

An adaptive approach of Feature Selection applied to Semi-Supervised Fuzzy Clustering

Wei Cai

School of Applied Mathematics
Guangdong University of
Technology
Guangzhou, China
caiwei-email@qq.com

Shengbing Xu[†]

School of Applied Mathematics
Guangdong University of
Technology
Guangzhou, China

[†]Corresponding Author. Tel: +8620
87081182; fax: +8620 87081182

Jiongzhi Liu

School of Management
Guangdong University of
Technology
Guangzhou, China
hdifhd@qq.com

Qingping Du

School of Environmental Science
and Engineering
Guangdong University of
Technology
Guangzhou, China
QPDU2008@126.com

Hefeng Chen

School of Applied Mathematics
Guangdong University of
Technology
Guangzhou, China
chenhf@gdut.edu.cn

Yinyun Lin

School of Applied Mathematics
Guangdong University of
Technology
Guangzhou, China
liny00@qq.com

ABSTRACT

Label information in the corresponding semi-supervised fuzzy clustering cannot be used efficiently due to feature redundancy. To address the problem, we propose an adaptive approach of feature selection applied to semi-supervised fuzzy clustering. There are three phases in our approach: 1) feature-score by fisher-score; 2) Min Mean Square Error of Feature Select criterion; 3) the number of features is selected by Min Mean Square Error of Feature Select criterion. We apply our approach to three semi-supervised fuzzy clustering methods. Experiments show that the adaptive approach of feature selection applied to semi-supervised fuzzy clustering can improve the clustering performance.

CCS CONCEPTS

• Computing methodologies~Machine learning~Learning settings~Semi-supervised learning settings

KEYWORDS

Adaptive Feature Selection, Semi-supervised Clustering, Fisher Score, Fuzzy Clustering

1 Introduction

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
EITCE 2020, November 6–8, 2020, Xiamen, China
© 2020 Association for Computing Machinery.
ACM ISBN 978-1-4503-8781-1/20/11...\$15.00
<https://doi.org/10.1145/3443467.3443843>

In many application scenarios, it is quite difficult for us to obtain labeled data information. In contrast to supervised learning algorithms which need a big amount of labeled data information [1][2][3][4], it's a good choice for us to use semi-supervised learning algorithms which only need a small amount of labeled data information [5].

Semi-supervised fuzzy clustering, which is widely used in pattern recognition, computer vision, speech recognition and other fields [6][7][8], is an important research topic based on fuzzy clustering in semi supervised learning [9]. Compared with unsupervised fuzzy clustering, which can't utilize the labeled data information, it can effectively utilize a small amount of labeled data information, such as: semi-supervised fuzzy C-means clustering (SFCM) [10], improved semi-supervised fuzzy C-means clustering algorithm (sSFCM) [11] and semi-supervised entropy regularized fuzzy c-means clustering (eSFCM) [12]; compared with hard clustering such as HCM and k-means [13][14], its corresponding algorithm sufficiently reflect the description of real-world data due to the memberships.

There are three kind of semi-supervised information which are utilized in semi-supervised fuzzy clustering: 1) labeled data information; 2) prior membership information; 3) pairwise-constraints information [15]. In contrast, the labeled data information is very common offered information can be used in semi-supervised clustering algorithm. However, in some real-world scenarios, there are redundancy features in some data. Feature-Redundancy leads to the lack of robustness and data interpretability, and makes the corresponding learning algorithms more prone to over fitting [16]. In general, there are two kind of approaches to solve the problem: Principal Component Analysis (PCA) and Feature-Selection.

PCA dimensionality reduction [17] proposed by Zhang Daoqiang, Zhi-Hua Zhou and other scholars is used to semi-supervised learning and PCA [18] achieves the purpose of

dimensionality reduction through non-negative matrix factorization [18]. PCA after dimension reduction is not the corresponding features, so the interpretability of PCA is not very strong. By contrast, feature selection can effectively overcome this kind of the interpretability problem [19-24]. Yet these approaches of feature selection can't still select adaptively features in clustering process.

Based on the above consideration, an adaptive feature selection approach applied to semi-supervised fuzzy clustering algorithm is proposed in our paper. We focus on improving the saliency of features through adaptive feature selection in order to realize fast and effective semi-supervised fuzzy clustering. In short, the main work of this paper could be summarized as the following:

- 1) We introduce semi-supervised fuzzy clustering and feature selection algorithms.
- 2) We construct adaptive selection approach by referring to Maximum inter-class Variance Method.
- 3) Adaptive feature selection is introduced into semi supervised fuzzy clustering.
- 4) We use data from UCI datasets and gesture datasets to build up numerical experiments to evaluate the performance of adaptive feature selection approach applied to clustering.

2 Related work

2.1 Semi-Supervised Fuzzy Clustering Algorithm

2.1.1 SFCMSFCM[10] is a semi-supervised fuzzy clustering method which makes the best of a small amount labeled data information. Given a sample dataset $X = \{x_1, x_2, \dots, x_n\}$ where $x_j \in \mathbb{R}^d (1 \leq j \leq n)$, SFCM groups X into c clusters by minimizing the following objective function:

$$J(U, V) = \sum_{i=1}^c \sum_{j=1}^n (u_{ij})^m \|x_j - v_i\|^2 + \alpha \sum_{i=1}^c \sum_{j=1}^n (u_{ij} - f_{ij} b_j)^m \|x_j - v_i\|^2 \quad (1)$$

Where $m(m > 1)$ denotes the degree of fuzziness, as m tends towards 1, SFCM approaches HCM. $\|\cdot\|$ indicates the Euclidean distance, $u_{ij} (0 \leq u_{ij} \leq 1)$ is the membership degree of the j -th sample x_j belonging to the i -th cluster whose centroid is $v_i, U = (u_{ij}), V = [v_1, v_2, \dots, v_c], 1 \leq i \leq c, 1 \leq j \leq n, 2 \leq c < n, u_j = (u_{1j}, \dots, u_{cj})^T$ and u_{ij} satisfies the following constraint condition:

$$\sum_{i=1}^c u_{ij} = 1, u_{ij} \geq 0 \quad (2)$$

m is the weighted index, which is an empirical value, the value is usually 2. Then we get the iterative formula of membership degree and clustering center:

$$u_{ij} = \frac{1}{(1+\alpha)} \left[\frac{1+\alpha(1-b_j \sum_{i=1}^c f_{ij})}{\sum_{i=1}^c \frac{1}{\|x_j - v_i\|^2}} + \alpha f_{ij} b_j \right] \quad (3)$$

$$v_i = \frac{\sum_{j=1}^n u_{ij}^2 x_j}{\sum_{j=1}^n u_{ij}^2} \quad (4)$$

where f_{ij} represents the true membership matrix of the labeled data information; b_j is a Boolean value, which equals 1 when sample x_j has a label; α is a predetermined suppression coefficient.

2.1.2 sSFCM Compared with SFCM, sSFCM[11] constructs two membership degree matrices: u_{ij}^L with labeled data information and u_{ij}^U with unlabeled data information. In order to adjust the influence of supervised information, sSFCM puts forward the parameter representing the weight of labeled data information points in the cluster center. Its objective function is the same as the objective function of SFCM. The iterative formulas of its membership matrix and clustering center are as follows:

$$u_{ij}^L = \frac{1}{(1+\alpha)} \left[\frac{1}{\sum_{k=1}^c \frac{1}{\|x_j - v_k\|^m}} + \alpha f_{ij} \right] = \frac{1}{1+\alpha} u_{ij}^{fcm} + \frac{\alpha}{1+\alpha} f_{ij} \quad (5)$$

$$u_{ij}^U = \frac{\left(\frac{1}{\|x_j - v_i\|^2} \right)^2}{\sum_{k=1}^c \left(\frac{1}{\|x_j - v_k\|^2} \right)^2} = u_{ij}^{fcm} \quad (6)$$

$$v_i = \frac{\sum_{x_k \in X^U} (\alpha u_{ij}^L) x_k + \sum_{x_k \in X^U} (u_{ij}^U) x_k}{\sum_{x_k \in X^U} (\alpha u_{ij}^L)^2 + \sum_{x_k \in X^U} (u_{ij}^U)^2} \quad (7)$$

2.1.3 eSFCMeSFCM [12] is a semi-supervised fuzzy clustering algorithm that utilize information entropy to express the relationship of the membership matrix and the prior membership matrix. Its objective function is as follows:

$$J(U, V) = \sum_{i=1}^c \sum_{j=1}^n u_{ij} \|x_j - v_i\|^2 + \lambda^{-1} \sum_{i=1}^c \sum_{j=1}^n (|u_{ij} - f_{ij} b_j| \ln |u_{ij} - f_{ij} b_j|) \quad (8)$$

$$v_i = \frac{\sum_{j=1}^n u_{ij} x_j}{\sum_{j=1}^n u_{ij}} \quad (9)$$

$$e^{-\lambda y_i} = \frac{1 - \sum_{i=1}^c u_{ij}}{e^{-1} \sum_{i=1}^c e^{-\lambda \|x_j - v_i\|^2}} \quad (10)$$

$$u_{ij} = f_{ij} + \frac{e^{-\lambda \|x_j - v_i\|^2}}{\sum_{i=1}^c e^{-\lambda \|x_j - v_i\|^2}} (1 - \sum_{i=1}^c f_{ij} b_j) \quad (11)$$

Among them, λ is an empirical parameter.

2.2 Fisher Score

Feature selection is used to select effective features and remove redundancy features. Fisher Score is a widely used method of feature selection. Given a data set $\{x_j, y_j\}, j = 1, 2, \dots, n$, where $x_j \in \mathbb{R}^d$ is the data feature, $y_j \in \{1, 2, \dots, c\}$ is the samples label, $\text{tr}(\cdot)$ represents the trace of the matrix. Then the Fisher Score of X is defined as:

$$FS(X) = \text{tr}\{(S_t)^{-1} S_b\} \quad (12)$$

between-class scatter matrix:

$$S_b = \sum_{i=1}^c n_i (v_i - v)(v_i - v)^T \quad (13)$$

total scatter matrix:

$$S_t = \sum_{j=1}^n (x_j - v)(x_j - v)^T \quad (14)$$

where n_i is the number of the i -th class data; $\{i\}$ is the set of the i -th class data subscripts; $v_i = \frac{1}{n_i} \sum_{j \in \{i\}} x_j$, which represents the

center of the i -th class data; $v = \frac{1}{n} \sum_{j=1}^n X_j$, which represents the center of all data.

3 An adaptive approach of Feature Selection applied to Semi-Supervised Fuzzy Clustering

3.1 Feature redundancy

Some data [25] have feature redundancy as shown in Figure 1. There are two clusters in the samples which are shown in Figure 1, and the two clusters can be distinguished by f_1 , while can't be distinguished by f_2 and f_3 . Existing research and experiments have shown that the data after feature selection will not weaken the representativeness of the data. As a method of dimension reduction and data quality improvement, feature selection can improve clustering performance and efficiency. From the perspective of clustering, deleting irrelevant features will not have any negative impact on the accuracy of clustering, and can reduce the storage-space and consuming-time.

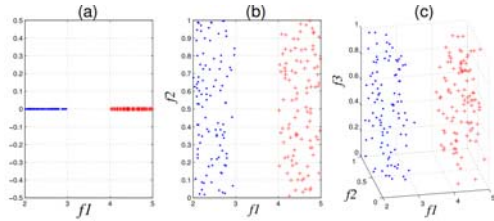


Figure 1: Why feature selection is needed before clustering

Based on the above discussion, an adaptive feature selection approach for semi-supervised fuzzy clustering is proposed.

3.2 An adaptive approach of feature selection

3.2.1 Feature-Score by Fisher-Score The importance of features can be ranked by the values which are calculated by Fisher Score method. We can distinguish invalid or even interfere with category information by Fisher Score method [26], so selecting an appropriate number of effective features is beneficial to improve the clustering effect to a greater extent. Therefore, the effective number of features varies data to data and needs to be selected for the corresponding data. For retaining features that have major impact on the labels, it is necessary to select features with larger Fisher-Score; in addition, the retention of insignificant features should be avoided as far as possible, so referring to the idea of Maximum inter-class Variance Method[27], there are two kind of features in the data: one is have a greater impact on the labeled data and one is have a lesser impact, and retain the features of the former category to complete the adaptive selection of feature number.

3.2.2 Min Mean Square Error of Feature Select criterion In order to determine the optimal number of retained features, we construct Mean Square Error of Feature Select index. First, features are sorted by descending order of Fisher-Score: $S = \{s_1, s_2, s_3 \dots s_d\}$. Secondly, in order to evaluate performance of feature selection, we

establish Mean Square Error of Feature Selection (MSEFS) index as follow:

$$\text{MSEFS}(t) = \sum_{i=1}^d \|s_i - s_t\|^2, 1 \leq t \leq d \quad (15)$$

On this basis, Min MSEFS criterion is established in our paper. When MSEFS of the objective function value is the smallest, t is the adaptive retained feature number.

3.2.3 The number of features is selected by Min Mean Square Error of Feature Select criterion The optimal number of retained features can be determined by Min Mean Square Error of Feature Select criterion:

$$\arg \min_t \text{MSEFS}(t) = \sum_{i=1}^d \|s_i - s_t\|^2, 1 \leq t \leq d \quad (16)$$

Then the corresponding retained features can be determined as follow:

$$s_1, s_2, s_3 \dots s_{t^*}$$

Where t^* is the optimal number of retained features by (16).

For example (as shown in Figure 2), when $d = 4$ and $t^* = 2$, it indicates that MSEFS value is the smallest, so the number of adaptive retained features is 2.

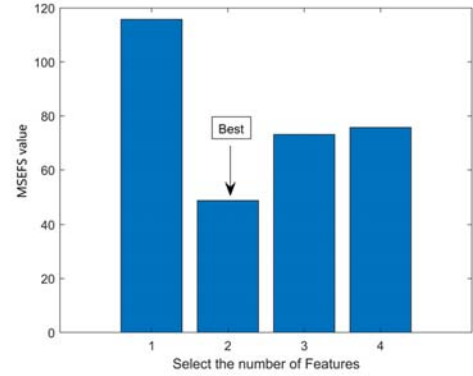


Figure 2: Min MSEFS criterion

3.3 An adaptive approach of feature selection applied in semi-supervised fuzzy clustering

Based on the analysis above, we propose to develop an adaptive approach of feature selection apply in semi-supervised fuzzy clustering:

Input: data feature $X^{d \times n} = (x_1, x_2, \dots, x_n)$, labeled data $L_1 = (l_1, l_2, \dots, l_p)$, unlabeled data $L_2 = (l_{p+1}, l_{p+2}, \dots, l_n)$.

Method:

1. Sort features by descending order of Fisher-Score;
2. Determine the optimal number of retained features by (16);
3. Given the number of clustering centers c , error-threshold E , suppression coefficient α , weighted power exponent m , currently iterative order $t=1$;
4. initialize: the clustering center v_i ;
5. Calculate u_{ij} on fixing v_i by (3), (5)(6), (11);
6. According to the clustering center iteration formula, update v_i by (4),(7),(9);

7. Calculate the objective function by (1),(8). When $t > 2$, calculate $e = J(t) - J(t - 1)$. If $e < E$, the algorithm ends; otherwise $t = t + 1$ and go back to step 5.

Output: Membership matrix U, clustering center V

4 Experiment

In this section, we present an experimental study to evaluate the adaptive approach of feature selection applied to semi-supervised fuzzy clustering.

4.1 Selection of Experimental Data

Table 1: Dataset information

Dataset	Number of Sample	Number of features	Cluster number
Iris	150	4	3
Wine	178	13	3
Wisconsin	569	30	2
Gesture Image	2062	4096	10

4.2 Parameter Setting and Initialization

According to [28], let $\alpha = 5$, $e = 10^{-5}$, the maximum number of iterations is 1000, $\lambda = 10$, and randomly initialize cluster centers v_i .

4.3 Evaluation Measures

There are many evaluation indices in fuzzy clustering [29]. In this algorithm, the partition coefficient V_{PC} , the partition entropy V_{PE} and xie-Beni index V_{XB} are selected as clustering performance evaluation indices.

$$V_{PC} = \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^N u_{ij}^2$$

$$V_{PE} = \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^N u_{ij}^2 \ln u_{ij}$$

$$V_{XB} = \frac{\frac{1}{N} \sum_{i=1}^c \sum_{j=1}^N u_{ij}^2 \|x_j - v_i\|^2}{\min_{i \neq j} (\|v_j - v_i\|^2)}$$

The above evaluation indices are often used for image segmentation based on fuzzy clustering. The partition coefficient V_{PC} and the partition entropy V_{PE} can reflect the fuzzy degree of membership matrix. The greater the value of V_{PC} , the less fuzziness of membership matrix is and the better the clustering effect is. The smaller the value of V_{PE} , the more accurate the classification of each element and the better the clustering effect. The smaller the value of V_{XB} , the better the clustering effect.

4.4 Algorithm Performance

4.4.1 Performance of the algorithm on clustering data There are three algorithms: SFCM, sSFCM and eSFCM algorithms as the comparison experimental algorithms and compare the effects of each algorithm on different percentage of labeled samples. The clustering algorithms with adaptive feature selection approach are named FS-SFCM, FS-sSFCM and FS-eSFCM. The experimental results on different datasets are shown in the following Figure 3 and Table 2.

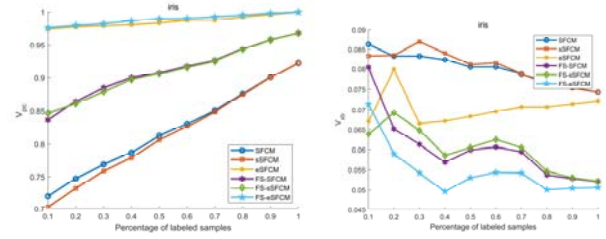


Figure 3: Experimental results of dataset iris

Table 2: Performance comparison of algorithms: 10% labeled samples

dataset	Algorithm	Vpc	Vpe	Vxb	Time(s)
Iris	SFCM	0.7166	0.5105	0.0875	0.0313
	sSFCM	0.6955	0.5403	0.0935	0.0203
	eSFCM	0.9708	0.0475	0.0650	0.0281
	FS-SFCM	0.8408	0.2978	0.0799	0.0344
	FS-sSFCM	0.8449	0.2890	0.0799	0.0234
	FS-eSFCM	0.9783	0.0368	0.0545	0.0094
Wine	SFCM	0.5232	0.8225	0.1237	0.0563
	sSFCM	0.5262	0.8169	0.1046	0.0328
	eSFCM	0.9852	0.0260	0.0756	0.0500
	FS-SFCM	0.6607	0.6138	0.1161	0.0187
	FS-sSFCM	0.6543	0.6216	0.1186	0.0328
	FS-eSFCM	0.9831	0.0287	0.0973	0.0516
Wisconsin	SFCM	0.8421	0.2635	0.1740	0.0391
	sSFCM	0.8379	0.2699	0.1979	0.0250
	eSFCM	0.9915	0.0133	0.1366	0.0437
	FS-SFCM	0.8625	0.2348	0.1069	0.0422
	FS-sSFCM	0.8696	0.2235	0.0999	0.0250
	FS-eSFCM	0.9953	0.0081	0.0886	0.0250

4.4.2 Performance of the algorithm on image data The gesture dataset1 has a total of 2062 gray images with a size of 64×64 , which are divided into 10 categories, as shown in Figure 4:



Figure 4: Ten tags of gesture dataset

¹From: <http://sykv.cn/m/view.php?aid=15498>

Each image is regarded as a sample, and stretched into 4096 features, then we start feature-selecting process and clustering process.

In this dataset, the maximum number of iterations is 10, and the value of λ in eSFCM is 1.

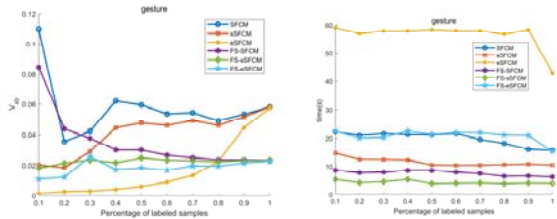


Figure 5: Experimental results of dataset gesture image

The result in Figure 5 shows that the clustering by using adaptive feature selection approach has better performance.

5 Conclusion

Removing redundancy features can avoid over fitting, and reducing the number of features has contributed to improving the robustness and interpretability of the model. Therefore, this paper proposes an adaptive feature selection approach for semi-supervised fuzzy clustering, which can get better clustering results by removing redundancy features.

ACKNOWLEDGMENTS

This work was supported in part by the Natural Science Foundation of China under Grant 61876044 and Grant 61672169, in part by the Scientific and Technological Innovation Strategy Special Fund Project of Guangdong Province under Grant pdjh2020b0182, in part by the Innovation Training Project of Guangdong University of Technology under Grant xj202011845507.

REFERENCES

- [1] Caruana R, Niculescu-Mizil A. An empirical comparison of supervised learning algorithms[C]//Proceedings of the 23rd international conference on Machine learning. 2006: 161-168.
- [2] Weinberger K Q . Distance Metric Learning for Large Margin Nearest Neighbor Classification[J]. Jmlr, 2009, 10.
- [3] Suykens J A K, Vandewalle J. Least squares support vector machine classifiers[J]. Neural processing letters, 1999, 9(3): 293-300.
- [4] Krizhevsky A , Sutskever I , Hinton G . ImageNet Classification with Deep Convolutional Neural Networks[J]. Advances in neural information processing systems, 2012, 25(2).
- [5] Zhu X, Goldberg A B. Introduction to semi-supervised learning[J]. Synthesis lectures on artificial intelligence and machine learning, 2009, 3(1): 1-130.
- [6] Mostafa J, Lam W. Automatic classification using supervised learning in a medical document filtering application[J]. Information Processing & Management, 2000, 36(3): 415-444.
- [7] Balcan M F, Blum A, Choi P P, et al. Person identification in webcam images: An application of semi-supervised learning[C]//ICML 2005 Workshop on Learning with Partially Classified Training Data. 2005, 2: 6.
- [8] Lin B. Online semi-supervised learning in contextual bandits with episodic reward[J]. arXiv preprint arXiv:2009.08457, 2020.
- [9] Bezdek J C , Ehrlich R , Full W . FCM: The fuzzy c-means clustering algorithm[J]. Computers & Geosciences, 1984, 10(2-3):191-203.

- [10] FAN Jiulun. A brief overview on suppressed fuzzy C-means clustering[J]. JOURNAL OF XI'AN UNIVERSITY OF POSTS AND TELECOMMUNICATIONS, 2014, 019(003):1-5.
- [11] Grira N , Crucianu M , Boujemaa N . Semi-Supervised Fuzzy Clustering with Pairwise-Constrained Competitive Agglomeration[C]// The 14th IEEE International Conference on Fuzzy Systems, 2005. FUZZ '05. IEEE, 2005.
- [12] Yasunori E , Yukihiro H , Makito Y , et al. On semi-supervised fuzzy c-means clustering.[C]// IEEE International Conference on Fuzzy Systems. IEEE, 2009.
- [13] Bora D J , Gupta A K . A Comparative study Between Fuzzy Clustering Algorithm and Hard Clustering Algorithm[J]. International Journal of Emerging Trends & Technology in Computer ence, 2014, 10(2):108-113.
- [14] Pawan K , Deepika S . Comparative analysis of FCM and HCM algorithm on Iris data set[J]. International Journal of Computer Applications, 2010, 5(2):33-37.
- [15] Śmieja M, Struski Ł, Figueiredo M A T. A classification-based approach to semi-supervised fuzzy clustering with pairwise constraints[J]. Neural Networks, 2020.
- [16] Jung D. Distributed Feature Selection for Multi-Class Classification Using ADMM[J]. IEEE Control Systems Letters, 2020, 5(3): 821-826.
- [17] Zhang, Daoqiang, Zhi-Hua Zhou, and Songcan Chen. "Semi-supervised dimensionality reduction." Proceedings of the 2007 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, 2007.
- [18] Lee, Hyekyoung, Jiho Yoo, and Seungjin Choi. "Semi-supervised nonnegative matrix factorization." IEEE Signal Processing Letters 17.1 (2009): 4-7.
- [19] Benabdeslem, Khalid, and Mohammed Hindawi. "Constrained laplacian score for semi-supervised feature selection." Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, Berlin, Heidelberg, 2011.
- [20] Gu, Quanquan, Zhenhui Li, and Jiawei Han. "Generalized fisher score for feature selection." arXiv preprint arXiv:1202.3725 (2012).
- [21] Zhao, Jidong, Ke Lu, and Xiaofei He. "Locality sensitive semi-supervised feature selection." Neurocomputing 71.10-12 (2008): 1842-1849.
- [22] Sugiyama, Masashi, et al. "Semi-supervised local Fisher discriminant analysis for dimensionality reduction." Machine learning 78.1-2 (2010): 35.
- [23] Zhang, Daoqiang, Songcan Chen, and Zhi-Hua Zhou. "Constraint score: A new filter method for feature selection with pairwise constraints." Pattern Recognition 41.5 (2008): 1440-1451.
- [24] CHEN Cheng, LIU Zhen-Yu . "Optimized PSO-FCM Cluster Algorithm Based on Principal Component Analysis ." Computer Systems & Applications 29.3 (2020): 213-217.
- [25] Pal N R , Bezdek J C . On cluster validity for the fuzzy c-means model[J]. IEEE Transactions on Fuzzy Systems, 2002, 3(3):370-379.
- [26] Koji Tsuda. Clustering with the Fisher Score[C]// International Conference on Neural Information Processing Systems. MIT Press, 2002.
- [27] DU Qi, XIANG Jianyong, YUAN Shengchun. An Improved Method of Maximum Classes Square Error[J]. Infrared Technology, 2003, 25(5):33-36.
- [28] Li C , Liu L , Jiang W . Objective function of semi-supervised Fuzzy C-Means clustering algorithm[C]// IEEE International Conference on Industrial Informatics. IEEE, 2008.
- [29] MENG Linghui, HU Chunhuan. Cluster Validity Index Based on Measure of Fuzzy Partition[J]. Computer Engineering, 2007, 33(011):15-17.