertebral2c and Vertebral3c: <http://archive.ics.uci.edu/ml/datasets/Vertebral+Column>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Key Laboratory of Financial Mathematics of Fujian Province University (Putian University) (no. JR201801).

Supplementary Materials

(1) The discriminant results of the abalone data set. (2) The discriminant results of the balance scale data set. (3) The discriminant results of the banknote authentication data set. (4) The discriminant results of the breast tissue data set. (5) The discriminant results of the cryotherapy data set. (6) The discriminant results of the iris data set. (7) The discriminant results of the vehicle silhouettes data set. (8) The discriminant results of the vertebral2c data set. (9) The discriminant results of the vertebral3c data set Run time of HDAM and MHDA: (1) Run time of the abalone data set; (2) Run time of the balance scale data set; (3) Run time of the banknote authentication data set; (4) Run time of the breast tissue data set; (5) Run time of the cryotherapy data set; (6) Run time of the iris data set; (7) Run-time of the vehicle silhouettes data set; (8) Run time of the vertebral2c data set; (9) Run time of the vertebral3c data set. (*Supplementary Materials*)

References

- [1] T. Hastie, R. Tibshirani, and A. Buja, "Flexible discriminant analysis by optimal scoring," *Journal of the American Statistical Association*, vol. 89, no. 428, pp. 1255–1270, 1994.
- [2] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Human Genetics*, vol. 7, no. 2, pp. 179–188, 1936.
- [3] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, 2000.
- [4] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: with Applications in R*, Springer, New York, NY, USA, 2013.
- [5] M. Fernandez-Delgado, E. Cernadas, S. Barro, and D. Amorim, "Do we need hundreds of classifiers to solve real world classification problems?," *Journal of Machine Learning Research*, vol. 15, pp. 3133–3181, 2014.
- [6] D. J. Hand, "Classifier technology and the illusion of progress," *Statistical Science*, vol. 21, no. 1, pp. 1–14, 2006.
- [7] A. Sierra, "High-order Fisher's discriminant analysis," *Pattern Recognition*, vol. 35, no. 6, pp. 1291–1302, 2002.
- [8] J. Yang, A. F. Frangi, and J.-Y. Yang, "A new kernel Fisher discriminant algorithm with application to face recognition," *Neurocomputing*, vol. 56, pp. 415–421, 2004.
- [9] W. Zheng, C. Zou, and L. Zhao, "Weighted maximum margin discriminant analysis with kernels," *Neurocomputing*, vol. 67, pp. 357–362, 2005.
- [10] E. K. Tang, P. N. Suganthan, X. Yao, and A. K. Qin, "Linear dimensionality reduction using relevance weighted LDA," *Pattern Recognition*, vol. 38, no. 4, pp. 485–493, 2005.
- [11] K. F. Lam and J. W. Moy, "Combining discriminant methods in solving classification problems in two-group discriminant analysis," *European Journal of Operational Research*, vol. 138, no. 2, pp. 294–301, 2002.
- [12] Z. Halbe and M. Aladjem, "Model-based mixture discriminant analysis-an experimental study," *Pattern Recognition*, vol. 38, no. 3, pp. 437–440, 2005.
- [13] C. Nazif and H. Erol, "A new per-field classification method using mixture discriminant analysis," *Journal of Applied Statistics*, vol. 39, no. 10, pp. 2129–2140, 2012.
- [14] M. J. Brusco, C. M. Voorhees, R. J. Calantone, M. K. Brady, and D. Steinley, "Integrating linear discriminant analysis, polynomial basis expansion, and genetic search for two-group classification," *Communications in Statistics—Simulation and Computation*, vol. 48, no. 6, pp. 1623–1636, 2019.
- [15] X. Wang and H. Wang, "Classification by evolutionary ensembles," *Pattern Recognition*, vol. 39, no. 4, pp. 595–607, 2006.
- [16] S. W. Ji and J. P. Ye, "Generalized linear discriminant analysis: a unified framework and efficient model selection," *IEEE Transactions On Neural Networks*, vol. 19, no. 10, pp. 1768–1782, 2008.
- [17] P. T. Pepler, D. W. Uys, and D. G. Nel, "Discriminant analysis under the common principal components model," *Communications in Statistics—Simulation and Computation*, vol. 46, no. 6, pp. 4812–4827, 2017.
- [18] Y. Wang and X. Wang, "Classification using semiparametric mixtures," *Journal of Applied Statistics*, vol. 46, no. 11, pp. 2056–2074, 2019.
- [19] K. E. Giles, M. W. Trosset, D. J. Marchette, and C. E. Priebe, "Iterative denoising," *Computational Statistics*, vol. 23, no. 4, pp. 497–517, 2008.
- [20] L. Huang and L. Su, "Hierarchical discriminant analysis and its application," *Communications in Statistics—Theory and Methods*, vol. 42, no. 11, pp. 1951–1957, 2013.
- [21] S. Chen and X. Yang, "Alternative linear discriminant classifier," *Pattern Recognition*, vol. 37, no. 7, pp. 1545–1547, 2004.
- [22] S. Chen and D. Li, "Modified linear discriminant analysis," *Pattern Recognition*, vol. 38, no. 3, pp. 441–443, 2005.
- [23] L. W. Huang, "Improvement distance discriminant analysis method," *Journal of Jiangnan University*, vol. 10, no. 6, pp. 745–748, 2011.
- [24] T. Hastie and R. Tibshirani, "Discriminant adaptive nearest neighbor classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 6, pp. 607–616, 1996.
- [25] J. DeVinney, C. E. Priebe, D. J. Marchette, and D. Socolinsky, "Random walks and catch digraphs in classification," in *Proceedings of the 34th Symposium on the Interface: Computing Science and Statistics*, vol. 34, pp. 1–10, Montreal, Canada, April 2002.
- [26] D. J. Marchette, E. J. Wegman, and C. E. Priebe, "Fast algorithms for classification using class cover catch digraphs," *Handbook of Statistics*, vol. 24, pp. 331–358, 2005.
- [27] D. Marchette, "Class cover catch digraphs," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 2, pp. 171–177, 2010.
- [28] D. Dua and E. Karra Taniskidou, *UCI Machine Learning Repository*, University of California, School of Information and Computer Science, Irvine, CA, USA, 2017, <http://archive.ics.uci.edu/ml>.
- [29] Y. T. Zhang and K. T. Fang, *Introduction to Multivariate Statistical Analysis*, Wuhan university press, Wuhan, China, 2013.