Acta Technica *62* (2017), No. 6B, 1251–1258

Validity evaluation of Ideological and political education in Colleges Based on fuzzy TOPSIS

Luyu Huang[1]

Abstract. As for the problem that data set structure information is not used and the problem of excessive removal of the noisy points in traditional clustering validity evaluation function and others, a new clustering validity evaluation function is proposed. The function is constituted by tightness measure and separation measure; add distance function in tightness measure to represent geometric structure of data set, to avoid the incompleteness brought by the single theory to the evaluation; to set distance critical value L for mutual constraint with the original membership critical value T in the separation measure, to reduce the removal number of the noisy points, to avoid the incorrectness due to the loss of data information to the evaluation result. Finally, compare the newly-constructed evaluation function with the original function in experiment; the experimental results show that the definition of the text display of e-book reader is related to its black and white contrast and its brightness of background white. In addition, the comfort level of the text reading is influenced by the color cast of the background white.

Key words. Electronic Book Evaluation Model, SVM Optimized, Grey Model, Clustering analysis.

1. Introduction

As a new generation of display device for reading, e-book reader has its unique technical performance; for example, most of them are light-reflective type, having relatively large field angel and low power consumption and etc. In addition to those peculiar advantages, the conventional performances like black and white contrast, the feature of background white (equivalent to the paper white of newspaper), gray level and etc. are the key points of its display quality [1]. As for this type of light-reflective display device, there is no fixed method for the measuring of its indexes like black and white contrast, background white performance and etc., as well as its perceived quality influence for the displaying of words. Through the subjective

[1]Heilongjiang Bayi Agricultural University, Daqing, Heilongjiang, 163319, China

evaluation experiment, the relationship between the perceptual clarity of the text display of the device and the objective value of the black and white contrast and the background white brightness is established in this experiment, to be able to predict the subjective perceived quality of the display device through the measurement of objective parameters.

Clustering validity evaluation is the method to evaluate the advantages and disadvantages of the clustering results; according to the different evaluation criteria of validity function, clustering validity evaluation method can be divided into internal evaluation method, correlation evaluation method and external evaluation method [1]. According to the different evaluation theories; clustering validity evaluation can be divided into the method based on fuzzy division of data set, the method based on the geometric construction of data set and the method based on data statistical information [2]. Clustering validity evaluation mainly takes the tightness within the class and the separation between the classes as the evaluation indexes. The greater the similarity of each element within the class is, the better the tightness within the class is; the less the similarity of each element within the class is, the better the separation within the class is; according to the characteristics of SVM, take the sample points that have sensitive classification and reduce the classification accuracy as the noisy points, which locate in the closest place of the foreign class. Due to the existence of noisy points which influence the effect of clustering evaluation, this paper improves the clustering validity function (hereinafter referred to as MPO function) proposed by literature [13], adds geometric structure of data set, adds distance critical value L in the tightness index and gives out its constraint condition, to build the new clustering validity evaluation index. The function effectively avoids the noisy points' influence on the testing of clustering results, reduces the number of removal noisy points, avoids the loss of data, and can obtain the noisy points between classes more accurately.

2. MPO Clustering Validity Evaluation Function

$Com(U, c)$: During the validity analysis on the noise environment, the noisy points are also included when considering the tightness and separation of the clustering, which makes the validity index sensitive to the noise and outliers. MPO clustering validity evaluation function is constituted by tightness measure and separation measure; tightness measure is fixed by fuzzy membership matrix and the clustering number c commonly, representing for the tightness degree within the class; the separation measure is defined as the distance between different fuzzy sets, representing for the separation degree between different classes. Through the improvement of the PC index, consider the monotonic trend problem with increase of the number of clustering c in the PC index, introduce $(\frac{c+1}{c-1})^{1/2}$ $(2 \leq c \leq n)$ to adjust the index, decrease the influence on the result from the changes of the number of the clustering, to get the tightness measure $Com(U,c)$:

$$Com(U,c) = \left(\frac{c+1}{c-1}\right)^{1/2} \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^{2} / u_M \,. \tag{1}$$

Among them, u_{ij} is the No. j element belonging to the membership of No. i class, $u_M = \min\limits_{1 \le i \le c} \sum\limits_{j=1}^{n} u_{ij}^2$, $\sum\limits_{j=1}^{n} u_{ij}^2 / u_M$ measure class i is corresponding to the tightness of the tightest class. $Com(U,c)$ represents the tightness degree of the data within the class, of which the greater the value is, the better the obtained fuzzy division is.

In order to get the correct division in the environment with noisy points and outliers, obtain the method of separation measure:

$$Sep(U,c) = \frac{1}{n}\sum_{j=1}^{n}\left(\sum_{a=1}^{c-1}\sum_{b=a+1}^{c} O_{abj}\left(U;c\right)\right). \tag{2}$$

Among them, $O_{abj}\left(U;c\right) = \begin{cases} 1-\left|u_{aj}-u_{bj}\right|, if \;\; \left|u_{aj}-u_{bj}\right| \ge T, a \ne b \\ 0, otherwise. \end{cases}$; u_{aj} and u_{bj} represent the membership values of No. j element to the class of a and b.

In $O_{abj}\left(U;c\right)$, use threshold value T to exclude the fuzzy data points in the clustering boundary [14]; the noisy point precisely belongs to this situation. $Sep(U,c)$ calculates and obtains the sum of the separation of all the data points through membership matrix. The smaller the value of $Sep(U,c)$ is, the more obvious the clustering division is.

MPO function is the difference of tightness measure and separation measure:

$$MPO(U,V) = Com(U,c) - Scp(U,c) \tag{3}$$

Through the contrast of multiple clustering invalidity indexes, including PC, FS, XB, OS, PACES, CO and W, it can be obtained that MPO can confirm the number of clustering in a good way, and can avoid the noisy points' influence on the data set. However, in the MPO function, the structure information of the data set is not used, insensitive to the distance. This paper mainly adds geometric structure feature of data set on the basis of MPO and avoids the influence brought by the single theory to the inspection result; in the separation measure; it needs to be very careful to remove the noisy points; if removing in a too hard way, it will cause the loss of data; so, this paper adds a new critical value L in $sep(c)$, which restrains the scope of the noisy points commonly with the original critical value T, to avoid the incorrectness to the result due to the loss of data information.

3. Construction of the new Clustering Validity Evaluation Function

3.1. Tightness measure

Considering that MPO function is not directly related to the geometric structure of the data set, and can not comprehensively evaluate the disadvantages of clustering effect, this paper integrates the geometric structure information which reflects the tightness degree inside the clustering to the tightness measure $Com(U,c)$, to get the new tightness measure index $Com'(U,c,d)$.

$$Com'(U,c,d) = \sum_{i=1}^{c} \frac{\sum\limits_{j=1}^{n} u_{ij}^2}{u_M d_M} \left(\frac{c+1}{c-1}\right)^{1/2}. \tag{4}$$

Among them, d_M is the Euclidean distance; the improved function includes the distance of object x_j to the clustering center v_i, and also includes x_j's membership function value to class i. The value of $Com'(U,c,d)$ reflects the tightness degree of each clustering data point; the greater the numerical result is, the tighter the elements in the cluster and the better the effect of the division is.

3.2. Separation Measure Based on the Dual-Constraint of the Noisy Points

In the clustering process, influence of noisy points and isolated points on the clustering result is particularly great; literature [13] mentioned that using the membership critical T of MPO function can make prefractionation of the noisy points under the condition of probability; however, using only the probability constraint to define the scope of the noisy points may cause the loss of data; in order to confirm the noisy points more accurately, on the basis of the membership critical value T, this paper adds the critical value L of the geometric structure; two critical values determine the noisy points commonly, which can effectively avoid the loss of data under the single theory. The newly-constructed separation measure based on probability condition and geometric construction is as follow:

$$Sep'(U,c,d) = \frac{1}{n\sum\limits_{j=1}^{n}\sqrt[3]{\frac{1}{c}\sum\limits_{i=1}^{c} d_{ij}}} \sum_{j=1}^{n}\left(\sum_{a=1}^{c-1}\sum_{b=a+1}^{c} W(U;c)\right). \tag{5}$$

Among them

$$W(U,c,d) = \begin{cases} 0, |u_{aj}-u_{bj}| \le T \& |d_{aj}-d_{bj}| \le L_j, a \ne b. \\ 1-|u_{aj}-u_{bj}|, otherwise \end{cases} \tag{6}$$

Set two critical values to exclude the fuzzy points in the boundary of the clustering;

according to the different distribution structures of clustering data, the value of T and L can be given by the expert, or determined by oneself according to the function. Through a large number of data experiments, this paper initially gives out the value $T = 0.01$; $L_j = \sqrt[3]{\frac{1}{c}\sum_{i=1}^{c} d_{ij}}$; d_{ij} represents the distance of No j element belonging to No. i class. $W(U; c, d)$ is the separation measure to the given data set; the smaller the value of $sep'(c)$ is, the more obvious the effect of the clustering division is.

3.3. Construction of Clustering Validity Function

Integrate the tightness measure and the separation measure, to get the new clustering validity evaluation function:

$$VS(U, V) = \frac{Com'(U, c, d)}{Sep'(U, c, d)}\,. \tag{7}$$

A good clustering demands large $Com'(U, c, d)$ and small $Sep'(U, c, d)$; $VS(U, V)$ defines validity function; the larger the value of $VS(U, V)$, the better the result of clustering division.

4. Algorithm Example and Result Analysis

4.1. Algorithm Example One

Use newly-constructed clustering validity function in the paper to perform effectiveness evaluation to fuzzy clustering result[15] of 12 months' average temperature of 31 provinces, cities, municipalities and direct-controlled municipalities in China.

Use MATLAB to gain class-center coordinate matrix c_{ij}, membership matrix u_{ij} and Euclidean distance d_M between object x_i and clustering center c_i. Respectively calculate MPO clustering validity evaluation function and VS clustering validity function and result gained is as is shown in Table 1 and Table 2.

Table 1. Result of $VS(U, V)$ Algorithm Exampleb

c	$com'(c)$	$sep'(c)$	$VS'(U,V)$
3	1.5433	2.362	0.653
4	1.5021	5.255	0.286
5	1.363	4.269	0.319

Table 2. Result of $MPO(U, V)$ Algorithm Example

c	$com(c)$	$sep(c)$	$VS(U,V)$
3	4.6741	2.06	2.614
4	6.076	5.082	0.994
5	6.8861	4.438	2.448

When $c = 3$, membership difference value of Beijing belonging to the first class and the second class is 0.0008, membership difference value of Taiyuan belonging to the second class and the third class is 0.01, and membership difference value of Hefei belonging to the first class and the third class is 0.0015, and they are all less than or equal to $T = 0.01$. But only difference value of distance of Beijing to the first class and the third class is less than L_j, therefore in the process of calculating $sep'(c)$, when $j = 1$, $a = 1$ and $b = 3$, $W(U; c, d) = 0$, all the other are value of $1 - |u_{aj} - u_{bj}|$. Substituted into formula, value of $sep'(3)$ is 2.362 at last.

Tianjin $u_{21} - u_{24} = 0.0055$, Jinan $u_{15,1} - u_{15,3} = 0.01$ and Kunming $u_{26,2} - u_{26,3} = 0.0006$, they are all less than or equal to$T = 0.01$. Only Jinan meets L_j constraint. When $j = 15$, $a = 1$ and $b = 3$ are gained, $W(U; c, d) = 0$, all the other are value of $1 - |u_{aj} - u_{bj}|$. Substituted into formula, value of $sep'(4)$ is 5.5255 at last.

When $c = 5$, membership difference value of every place belonging to the second class and the fifth class is the smallest, most of values are less than $T = 0.01$. Simultaneously consider critical value T and L, and there are only 15 situations where $W(U; c, d) = 0$. The phenomenon that loss of data brings about inaccuracy to result is avoided.

According to result calculated, when $c = 3$, value of $VS(U, V)$ is the greatest and partition effect is the most obvious. Compared with index of original $MPO(U, V)$, $VS(U, V)$ validity function added to critical value of distance L is the same with original $MPO(U, V)$ function, and it is ensured that the most optimal number of clustering is 3, and partition effect is more obvious. The phenomenon that excessive deleting of noisy point brings about loss of data information is avoided, fuzzy point between class and class can be gained more correctly.

4.2. Algorithm Example Two

Perform effectiveness evaluation to fuzzy clustering result of scientific research topic ability of 31 provinces, cities, municipalities and direct-controlled municipalities in China. Imitate steps in algorithm example one, and result gained is as is shown in Table 3 and Table 4.

Table 3. Result of $VS(U, V)$ Algorithm Example

c	$com'(c)$	$sep'(c)$	$VS'(U,V)$
3	3.9164	0.4406	8.8888
4	2.6107	1.4854	1.7576
5	2.5778	2.348	1.0979

Table 4. Result of $MPO(U, V)$ Algorithm Example

c	$com(c)$	$sep(c)$	$VS(U,V)$
3	15.0506	1.7066	8.019
4	13.2304	3.9535	3.3465
5	16.1657	7.3106	2.2113

When $c = 3$, only membership difference value of Shanghai belonging to the first class and the second class is 0.01 and is less than or equal to $T = 0.01$. But, difference value of distance of the first class and the second class is less than L_j. Therefore, when $c = 3$, there is no noisy point deleted. All the other are value of $1 - |u_{aj} - u_{bj}|$. Substituted into formula, value of $sep'(c)$ is 0.4406 at last.

When $c = 4$, Sichuan $u_{23,1} - u_{23,2} = 0.0079$ simultaneously meets constraint. Value of $W(U; c, d)$ is 0. All the other are value of $1 - |u_{aj} - u_{bj}|$, and value of $sep'(4)$ is 1.4854 at last.

When $c = 5$, the points which meets $|u_{aj} - u_{bj}| \leq T$ are 19. After being added to $|d_{aj} - d_{bj}| \leq L_j$ constraint, the points of which value of $W(U; c, d)$ is 0 are 4.

Values of $VS(U, V)$ and $MPO(U, V)$ are gained through calculation and the most optimal number of clustering is ensured as 3. Effectiveness of function is verified again.

5. Conclusion

On the basis of the analysis of the advantages and disadvantages of various clustering validity evaluation functions, this paper improves the MPO evaluation function of them, to get the new clustering validity function. Through the two theories of distance and membership combining fuzzy division of geometric structure domain, the function constructs the tightness measure and separation measure, introduces distance critical L on the basis of the original membership critical value T and decreases the influence from the noisy points between the classes on the clustering evaluation result. Through the algorithm examples, the rationality and validity of the method are illustrated.

6. Acknowledgement

The 12th five-year plan of education science in Heilongjiang province “The practice of general secretary xi’s speech in the form of classroom teaching organization under” (GJB1215051); Research on party building theory of Heilongjiang Bayi Agricultural University“Study on core values of college students under the Socialist ideology”(XDJ16106).

References

[1] Ying Liang, Yingying Yi, Qiufen Sun: (2014). The Impact of Migration on Fertility under China’s Underlying Restrictions: A Comparative Study Between Permanent and Temporary Migrants, Social Indicators Research, 116(1): 307-326.

[2] Liu Y, Yang J, Meng Q: Stereoscopic image quality assessment method based on binocular combination saliency model[J]. Signal Processing, (2016), 125: 237-248.

[3] Li Su, Hanwei Wang, Jingwei Miao ,Ying Liang: (2015).Clinicopathological Significance and Potential Drug Target of CDKN(2)A/p(16) in Endometrial Carcinoma, Scientific Reports, 5:13238, (2015). 1-7.

[4] Jiang D, Xu Z, Lv Z: A multicast delivery approach with minimum energy consumption for wireless multi-hop networks[J]. Telecommunication Systems, (2015): 1-12.

[5] Lin Y, Yang J, Lv Z: A self-assessment stereo capture model applicable to the internet of things[J]. Sensors, (2015), 15(8): 20925-20944.

[6] Ying Liang, Wanyi Lu, Wei Wu: 2014. Are social security policies for Chinese landless Ying He, J., Geng, Y., Wan, Y., Li, S., and Pahlavan, K. (2013). A cyber physical test-bed for virtualization of RF access environment for body sensor network. Sensors Journal, IEEE, 13(10), 3826-3836.

[7] Geng, Y., Chen, J., Fu, R., Bao, G., and Pahlavan, K.: Enlighten Wearable Physiological Monitoring systems: On-Body RF Characteristics Based Human Motion Classification Using a Support Vector Machine. PP(99), 1-16.

[8] Ying Liang: Satisfaction With Economic and Social Rights and Quality of Life in a Post-Disaster Zone in China: Evidence From Earthquake-Prone Sichuan, Disaster Medicine and Public Health Preparedness, 9(2):111-118.

[9] Lv Z, Li X, Zhang B.: Managing Big City Information Based on WebVRGIS. IEEE Access, (2016), 4: 407-415.

[10] Yang J, Lin Y, Gao Z.: Quality Index for Stereoscopic Images by Separately Evaluating Adding and Subtracting[J]. PloS one, (2015), 10(12): e0145800.

[11] Jiang D, Xu Z, Lv Z.: A multicast delivery approach with minimum energy consumption for wireless multi-hop networks[J]. Telecommunication Systems, (2015): 1-12.

[12] Ying Liang, Runxia Cao.:(2015). Employment assistance policies of Chinese government play positive roles! The impact of post-earthquake employment assistance policies on the health-related quality of life of Chinese earthquake populations, Social Indicators Research, 120(3): 835–857.

[13] Liu Y, Yang J, Meng Q.: Stereoscopic image quality assessment method based on binocular combination saliency model[J]. Signal Processing, (2016), 125: 237-248.

[14] Song, X., and Geng, Y.: Distributed Community Detection Optimization Algorithm for Complex Networks. Journal of Networks, 9(10), 2758-2765.

[15] Pahlavan, K., Krishnamurthy, P., and Geng, Y. : Localization Challenges for the Emergence of the Smart World, Access, IEEE, 3(1), 1-11.

[16] Ying Liang, Xiukun Wang.: (2013). Developing a new perspective to study the health of survivors of Sichuan earthquakes in China: a study on the effect of post-earthquake rescue policies on survivors' health-related quality of life, Health Research Policy and Systems, 11:41,1-12.

[17] Li X, Lv Z, Hu J.: Traffic management and forecasting system based on 3d gis[J]. Cluster, Cloud and Grid Computing (CCGrid), 2015 15th IEEE/ACM International Symposium on, (2015): 991-998.

Received May 7, 2017