Digital Management of Teaching Cases in Colleges and Universities Based on Cluster Analysis

https://doi.org/10.3991/ijet.v18i10.40241

Run Wu, Jinping Wang(✉)
School of Literature, Cangzhou Normal University, Cangzhou, China
wjpwxy@caztc.edu.cn

Abstract—The necessity of digital management of teaching cases is mainly reflected in the aspects of improving teaching quality, facilitating retrieval and storage, achieving cross-platform sharing, conducting real-time updating, saving resources, and carrying out data analysis and evaluation. However, there are still some defects in the existing management models or methods, which lead to difficulties in data storage and retrieval, and affect the utilization efficiency of teaching resources. To this end, this article takes the Chinese language and literature major as an example, and studies the digital management of teaching cases in colleges and universities based on cluster analysis. First of all, it quantifies the quality of digital management of teaching cases, enables resource demanders to accurately select digital resources of teaching cases that meet their teaching or learning needs through the evaluation index of service quality when faced with diversified digital resources of teaching cases in colleges and universities. The clustering algorithm is used to mine potential topics and patterns in teaching cases, which improves the classification efficiency of teaching cases and enables educators to have a deeper understanding of teaching content and educational needs. It uses "absolute index", "incremental index" and "fluctuation index" to construct the similarity measurement distance function of the basic attributes of teaching cases in colleges and universities and uses the Ward method based on variance analysis to classify the characteristics of teaching cases in colleges and universities. Experimental results verify the effectiveness of the proposed method.

Keywords—cluster analysis, teaching cases in colleges and universities, topic of digital management teaching case

1 Introduction

The implementation of digital management of teaching cases in colleges and universities can enable teachers to easily find and share high-quality teaching cases, thereby improving the quality of teaching. At the same time, teaching cases can be stored in electronic form, which is convenient for teachers and students to refer to at any time [1–5]. Teaching cases in colleges and universities shared on different platforms can also facilitate communication and cooperation among teachers, students and other

educators [6–11]. This helps to promote the fair distribution of educational resources and realize the popularization of high-quality educational resources. Teachers can also adjust the cases according to the students' feedback and their own teaching needs, so that the content of the case can be updated quickly to adapt to the ever-changing educational environment and technological progress [12–21]. Clustering technology can help build a structured teaching case knowledge base, which is convenient for educators to manage and maintain teaching case resources. Therefore, the research on the digital management of teaching cases in colleges and universities based on cluster analysis has high research value and significance.

Xiang et al. [22] proposed three task-driven combined case teaching strategies for college teaching in STEAM education. Taking the wooden arch bridge intelligent monitoring project as an example, it expounded the design and application of three task-driven combined case teaching strategies in teaching practice and used classroom observations and interview surveys to demonstrate the positive effects of three different combined case strategies on academic performance or ability. The interactive case teaching method is a good way to realize "silent and soft" ideological and political teaching. Jiang and Song [23] proposed to infiltrate moral education into the teaching of professional course knowledge and a series of activities through interactive case teaching, forming a course teaching combining "intellectual education" and "moral education", combining value shaping, knowledge imparting and ability training. In this way, the fundamental task of fostering virtues will be fully implemented. With the rapid development of social informatization, higher requirements are placed on students' computer application ability.

It can be found from the previous research that domestic and foreign scholars focus on how to better use the existing teaching cases, and the research on the centralized management and clustering of teaching cases in colleges and universities is relatively rare. Moreover, the existing digital management mode or method of teaching cases in colleges and universities has greatly improved the availability and availability of teaching case resources. However, there are still some defects in these models or methods, which limit the further optimization of teaching resource management in colleges and universities. As teaching case formats and structures may vary across disciplines, majors, and faculty, existing management models may not fully accommodate this diversity. This may lead to difficulties in data storage and retrieval, affecting the utilization efficiency of teaching resources. Existing digital management methods often only rely on keyword search and catalog classification. This method cannot make full use of the potential information in teaching cases, such as topics and knowledge points. Therefore, the efficiency in resource retrieval, recommendation and analysis needs to be improved. To this end, this article takes the Chinese language and literature major as an example, and studies the digital management of teaching cases in colleges and universities based on cluster analysis. First of all, in Chapter 2, the article quantifies the quality of teaching case digital management and enables resource demanders to accurately select digital resources of teaching cases that meet their teaching or learning needs through the evaluation index of service quality when faced with diversified digital resources of teaching cases in colleges and universities. In Chapter 3, the article uses clustering algorithm to mine potential topics and patterns in teaching cases, which improves the classification efficiency of teaching cases and enables educators to

have a deeper understanding of teaching content and educational needs. In Chapter 4, the article adopts "absolute index", "incremental index" and "fluctuation index" to construct the similarity measure distance function of the basic attributes of teaching cases in colleges and universities and uses the Ward method based on variance analysis to divide characteristics of teaching cases in colleges and universities. Experimental results verify the effectiveness of the proposed method.

2 Quantification of the quality of digital management of teaching cases

As digital resources of teaching cases in colleges and universities in the university teaching case database are enormous, there will inevitably be a large number of resources with similar content. Clustering case resources by content, scope of application, etc. can organize case resources with similar content into clusters. Case resource demanders such as teachers, students or other educators face diversified digital resources of teaching cases in colleges and universities. Among these digital resources of teaching cases in colleges and universities with similar content, it's possible to further screen the digital resources of teaching cases in colleges and universities that meet our own teaching or learning needs through the evaluation index of case call service quality.

The response time of case-called service, denoted by R_{TI} is the average time between a request and a service response by a teacher, student, or other educator. The service reliability is represented by R_{RY}, which is the ratio of the number of successful case call executions to the number of case calls. The total number of calls by teachers, students or other educators is represented by R_{CO}, the user's evaluation of the service is represented by R_{RN}, and the comprehensive quality index of the case call service defined herein is represented by R_{MU}, then the quintuple of case call service quality can be defined is $R_w = <R_{TI}, R_{RY}, R_{CO}, R_{RN}, R_{MU}>$.

Assume that the time interval between a teacher, student or other educator initiating a service call request and receiving the service response is represented by o_i, and the total number of calls of the service is represented by R_{CO}, the calculation formula of R_{TI} is as follows:

$$R_{TI} = \frac{\sum_{i=1}^{R_{CO}} o_i}{R_{CO}} \tag{1}$$

Assume that the cumulative number of successful service executions is represented by R_{SU}, R_{RY} is the ratio of R_{SU} to R_{CO}, the calculation formula is as follows:

$$R_{RY} = \frac{R_{SU}}{R_{CO}} \tag{2}$$

R_{RN} is only for teachers, students or other educators to evaluate the service after calling, that's, the evaluation interval is [0, 10]. Assuming that the resource demander's

evaluation is represented by SC_i, and the total number of evaluations of the service is represented by m, the calculation formula is as follows. Not all teachers, students or other educators may evaluate resources after they are obtained, so m is often not equal to R_{CO}.

$$R_{RN} = \frac{\sum_{i=1}^{m} SC_i}{m} \tag{3}$$

R_{MU} is a combination of R_{TI}, R_{RY} and R_{RN}, where R_{RY} and R_{RN} are ratio indexes with fixed value ranges. R_{RN} can be normalized to R'_{RN}:

$$R'_{RN} = \frac{R_{RN}}{10} \tag{4}$$

While R_{TI} is a numerical case call service quality index, this article uses a conversion method for R_{TI} to eliminate the impact of orders of magnitude, so that the value of R_{TI} after normalization can converge to an ideal range, and the relevant formula is given by the following formula:

$$R'_{TI} = \begin{cases} \dfrac{R_{TI}^{\max} - R_{TI}}{R_{TI}^{\max} - R_{TI}^{\min}}, if \quad R_{TI}^{\max} \neq R_{TI}^{\min} \\ 1, if \quad R_{TI}^{\max} = R_{TI}^{\min} \end{cases} \tag{5}$$

R_{MU} takes the weighted sum of R'_{TI}, R_{RY} and R_{RY}, where the weight of R'_{TI} is λ, the weight of R_{RY} is α, and the sum of the three weights is 1, then:

$$R_{MU} = \lambda * R'_{TI} + \alpha * R_{RY} + (1 - \lambda - \alpha) * R'_{RN} \tag{6}$$

3 Cluster-based teaching case topic discovery

The clustering algorithm applied to the digital management of teaching cases in colleges and universities can automatically classify similar teaching cases into one group to help teachers and students quickly find relevant teaching resources. Resource demanders do not need to spend a lot of time on manual classification, which will improve retrieval efficiency. In order to improve classification efficiency and enable educators to gain a deeper understanding of teaching content and educational needs, clustering algorithm needs to mine potential topics and patterns in teaching cases. In order to improve the representation ability of feature words in teaching cases, this article builds a corresponding topic model based on them, which uses Dirichlet distribution to represent topics and documents.

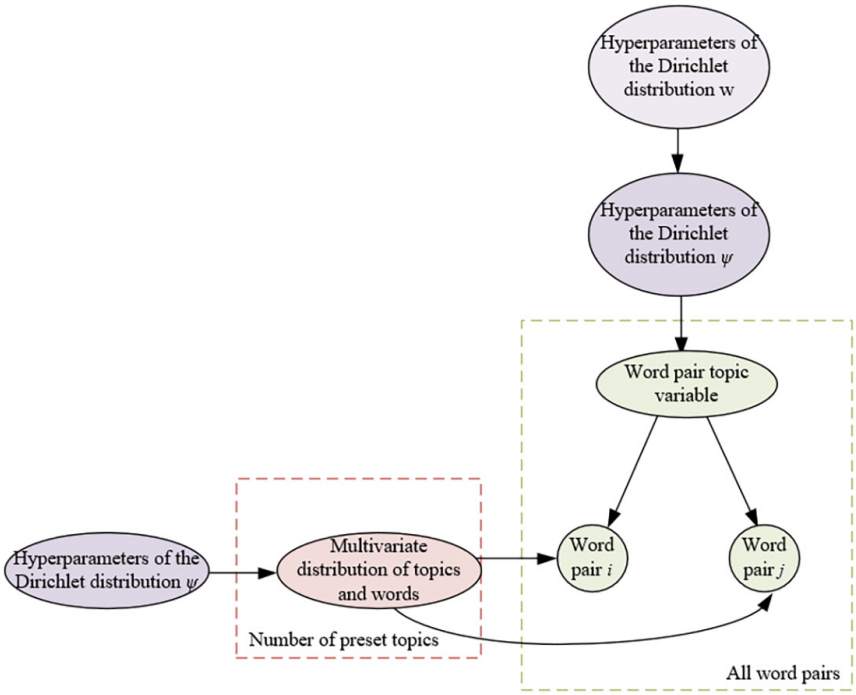

Fig. 1. Teaching case topic model generation process

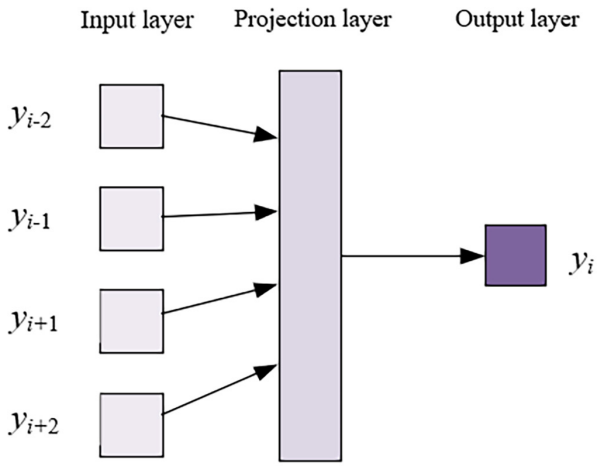

Fig. 2. *CBOW* model structure diagram

Figure 1 shows the process of generating the topic model of the teaching case. The steps of topic model generation based on word vectors trained by *word2vec* are described in detail as follows.

Step 1: Capture the context information and word meaning of the teaching case according to the *CBOW* model. Figure 2 shows the structure diagram of the *CBOW* model, and calculate the semantic distance e of each word pair $y = (y_i, y_j)$ based on the obtained training results.

Step 2: Assume that each case text corresponds to a topic distribution, which is a Dirichlet distribution. Then the term distribution $\psi_c \sim Dir(\gamma)$ under the topic c of each case can be sampled from it.

Step 3: In the teaching case resource description library K, sample the word pair-topic distribution $\omega_c \sim Dir(x)$.

Step 4: For each word pair $y = (y_i, y_j)$ in all word pairs $|Y|$ in K, sample a teaching case resource topic value c, enabling it to conform to $c \sim Mult(\omega)$.

Step 5: Sample word pair y_i, y_j from c that conforms to $y_i, y_j \sim Mult(\psi_c)$, and the number of sampling and the way of topic update are determined according to e.

The joint probability distribution of the word pair $y = (y_i, y_j)$ in c can be further calculated based on the following formula:

$$LG(y) = \sum_c LG(c)LG(y_i \mid c)LG(y_j \mid c) = \sum_c \omega_c \psi_{i|c} \psi_{j|c} \tag{7}$$

The probability of K can be calculated by the following formula:

$$LG(Y) = \prod_{(i,j)} \sum_c \omega_c \psi_{i|r} \psi_{j|c} \tag{8}$$

In order to realize the alternate conditional sampling of model parameters, this article introduces the Gibbs sampling method. Firstly, the initial status of a Markov chain is randomly determined, and then the conditional probability $TG(c|c_y, Y, x, \gamma)$ of the teaching case resource feature word pair $y = (y_i, y_j)$ is calculated to obtain the conditional probability of K. Assuming that the number of topics is represented by m_c, the total number of all words in the teaching case resource description database is represented by N, and the number of occurrences of the teaching case resource word pair γ under the topic word c of teaching case resources is represented by $m_{y|c}$, then the calculation formula is as follows:

$$TG(c \mid c_{-y}, Y, \beta, \gamma) \propto (\beta + m_c) \frac{(m_{y_i|c} + \gamma)(m_{y_j|c} + \gamma)}{(\sum_y m_{y_c} + N\gamma)^2} \tag{9}$$

If a teaching case resource feature word pair $b = (b_i, b_j)y = (y_i, y_j)$ is assigned to topic c, then y_i and y_j also belong to topic c respectively. Assuming that the distribution parameters of terms under each teaching case resource topic are represented by ψ, and the distribution parameters of K's teaching case resource topics are represented by ω, the estimation formulas of ψ and ω are given in the following formula:

$$\omega_c = \frac{m_c + \beta}{|Y| + L\beta} \tag{10}$$

$$\psi_{y|c} = \frac{m_{y|c} + \gamma}{\sum_{y} m_{y|c} + N\gamma} \tag{11}$$

Based on the above steps, a teaching case document – topic matrix with n rows and m columns can be obtained. The rows represent n teaching case resource description documents, and the columns represent m-dimensional teaching case topic vectors.

4 Cluster-based digital management pattern recognition of teaching cases

The cluster analysis for the digital management of teaching cases in colleges and universities can help educators find high-quality teaching cases and their shortcomings. By comparing different class of cases, educators can understand which teaching methods and content are more popular with students, so as to optimize and improve the teaching cases. Next, this article will reclassify the teaching cases of colleges and universities based on the cluster analysis method, identify and analyze the existing digital management models of teaching cases in colleges and universities, explore the differences in the digital management models of teaching cases in colleges and universities, and analyze the primary characteristics of the digital management model of teaching cases in colleges and universities and the differences in each link.

In order to meet the requirements of systematic analysis and ensure less information, this article adopts "absolute index", "incremental index" and "fluctuation index" to construct the similarity measure distance function of the basic attributes of teaching cases in colleges and universities and uses the Ward method based on variance analysis to classify features of teaching cases in colleges and universities.

The Ward method based on variance analysis is a hierarchical clustering method, and its core idea is to cluster by minimizing the sum of squares within each group. The method attempts to find those data points or clusters that, after merging, result in the smallest increase in variance within the group for merging. This approach is able to generate relatively compact and uniformly sized clusters. Point distance and class distance are the main ways of this method to describe the distance of data points or clusters.

Assume that the case samples in the teaching case resource description library is represented by a_{ilo}, then $i = 1, 2, \ldots M$, $L = 1, 2, \ldots n$, $o = 1, 2, \ldots O$. The SD of the l-th index in the o period is represented by R_{lo}, and the distance between any two case samples i and j_{is} represented by e_{ij}, then the absolute distance between any two case samples can be defined as $e_{ij}(JD)$:

$$e_{ij}(JD) = \left[\sum_{l=1}^{n} \sum_{t=1}^{O} (a_{ilo} - a_{jlo})^2 \right]^{1/2} \tag{12}$$

Let $\Delta a_{ilo} = a_{ilo} - a_{ilo-1}$ and $\Delta a_{jlo} = a_{jlo} - a_{jlo-1}$, the absolute quantity difference between two adjacent periods is determined by Δa_{ilo} and Δa_{jlo}, then the incremental distance $e_{ij}(ZL)$ between any two case samples can be defined:

$$e_{ij}(ZL) = \left[\sum_{l=1}^{n} \sum_{p=2}^{O} \left(\frac{\Delta a_{ilo}}{a_{ilo-1}} - \frac{\Delta a_{jlo}}{a_{jlo-1}} \right)^2 \right]^{1/2} \tag{13}$$

Assuming that the average value of the l-th index observation value of case sample i in the total period O is represented by $\overline{a_{il}}$, it can be calculated by $\overline{a_{il}} = 1/O \sum_{o}^{O} = 1 a_{ilo}$. The standard deviation of the l-th index observation value of case sample i in the total period O is represented by R_{il}, which can be expressed by $R_{il} = 1/O - 1 \sum_{t=1}^{O} (a_{ilo} - \overline{a_{il}})^2$, the fluctuation distance $e_{ij}(BD)$ between any two case samples can be defined:

$$e_{ij}(BD) = \left[\sum_{l=1}^{n} \left(\frac{\overline{a_{il}}}{r_{il}} - \frac{\overline{a_{jl}}}{r_{jl}} \right)^2 \right]^{1/2} \tag{14}$$

Assuming the normalized value of the distance $d_{ij}(JD)$ is denoted by $zd_{ij}(JD)$, there is an "integrated" distance $e_{ij}(ZH)$ between any two case samples.

$$e_{ij}(ZH) = \theta_1 \cdot ce_{ij}(JD) + \theta_2 \cdot ce_{ij}(ZL) + \theta_3 \cdot ce_{ij}(BD) \tag{15}$$

Assuming that the center of gravity of H_k is represented by $\overline{A_k}$, the index value of the $i(i = 1, 2, ..., M_k)$-th case sample in i is represented by A_{ik}. If the case samples are divided into l classes, the sum of squared deviation Q_k of the case samples in the H_k class can be calculated by the following formula:

$$Q_k = \sum_{=1}^{M_k} (A_{ik} - \overline{A}_l)^T (A_{ik} - \overline{A}_k) \tag{16}$$

The sum of squares of total deviations of L classes can be calculated by the following formula:

$$Q = \sum_{k}^{l} Q_k \tag{17}$$

When the number l of case sample classes is constant, if Q has a minimum value, the sum of squared deviations of case samples of H_k in the case sample classification in the teaching case resource description library can be calculated by the following formula. Let A_{ik}, B_{ik} and C_{ik} be n-dimensional vectors. The value of the j-th index of case sample i in o period is represented by A_{ijo}, and the increment speed of the j-th index of case sample i in o period is represented by $B_{ijo} = \Delta a_{ijo}/\Delta a_{ijo-1}$, the absolute quantity difference between the case sample i and the j-th index in o period and o–1 period in H_k is expressed by $\Delta a_{ijo} = A_{ijo} - A_{ijo-1}$, the variation degree of the j-th index of the case sample i in H_k is expressed by $C_{ij} = \overline{a_{ij}} / r_{ij}$, then the calculation formula is as follows:

$$\begin{aligned} Q_k^* = \sum_{i=1}^{M_k} \sum_{o=1}^{O} \left[\theta_1 (A_{ik} - \overline{A}_k)^T (A_{ik} - \overline{A}_k) \right] + \sum_{o=2}^{O} \left[\theta_2 (B_{ik} - \overline{B}_k)^T (B_{ik} - \overline{B}_k) \right] \\ + \sum_{i=2}^{O} \left[\theta_1 (C_{ik} - \overline{C}_l)(C_{ik} - \overline{C}_k) \right] \end{aligned} \tag{18}$$

Assuming that the sum of squares of total deviations of M_k case samples in H_k is represented by Q_k, the formula for calculating the sum of squares of total deviations is as follows:

$$Q^* = \sum_{k=1}^{l} Q_k^* \tag{19}$$

The sum of the squared deviations of any two case sample classes after merging can be regarded as the square distance between classes in the algorithm, and classes of any two case sample H_c and H_w are further combined into a new class H_s, then the distance between class H_c and H_w can be calculated by the following formula:

$$C_{ls}^2 = Q_s^* - (Q_c^* + Q_w^*) \tag{20}$$

Let the number of case samples corresponding to H_c, H_w, H_s and H_l be denoted by M_c, M_w, M_s and M_l respectively, and the following formula gives the distance recursion formula between H_l and H_c:

$$E_{ls}^2 = \frac{M_c + M_l}{M_s + M_l} E_{ls}^2 + \frac{M_w + M_l}{M_s + M_l} E_{lw}^2 + \frac{M_s}{M_s + M_l} E_{cw}^2 \tag{21}$$

Similarly, for a fixed l, the choice ensures that E_{ls}^2 achieves minimal classes of case samples.

5 Experimental results and analysis

Table 1. The final cluster centers expressed by the original attributes

	First Class	Second Class	Third Class	Fourth Class	Fifth Class
Name	3.517261	1.876455	1.7/6764	4.461523	2.879654
Key words	10.135214	7.415346	5.349761	5.916744	6.315794
Subject class'	0.7962512	0.516974	2.159751	0.114795	1.613548
Background description	2.464687	2.646547	2.156647	1.646578	1.547684
Problem statement	955.652375	826.157613	271.671285	838.482715	382.138564
Theoretical framework	0.057134	0.846134	0.063214	0.161483	0.716424
Data and analysis	0.024613	0.025144	0.026131	0.091624	0.517842
Solution	112.646545	664.131685	513.154464	319.746864	134.345646
Discussion	1.646513	1.214467	3.465791	2.846153	2.754132
Reflection	38.464561	–13.646845	46.754646	35.165473	–13.754667
Conclusion	0.746561	0.014957	–0.027464	–0.0054341	–0.024679
Enlightenment	0.016465	0.416498	0.213565	0.0146879	0.246787
Further reading and resources	1008.135465	959.654654	910.749879	990.465671	1022.468794

According to authenticity, relevance, complexity, enlightenment, operability, challenging, and expandability, teaching cases in colleges and universities are divided into five classes with different scoring levels. Authenticity can ensure the authenticity and credibility of the case, and this attribute reflects whether the teacher can combine theoretical knowledge with practical application based on the case to improve the practical ability of classroom teaching. Relevance reflects whether the teaching case is closely related to the topic of the course, so that students can better understand and master the course knowledge. Complexity ensures the development of students' ability to analyze, judge and solve problems. Enlightenment makes the teaching case stimulate students' thinking and interest, thus prompting them to actively participate in classroom discussions and learning. The teaching case should have a certain degree of operability, that's, the problems and tasks in the case can guide students to take specific actions and measures. Operability helps to improve students' hands-on ability and practical ability. Challenges can come from the difficulty of the problem, the innovativeness of the solution, and so on. Expandability ensures that teaching cases can be applied and promoted in different topic areas and courses.

According to the five classes of teaching cases in colleges and universities determined by the final clustering center of the original attribute expression given in Table 1, the corresponding first to fifth class of teaching cases in colleges and universities are defined as knowledge-based teaching cases in colleges and universities and skill-based teaching cases in colleges and universities respectively, thinking-oriented teaching cases in colleges and universities, emotional-attitude-based teaching cases in colleges and universities, and field investigation-based teaching cases in colleges and universities.

In the above, this study has fully considered that the point distance and class distance are the classification results when the method describes the distance of data points or clusters. The data can be eliminated and the original attributes can be analyzed, and then the impact of other case attributes in colleges and universities can be explored.

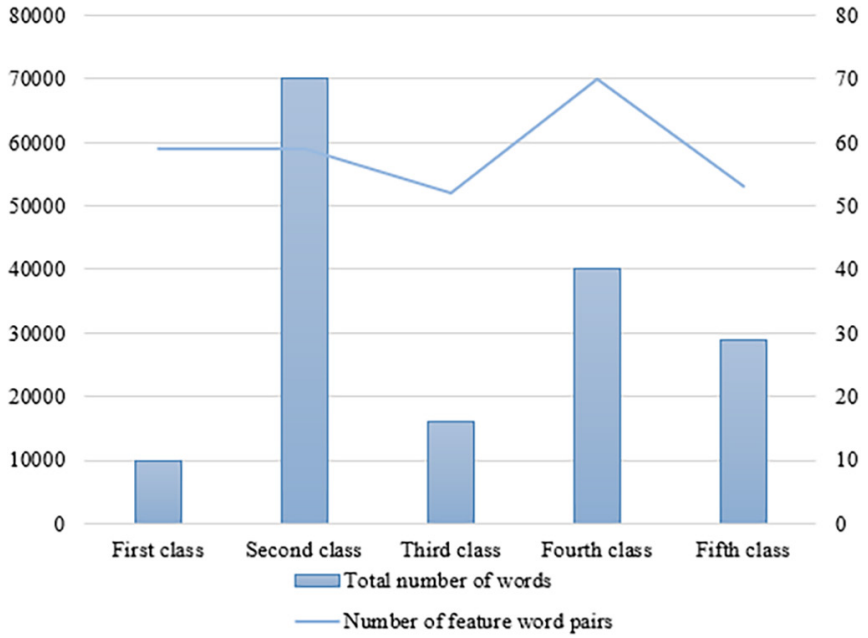

Fig. 3. Relationship between the total number of words and the number of feature word pairs in five classes teaching cases in colleges and universities

Knowledge-based teaching cases in colleges and universities are mainly used to teach theoretical knowledge and concepts. Skill-based teaching cases in colleges and universities are mainly used to cultivate students' practical operation and application skills. Thinking-oriented teaching cases in colleges and universities are mainly used to cultivate students' abilities in analysis, judgment, problem-solving and innovation. Emotional-attitude-based teaching cases in colleges and universities are mainly used to cultivate students' values, emotions and attitudes. Field investigation-based teaching cases in colleges and universities require students to go to actual places for investigation and research. Figure 3 shows the relationship between the total number of words and the number of feature word pairs in five classes teaching cases in colleges and universities. Since skill-based cases usually provide specific operation steps, methods and skills, they are often used in the teaching process of teachers, so the total number of words is relatively large. Emotional-attitude-based cases are mainly used to cultivate students' values, emotions and attitudes, and the purpose is relatively clear. Teachers use this class of case discussion to guide students to reflect and introspect, so as to improve their sense of social responsibility and professionalism. Therefore, there are relatively many feature word pairs in this class of cases.

Table 2. The number of teaching cases in colleges and universities and the proportion of use frequency

Class	Number	Quantity Ratio	Use Frequency	Use Ratio
First class	9	20	9.44	21
Second class	12	17	2.81	51.7
Third class	14	24	0.61	4.6
Fourth class	10	22	0.50	3.7
Fifth class	15	17	0.77	19
Total	60	100	1 4.13	100

Table 2 is obtained according to the classification results and the initial data of teaching cases in colleges and universities. It can be seen from the table that the second class of teaching cases in colleges and universities accounts for 21%, and the frequency of use by teachers accounts for 51.7% of the profits, which is in line with the 80/20 rule, so it is reasonable to divide teaching cases in colleges and universities into five classes. The following are the results of the experimental analysis taking the Chinese language and literature major as an example:

(1) Knowledge-based teaching cases in colleges and universities enable students to better understand theoretical knowledge related to Chinese language and literature, such as literary genres, literary history and rhetoric. Through case studies, students can apply the theoretical knowledge they have learned to the analysis of actual literary works, enhance the ability to combine theory with practice, and help to increase students' interest in related knowledge and enhance their enthusiasm for learning. At the same time, they enable teachers to impart knowledge in a more intuitive and vivid way, provide a variety of teaching methods, and help to improve the attractiveness and fun of classroom teaching. It is also necessary to pay attention to its potential negative effects, such as emphasizing theory and heavy preparation work.
(2) Skill-based teaching cases in colleges and universities enable students to master practical skills related to Chinese language and literature through practice, such as writing literary reviews and sorting out ancient books. They can improve students' ability to analyze and solve problems, and enhance their ability of independent learning and exploration. At the same time, teachers can combine theoretical knowledge with practical operation to improve teaching quality. It is also necessary to pay attention to its potential negative effects, such as time and energy investment and heavy preparation work.
(3) Thinking-based teaching cases in colleges and universities enable students to exercise their analytical ability through the analysis of the topic and style of literary works, so that they can make correct decisions in the face of complex problems, which can stimulate students' thinking and encourage them to develop unique perspectives and insights that foster innovation. Such teaching cases may make some students feel frustrated and confused, requiring students to invest more time and energy, and some students may feel greater pressure.

(4) Emotional-attitude-based teaching cases are guided by emotional experience and values in literary works to help students develop humanistic qualities and aesthetic tastes, stimulate students' interest in learning Chinese language and literature, and help students understand the psychology of characters in different cultural backgrounds and historical periods, thereby enhancing empathy. However, such teaching cases may cause students to pay too much attention to emotional experience, leading to overflowing emotions, ignoring the learning and mastery of theoretical knowledge and practical skills, and affecting the rigor of academics and learning.

(5) Field investigation-based teaching cases in colleges and universities enable students to have an in-depth understanding of the actual background of Chinese language and literature, and expand their knowledge horizons, such as visiting historical sites and understanding the background of literary works, which will help students apply the theoretical knowledge they have learned to actual scenarios, improve their practical ability and analytical skills, and develop teamwork and communication skills. However, this class of cases may require students to invest more time and energy, and some students may feel greater pressure, and may also face certain safety risks, such as traffic accidents and bad weather.

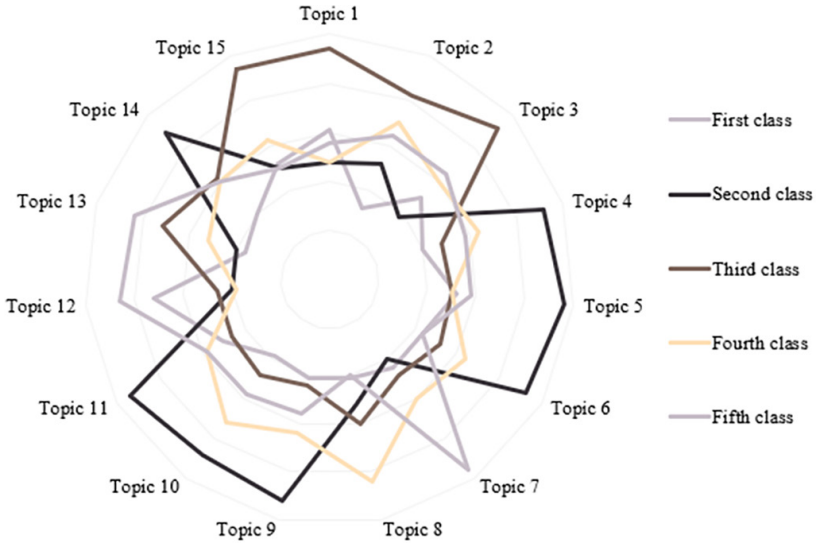

Fig. 4. Radar chart of topic differences in five classes of teaching cases in colleges and universities

Clustering results classify the teaching cases in colleges and universities from different sources investigated herein into five classes, and the teaching cases of the five classes show different degrees of differences in 15 topics. The radar chart can further visualize the differences between the five classes of teaching cases, and enhance the intuitive feeling of the differences between the five classes of teaching cases, see Figure 4 for details. The specific 15 topics include the ancient literature appreciation,

modern literature review, literary creation, ancient book collation, language research, language teaching, drama research, folk literature, cultural heritage, literature and society, literature and psychology, literature and history, literature and philosophy, and literature and film. Specifically, the six topics of ancient literature appreciation, modern literature review, literary creation, poetry appreciation, drama research, and folk literature belong to the core topics of literature research and appreciation, and the three topics of ancient book collation, language research, and language teaching belong to the core of language research and teaching topic. Cultural inheritance, literature and society, literature and psychology, literature and history, literature and philosophy, literature and film are the core topics of literature and interdisciplinary research.

6 Conclusion

This article takes the Chinese language and literature major as an example, and studies the digital management of teaching cases in colleges and universities based on cluster analysis. First of all, it quantifies the quality of digital management of teaching cases, enables resource demanders to accurately select digital resources of teaching cases that meet their teaching or learning needs through the evaluation index of service quality when faced with diversified digital resources of teaching cases in colleges and universities. The clustering algorithm is used to mine potential topics and patterns in teaching cases, which improves the classification efficiency of teaching cases and enables educators to have a deeper understanding of teaching content and educational needs. It uses "absolute index", "incremental index" and "fluctuation index" to construct the similarity measurement distance function of the basic attributes of teaching cases in colleges and universities and uses the Ward method based on variance analysis to classify the characteristics of teaching cases in colleges and universities. Combined with experiments, five classes of teaching cases in colleges and universities determined by the final clustering centers expressed according to the original attributes, this article explores the relationship between the total number of words and the number of feature word pairs in five classes of teaching cases in colleges and universities, investigates the number of teaching cases in colleges and universities and the proportion of use frequency, and gives the experimental analysis results taking Chinese language and literature majors as an example. It draws the topic difference radar chart of five classes of teaching cases in colleges and universities, carries out the clustering comparison of teaching cases under different core topics, analyzes the differences of teaching case management modes in five classes, provides the analysis results and verifies the effectiveness of the proposed method.

7 Acknowledgment

Research on the construction of Chinese language and character curriculum system in local normal universities under the background of new liberal arts construction. (Grant No.: SD2021117).

8 References

[1] Lv, Z., Wang, J. (2020). The exploring on the university student honor management system based on enterprise refined management under the background of big data. In Journal of Physics: Conference Series, 1453(1): 012164. https://doi.org/10.1088/1742-6596/1453/1/012164

[2] Ong, M.I.U., Ameedeen, M.A. (2020). A design theory for student self-service university management system. In Journal of Physics: Conference Series, 1529(5): 052052. https://doi.org/10.1088/1742-6596/1529/5/052052

[3] Fei, G. (2023). On the design of student employment module in higher education management system based on genetic algorithm. In Application of Big Data, Blockchain, and Internet of Things for Education Informatization: Second EAI International Conference, BigIoT-EDU 2022, Switzerland, pp. 232–242. https://doi.org/10.1007/978-3-031-23947-2_25

[4] Aljarallah, N.A., Dutta, A.K., Alsanea, M., Sait, A.R.W. (2023). Intelligent student mental health assessment model on learning management system. Computer Systems Science and Engineering, 44(2): 1853–1868. https://doi.org/10.32604/csse.2023.028755

[5] He, S., Li, Y. (2023). Design and application of college student management system based on big data technology. Wireless Communications and Mobile Computing, 2023: Article ID: 5752949. https://doi.org/10.1155/2023/5752949

[6] Kurniawan, Y., Bhutkar, G., Alfarrel, J.F., Theo, R.R., Denanto, R.K., Nangka, T.P.P., Jericho, I. (2020). Unlocking student's preference on two BINUS mobile learning management system. In 2020 International Conference on Information Management and Technology (ICIMTech), Bandung, Indonesia, pp. 972–976. https://doi.org/10.1109/ICIMTech50083.2020.9211253

[7] Darko, C., Jagger, D. (2020). How can we manage the blackboard learning management system to enhance student's learning. In Proceedings of the 2020 4th International Conference on Deep Learning Technologies, Beijing China, pp. 49–54. https://doi.org/10.1145/3417188.3417205

[8] Song, Y.N., Luan, Z.Q. (2020). Function design optimization of learning management system (LMS) based on student perspective-case study of canvas application university of Colorado Denver. In Journal of Physics: Conference Series, 1621(1): 012058. https://doi.org/10.1088/1742-6596/1621/1/012058

[9] Mozahem, N.A. (2020). Using learning management system activity data to predict student performance in face-to-face courses. International Journal of Mobile and Blended Learning (IJMBL), 12(3): 20–31. https://doi.org/10.4018/IJMBL.2020070102

[10] Omar, N., Ariff, M.A.M., Shah, A.F.I.M., Mustaza, M.S.A. (2020). Estimating synthetic load profile based on student behavior using fuzzy inference system for demand side management application. Turkish Journal of Electrical Engineering and Computer Sciences, 28(6): 3193–3207. https://doi.org/10.3906/elk-1912-101

[11] da Costa, R.D., de Souza, G.F., de Castro, T.B., de Medeiros Valentim, R.A., de Pinho Dias, A. (2020). Identification of learning styles in distance education through the interaction of the student with a learning management system. IEEE Revista Iberoamericana de Tecnologias del Aprendizaje, 15(3): 148–160. https://doi.org/10.1109/RITA.2020.3008131

[12] Mijić, D., Bjelica, O., Durutović, J., Ljubojević, M. (2019). An improved version of student attendance management system based on RFID. In 2019 18th International Symposium Infoteh-Jahorina (INFOTEH), East Sarajevo, Bosnia and Herzegovina, pp. 1–5. https://doi.org/10.1109/INFOTEH.2019.8717750

[13] Rahhou, A., Talbi, M. (2019). Information system of performance management and reporting at university: Example of student scorecard. In Information Systems and Technologies to Support Learning: Proceedings of EMENA-ISTL 2018 2, Marrakech, Morocco, pp. 67–76. https://doi.org/10.1007/978-3-030-03577-8_8

[14] Dorobăț, I., Corbea, A.M.I., Muntean, M. (2019). Integrating student trust in a conceptual model for assessing learning management system success in higher education: An empirical analysis. IEEE Access, 7: 69202–69214. https://doi.org/10.1109/ACCESS.2019.2919100
[15] Jain, N., Gupta, V. (2019). The impact of knowledge management system on student performance: A case study of the University of Delhi. VINE Journal of Information and Knowledge Management Systems, 49(1): 115–135. https://doi.org/10.1108/VJIKMS-07-2018-0065
[16] Nasrulloh, M.Y., Putra, Y.H. (2019). Student data management information system using the Zachman framework. In IOP Conference Series: Materials Science and Engineering, 662(2): 022101. https://doi.org/10.1088/1757-899X/662/2/022101
[17] Brabete, V., Goagără, D. (2022). Digitalization – a danger to accounting professionals? Journal of Corporate Governance, Insurance, and Risk Management, 9(S1): 25–48. https://doi.org/10.51410/jcgirm.9.1.3
[18] Songsom, N., Nilsook, P., Wannapiroon, P., Fung, L.C.C., Wong, K. (2019). System architecture of a student relationship management system using internet of things to collect digital footprint of higher education institutions. International Journal of Emerging Technologies in Learning (iJET), 14(23): 125–140. https://doi.org/10.3991/ijet.v14i23.11066
[19] Fosu, M.A., Gupta, T., Michael, S. (2019). Social media in chemistry: Using a learning management system and twitter to improve student perceptions and performance in chemistry. In Technology Integration in Chemistry Education and Research (TICER), pp. 185–208. https://doi.org/10.1021/bk-2019-1318.ch013
[20] Ateş Çobanoğlu, A. (2018). Student teachers' satisfaction for blended learning via Edmodo learning management system. Behaviour & Information Technology, 37(2): 133–144. https://doi.org/10.1080/0144929X.2017.1417481
[21] Cheng, M., Yuen, A.H.K. (2018). Student continuance of learning management system use: A longitudinal exploration. Computers & Education, 120: 241–253. https://doi.org/10.1016/j.compedu.2018.02.004
[22] Xiang, X., Xue, T., Jing, Z. (2020). The design of college students' ideological dynamic visualization management information system under the background of big data. In 2020 International Conference on Computers, Information Processing and Advanced Education (CIPAE), Ottawa, ON, Canada, pp. 302–305. https://doi.org/10.1109/CIPAE51077.2020.00082
[23] Jiang, T., Song, Y. (2020). The information grading management system of college students based on deep learning. In e-Learning, e-Education, and Online Training: 6th EAI International Conference, eLEOT 2020, Changsha, China, pp. 308–317. https://doi.org/10.1007/978-3-030-63952-5_26

9 Authors

Run Wu is a lecturer in the school of literature, Cangzhou Normal University, graduated from Hebei University, research areas including Literature theory, aesthetics and education and teaching, has published 8 papers, presided over a number of scientific research projects. Currently studying for a PhD at Jeonju University in South Korea. Email: w09051990@163.com. ORCID: https://orcid.org/0000-0002-7474-4630.

Jinping Wang is an associate professor of Chinese at Cangzhou Normal University. She graduated from Hebei University with a PhD in Chinese philology, her research interests include language standardization, language teaching, college teaching reform, etc. She has published 5 papers in core journals and 3 monographs. Email: wjpwxy@caztc.edu.cn. ORCID: https://orcid.org/0000-0001-5727-3032.

Article submitted 2023-02-28. Resubmitted 2023-04-07. Final acceptance 2023-04-11. Final version published as submitted by the authors.