

Fuzzy C-Means Clustering Algorithm in College Students' Mental Health Test

Caixia He*

Department of Ideological and political theory, Hunan
Software Vocational and Technical University, Xiangtan,
411201, Hunan, China
E-mail: hecai1314@126.com
*corresponding author

Rencai Luo

Tensioner segment
Schaeffler (XiangTan) Co.Ltd, Xiangtan, 411001, Hunan,
China
E-mail: 252216130@qq.com

Abstract—Human resources have generated massive and diverse data. When it comes to big data, the most important feature of data mining (DM) is their uniqueness. This article aims to study the application of fuzzy C-means clustering algorithm in college students' mental health test. This article describes the development and application of aggregation analysis algorithms, identifies the most common distance types and aggregation tasks in aggregation, and describes some classic aggregation analysis algorithms in detail. On the basis of explaining the advantages and disadvantages of the speed analysis algorithm and the algorithm compilation process, we propose a two-step concept, discuss the algorithm flow of the aggregation model in detail and discuss the diagram, using the calculation algorithm to analyze the mental health assessment tools such as UPI Student mental health data created to create DM models and mine student groups and their components, research about student psychology, summarize student and mental health, recovery and anxiety, introversion and extroversion, emotion and calm and cautious, fear and boldness The relationship between. In order to facilitate the school to formulate new methods to prevent students' psychological problems, it will create a positive and healthy environment for college students to study and live. Experiments have shown that in the results of the college students' mental health test analyzed in this article, the psychological problems of male college students are relatively more prominent, while 80% of college students have normal mental health.

Keywords—DM, Cluster Analysis, Mental Health, Personality Factors

I INTRODUCTION

Many health care and medical institutions at the University have done a lot of work in developing a mental health education system for college students, technology. However, the current situation of this work in many colleges and universities in our country is unbalanced, and many mental health education programs are not perfect. [1-2].

The research on the application of the fuzzy C-means clustering algorithm in the mental health test of college students, many scholars have studied it and achieved good results. For example, Kuo RJ proposed the fuzzy set theory. Since then, people have used the fuzzy form to The clustering is handled and named as fuzzy cluster analysis, which describes the uncertainty level that each object belongs to all classes [3]. Concepción Violán proposed the Fuzzy C-Means (FCM) algorithm, which is the first and most widely used algorithm because it has strong fuzzy properties and can retain more information than hard partitioning [4].

This article describes the development and application of aggregation analysis algorithms, identifies the most common distance types and aggregation tasks in aggregation, and describes some classic aggregation analysis algorithms in detail. On the basis of explaining the advantages and disadvantages of the speed analysis algorithm and the algorithm compilation process, we propose a two-step concept, discuss the algorithm flow of the aggregation model in detail and discuss the diagram, using the calculation algorithm to analyze the mental health assessment tools such as UPI Student mental health data created to create DM models and mine student groups and their components, research about student psychology, summarize student and mental health, recovery and anxiety, introversion and extroversion, emotion and calm and cautious, fear and boldness The relationship between. In order to facilitate the school to formulate new methods to prevent students' psychological problems, it will create a positive and healthy environment for college students to study and live.

II ALGORITHMS

A. Fuzzy C-Means Clustering Algorithm

Fuzzy clustering (FCM) extends the idea of traditional clustering, by using membership functions, it is possible to assign each data to all clusters. The idea of FCM algorithm: Let the clustering target $X=\{x_1, x_2, x_3, \dots, x_n\}$ be a finite sample data set in the real number space, and n is the number of elements in the data set. The FCM algorithm aggregates the sample data into c fuzzy classes. Let c_i be the cluster

center of each class i , and $d_{ij} = \|c_i - x_j\|$ be the Euclidean distance between the i -th cluster center and the j -th data point, the weighted index $m \in [1, \infty]$, and the degree of membership. $U=\{u_{ij}\}$ is an $n \times c$ -dimensional matrix that satisfies the following formula:

$$\sum_{i=1}^c u_{ij} = 1, j = 1, 2, \dots, n \quad (1)$$

$$J(U, c_1, \dots, c_e) = \sum_{i=1}^e j_i = \sum_{i=1}^e \sum_{j=1}^n j_{ij}^m d_{ij}^2 \quad (2)$$

The first formula states that the participating points of each segment belonging to each category sum to 1; the second formula is the field function used to determine the

sum of distances. The principle that confuses the c-mean compound algorithm is to find a solution when the second formula gets its minimum value under the constraints of the first formula. Where the number of classes c is pre-given c [9-10].

B. The Requirements of College Students' Mental Health Analysis for Cluster Analysis Algorithm

The clustered data may be hundreds or tens of thousands. This requires the clustering algorithm to process data with scalability, and many existing clustering algorithms are very effective in clustering small data sets with few data objects. However, for large-scale databases, the results of clustering are not optimistic and time-consuming. This requires the algorithm to have strong scalability, and the time complexity should not be too high.

The weak dependence of input parameters on domain knowledge. Most clustering algorithms require users to set various parameters, such as the number of clusters, the support degree of the result, the confidence degree, etc. The values of these parameters have a great impact on the results of cluster analysis. However, the setting of these parameters is often closely related to the professional knowledge of the sample itself, and it is difficult for ordinary users to set these parameters well. Therefore, a good algorithm should have little dependence on these parameters.

The results are independent of the order of input records. The results obtained by a good clustering algorithm should be independent of the input order of the data. The ability to handle high-dimensional data, some algorithms perform well when dealing with datasets with relatively few dimensions, but the clustering results for high-dimensional data are not so ideal. High-dimensional data itself contains a large amount of information, which is more troublesome to process, so the cluster analysis of high-dimensional data is very challenging.

Constraint-based clustering often has many additional constraints in the specific application of the algorithm, and a good clustering algorithm can take these constraints into account. When adding constraints, there is still a better performance [11-12].

III DESIGN AND EXPERIMENT

A. College Student Personality Questionnaire

The process of implementing a college student mental health program is to collect data on student health. The main data source is the UPI panel. The Student Personnel Market (UPI) is a very important measure to measure the mental health of college students, and it is also one of the important skill tests for counselors in vocational training. The College Student Behavioral Questionnaire is mainly used to assess students' mental health status, assess students' mental health problems, establish students' brain records, and discover dynamic observations of mental health. It can help scientists or counselors quickly identify problem cases, enabling schools to provide timely counseling services to students with mental health issues, and to take appropriate actions to intervene in a crisis, after the selection and refinement of character data. The relevant data is installed in a complete database of scientific tests, the data is analyzed using a fuzzy clustering algorithm, and the analysis generates compiled results that can assess whether students are mentally healthy.

B. Experimental Design

This article mainly conducts fuzzy cluster analysis for the survey results of the college students' personality questionnaire constructed in this article, first of all, the analysis of various psychological characteristics and variance of male and female college students. The second is to judge whether the mental health of college students is universal or not through the analysis of the results of the questionnaire survey of different universities.

IV EXPERIMENTAL ANALYSIS

A. Questionnaire Analysis

This article mainly conducts the UPI College Student Personality Questionnaire survey on the 2019 grade students of the Finance Department of a university, analyzes the differences in various psychological characteristics between male and female college students, and analyzes the results of the questionnaire. The data obtained are shown in Table 1.

Boys are more decisive than girls. Better than boys, boys have more prominent personality problems, mainly indifference and stubbornness, low trust, suspicious and stubborn, and poor self-discipline.

TABLE 1. VARIABLE DIFFERENCES OF MALE AND FEMALE

secondary personality factors	Male college student		Female college student		T	P
	Mean	Std. Deviation	Mean	Std. Deviation		
adaptability	5.17	1.66	4.47	1.66	8.28	0.000
extraversion	5.30	1.78	5.71	1.71	4.52	0.000
decisive	5.13	1.42	4.29	1.46	11.43	0.000
emotional	5.19	1.41	4.48	1.46	9.73	0.000

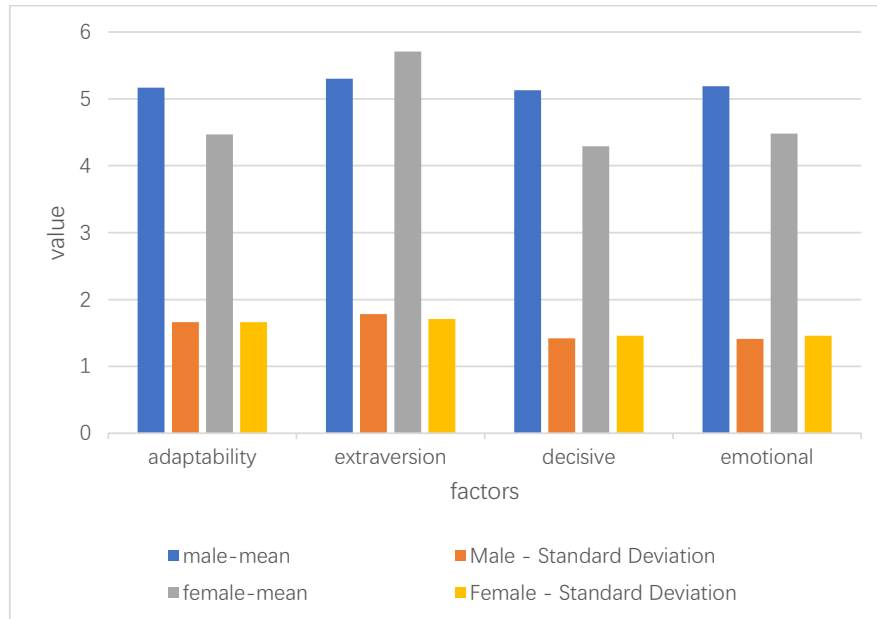


Figure 1. Variable differences of Male and Female

B. Statistical Analysis of College Students' Mental Health Test

In order to test whether the psychological analysis results of college students obtained in this article are universal, this article conducts a survey on the mental health of all freshmen in two universities, and summarizes the obtained data, as shown in Table 2.

TABLE 2. UPI STATISTICAL ANALYSIS

	normal	General Inquiries	Focus counseling treatment on or
Area1	79.9%	11.61%	8.40%
Area2	81.00%	12.31%	6.69%

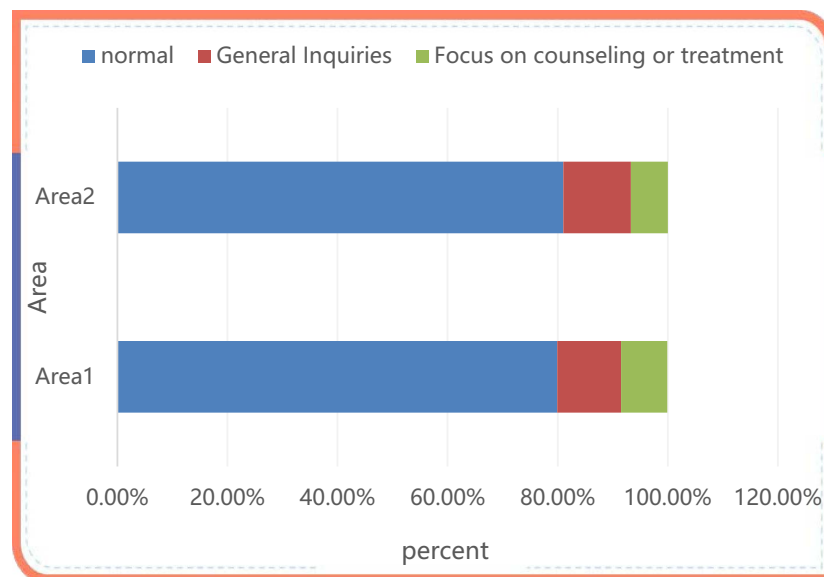


Figure 2. UPI Statistical Analysis

From Figure 2, we can clearly see that the mental health of freshmen in the two universities is not much different, and the structure is similar. Therefore, the mental health of college students analyzed in this article is universal, and about 80% of freshmen the mental health is good, about 12%

of the freshmen need general psychological counseling, and about 8% of the freshmen need focused counseling or treatment.

V CONCLUSIONS

Using the method of clustering analysis to mine and researching the main factors of the mental health of college students is an interesting discussion. Through clustering analysis, we can explore the mental health laws of college students, and provide a reliable basis for student management and training. Results of the first study of this book: First, it is a useful attempt to analyze the mental health of college students using the method of clustering analysis. The second is to develop the mental health data of college students and the mining model of college students' mental health. Looking forward to the future, for the collection of mental health data of college students, professional intelligence should be used further to identify factors related to mental health, in order to achieve accurate predictions. In the end, mental health testing plays a role in finding and finding the cause, but it also has some limitations. As long as we can address these limitations and consider avoiding test negligence due to its limitations in practice, it can provide us with valuable information on mental health testing to be an effective mental health test for college student.

REFERENCES

- [1] Tang Q, Zhao Y, Wei Y, et al. Research on the Mental Health of College Students Based on Fuzzy Clustering Algorithm. *Security and Communication Networks*, 2021, 2021(3):1-8.
- [2] Lei X, Ouyang H, Xu L. Kernel-Distance-Based Intuitionistic Fuzzy c-Means Clustering Algorithm and Its Application. *Pattern Recognition and Image Analysis*, 2019, 29(4):592-597.
- [3] Kuo R J, Lin J Y, Nguyen T. An application of sine cosine algorithm-based fuzzy possibilistic c-ordered means algorithm to cluster analysis. *Soft Computing*, 2021, 25(11):1-16.
- [4] Concepción Violán, Quintí Foguet-Boreu, Sergio Fernández-Bertolin, et al. Soft clustering using real-world data for the identification of multimorbidity patterns in an elderly population: cross-sectional study in a Mediterranean population. *BMJ Open*, 2019, 9(8):e029594.
- [5] WEI, Meng, LIU, et al. The Application Status of Horticultural Therapy in the Study of College Students' Physical and Mental Health. *Journal of Landscape Research*, 2020, v.12(03):108-110.
- [6] Mensah I K, Tianyu Z, Zeng G, et al. Determinants of the Continued Intention of College Students in China to Use DiDi Mobile Car-Sharing Services:. *SAGE Open*, 2019, 9(4):375-379.
- [7] Barida B, Chinagolum I, Shedrack M. An Enhanced Application of Fuzzy C-Mean Algorithm in Image Segmentation Process. *The Journal of Scientific and Engineering Research*, 2019, 6(1):193-197.
- [8] Al-Dabagh M. Automated tumor segmentation in MR brain image using fuzzy c-means clustering and seeded region methodology. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 2021, 10(2):284-290.
- [9] Cardone B, Martino F D. A Novel Fuzzy Entropy-Based Method to Improve the Performance of the Fuzzy C-Means Algorithm. *Electronics*, 2020, 9(4):554.
- [10] Ning Y, Shi X, Yin J, et al. Application of fuzzy C-means clustering method in the analysis of severe medical images. *Journal of Intelligent and Fuzzy Systems*, 2020, 38(8):1-11.
- [11] Mahesa R, Wibowo E P. OPTIMIZATION OF FUZZY C-MEANS CLUSTERING USING PARTICLE SWARM OPTIMIZATION IN BRAIN TUMOR IMAGE SEGMENTATION. *Journal of Theoretical and Applied Information Technology*, 2020, 98(19):3055-3066.
- [12] Yang Y, Li M, Ma X. A Point Cloud Simplification Method Based on Modified Fuzzy C-Means Clustering Algorithm with Feature Information Reserved. *Mathematical Problems in Engineering*, 2020, 2020(4):1-13.