Exploring the Evaluation of College Students' Learning Effect Under SPOC Teaching Mode Based on Cluster Analysis

Yuhan Shi(✉)

Flight Academy, Civil Aviation University of China, Tianjin, China
yh_shi@cauc.edu.cn

Abstract. SPOC (Small Private Online Course) is an online and offline blended teaching mode. Many colleges and universities have adopted this teaching mode to improve teaching quality. However, due to the particularity of teaching mode, the way of evaluating students' learning effect under this mode needs to be further explored. In this study, a course with SPOC teaching mode was selected as an example and K-means algorithm was applied to conduct cluster analysis on students' learning behaviours and grades, so as to form an evaluation of students' learning effects by classification. The study also predicted the importance of various factors that might influence learning outcomes. The research results verify the effectiveness of K-means clustering method in evaluating students' learning effect. The application of this evaluation method can enable students to improve themselves in a targeted way, and also help teachers find potential problems in SPOC teaching mode, so as to further improve the teaching quality.

Keywords: SPOC · Blended Teaching · Evaluation of Learning Effect · Cluster Analysis · K-means Algorithm

1 Introduction

China has always attached great importance to the development of education and has been continuously promoting the development of informatization education. From the rise of distance education in the 1990s, to the recorded online courses, ubiquitous education, then to the emergence of MOOC (Massive Online Open Course) in the 21st century, with the development of Internet technology, the form and content of online education are constantly upgraded and improved. Online education breaks the limitations of time and space in traditional teaching and creates an environment for learners to learn anytime, anywhere, and repeatedly. It meets the personalized learning needs of learners with different identities and levels. However, online courses such as MOOCs also have their drawbacks [2, 4]. For example, since the audience of MOOC courses is very wide, the courses provide undifferentiated educational services, and it is impossible to teach students in accordance with their aptitude. Beyond that, MOOC is not suitable for experimental and practical courses due to its natural nature. In addition, although

Z. Zhan et al. (Eds.): EIMSS 2022, AHCS 7, pp. 495–503, 2023.
https://doi.org/10.2991/978-94-6463-024-4_52

most MOOCs have opened discussion areas along with courseware, most of the learners are basically independent learning. There is a lack of sufficient teaching interaction between teachers and students.

Therefore, educators in various countries have been trying to find a new teaching mode that combines the interactive advantages of traditional teaching models with the convenience of online learning, so as to complement each other. Professor Armando Fox of University of California, Berkeley [3] proposed the concept of SPOC (Small Private Online Course), which corresponds to MOOC. SPOC is a blended teaching mode that combines the advantages of classroom and online teaching [7, 8]. Compared with MOOC, which are aimed at learners in the whole society, SPOC mainly targets students in school, which facilitates interaction between teachers and students. SPOC also convenient for teachers to teach students in accordance with their aptitude and carry out experimental and practical courses.

With the continuous development and popularization of the concept of SPOC teaching mode, an increasing number of colleges and universities have begun to adopt SPOC teaching mode to improve teaching quality. Due to the differences between SPOC, traditional and MOOC classes, the evaluation method of students' learning effect cannot be copied from the traditional teaching model and MOOC model. How to evaluate students' learning effect in SPOC teaching mode has become an important topic for educators to study.

This study will take the course "Flight Human Factors and Crew Resource Management" as an example to explore and discuss this issue.

2 Course Introduction and Evaluation Principle

2.1 Course Introduction

"Flight Human Factors and Crew Resource Management" is a professional compulsory course for flight technology majors, with a total of 40 credit hours, including 24 credit hours for the theoretical part and 16 credit hours for the experimental part. The course adopts SPOC mode for teaching, in which the release of online resources relies on the Zhihuishu platform. Teachers publish learning resources on the platform and carry out online teaching and students need to complete all courses both online and offline. In addition, Offline teaching is the mutual supplement of students' online learning, so as to timely find the loopholes in students' knowledge and make up for them. Furthermore, teachers conduct offline learning and case analysis of the experimental part of the course to increase interaction with students.

2.2 Evaluation Principle

2.2.1 Diversity Principle

The diversity of evaluating students' learning effects in the SPOC teaching mode is reflected in the following two aspects: First, the diversity of participants in the evaluation. In the SPOC mode, the learning effect of students is not only evaluated by teachers, but also by learning platforms, students themselves or their peers. The second is the diversity

of evaluation content. Students' learning effects and self-development are the result of the joint action of many factors. Therefore, it is necessary to eliminate one-sided evaluations that only focus on students' test scores while ignoring factors such as students' abilities, emotions and moral values.

2.2.2 Scientific Principle

In the process of evaluating students' learning effect, it must be based on facts, using scientific and objective evaluation methods to quantify the learning effect and avoid subjective assumptions. In addition, the evaluation of students' learning effect should also be based on the students' learning needs and future career development.

2.2.3 Subjectivity Principle

The modern educational philosophy generally holds that students are the main body of cognition and development in teaching, and educational activities cannot be separated from learners' own absorption and internalization of knowledge. Therefore, the evaluation system of SPOC teaching mode should not only help to reflect and guide students to play the main role, but also give full play to the self-active role of students.

3 Materials and Methods

3.1 Materials

3.1.1 Project Design

Objects: The learning process data of 90 sophomore students majoring in flight technology.

Objective: To explore the learning process and score distribution of students in SPOC teaching mode, and analyse the influence of various factors on students' learning effect.

Data: student name, student ID, class, student attendance times, online learning duration, daily homework scores, experimental scores and final exam scores.

Tools: IBM SPSS Modeler 18.0.

3.1.2 Data Collection and Pre-processing

This study analyses the learning data of 90 students from two administrative classes of 2020 undergraduates majoring in Flight Technology in Civil Aviation University of China. The original data of students' attendance times, online learning duration, daily homework scores, experimental scores and final exam scores are derived from the Zhihuishu platform. Since K-means is only applicable to data sets with continuous attributes, that is, numerical data, this study removes non-quantifiable data information to ensure that the data for cluster analysis are all quantifiable data. Information on students with special circumstances, such as dropping out of school and delaying the test, was also removed. Finally, due to the different orders of magnitude and dimensions of the collected data, all data are normalized to [0,1] for subsequent cluster analysis in order to avoid data redundancy and improve the convergence speed of the algorithm.

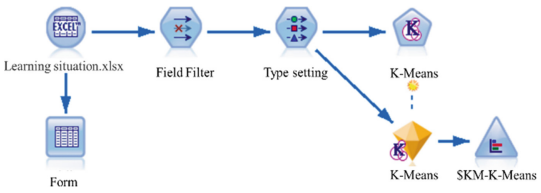

Fig. 1. K-means algorithm clustering analysis model.

3.1.3 Model Building

In SPSS Modeler, K-means algorithm is used to establish the clustering analysis model, as shown in Fig. 1. After field filtering and type setting, the data source file of students' learning status is connected to the K-means algorithm model. After running the algorithm, the results are analysed to obtain the clustering distribution of students' learning effects and the influence of various factors on students' learning effects.

3.2 Methods

In order to realize the diversity, scientific and subjectivity principle of evaluation, a scientific method that is not easily disturbed by inherent thinking and subjective factors must be selected to evaluate students' learning effects. Cluster analysis [5] is a typical unsupervised learning method in machine learning, which can classify a large number of data samples without giving any additional prompts to achieve the effect of classification evaluation. Details are given below.

3.2.1 Cluster Analysis

Cluster analysis is also known as group analysis or data segmentation. Its essence is to group or divide a data set into several subsets or "groups" so that the data objects within each group are more related or have shorter distances than those assigned to different groups. In short, clustering is the process of clustering all target data sets into different groups when the number of categories in the data set is unknown.

Clustering algorithm includes many types, such as partitioning clustering, hierarchical clustering, grid-based clustering, density-based clustering and fuzzy clustering, among which the former two are widely used.

3.2.2 K-Means Clustering Algorithm

This study adopts the K-means clustering algorithm in partitioning clustering, which was proposed by J.B. Macqueen in 1966 [6]. It has also been widely used in the field of college student performance analysis [1, 9, 10]. Its flowchart is shown in Fig. 2. In this flowchart, first determining the number of clusters K according to the input data set N, and then preliminarily determining the cluster centre; selecting the commonly used Euclidean distance as the formula for calculating distance, and determining the distance d from each object to the cluster centre, then dividing into the nearest cluster; using iterative method to continuously update the centre point, until the centre point position

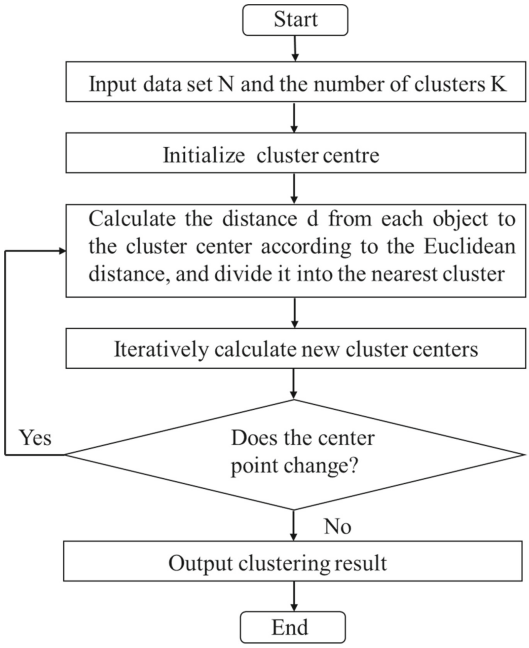

Fig. 2. Flowchart of K-means algorithm.

does not change, output the clustering result. It's worth noting that the selection of K value is very important, which directly affects the clustering effect.

4 Results and Discussion

4.1 Determination of K Value

In K-means algorithm, the value of K is the only adjustable parameter and also the key to cluster analysis. There are many methods to determine K value, such as contour coefficient method and elbow method, etc., which usually need to select K value based on the background knowledge of the application field or the actual situation. This study uses elbow method to determine K value. Set the number of clusters from 2 to 5, a total of four K values, run the algorithm respectively, and the obtained clustering quality is shown in Table 1.

According to the rule of the elbow method, it can be seen from Fig. 3 that when the algorithm is run for the third time, that is, when the K value is 4, the downward trend of the clustering quality line tends to be slow, so the number of clusters can be set to 4, that is, $K = 4$.

4.2 K-Means Clustering Results

Bringing $K = 4$ into the K-means algorithm and running it, we will get the result shown in Fig. 4. The number of students in cluster 1 is 50, which is more than half of the total

Table 1. Clustering quality under different K values.

Time	1	2	3	4
K value	2	3	4	5
Clustering quality	0.55	0.53	0.5	0.5

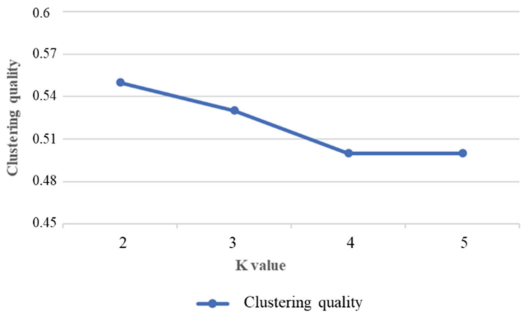

Fig. 3. Clustering quality change trend.

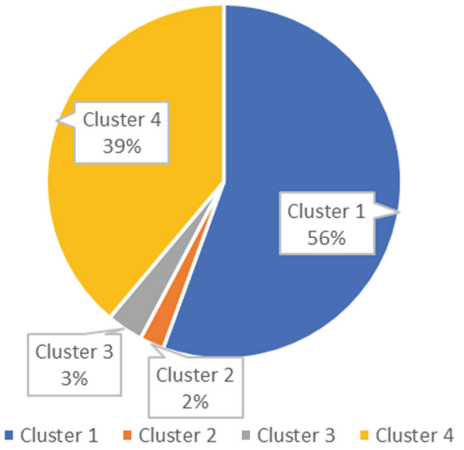

Fig. 4. K-means clustering results.

number of students, reaching 56%. Generally speaking, the learning effect of this group of students is in the upper middle level of the class. There are 2 students in cluster 2, accounting for 2% of the total number of students. These students belong to the category with poor overall learning effect. The total number of students in cluster 3 is 3, accounting for 3% of the total number of students. The overall learning effect of these students is very excellent. There are 35 students in cluster 4, accounting for 39% of the total number of students, and their learning effect is of medium level.

Table 2. The mean of clustering results of students' learning behaviours and grades.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Attendance times	0.98	0.0	1.0	0.1
Online learning duration	0.20	0.0	0.90	0.89
homework scores	0.95	0.28	1.0	0.93
Experimental scores	0.71	0.52	0.8	0.39
Final exam scores	0.77	0.35	0.79	0.5

Table 2 shows the five learning behaviours and results of students, including attendance times, online learning duration, daily homework scores, experimental scores, and final exam scores. The value in each cluster is the mean value of the behavioural characteristics, through which the learning behavioural characteristics and learning effects of students in each cluster can be analysed. Students in cluster 1 have good overall attendance, relatively high homework and experiment scores, short online learning time and average scores in final exams are above average. Such students perform well in offline learning but show little interest in online courses, accounting for 56% of the total number of students, indicating that about half of the students may still be more accustomed to the traditional teaching mode and have not fully utilized the advantages of online courses in SPOC teaching mode.

A total of 2 students belong to cluster 2, and these two students are at a poor level in terms of learning behaviour and performance. It can be seen that such students have poor learning behaviour and low learning interest. Besides, their final exam scores are not satisfactory. In future teaching, teachers should focus on cultivating their learning interest and improving their learning initiative.

The students in cluster 3 are in the leading position in the class in terms of attendance, online learning duration, daily homework, experimental scores and final exam scores. They can make full use of the advantages of online and offline courses in the blended teaching mode and improve their autonomous learning ability. It shows that in the SPOC teaching mode, this kind of students have achieved a good effect.

Students in cluster 4 have a slightly lower attendance rate and perform better in online learning duration and daily homework. However, they are slightly worse in experiments, with a mid-term final grade. Such students like online learning, but their practical application ability is weak, so they should pay more attention to the cultivation of their practical ability in future learning, and strengthen communication and interaction with teachers and cooperation with group students.

It is worth noting that the gap between the final exam scores of students in cluster 1 and cluster 3 is not large. In the traditional teaching mode, these two types of students may trend to be classified as "good grades" students, while ignoring the differences between the two classes of students in autonomous learning, practical exploration and ability development, resulting in teachers not being able to discover the shortcomings of some students in time. Fortunately, through the clustering analysis algorithm, teachers can quickly distinguish the two types of students and better implement teaching in

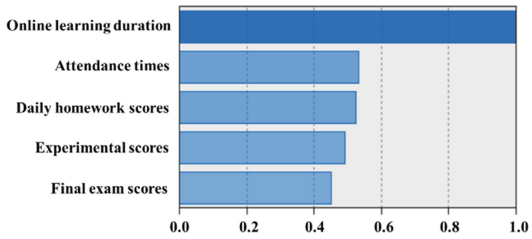

Fig. 5. The importance of the five variables predicted by K-means to the learning effect.

accordance with their aptitude. This is precisely the significance of applying cluster analysis to explore students' learning effects.

4.3 Prediction of Influencing Factors of Learning Effect

By using K-means algorithm, this study also conducts a predictive analysis on the importance of the influence of five variables on students' learning results, and the results are shown in Fig. 5. It can be seen that in the SPOC teaching mode, final exam scores become the least important variable, while the importance of online learning duration is far ahead of other variables and becomes the most important factor affecting the evaluation of students' learning effect. This result is highly consistent with the original intention of building SPOC teaching mode, and further proves the effectiveness of K-means algorithm in evaluating students' learning effect under SPOC mode.

5 Conclusions

This study introduced the concept, advantages and differences of SPOC teaching mode from other teaching modes firstly. After that, the necessity and principles of evaluating the learning effect of college students under the SPOC teaching mode was expounded. In this study, the course "Flight Human Factors and Crew Resource Management", which has implemented the SPOC teaching model, was selected as an example, and the K-means clustering algorithm was adopted to analyse the students' learning effect and various factors that may affect the learning effect. The results verified the validity of K-means clustering algorithm in evaluating students' learning effect. The application of this evaluation method can enable students to improve themselves in a targeted manner and it can also help teachers to discover potential problems in the SPOC teaching mode, thereby further to improve the quality of teaching.

Acknowledgements. The research is supported by the Fundamental Research Funds for the Central Universities (3122021113).

References

1. Chang W, Ji X, Liu Y, Xiao Y, Chen B, Liu H, Zhou S (2020) Analysis of university students' behavior based on a fusion K-means clustering algorithm. Appl Sci 10(18):6566
2. Chen H (2015) The causes of mooc rapid development and its advantages and disadvantages analysis. China Mod Educ Equip
3. Fox A, Patterson DA, Ilson R et al (2014) Software engineering curriculum technology transfer: lessons learned from MOOCs and SPOCs. UC Berkeley Technical Report
4. Guo P (2017) MOOC and SPOC, which one is better? Eurasia J Math Sci Technol Educ 13(8):5961–5967
5. Jain AK, Dubes RC (1988) Algorithms for clustering data. Prentice-Hall, Inc
6. Macqueen JB (1966) Some methods for classification and analysis of multivariate observations
7. Nahornyi VV, Nahorna NO, Murzina OA, Tyrsa TO (2021) Small private online courses (SPOC) as a new approach to higher education
8. Ruiz-Palmero J, Fernández-Lacorte J-M, Sánchez-Rivas E, Colomo-Magaa E (2020) The implementation of small private online courses (SPOC) as a new approach to education. Int J Educ Technol High Educ 17(1):1–12
9. Wang J, Zhang Y (2019) Clustering study of student groups based on analysis of online learning behavior. In Proceedings of the 2019 International Conference on Modern Educational Technology (pp 115–119)
10. Zhang G, Zhang Y, Ran J (2018) Research on clustering mining and feature analysis of online learning behavioral data based on SPOC. In 2018 13th international conference on computer science & education (ICCSE) (pp 1–6). IEEE