Eurasia Journal of Science and Technology	ISSN: 2663-1024	Volume 1, Issue 1,pp 18-19	March 2019

ANALYSIS OF THE EXAM SITUATION BASED ON CLUSTERING ALGORITHM NORMAL COLLEGE OF BEIJING UNION UNIVERSITY

Feng Han
Teachers' College of Beijing Union University, Beijing, 100011, China

Abstract:
In this paper, k-means clustering algorithm is applied to the analysis and research of students' final examination of advanced mathematics, and students' scores of different types of questions are taken as indicators for clustering. We should carry out stratified teaching for different types of students and properly group cooperative learning, so as to better grasp the learning situation of students.
Key words: K-means clustering algorithm; Examination analysis; Cooperative learning.

Nowadays, with various teaching modes, educators increasingly find the importance of teaching result feedback, especially the analysis of students' examination results [1]. Usually method of test case analysis often according to the analysis of scores of students, such as proficiency, pass rate, but that is kind of In order to better understand the situation of the students, students and will be similar, timely grasp the overall situation of students, especially for this big class has the vital significance. We select the score of each question in the examination paper as the index to study, and group the students according to this index, so as to better grasp the students' learning dynamics.

When there are large of students, we need to use modern statistical methods. Classification is carried out according to index conditions, often by means of clustering. There are many clustering methods [2]. The d method (DBSCAN) needs to determine the distance r and min Points. Gaussian mixture model (GMM) is used to cluster the maximum expected value (EM), but GMM is used to cluster the first assumption is that the data poi. The advantage of k-means [3] is its rapidity and simplicity, especially for mufti-index discrete variables, although the number of clustering categories needs to be specified in advance.

1. PRINCIPLE AND IMPLEMENTATION OF K-MEANS CLUSTERING ALGORITHM

K-means clustering algorithm is realized by R software. The basic steps and principles of clustering algorithm are as follows:

(1) randomly create k points as barycenter;

(2) calculate the distance between the center of mass and data points for each data point in the data set, and the default distance in R software is Euclidean distance;

(3) find the nearest center of mass for each point and assign it to the cluster corresponding to the center of mass;

(4) after distribution, the center of mass will change, and the new center of mass and E value will be calculated;

(5) repeat (2), (3) and (4) until the maximum number of iterations or the difference between the newly calculated E value and the last iteration E value is less than a given threshold.

2. AN ANALYSIS OF STUDENT TEST SCORES

2.1 Data Description

The data were selected from the final examination results of students in There were 17 small questions in the paper, and the number of students was 197. We chose the scores of these 17 questions as clustering indicators for clustering. In order to better grasp the dynamics of students, we set the number of clustering as 6.

2.2 Description of Data Analysis Results

We use R software to realize the clustering process, which is specifically completed by the function command k means in the R program. We preset 6 categories, and the clustering results were divided into 6 categories, including 56 students in category 1, 31 students in category 2, 28 students in category 3, 27 students in category 4, 25 students in category 5, and 30 students in category 6.

We analyze students' scores from several typical questions we pay attention to, and choose the four questions of subjective questions 7, 8, 9 and 10. Table 1 below shows the mean distribution of the scores of the four questions for the six types of students.

Table.1 6 table of mean distribution of 7-10 scores in class

Question number	MEAN1	MEAN2	MEAN3	MEAN4	MEAN5	MEAN6
zg_7	7.6	3.4	6.6	6.6	6.6	6.9
zg_8	6.9	3.7	6.9	5.5	5.4	6.3
zg_9	7.2	4.0	6.1	7.0	6.3	6.5
zg_10	7.0	1.2	2.9	2.6	5.4	7.2

In order to better display the research results, we drew the scores of questions 7-10 in 6 groups in the figure below, as shown in figure 1 below.

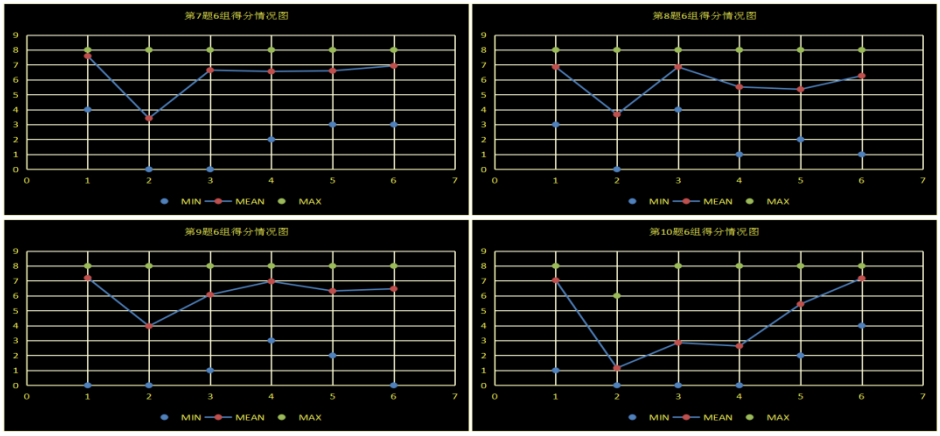

Fig.1 6 distribution of mean scores of 7-10 questions in class

The score of question 7 of the students in the first category was between 4-8, with an average of 7.6. Students in category 2 scored between 0 and 8 on the 7th question, with an average of 3.4. Students in category 3 scored between 0 and 8 in question 7, with an average of 6.6. Students in category 4 scored between 2-8 in question 7, with an average of 6.6. Students in category 5 scored between 3 and 8 in question 7, with an average of 6.6. Students in category 6 scored between 3 and 8 in question 7, with an average of 6.9. It can be seen that the students of the second group scored low in the, and the related knowledge points need to be paid attention to. The students of the first group scored high, so the two groups can cooperate and communicate with each other.

Students in category 1 scored between 3 and 8 in question 8, with an average of 6.9. Students in category 2 scored between 0 and 8 in question 8, with an average of 3.7. Students in category 3 scored between 4 and 8 in question 8, with an average of 6.9. Students in category 4 scored between 1 and 8 in question 8, with an average of 5.5. Students in category 5 scored between 2-8 on the 8th question, with an average of 5.4. Students in category 6 scored between 1 and 8 on question 8, with an average of 6.3. It can be seen that the students of the second group scored low in the eighth question, and the related knowledge points need to be focused on. The students of the first and third groups scored high, and these groups can conduct inter-group cooperation and communication.

The score of the 9th question of the students in the first class is between 0 and 8, and the average is 7.2. Students in category 2 scored between 0 and 8 in question 8, with an average of 4.0. Students in the third category scored between 1 and 8 in question 8, with an average of 6.1. Students in category 4 scored between 3 and 8 in question 8, with an average value of 7.0. Students in category 5 scored between 2 and 8 in question 8, with an average of 6.3. Students in category 6 scored between 0 and 8 on question 8, with an average of 6.5. It can be seen that the students of the second class scored low in the ninth question, and the related knowledge points need to be focused on. The students of the first class and the fourth class scored high, and these groups of students can carry out inter-group cooperation and communication. At the same time, in particular, there is polarization in the first class, which requires intra-group cooperation and communication.

Category 1 students scored between 1 and 8 on question 10, with an average of 7.0. Students in the second category scored between 0 and 6 in the 10th question, with an average value of 1.2. Students in the third category scored between 0 and 8 in question 10, with an average of 2.9. Students in category 4 scored between 0 and 8 in question 10, with an average of 2.6. Students in category 5 scored between 2-8 in the 10th question, with an average of 5.4. Students in category 6 scored between 4 and 8 on question 10, with an average score of 7.2. It can be seen that the students of the second, third and fourth categories have relatively low scores in the 10th question, and the relevant knowledge points need to be focused on. The students of the first and sixth categories have relatively high scores, and these groups of students can conduct inter-group cooperation and communication. At the same time, in particular, there is polarization in the first category, which requires intra-group cooperation and communication.

3. CONCLUSION

We apply the clustering algorithm to the analysis of students' examination results, so as to provide some methods for us to carry out cooperative and exchange learning, especially to better understand students' learning situation. In practice, we can select the appropriate clustering algorithm according to the actual characteristics of the data. This paper can be used for reference to analyze the examination situation quickly.

REFERENCE

[1]. Hao Linqing. Research on the evaluation system of classroom teaching quality in colleges and universities [D]. China University of geosciences (Beijing), 2017.

[2]. He Yulin, Hua Zhexue. Journal of shenzhen university (science and technology edition), 2019, 36(01):4-17. (In Chinese).

[3]. Zhu Jianyu. Research and application of k-means algorithm [D]. Dalian university of technology, 2013.