International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 6977-6987

https://doi.org/10.70517/ijhsa463598

K-mean clustering based education quality assessment method and intelligent learning path design

Yuting Wang[1,*]
[1] Applied Technology College, Anshan Normal University, Anshan, Liaoning, 114000, China
Corresponding authors: (e-mail: wangyuting34@126).

Abstract Educational quality assessment faces the challenge of diversifying students' needs, and traditional assessment methods are difficult to accurately cope with individual differences. This study proposes an educational quality assessment and learning path design method that combines K-mean clustering and reinforcement learning. First, we construct a student user profile and establish a teaching quality evaluation system with four primary indicators and eighteen secondary indicators; then, we classify the student group by K-mean clustering, and determine the optimal number of clusters based on the sum of squared intra-group distances; finally, we use reinforcement learning algorithms to design a personalized learning path recommendation system. The empirical study collected the online education assessment scores of 30,000 students, and the cluster analysis showed that four groups of students were clearly characterized: the first group (10,144) had moderate learning pressure but poor memory; the second group (6,845) had high learning pressure but lacked motivation; The third category (7168) has extreme learning pressure and poor emotional control; the fourth category (5843) has good grades with little learning pressure but poor emotional control. Reinforced learning path recommendation experiments show that the system reaches a stable learning gain after 45 iterations and can generate different learning paths according to learners' individualized needs. The results prove that this method can effectively identify the characteristics of student groups, provide learning paths that meet individual needs, and provide a feasible solution for intelligent and personalized teaching.

Index Terms K-mean clustering, education quality assessment, reinforcement learning, learning path, user profiling, personalized teaching

I. Introduction

The K-mean clustering algorithm is a common unsupervised learning method for dividing a data set into K distinct clusters [1], [2]. The goal of the algorithm is to find groups in the data labeled by the variable K. The algorithm works iteratively to assign each data point to one of the K clusters based on the features provided, and many factors need to be considered in determining the ideal number of clusters, K, including the characteristics of the data, the purpose of the clustering, and the computational resources [3]-[6]. With the application of artificial intelligence in education, the application of K-mean clustering algorithms in education quality assessment and intelligent learning path design has been gradually emphasized [7], [8].

Education quality assessment is to promote the improvement and enhancement of education by evaluating the education process and education results [9], [10]. The objectives of assessment usually include aspects of students' knowledge, skills, abilities and potential [11], [12]. Educational quality assessment can help educational institutions and education departments to understand the current status and problems of education, and provide a basis for formulating effective educational policies and improving educational methods [13], [14]. Compared with traditional assessment methods, the advantage of K-mean clustering in education quality assessment is that it can realize the intelligent analysis and classification of multi-dimensional indicators through data-driven, and it can assess the quality of education more objectively [15]-[17]. In modern education, intelligent learning path design has become an important way to improve teaching effectiveness, and the intelligent learning path design of K-mean clustering can classify learning resources and design personalized learning paths based on students' needs, which is of great significance for improving teaching quality [18]-[21].

Education quality assessment and personalized learning path design are important issues facing contemporary education. In the current educational environment, learners have different backgrounds and different levels of ability, and the traditional unified teaching mode is difficult to meet the diversified learning needs. In the era of big data in education, analyzing learners' behavioral data, building user profiles, and then realizing data-driven education quality assessment and personalized learning path recommendation have become a key way to improve the

effectiveness of education. At present, education quality assessment research mainly focuses on the construction of the evaluation index system, but there is less research on the identification and categorization of the differences in the student population; and in terms of learning path recommendation, although there are explorations based on methods such as knowledge graph and collaborative filtering, there is a lack of programs that organically combine student characteristics with path design. The assessment of education quality under diversified learning needs needs to take into account the individual differences of students, including multi-dimensional factors such as learning level, learning habits, and learning motivation. Traditional assessment methods often use a single dimension or a simple weighted average, making it difficult to fully reflect the complexity of education quality. Meanwhile, the design of learning paths mostly relies on experts' experience and lacks data support, making it difficult to respond to the dynamic changes in learners' needs in a timely manner. K-mean clustering, as a classical unsupervised learning algorithm, can naturally classify student groups based on multidimensional features and discover potential group patterns; while reinforcement learning techniques continuously optimize decision-making strategies through trial and error to provide learners with personalized learning paths. The combination of the two provides new ideas to solve the problems of education quality assessment and personalized learning path design.

Based on this background, this study proposes a combination of K-mean clustering and reinforcement learning for education quality assessment and intelligent learning path design. Firstly, we construct a student user profile to characterize students from multiple dimensions; secondly, we establish a teaching quality evaluation index system, classify the student group through K-mean clustering algorithm, and explore the characteristics and educational needs of different categories of students; then we construct a personalized learning path recommendation model based on reinforcement learning, and take learners as intelligent bodies to optimize the learning strategy through continuous interaction; finally, we validate the effectiveness of the method through actual teaching cases. Finally, the effectiveness of the method is verified through actual teaching cases. This study hopes to accurately identify the characteristics of student groups and provide personalized learning path recommendation for different types of students through the combination of data mining and artificial intelligence technology, so as to improve the scientific nature of education quality assessment and the relevance of learning effects, and provide theoretical support and technical solutions for intelligent and personalized teaching practice.

Table 1: The weight of the teaching quality index

Criterion layer	Index layer
Teaching management X1	Teaching time arrangement Y1
	Extracurricular practice Y2
	Performance rating system Y3
Teacher X2	Teaching attitude Y4
	Instructional language Y5
	Student interaction Y6
	Lecture questions Y7
	Contact practice Y8
	Classroom organization Y9
Student X3	Learning interests and motivation Y10
	Learning habits Y11
	Learn spontaneity and initiative Y12
	Learning method Y13
	Application capacity Y14
Course setting X4	Applicability of material Y15
	The interest of the course Y16
	Course difficulty Y17
	Curriculum value Y18

II. K-mean clustering based education quality assessment methods

II. A. Analysis of students' diversified learning needs

II. A. 1) Profiling student users

The drawing of student user profiles [22] is based on various types of student-related data, and most of them come from the application systems such as academic affairs and attendance, student management and so on. The behavior represented by each data is classified to build a user profile label, which can be used to describe and

portray students from different dimensions and build a vivid and clear student profile. In the process of drawing student user profiles, this paper selects nine labeling indicators with significant selection value in terms of the content of inquiry, such as “country”, “grades (average score in major courses)”, “study habits”, etc., and conducts standardized numerical evaluation of some highly subjective data to finally output a visual user profile, which helps explore the learning status and needs of the student more intuitively. In the process of creating the user profile, this paper selects nine label indicators with significant selection value in terms of the content of the inquiry, and carries out standardized numerical evaluation of some highly subjective data, and finally outputs the visual user profile, which is helpful for more intuitively exploring the current status of the student's learning and his/her needs.

II. A. 2) Establishment of a teaching quality evaluation index system

Through the analysis and research of student teaching, this paper has sorted out many factors affecting teaching quality. In the hierarchical structure of teaching quality evaluation, four first-level indicators and eighteen second-level indicators are identified, and the weights of the indicators are calculated by combining the results of the expert sorting method on the importance of the indicators, and the results are shown in Table 1.

II. B.K-means mean clustering algorithm

Cluster analysis is one of the important techniques in the field of data mining, cluster analysis deals with data objects that are unknown, the process of cluster analysis is the process of grouping the data objects in a given set into multiple clusters, the data objects within the same cluster have strong similarity, while the data objects between different clusters have dissimilarity, the data objects in the set are grouped into multiple clusters, this process is called cluster analysis. Cluster analysis is widely used in meteorological analysis, financial analysis, experimental analysis and other data processing K-means clustering [23] algorithm is the most popular algorithm in cluster analysis, based on the K-means clustering algorithm and derived from many algorithms, such as standard K-means, Scalable K-means, EM, etc., the purpose of the cluster analysis to make the results can be optimized. In the K-means mean clustering algorithm, K is used as a parameter for the number of clusters, i.e., n data objects are divided into K clusters, so that the data objects within the clusters have a high degree of similarity, while the data objects between the clusters have the lowest degree of similarity.

In the dataset B containing n data points, i.e., $B(x_1, x_2 \cdots x_n)$, k is the value of the number of expected clusters, and $M_1, M_2, M_3 \cdots M_K$ is a subset of K of M, and $M_i \cap M_i = \Phi$ In the K-means clustering algorithm, the Euclidean distance is used to obtain the center of mass C_j nearest to the point x_i.

Algorithm inputs: (1) the number of clusters K; (2) a dataset with n data objects.

Algorithm function: K-means algorithm calculates the mean value based on the objects in the clusters

Algorithm output: form K clusters such that the mean value of the distance between data objects in each cluster is minimized.

K-means mean clustering algorithm steps are as follows:

1) Among the n data objects in the data set, select K random data objects as the center of mass of the K clusters

2) Assign the remaining data objects to the class belonging to the center of mass closest to them, forming K classes using the clustering criterion function:

$$E = \sum_{i=1}^{k} \sum_{0 \in C_i} \| O - M_i \| \tag{1}$$

The calculation of the distance is performed by this clustering criterion function, where the value of K is the number of classes, and M_i is the representative of class C_i, which is the square of the Euclidean distance between the remaining data objects 0 in class C_i and the representative of class C_i, M_i.

3) Calculate the centers of gravity of each class, and use these centers of gravity as the representatives of each class, using the algorithm:

$$a_j = \frac{\sum_{i=1}^{n} 1\{C^{(i)} = j\} x^{(i)}}{\sum_{i=1}^{n} 1\{C^{(i)} = j\} x^{(i)}} \tag{2}$$

4) Repeat steps (2) and (3) continuously until all data objects are no longer assigned or the maximum number of iterations is reached, or function convergence is reached, which results in the division of K clusters leading to the minimum criterion function.

The K-means mean clustering algorithm is implemented as follows:

1) K different data objects are randomly selected from dataset B as representatives $H_1, H_2, H_3 \cdots H_k$ of K classes $C_1, C_2, C_3 \cdots C_k$.

2) repeat for each data object 0 in B.

$\| 0 - H_i \| = \min\{\| 0 - H \| 1 \le j \le k\}$ //Find the P_i that has the closest distance to 0. Assign O to the C_i to which H_i belongs to, and for each class $C_j (1 \le j \le k)$, $H_j = M_j = \frac{\sum_{0 \in C_i}^{n} x}{| C_j |} / /$ the center of gravity of the class is chosen as the representative of the class $E = \sum_{i=1}^{k} \sum_{0 \in C_i} \| O - M_i \|$ Compute the clustering criterion function E.

3) until K class representatives do not change anymore.

The general clustering effectiveness is measured by the average quantization error $q(c)$ as follows:

$$q(c) = \frac{1}{n} \sum_{i=1}^{n} (X_i, C_j) \tag{3}$$

It works better when $x_i \in D_i$ and the value of $q^{(c)}$ is small.

On the basis of portraying individual student portraits, student groups will be further divided and their characteristics analyzed through clustering modeling to achieve a stratified group portrait of students. In this paper, the mean clustering algorithm in K-means clustering is used in the classification of the model, and the cluster analysis is characterized by simple calculation, and the results presented are relatively intuitive and easy to understand. When used in the field of teaching, it can effectively analyze student learning.K-means clustering analysis is an analysis method that selects the initial center and carries out several iterations, calculates the size of the distance between each body and the clustering center, and finally divides the samples into different categories.

The main tasks in the data preparation phase include the collection, extraction and transformation of raw data. Since students' learning differences are affected by internal and external factors such as cultural background, language foundation, learning motivation, and individual learning ability to varying degrees, this paper extracts six typical characteristic dimensions that can reflect the student's learning level, willingness to learn, and ability to learn as labeling indexes for modeling and stratification in the student's user profile.

According to this, the collected indicators are "learning level", "average grade of professional courses", "number of attendance", "whether the learning objectives are clear", "learning habits tendency" and "one week of extracurricular learning time". In order to achieve hierarchical measurement, it is necessary to transform the string type data accordingly. In the label indicator "Chinese Proficiency", the value of "not at all understandable" is 1, and the value of "can listen and speak simply" is 2. Specifically, refer to the "Evaluation Criteria" column in Table 1 to realize the label data preprocessing in turn.

This modeling randomly selected 50 students' information data in the past year as a sample and exported to Excel to save, in order to avoid the phenomenon of null values or duplication of some data due to missing exams or inaccurate collection, the invalid sample data with missing, heavy null values and other situations were timely eliminated. And SPSS software was used to achieve K-means cluster analysis of the data, after the completion of the data import, in order to solve the problem of inconsistency of the scale and other issues also need to standardize the data in order to weaken the order of magnitude of the differences.

III. Empirical evidence for the assessment of the quality of education

In this paper, the scores of online education assessment of a city's 10 districts and counties, totaling 100 schools and 30,000 students during the period of May 2018 to October 2019, are collected through the method of education assessment.

Before clustering students, a reasonable number of student categories is first determined by minimizing the sum of squares of distances within the clustered groups. After determining the number of clusters, we need to determine the initial class center of each class, we take the contribution rate of each principal component as the weight to find out the point with the farthest distance as the initial class center, and cluster the data by K-Means.

III. A. Determination of the number of classes

Before clustering the assessment scores with K-Means, the first step is to determine the number of classes. Determining the appropriate number of classes has a great impact on the clustering results. In this paper, the number of suitable classes is determined by minimizing the sum of squared distances within the clustered groups.

The basic idea is to make the difference between sample points of the same class after clustering as small as possible, and the difference between sample points of different classes as large as possible. The main steps are as follows:

① Set the samples to be clustered have P variables, respectively, in order to cluster the samples into 1 to P classes.

② For each clustering result, calculate the sum of squared distances within the group.

③ Compare the sum of squares of intra-group distances under different clustering results, and select the appropriate number of clusters.

Figure 1 for the sample data in different clusters under the intra-group distance sum of squares curve, from the figure we can find: when the number of clusters in the 4, the intra-group distance sum of squares tends to be almost unchanged, i.e., to increase the number of classes will not be significantly less than the intra-group distance sum of squares. Therefore, we first consider clustering the samples into 4 classes.

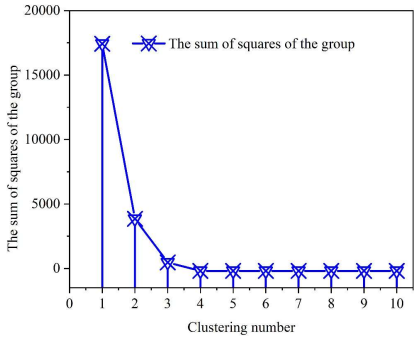


Figure 1: Sum of squares in the group of different clustering number

III. B. Clustering results

By minimizing the sum of squares of the distances within the clustered groups, we determined that we first clustered the dimensionality-reduced student assessment scores into 4 classes, and the results are shown in Fig. 2, where the numbers represent the location of the center of each cluster. From the figure, we found that the distance between the centers of classes 1, 2, 3, and 4 is farther, and the samples are better differentiated. But the distance of the clustered part of the sample is closer, in order to explore whether there is a more appropriate number of clusters, this paper clusters the samples into 5 and 6 classes, and the clustering results are shown in Fig. 3 and Fig. 4, respectively.

Figure 3 shows the results of clustering students' assessment scores into 5 classes, and it is found that the centers of 1, 2, 3, and 5 classes are farther away, and can distinguish the data better. The center of 4 classes is close to the center of 1 class, and the distinguishing effect is poor.

Figure 4 shows the results of clustering the students' assessment scores into 6 categories, and it is found that the centers of 1, 2, 3 and 5 categories are relatively far away from each other, but the samples of 4 and 6 categories overlap more, and the samples are poorly differentiated. Therefore, we may wish to cluster the sample data into 4 classes, and the next analysis are based on clustering the samples into 4 classes.

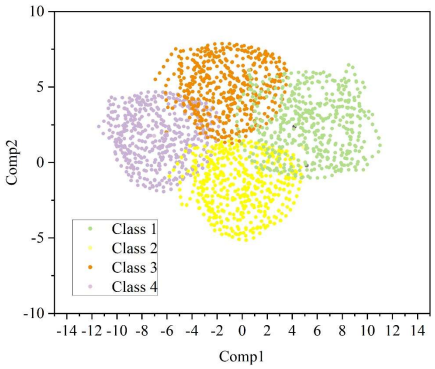


Figure 2: Rendering of 4 clusters

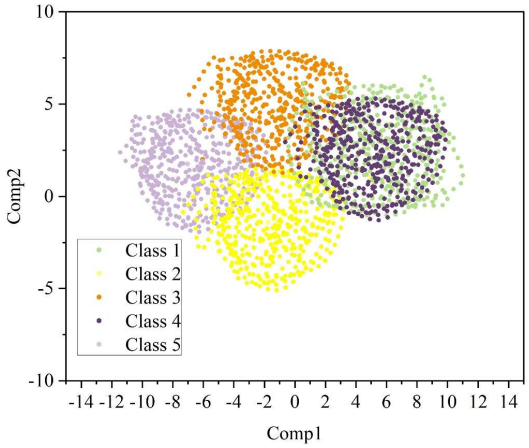


Figure 3: Rendering of 5 clusters

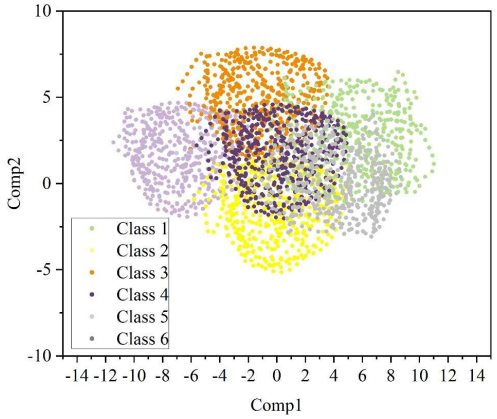


Figure 4: Rendering of 6 clusters

The students' assessment scores were clustered into four categories with sample sizes of 10144, 6845, 7168 and 5843 for each category. it was found that category 1 had the highest number of students, category 4 had the lowest number of students, and the number of students in categories 2 and 3 were close to each other. The coordinates of each cluster center are shown in Table 2.

The own, horizontal and vertical comparison of each class of students affected by each principal component revealed that for class 1 students, the second, sixth and seventh principal components had a positive effect and the first, third, fifth and eighth principal components had a negative effect. This reflects that these students have moderate learning pressure, get along well with their teachers, parents and classmates in their daily life, and can control their emotions well, but their memory performance is poor, and their learning methods need to be improved.

For Category 2 students, this category accounts for 22.82% of the total. The main characteristics are that they are under greater pressure to study, they get along well with their friends, but their relationship with their parents is not satisfactory, and they are optimistic, but lack motivation. For this group of students, teachers and parents should communicate with them, understand their inner thoughts, play a guiding role, and help them establish a correct concept of learning; students themselves should open their hearts to teachers and parents, reflect their problems to parents and teachers in a timely manner, learn to plan for themselves, and improve their ability to execute.

For Category 3 students, this category accounts for 23.89% of the total. They are mainly characterized by very high academic pressure, poor relationships with teachers, parents and classmates, and poor control of their emotions. These students are the ones with relatively bigger problems. This kind of students is often called “problem students”, so teachers and parents should focus on concern, specifically analyze the student's problems and the root causes of the problem, to find appropriate solutions; students themselves should learn to combine work and rest in the learning process, calm down when things go wrong, learn to talk to others and learn to control their own emotions.

For Category 4 students, this category accounts for 19.48% of the total. The main characteristics of these students are that they are under less pressure to learn, have good learning methods, have good academic performance,

have good relationships with teachers, parents and classmates, but have poor emotional control. Teachers and parents should cultivate students' ability to control their emotions; students themselves should find good channels to vent their emotions, cultivate hobbies and interests, and achieve all-round development. Teachers should also analyze the learning habits and characteristics of these students to better help other students.

Table 2: Class center coordinates of 4 clusters

Categories	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6	Comp.7	Comp.8
1	-0.76	0.64	-0.31	-0.09	-0.08	0.05	0.1	-0.07
2	1.06	-1.69	-0.04	-0.05	0.01	0	0.05	-0.04
3	3.18	-0.75	0.25	0.11	0.01	-0.02	-0.09	0.04
4	-4.36	0.09	0.17	0.1	0.12	-0.04	-0.12	0.06

IV. Intelligent Learning Path Recommendation Based on Reinforcement Learning

IV. A. Learning path recommendation method

There are three main research areas in learning path recommendation. The first is learner modeling, which involves the characterization problems and extraction methods of learners' ability level, psychological state, style interest, etc.; the second is learning object modeling, which involves the discovery of the correlation information between the learning recommendation object and the learner's personalized parameters; and the third is the design of recommendation algorithms, which involves the problem of strategy selection and optimal matching between the learner and the object. According to related research, the mathematical description of the personalized learning path recommendation problem takes the following form. Given a learning objective g, knowledge points kp related to the learning objective, learning resources $r=(kp_1,kp_2,\cdots,tp,s)),s.t.kp_{i+1}=f(kp_i)$, r is an ordered vector of knowledge points kp, where tp $\in$ {text, image, video} represents the type of r, s $\in$ {course, chapter, knowledge unit, knowledge point} represents the granularity of r related to the hierarchy of the objective g, and $f(\cdot)$ is a conversion function representing the knowledge points kp_i is a prerequisite learning resource for kp_{i+1}. The vector e is generally used to describe the a priori features of each learner, and the feature vectors are generally used as the initial input values of various recommendation algorithms. The learning path is denoted as pn and has $pn'=\{e',r'\}$, where e and r denote the learner's state characteristics and the set of learning resources corresponding to time t, respectively. Thus, the learning path L_p is an ordered sequence of pn nodes with g as the goal associated with the specified learner, $L_p=\{pn^0,pn^1,\cdots,pn^t,\cdots,pn^m \mid pn^t=f(pn^{t+1}),g=f(L_P)$.

The mainstream learning frameworks for personalized learning include three types: machine learning-based, evolutionary computation-based, and knowledge graph-based. Machine learning frameworks convert path recommendation into a prediction problem, which is categorized into supervised and unsupervised. Evolutionary computation generally uses genetic algorithm, ant colony algorithm, etc. to solve the path search problem, and currently population computation and population intelligence are its key research directions. Knowledge mapping is based on knowledge engineering and ontological methods, and the path search method with constraints is used to find out the best recommended path on the basis of constructing a domain map.

Reinforcement learning has become a hot research topic in machine learning and gradually evolved into a popular branch, and its upgraded version, AlphaGo zero, based on the deep reinforcement learning framework [24], continues to outperform human players in the World Go Championships and competitive gaming competitions. Reinforcement learning is particularly suitable for solving sequential decision optimization problems, and can achieve better results in personalized learning path recommendation.

IV. B. Learning path recommendation based on reinforcement learning

Reinforcement learning treats the learner as an intelligence and guides it to learn autonomously through continuous “trial and error”. The classical learning model is based on a Markovian decision-making process, where the learner continuously learns new knowledge points and then uses the reward received to guide the suitability of the learning behavior to maximize the cumulative reward to achieve a specific goal. This process can be simply described by the 5-tuple (S,A,P,R,γ), where S is a finite set of states, A is a finite set of actions, P is the state transfer probability, R is the reward function, and γ is the discount factor used to compute the cumulative reward.

The policy π is a mapping of states S to actions A, and the policy π assigns an action probability $\pi(a \mid s)=p(A_t=a \mid S_t=s)$ to each state s. The goal of reinforcement learning is to discover an optimal policy π^r for the learner that maximizes the sum of the expected discounted rewards received by the learner, i.e.,:

$$V^{\pi}(s)=\sum_{a}\pi(s,a)\left[R(s,a)+\gamma\sum_{s}Pr(s/s,a)V^{\pi}(s)\right] \tag{4}$$

$R(s,a)+\gamma\sum_{i}Pr(s^{'}\mid s,a)V^{\pi}(s)$ represents the cumulative future rewards for each possible decision in the current s state. The learner chooses actions in the current state by "trial and error", and cycles through this sequence of state→action→reward to maximize the benefit of the learner's specified goal.

User feedback. For a learner, assume that at a certain moment t the model selects the current transition state as s based on the preceding state, and the user can give a positive and negative label l^{+},l^{-} to this state, then for all state sequences S, there are:

$$C=\{(s_1,T_1,T_1,\cdots,(s_1,T_1,T_1,\cdots,(s_n,T_n,T_n)\},where\ T_i,T_i\in(0,\cdots,1) \tag{5}$$

The formulas are further integrated to form a parameterized policy optimization framework based on user feedback:

$$G_{\eta}(C,V)=\eta\cdot T(\theta,V)-(1-\eta)\cdot L(\theta,C) \tag{6}$$

The T function represents the maximum discount reward based on the parameter θ, L is the loss assessment function, which represents the gap between the labels chosen by the learner and the realized state transitions, and η serves as a moderator to balance the weights of the maximum discount reward and the loss function; in other words, when η is taken to be l, the formulation degenerates into a traditional strategy-based optimization algorithm. $T(\theta,V)$ can be solved using a gradient descent method, specifically, assuming that the learner l obtains the learning path recommendation with maximum reward π_{θ}^{*} with probability $p_{\theta}(T_l)$, then there are:

$$\begin{aligned} T(\theta,V)&=\frac{1}{W(V)}\sum_{v_j<i}w(V_j)R_i \\ \text{where}\ W(V_i)&=\frac{P_s(T_i)}{P_{s_i}(T_i)} \\ \text{and}\ W(V)&=\sum_{v_i<i}w(V_i)R_i \end{aligned} \tag{7}$$

where R_l denotes a particular learning path computed by learner l.

$$L(\theta,C)=-\sum_{i=1}^{M}\log(E[\pi_{\theta}(O(s_i)\mid l_i^{*},l_i^{-})]) \tag{8}$$

where $(O(s_i)\mid l_i^{*},l_i^{-})$ 1 is a conditional probability on the user's labeling and assumes that each learner's labeling feedback is independently homoscedastic.

V. Reinforcement Learning Based Intelligent Learning Path Generation Effect Validation

In order to verify the effectiveness of the learning path generation method proposed in this paper, the study selects real learners, learning resources and the fit relationship between them (questionnaire and measurement scale), and takes the content of "Web Attack" and "Web Defense" in the elective course "Network Attack and Defense Technology" of a university as an example to evaluate and test the proposed learning path generation mechanism. At the same time, the learning path recommendation network platform was preliminarily designed and implemented.

V. A. Data sources

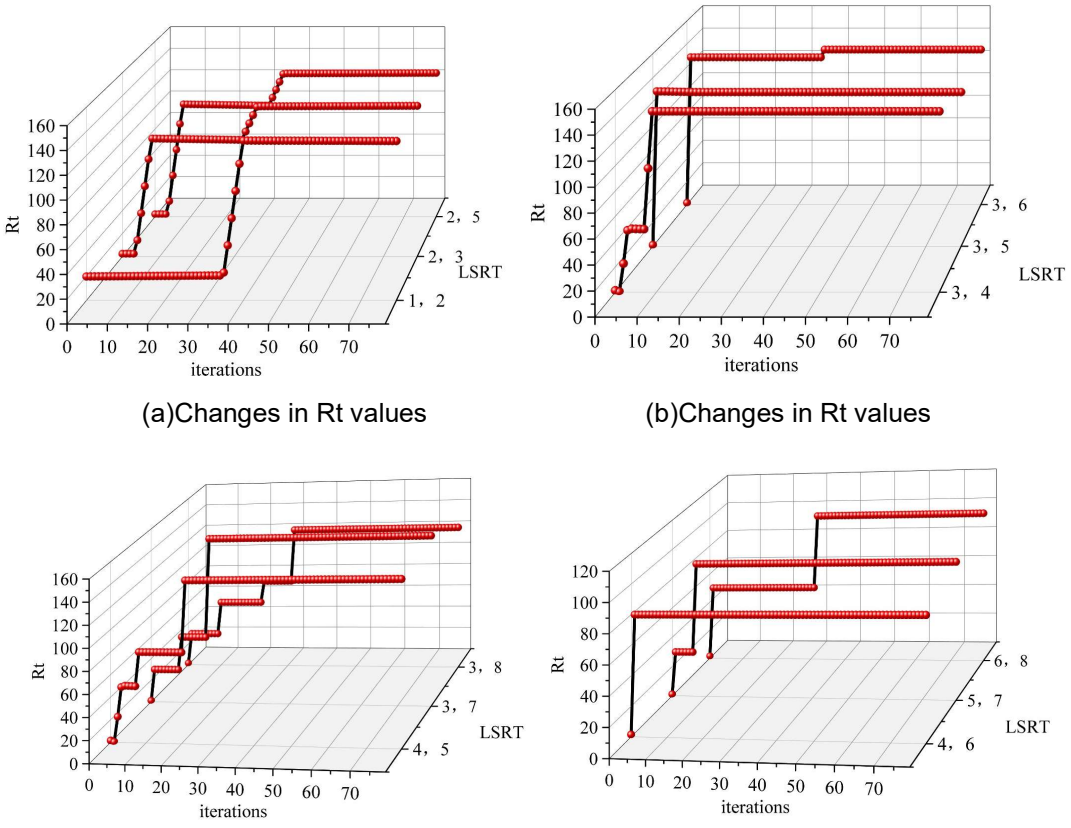
The learner data comes from the students who take the course of "Web Attack and Defense Technology" in the College of Computer Science and Technology of a university, totaling 50 people. Learning resources data from a university national computer information security and network attack and defense virtual simulation experimental teaching platform in the "Web attack" and "Web defense" two units of virtual simulation experiments, including a total of eight virtual simulation experiments, including Four principles of demonstrative experiments (Web application architecture, Web application information collection, Web application security threats and Web Trojan detection and analysis), three design interactive experiments (SQL injection, XSS cross-site scripting attacks and phishing techniques), a comprehensive experiment (Web penetration attack experiment).

V. B. Experimental Procedures

Learners need to set two variables according to their own needs, i.e., the learning target value D and the fitness weights. The learning target value can be set arbitrarily according to their own situation. When the learning target value is larger than all the learning resource gain values, the learning process will stop after exceeding the maximum number of iterations, and the generated learning path recommendation queue will be the largest gain value queue under the current demand. When the learning target value is smaller than all the learning resources benefit value, the learning process will stop when the learning target value is reached, and the generated learning path recommendation queue will be the largest benefit value queue to meet the current demand.

When we set the learning benefit weights to 4, education weights to 2, group weights to 4, technology weights to 1, submitted after 2000 iterations, the generated learning path is 0. Start → 1. Web application architecture → 2. Web application information collection → 3. Web application security threats → 8. Web Trojan Detection and Analysis → 9. End.

When we set the four weights to (2, 2, 2, 2), observe the change of learning selection gain LSRT, the resulting number of iterations and Rt values are shown in Fig. 5, Figs. (a)-(d) are all the display of effective items in the LSRT matrix, shown in Fig. 5 for the 80 iterations within the LSRT matrix of the effective items, i.e., the value of the item changes, the other value of the unchanged item on the learning path has no The other terms with unchanged values have no effect on the learning path and can be regarded as invalid terms. We can see the change of Rt value during the iteration process, the values of each value of the learning selection gain table converge after 45 iterations, which indicates that under the current parameters, the learning gain reaches stability after 45 iterations, and the learning path is generated due to the size of the learning problem is not too large, and the effective terms basically complete convergence in 80 iterations, i.e., the intelligence of the reinforcement learning has already obtained a stable solution, and a stable and optimal learning path has been generated. From this process, we can see that the reinforcement learning intelligent body carries out human-like selection operation, and through hundreds of trial and error and continuous accumulation of experience, it finds the learning path that meets its own needs and maximizes the learning gain. This part of the experiment proves the effectiveness of the adaptive learning path method based on reinforcement learning.



(a)Changes in Rt values (b)Changes in Rt values

(c)Changes in Rt values (d)Changes in Rt values

Figure 5: The change in rt values of the valid items in the LSRT matrix

V. C. Analysis of results

We generated the learning paths of 50 students with valid questionnaires, of which 10 students' learning paths are shown in Table 3, from which we can see that after students choose different weights, the wisdom of enhanced learning will be based on the students' individualized needs, combined with the questionnaire's fit information, to generate personalized learning paths, such as the table of the student 1's learning gains weighted as 5, education weighted as 1, group weight is 3, technology weight is 5, and the corresponding learning path is 0→1→2→5→6→7→8, while student 2's learning gain weight is 4, education weight is 2, group weight is 4, technology weight is 5, and the corresponding learning path is 0→1→2→5→9.

These two types of personalized choices bring about differences in learning paths. As a result, the intelligence based on reinforcement learning is able to generate customized learning paths according to students' individualized needs. In addition, we can also see from the table that other students also choose different weights according to their own needs and obtain different learning paths, here the first four of the learning paths have been fixed due to the knowledge-supporting relationship of the learning resources, so that the learning paths in the table seem to have fewer change pairs. This part of the experiment proves the practicality of the adaptive learning path method based on reinforcement learning.

Table 3: Student choice values and corresponding generation learning paths

Student	FW	[EW,SW,TW]	Learning path
1	5	[1,3,5]	0→1→2→5→6→7→8
2	4	[2,1,5]	0→1→2→5→9
3	8	[2,4,6]	0→1→2→5→8
4	9	[2,5,5]	0→1→2→5→7
5	1	[7,5,1]	0→1→2→5→6→8→9
6	5	[2,5,6]	0→1→2→5→6→7
7	4	[7,5,8]	0→1→2→5→6→8
8	3	[1,4,5]	0→1→2→5→6→9
9	6	[1,2,7]	0→1→2→5→6
10	7	[4,5,6]	0→1→2→5→6→7→9

VI. Conclusion

The K-mean clustering combined with reinforcement learning approach to education quality assessment and intelligent learning path design has demonstrated significant results in practice. By analyzing the assessment data of 30,000 students, four clusters were determined to be the optimal classification scheme, and the sum of squares of intra-group distances stabilized after four classes. The clustering results successfully categorized students into four classes, accounting for 33.81%, 22.82%, 23.89%, and 19.48% of the overall population, with each class exhibiting unique learning characteristics and demand patterns. Based on this classification, the reinforcement learning algorithm generates a personalized learning path through iterative optimization, and the learning gain reaches a stable state after 45 iterations, providing customized learning solutions for students with different learning characteristics.

The experiment proves that even if learners choose different combinations of weights, such as learning gain weights, education weights, group weights and technology weights, the system is able to generate optimal learning paths that meet their individualized needs. This approach overcomes the limitations of traditional educational assessment, and is able to accurately identify the differences in student group characteristics, and provide targeted teaching support and learning path design based on this.

Future work will further optimize the clustering feature selection and reinforcement learning algorithm parameters, expand the experimental sample size and subject coverage, explore the possibility of applying this method to more complex educational scenarios, and provide technical support for the construction of a more intelligent and personalized educational ecosystem.

References

[1] Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. IEEE access, 8, 80716-80727.
[2] Chong, B. (2021). K-means clustering algorithm: a brief review. Academic Journal of Computing & Information Science, 4(5), 37-40.
[3] Yuan, C., & Yang, H. (2019). Research on K-value selection method of K-means clustering algorithm. J, 2(2), 226-235.
[4] Ashabi, A., Sahibuddin, S. B., & Salkhordeh Haghighi, M. (2020, December). The systematic review of K-means clustering algorithm. In Proceedings of the 2020 9th International Conference on Networks, Communication and Computing (pp. 13-18).

[5] Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. Electronics, 9(8), 1295.
[6] Dhanachandra, N., Manglem, K., & Chanu, Y. J. (2015). Image segmentation using K-means clustering algorithm and subtractive clustering algorithm. Procedia Computer Science, 54, 764-771.
[7] Liu, R. (2022). Data Analysis of Educational Evaluation Using K‐Means Clustering Method. Computational Intelligence and Neuroscience, 2022(1), 3762431.
[8] Chi, D. (2021, January). Research on the application of k-means clustering algorithm in student achievement. In 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE) (pp. 435-438). IEEE.
[9] Noaman, A. Y., Ragab, A. H. M., Madbouly, A. I., Khedra, A. M., & Fayoumi, A. G. (2017). Higher education quality assessment model: towards achieving educational quality standard. Studies in higher education, 42(1), 23-46.
[10] Liu, S., & Liu, S. (2016). Higher education quality assessment and university change: A theoretical approach. Quality Assurance and Institutional Transformation: The Chinese Experience, 15-46.
[11] Gerritsen-van Leeuwenkamp, K. J., Joosten-ten Brinke, D., & Kester, L. (2017). Assessment quality in tertiary education: An integrative literature review. Studies in Educational Evaluation, 55, 94-116.
[12] Pettersen, I. J. (2015). From metrics to knowledge? Quality assessment in higher education. Financial Accountability & Management, 31(1), 23-40.
[13] Vykydal, D., Folta, M., & Nenadál, J. (2020). A study of quality assessment in higher education within the context of sustainable development: A case study from Czech Republic. Sustainability, 12(11), 4769.
[14] Lucander, H., & Christersson, C. (2020). Engagement for quality development in higher education: a process for quality assurance of assessment. Quality in Higher Education, 26(2), 135-155.
[15] Li, Y., & Zhang, H. (2024). Big data technology for teaching quality monitoring and improvement in higher education-joint K-means clustering algorithm and Apriori algorithm. Systems and Soft Computing, 6, 200125.
[16] Zhang, D. (2025, January). Quality Evaluation of College English Classroom Teaching based on K-Means Clustering Algorithm. In 2025 International Conference on Intelligent Systems and Computational Networks (ICISCN) (pp. 1-6). IEEE.
[17] Lang, W. (2022). Research on College English Teaching Quality Assessment Method Based on K‐Means Clustering Algorithm. Mathematical Problems in Engineering, 2022(1), 4134827.
[18] Li, F., & Li, Q. (2025). Research on Intelligent Learning Path Planning and Recommendation Algorithm in OMO Teaching Mode Based on Artificial Intelligence Grand Modeling. J. COMBIN. MATH. COMBIN. COMPUT, 127, 6927-6942.
[19] Zhang, J. (2025, March). Personalized learning pathway design and course recommendation based on K-means, Apriori, and genetic algorithm. In Second International Conference on Big Data, Computational Intelligence, and Applications (BDCIA 2024) (Vol. 13550, pp. 1039-1045). SPIE.
[20] Obaid, T., Aghajani, H., & Linn, M. C. (2023). Using optimized clustering to identify students' science learning paths to knowledge integration. STEM Education Review, 1.
[21] Hong, J. (2021). Research on the Development of Innovation Path of Ideological and Political Education in Colleges and Universities Based on Cloud Computing and K‐Means Clustering Algorithm Model. Scientific Programming, 2021(1), 4263981.
[22] Ma Lina. (2024). Precise Implementation of Mental Health Thinking Measures in Universities Based on Student Portrait Analysis. International Journal of e-Collaboration (IJeC),20(1),1--1.
[23] Masao Ueki. (2025). A deflation-adjusted Bayesian information criterion for selecting the number of clusters in K-means clustering. Computational Statistics and Data Analysis,209,108170-108170.
[24] Keyu Wu,Mahdi Abolfazli Esfahani,Shenghai Yuan & Han Wang. (2018). Learn to Steer through Deep Reinforcement Learning. Sensors,18(11),3650-3650.