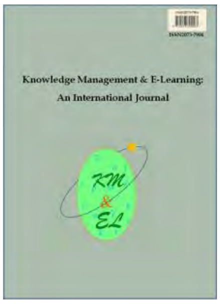
A study on the assessment of learner ability based on the combination of auto-encoder and quadratic k-means clustering algorithm

Junfeng Man
Hunan University of Technology, China
Hunan First Normal University, China
Rongke Zeng
Hunan University of Technology, China
Xiangyang He
Hua Jiang
Hunan First Normal University, China



Knowledge Management & E-Learning: An International Journal (KM&EL)
ISSN 2073-7904

Recommended citation:
Man, J., Zeng, R., He, X., & Jiang, H. (2024). A study on the assessment of learner ability based on the combination of auto-encoder and quadratic k-means clustering algorithm. *Knowledge Management & E-Learning, 16*(4), 697–715. https://doi.org/10.34105/j.kmel.2024.16.032

Knowledge Management & E-Learning, 16(4), 697–715

A study on the assessment of learner ability based on the combination of auto-encoder and quadratic k-means clustering algorithm

Junfeng Man

College of Computer Science
Hunan University of Technology, China
Hunan Provincial Key Laboratory of Informatization Technology for Basic Education
Hunan First Normal University, China
E-mail: mjfok@qq.com

Rongke Zeng

School of Computer Science
Hunan University of Technology, China
E-mail: 1119884984@qq.com

Xiangyang He

School of Education
Hunan Provincial Key Laboratory of Informatization Technology for Basic Education
Hunan First Normal University, China
E-mail: sunhexy@163.com

Hua Jiang*

School of Education
Hunan First Normal University, China
E-mail: 99245342@qq.com

*Corresponding author

Abstract: At present, the widespread use of online education platforms has attracted the attention of more and more people. The application of AI technology in online education platform makes multidimensional evaluation of students' ability become the trend of intelligent education in the future. Currently, most existing studies are based on traditional statistical methods to rank and evaluate students' achievement, but this will lead to problems of single type of data and the inability of intra-class evaluation. In order to solve above problems in traditional statistical methods, a multidimensional learning ability evaluation method is proposed in this paper, which is based on auto-encoder and quadratic K-means clustering algorithm. It will be applied to the domain of intelligence education to evaluate students' multidimensional learning ability. First, this method uses auto-encoder (AE) to reconstruct the students' learning behaviour features in order to improve the clustering effect, then performs k-means clustering twice on reconstruction data. By using clustering to address the issue that cannot be addressed within the category, it ranks and evaluates students. This research employs a real data set of a particular platform for comparative studies in order to assess the performance of this strategy on various data sets. The results of the experiments demonstrate that this method performs much better than both the conventional clustering algorithm and the PCA-based reconstruction clustering method.

Keywords: Intelligence education; Learner competence assessment; Auto-encoder; K-means clustering

Biographical notes: Dr. Junfeng Man is the dean of the School of Computer Science in Hunan First Normal University. His research interest includes but are not limited to: artificial intelligence, big data, and educational technology.

Mr. Rongke Zeng is a graduate student at the School of Computer Science, Hunan University of Technology. His research interest includes but are not limited to: natural language processing, big data, and educational technology.

Dr. Xiangyang He is the vice dean of the School of Computer Science in Hunan First Normal University. His research interest includes big data, educational technology.

Miss. Hua Jiang is a teacher of the School of Computer Science in Hunan First Normal University. Her research interest includes big data, educational technology and association rules.

1. Introduction

The recent rapid development of information technology has presented a number of fresh market issues and industrial strategic opportunities to a variety of sectors. The same is true for the education sector, but it has been challenging for the traditional educational model to satisfy the needs of students. Currently, an increasing number of people prefer to improve themselves through online learning (Dong et al., 2023). As artificial intelligence technology advances, intelligence education has been proposed and defined as a new educational model based on a new generation of information technology. Through the use of educational big data, including resources, behaviours, situations, and management involved in teaching, management, evaluation, and decision-making (Zheng et al., 2019), it seeks to mine, analyze, integrate, and create a learning ecosystem with the characteristics of intelligent learning guidance, accurate recommendation, and fine evaluation. Colleges and universities have steadily gathered a significant amount of educational data resources as a result of the development of smart campuses and education informatization. In order to better support education and teaching management, it is essential to extract useful information from these enormous educational information sets. In recent years, educational data mining has drawn increasing interest as a developing research area, particularly when used to assess students' multidimensional learning capacity in a particular course (Yang et al., 2021).

The examination of students' learning ability is a procedure that focuses on students' learning. We can construct a process that captures the dynamic development characteristics of students' learning ability by mining the relationship between learning content and learning efficiency. It is vital to investigate the intelligent assessment method of learners' multidimensional learning ability in order to conduct refined and whole-process evaluation of learners, teachers, environment, and other aspects. There are currently few research on the methods used to assess students' learning abilities. Their student evaluation research is based solely on simple data mining technologies. For example, they only utilize data analysis software and standard clustering algorithms such as k-means to evaluate and analyze student skills, or they use advanced clustering algorithms to assess and analyze students' scores for the features of a large number of learning behaviour data provided by students. Furthermore, the data type is limited, which is insufficient for assessing students' abilities.

Image denoising (Guleria et al., 2023) and recommendation fields (Dong et al., 2023) are two areas where auto-encoder technology has shown great success. Auto-encoder technology can transform an object from its original space to a low-rank space and then map the object back to its original space. During this secondary mapping (data reconstruction) process, the original object's dimensions might be assigned once more. This research uses deep learning auto-encoder technology to reconstruct the features of students' multidimensional learning capacity in the area of education. That is, we may obtain the low-rank code of students' learning ability space by performing a nonlinear transformation on the space of learning data created by students during the learning process. To achieve the goal of feature reconstruction, the low-rank students' learning ability is then decoded and rebuilt into the space of students' learning ability.

The k-means method (Yuan et al., 2023) is an unsupervised learning algorithm with simple implementation and great results. Unlike the samples in other clusters, the samples in clustered clusters are comparable to one another. The samples in clustered clusters are similar to each other but not to the samples in other clusters. Despite the fact that the k-means clustering algorithm has been proposed for more than 50 years, it is still one of the most widely used partition clustering algorithms. Successful case studies and experience are the main reasons for its popularity (Hu et al., 2023). Students who have significantly different grades are clustered into different categories by the clustering algorithm, which can group students with similar grades together into the same category. The average score of each category is then determined in turn, and the average scores are sorted. The category with the highest average score receives a high score, while the one with the lowest receives a poor score. Based on this, the k-means algorithm is adopted to evaluate the learning behaviour data generated by students in this paper to obtain the student's ability assessment score.

Therefore, this study proposes a multidimensional learning ability evaluation method based on auto-encoder and quadratic k-means clustering algorithm. This method reconstructs the students' learning behaviour features using an auto-encoder to improve the clustering effect, then performs k-means clustering twice on dimension reduction data. It scores and assesses students by using clustering twice to address the question that cannot be assessed within the category. This paper uses a real dataset from a platform for comparative experiments to evaluate the performance of this method on multiple datasets.

With the development of intelligent education, better algorithms built on machine learning and deep learning have been used for tasks like course and exam recommendation. For example, Song and He (2022) propose an online course recommendation model based on an enhanced auto-encoder. They enhance the auto-encoder by using the long short-term memory network. This enables the model to extract the data's temporal features. In the course recommendation task, it performs better than the traditional auto-encoder algorithm. However, the disadvantage is that numerous LSTM network parameters need to be manually set up. It is necessary to investigate how to find the best parameter algorithm automatically. Xiong et al. (2019) propose a personalized test item recommendation method based on deep auto-encoder and secondary collaborative filtering. They integrate item response theory and deep auto-encoder score prediction into the recommendation algorithm. It solves the problems of incorrect positioning of recommended test questions and the lack of attention to the knowledge point level in conventional personalized recommendation algorithms. Through comparative experiments, they demonstrated that the results of the proposed recommendation method are more personalized and accurate than traditional test item recommendation methods. It completes the task of personalized test item recommendation. As a result, good results have been achieved in some areas of intelligent education based on machine learning and deep learning algorithms.

1.1. Related work

At the present time, there are also relevant studies on the multi-dimensional assessment of the learning ability of students. Peng et al. (2020), for example, use an algorithm that combines k-means clustering and SVM to evaluate the score data of each sub-question in 200 advanced mathematics student examination papers. They investigate eight student ability factors. Based on student performance, it indicates that the cluster SVM semi-supervised learning model performs well in classification tasks. However, it is obviously not enough to evaluate students' comprehensive ability only based on test scores. To extract multidimensional features from learner behaviour data, Li et al. (2022) propose an improved PCA-GRBM algorithm. They then use the DBSCAN algorithm for multiple clustering to the creation of learner profiles. Based on the data generated during the learning process, they divide students into three groups: excellent, low-level, and high-risk learners. Although their proposed algorithm outperforms traditional algorithms in terms of performance, there is still space for improvement. To complete the evaluation of the comprehensive quality of student groups, Xie et al. (2021) propose SOM neural network to create a classification model of student groups. They categorize the student population as exceptional, average, or poor in order to compare the students in each stratum. However, the interior of each layer is not comparable. According to class rankings, Chen (2018) divide student marks into A, B, and C categories. To address the issue that the k-means algorithm is unable to cluster non-numerical data, they propose the SimRank algorithm. The Apriori method for training program data mining is then introduced. They perform analyses of confidence based on course and professional performance as well as analyses of ability assessment based on the results of various courses undertaken by the group. It solves the problem that the present method cannot evaluate students in real-time. However, the assessment of the students' abilities is still done solely in terms of their academic achievement. The data format is relatively simple. Ji et al. (2017) employ the k-means algorithm to investigate the variables that influence students' ability based on the association rule model. However, they only conduct analysis and investigation on the ability of the students' grades. Chen (2015) evaluates the physical science academic ability of high school students in high school physics courses using the data analysis function in Excel software, and creates a performance ability radar chart of the students' physics knowledge points. Then he analyzes and diagnoses the subject's academic achievement using the students' academic level radar chart. He also employs the radar chart to evaluate students' abilities. This enables us to quickly assess the student's performance. To examine how teachers' educational technology skills develop and promote, Du et al. (2010) develops a radar map of those talents. A theoretical research reference for the linked literature on the creation of the student ability radar map is provided by the analysis of this article.

From the above papers, it can be seen that curriculum recommendations, test question recommendations, score analysis, and other aspects of students' education have gradually shifted toward informatization and intelligence. More research is also needed to assess students' multidimensional learning ability. Currently, the majority of researchers are evaluating and analyzing the scores of students' abilities in the context of exams using improved clustering algorithms. Although the categorization effect is achieved by classifying students as exceptional, medium, or qualified, this method cannot be examined within each category. Furthermore, under a specific course of the present widely used online education platform, there hasn't been any extensive research on how to assess students' learning ability using the data generated by their learning behaviour. This study suggests using auto-encoder and quadratic k-means clustering method to assess students' multidimensional learning ability in order to address the aforementioned issues.

The innovations of this paper are:

First, in order to create a general multidimensional index to assess students' learning ability in a course, we examine a dataset of learning behaviour features produced by learners during the learning process. We may create indicators for knowledge skill scores, learning attitude scores, course interest scores, and overall scores using the quadratic k-means clustering algorithm to show the various dimensions of students' learning abilities for a particular course. It offers a solution for the issue of single data type and single ability evaluation type. After secondary clustering, each student's score index is unique, resolving the intra-class evaluation problem that traditional methods cannot.

Second, because students demonstrate a diversity of learning behaviour features within a specific course, data reconstruction is required to improve the clustering impact. In this paper, we look at the effect of clustering algorithm on the reconstruction of learning behaviour features. The auto-encoder technique is used to reconstruct data and is contrasted to the traditional PCA data reconstruction method. The effectiveness of the algorithm used in this research for assessing students' multidimensional learning ability is then explored utilizing a variety of clustering techniques in comparison experiments.

Third, we assess students' multidimensional learning ability using the auto-encoder and quadratic k-means clustering approach. For teachers and students in the school to use as a learning reference, we also create a radar map of the students' multidimensional learning ability. It offers theoretical and statistical basis for the subsequent study on precise learning diagnosis, tracking of the learning process, and assessment of students' learning efficiency. In this way, it accomplishes the requirement of intelligent education, which is refinement and intelligence in the evaluation of students' learning ability.

The remainder of the paper is structured as follows: The necessary theoretical basis and overall methodology of the capability evaluation approach are described in Section 2 of this paper. The suggested model and the comparison model are applied to experimental investigation in Section 3 using actual data. Section 4 offers a brief summary of this study and suggests directions for additional research.

2. Methods

Because learning behaviour has numerous aspects and the learning behaviour data collected by each student during the learning process is not significantly different. This study first employs the auto-encoder to extract features and reconstruct the data and the secondary k-means clustering approach is then used to score students' ability evaluation in order to effectively score and evaluate students' various ability dimensions.

2.1. Auto-encoder

Rumelhart proposed the Auto-Encoder in 1986 as a design for an unsupervised neural network. Its primary objective is to learn a distributed feature representation that is compressed for a given data set (Li et al., 2023). A typical three-layer neural network serves as the auto-encoder. Between the input layer and the hidden layer, coding occurs, and between the hidden layer and the output layer, decoding occurs. By encoding the input data, it creates a coded representation of the data. Additionally, it reconstructs the input data by decoding the hidden layer's encoded representation. The learning impact of the auto-encoder algorithm is measured using the reconstruction loss function.

The basic structure of the auto-encoder used in this research is depicted in Fig. 1. The input layer in the construction of knowledge skill features is 4-dimensional features, the hidden layer is reduced to 3-dimensions, and the final output layer is 4-dimensional features to accomplish data feature reconstruction.

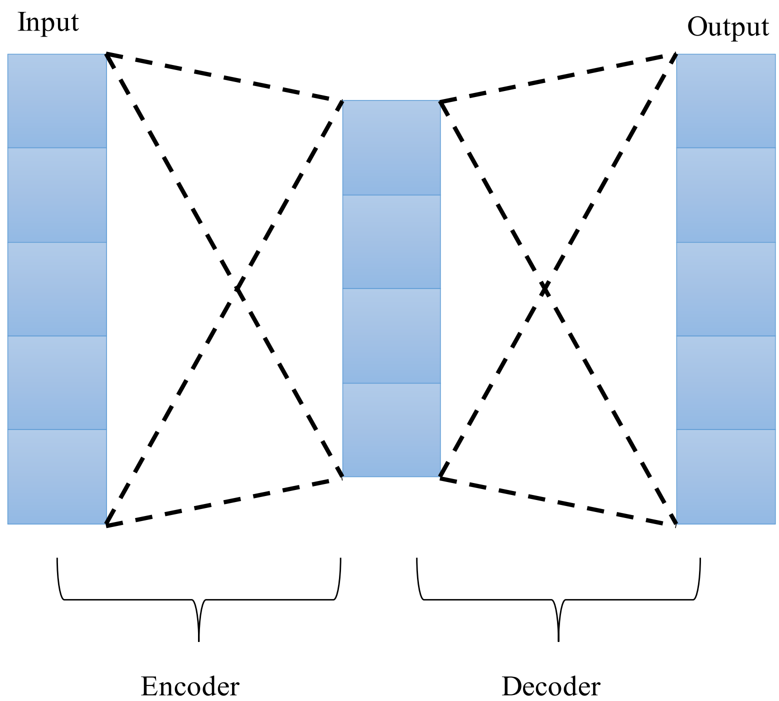


Fig. 1. The basic structure of the auto-encoder

The auto-encoder, as an unsupervised neural network, aims to produce output data that is as close to the input data as possible. The auto-encoder's encoding process is to encode the original data x into the encoder function of extracting feature z (z = f_ϕ (x)). The decoding process consists of a decoder function that converts the extracted feature z into output data r (r = g_θ (z)). The reconstruction loss function is used to calculate the difference between these two parts, with the aim of retaining as much original data information as possible in the extracted features (Tao et al., 2022). The formal meaning is as follows:

$$L = min\frac{1}{n}\sum_{i=1}^{n}\left(x_i - g_\theta(f_\phi(x))\right)^2 \quad (1)$$

$f_\phi(x)$ and $g_\theta(z)$ are both sigmoid functions, ϕ and θ represent the parameters of the encoder and decoder, respectively, *n* is the sample size, x_i is the input sample. The sigmoid function as formula (2):

$$Sigmoid(x) = \frac{1}{e^{-x}} \quad (2)$$

2.2. K-means clustering algorithm

The k-means method, the most widely used clustering algorithm, was first used by MacQueen (1967) in 1967. It is the most well-known clustering method in large data processing technology. K-means is a division hard clustering method as well (Qu et al., 2019). In comparison to other clustering algorithms, the k-means algorithm has been extensively used and explored in industry and scientific study areas due to its excellent effect and simple idea. The k-means algorithm works on the following principle: in training data set D, after inputting the necessary number of clusters K, choose at random the data vector with the same number of clusters from D as the initial cluster center. Then, using the minimal distance principle, the samples are aggregated by calculating the distance between each sample and the cluster center (Wang et al., 2020). Next, a new cluster center should be updated by averaging each cluster. Repetition and cycling are continuing. When the error sum of squares function value is steady at the lowest value, the iteration is finally terminated.

In the K-means algorithm, the Euclidean distance is used to calculate the distance between each sample in the data. It has lower algorithmic complexity and is suitable for large data sets. The given data D = {x_1,x_2,⋯,x_n } is used to calculate the Euclidean distance (Zhao et al., 2021) between any two sample points.

$$d(x_p \text{, } x_q) = \sqrt{\sum_{i=1}^{m}\left(x_{pi} - x_{qi}\right)^2} \quad (3)$$

In formula (3): $x_p = \{x_{p1}, x_{p2}, \cdots, x_{pm}\}$, $x_q = \{x_{q1}, x_{q2}, \cdots, x_{qm}\}$, *m* are the dimensions of the sample elements.

In the traditional k-means algorithm, the K value and the initial clustering center are randomly selected. To ensure the precision, objectivity, and validity of the clustering results, the appropriate K value should be determined from the data itself. The details of the data separation are more clearly seen thanks to K-means clustering. The SSE(the sum of squared errors) value decreases as the number of categories increases. On the SSE and K value relationship diagram, it is shown as follows: the change trend of the SSE is obvious when the K value is close to the true value of the number of clusters. When the K value is equal to the real value, the change trend of SSE slows down as the K value increases (Chen et al., 2022). In this study, the elbow technique is used to calculate the optimal number of K clusters. It is based on the relationship between the number of clusters and the sum of squared errors. The formula for the sum of squared errors function is as follows:

$$SSE = \sum_{i=1}^{K} \sum_{j=1}^{r_i} \left(x_j - v_i\right)^2 \tag{4}$$

In formula (4): K is the number of clusters, r_i is the number of sample elements in the ith cluster, x_j is the sample element in the ith cluster, v_i is the data mean of all sample elements in the ith cluster. The main process of K-means clustering is shown in Algorithm 1:

Algorithm 1: K-means clustering algorithm
Input:D $= \{x_1, x_2, \cdots, x_n\}$ training data, K is the number of categories
Output: clustering results
1 Arbitrarily select k objects as the initial cluster center
2 REPEAT
For j = 1 to n do
calculate $d(x_p, x_q)$ according to Eq. (3)
assign each x_j to the closest clusters;
For i = 1 to k do
Update each cluster average
calculate SSE according to Eq.(4)
UNTIL SSE no longer changes significantly

2.3. The overall process of capacity assessment clustering

As shown in Fig. 2, the clustering algorithm process of student ability assessment presented in this work is primarily separated into three steps:

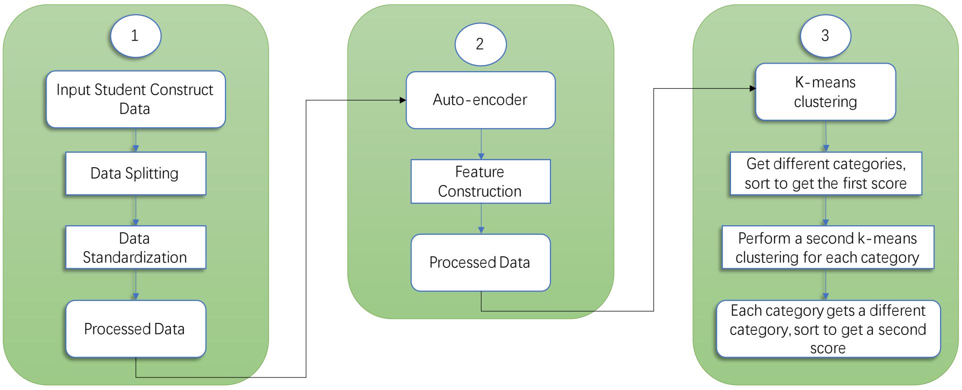


Fig. 2. The overall clustering process of student ability assessment scoring

1) Divide the data set into four parts and standardize the data to enable the preprocessing function.

2) Reconstructing the feature of the student learning behaviour data by auto-encoder. The reconstruction loss is utilized as the loss function.

3) First, the K-means algorithm is used to first cluster these data to obtain each category results for the first time. Second, calculate the average of each category and sort them to assign scores, as the single digit of the ability assessment score. Finally, use the categories from the first clustering as inputs for the second k-means clustering. Calculate the average score for the categories in the second k-means clustering and assign the score as the decile score. The single-digit and decile-digit scores represent the multifaceted assessment of the student's ability for learning in the end. The average score for each category is calculated using the following formula:

$$The\ \text{average score} = \frac{Add\ the\ grades\ of\ the\ students\ in\ the\ category}{Number\ of\ students\ in\ this\ category}$$

The following is the process of obtaining Li's knowledge skill ability evaluation score:

Set k to 10. Student Li is classified as Class A after the first clustering. Calculate the average scores of the students in the ten categories at this time, then sort the ten categories' average scores. Provide Class A students with 9 points if their average score is the highest. Continue the second clustering in Class A to obtain ten categories, and calculate the average score of them. If Li's category gets the highest average score, he will receive 9 points. Finally, the knowledge skill ability evaluation score of Li student is obtained by adding the first score and the second score * 0.1, which is 9+9 * 0.1 = 9.9 points. The result is scaled to 0.0–10.0 points if k is not set to 10. Repeat the above process to calculate four sets of ability scores.

The main flow of the method in this paper is shown in Algorithm 2. The learner's learning behaviour feature matrix is used for the input, and the learner's scores for their learning abilities are used for the output. Lines 1-4 are for data preprocessing and setting model parameters. Lines 5-6 show that the data is input into the auto-encoder algorithm for feature reconstruction to obtain output data. Lines 7-11 show how quadratic k-means clustering algorithms are used to determine the learner's ability scores from the feature reconstruction. Each student receives four sets of scores after this algorithm has been applied, including knowledge skills, learning attitude, course interest, and total score, and nearly no two students receive the exact same ability score.

Algorithm 2: AE+ Quadratic K-means

Input: learner learning behaviour features

Output: Learner’s scores for each learning ability

1 Loading the ratings dataset

2 dataset ← normalized learner learning behaviour features

3 Seting the model’s parameters

4 (epoch, batches, lr) ← Initialize(parameters)

5 $hidden_i$ ← using X and auto-encoder to calculate $f_\phi(x)$

6 $output_i$ ← using $hidden_i$ to calculate $g_\theta(z)$

7 For i = 1 in $output_i$ do

using $output_i$ and k-means clustering algorithm to calculate $result1_i$

8 Score1 ← Sort the average score of $result1$ and assign scores

9 For j = 1 in $result1$ do

using $result1_j$ and k-means clustering algorithm to calculate $result2_j$

10 Score2 ← Sort the average score of $result2$ and assign scores

11 Final score ← Score1 + Score2 * 0.1

3. Cluster analysis of student ability evaluation based on AE reconstruction

In this study, learning behaviour data from students is analyzed to produce various assessment of students’ abilities. We decompose the data features and cluster them into four dimensions of student ability assessment: knowledge and skills, learning attitude, course interest, total score.

3.1. Experimental dataset and data preprocessing

Dataset: This data set is provided by a teacher under the current mainstream online education platform. It records the learning situation of 152 students majoring in ideological, political and musical majors at a university in Hunan, China, from September 2022 to December 2022, who are studying modern educational technology application courses.

Data preprocessing: Due to the time limits of the learning assignments, some students may miss the deadline and not finish, producing learning data of 0. To ensure the truthfulness of the data collection, these data are not subjected to procedures like data deletion. Therefore, the following two tasks must be completed as part of the data preprocessing in this work.

3.2. Data splitting

We select 11 features that reflect students' learning situation. And we combine them to build new features to represent the assessment dimension of students' learning ability: the task point completion percentage, the course video progress, the chapter quiz progress and the video watching time to construct learning attitude scores, the course video scores, the chapter test scores, the homework scores and the test scores to construct knowledge and skill scores, the number of discussions, the number of chapters studied and the number of replies to the topic construct the course interest score, and all learning situation features construct the total ability scores. Therefore, in order to acquire the relevant student ability assessment scores, each of the four sections of the data set must be clustered.

To assess the performance of our fusion algorithm on real datasets, we do experiments and compare our fusion algorithm with a number of conventional clustering methods and the clustering approach following PCA reconstruction (Mo et al., 2022).

3.3. Data standardization

The proposed clustering in this study aims to determine the students' multidimensional learning capacity evaluation score. Data standardization could remove the impact of dimensional differences in student learning data on clustering results in order to accelerate convergence and shorten the training time for the auto-encoder due to significant differences in the dimensions of the dataset. In this paper, the z-score standardization approach is used to standardize the data of each learning behaviour feature (Yang et al., 2023). The standardization formula is as follows:

$$X_{new} = \frac{X_i - Mean}{Std} \tag{5}$$

X is the original data set, $Mean$ is the mean of the original data set, and Std is the standard deviation of the original data set.

3.4. Auto-encoder realizes feature reconstruction

We first utilize the auto-encoder to reconstruct features from the data set. The number of neurons used in each layer of the auto-encoder to create the different student ability assessment dimensions is shown in Table 1.

Table 1
The number of neurons in each layer of different data set

The number of layers of the auto-encoder	Input layer	Hidden layer	Output layer
Knowledge skills	4	3	4
Learning attitude	4	3	4
Course interest	3	2	3
Total score	11	5	11

The choice of learning rate is related to the convergence speed of our model. On the knowledge-skill dataset, we test using the auto-encoder model, with the learning rate set to 0.075, 0.01, 0.1, 0.25, 0.5 and 1 respectively. The comparison effect of each learning rate is shown in Fig. 3(a)(b). In Fig. 3(a), different learning rates make the model converge after 1500 epochs. It can be observed that the loss of model with a learning rate of 0.01 and 0.1 decreases very slow. Therefore, we remove the learning rate of 0.01 and 0.1 in Fig. 3(b). We can clearly see that the loss function of the model decreases fastest when the learning rate is 1, and the loss of the model is the smallest at that time. The loss of the model under each learning rate is shown in Table 2.

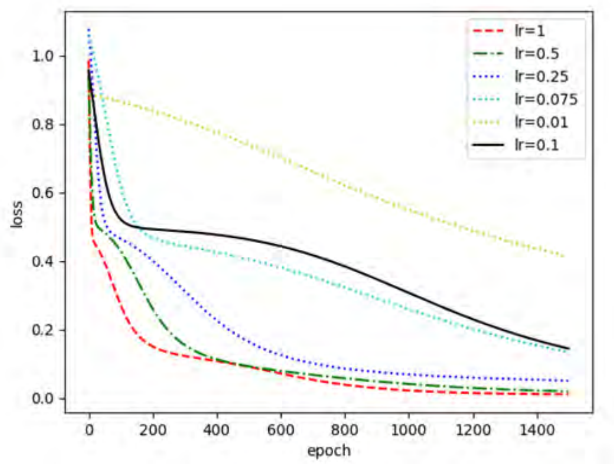


Fig. 3(a). Comparison chart of each learning rate of auto-encoder

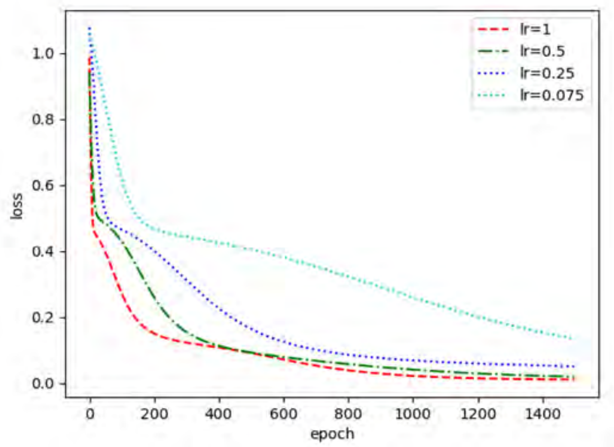


Fig. 3(b). Comparison chart of each learning rate of auto-encoder

Table 2
Comparative experiment of learning rate

learning rate	0.01	0.075	0.1	0.25	0.5	1
Loss	0.3115	0.1582	0.1271	0.0240	0.0492	0.0106

We set the learning rate to 1 and observe at the change in the loss function value with the number of epochs in the of four datasets to determine the number of epochs for the auto-encoder model for each dataset. The experimental finding that the model steadily converges after 1500 iterations is shown in Fig. 4. Our choice for the model parameter is 1500, for this reason.

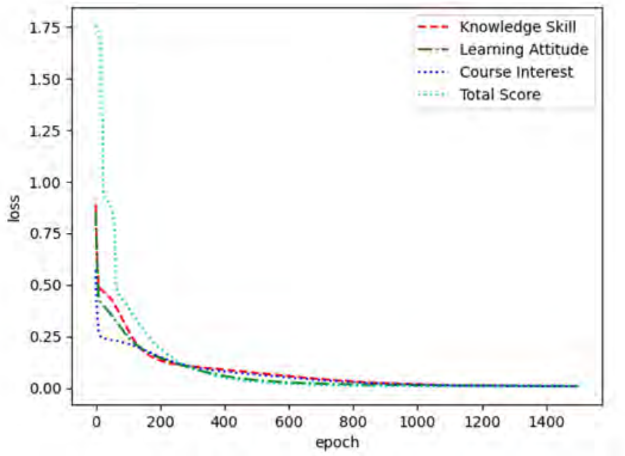


Fig. 4. Four sets of data training loss changes

3.5. Selection of the number of clusters

We use k-means clustering to obtain the evaluation scores of the four abilities after data preprocessing. The four clustering model results would be obtained respectively. In the K-Means algorithm, the number of clusters needs to be determined in advance. However, due to inexperience, it is frequently challenging to determine the value of k in practice. If the k value is too small, the data objects within the same cluster will be very different. The difference between clusters won't be significant if the k value is too large. The final clustering result will fall into a local optimum if the k value is not chosen properly (Yang & Zhao, 2019). In order to determine the optimal number of cluster categories based on the relationship between the number of clusters and the sum of squared errors, this study compares the SSE index under various numbers of clusters. In all data sets, using the K-means algorithm, the connection between SSE and K numbers is shown in Fig. 5. The learning attitude model's inflection point occurs when the number of clusters reaches 5, as can be observed, and this is when the SSE index decreases the most rapidly. It demonstrates a stronger clustering result since the data in the clustering is rather compact. Hence, it is appropriate to select 5 clusters for the learning attitude model. This is used to determine the number of clusters in other capability models.

3.6. Evaluation indicators

The clustering method used in this experiment is being assessed with three commonly used evaluation indicators. They are silhouette coefficient, CH score and DBI value.

Silhouette Coefficient: The silhouette coefficient is an evaluation method for clustering effects. It was first proposed by Rousseeuw (1987) in 1987. It comprehensively considers cluster agglomeration and separation. Successful clustering results usually have lower cohesiveness and higher cluster separation. To obtain the overall silhouette coefficient of the clustering result, we average the silhouette coefficients across all vectors. The higher the total silhouette coefficient value, the better the grouping result (Wu et al., 2021). The following is the formula for calculating the maximum average silhouette coefficient:

$$s = \frac{(b - a)}{\max(a, b)} \tag{6}$$

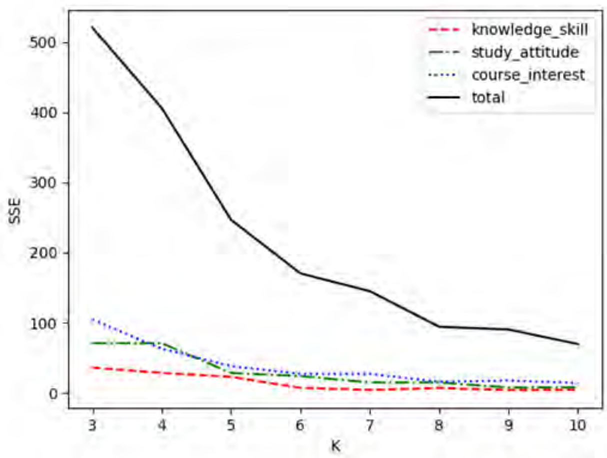


Fig. 5. Four sets of data training loss changes

Compute the two distances in formula (6): a is the average distance between the current sample point and other sample points of the same type, b is the average distance between the current sample point and other sample points of the closest class, the silhouette coefficient score of the current sample point is s.

Calinski Harabasz Score (CH score): It calculates a score by evaluating between-class variance and within-class variance. The larger the score, the better. The following is the formula: k represents the number of cluster categories, N represents the number of all data, SSB is the between-class variance and SSW is the within-class variance.

$$s = \frac{SSB}{k - 1} \Big/ \frac{SSW}{N - K} \tag{7}$$

Davies-Bouldin Index (DBI value): It is a clustering algorithm evaluating the metric, the lower the value the better. $\overline{S_\iota}$ is the average Euclidean distance from the ith class sample to its class center, $||w_i - w_j||_2$ is the Euclidean distance between the class centers of the i and j classes. The following is the formula:

$$DBI = \frac{1}{N} \sum_{i=1}^{N} max \frac{(\overline{S_\iota} + \overline{S_J})}{||w_i - w_j||_2} (j \neq i) \tag{8}$$

3.7. Comparative experiment

We construct knowledge skills, course interest, learning attitude, and total score to evaluate students' learning capacity by analyzing the data set. Before categorizing the data, we first utilize the auto-encoder to reconstruct features from the data. Take the average score for each category and assign the score after sorting. Then, for each category, run k-means clustering again and take the average score for sorting and assigning scores. Finally, we can obtain students' multidimensional learning ability assessment scores.

For comparative experiments, we choose a number of traditional clustering algorithms, the k-means algorithm after PCA reconstruction, and the k-means algorithm based on auto-encoder. This section shows and analyzes the results of different algorithms on four different data sets: knowledge skill score, course interest score, learning attitude score and total score.

Using the elbow technique, we select the K value in each data set at the point of maximum gradient change of SSE, where knowledge skills k = 4, learning attitude k = 5, course interest k = 4, and total score k = 8. We set the dimension of the PCA to be the same as the auto-encoder.

The silhouette coefficients compared by various algorithms are shown in Table 3. The CH scores are shown in Table 4, and the DBI values are shown in Table 5.

Table 3
Comparative experiment of learning rate

	Knowledge skills	Learning attitude	Course interest	Total score
Fuzzy C clustering	0.528	0.396	0.360	0.146
Hierarchical clustering	0.538	0.416	0.383	0.218
K_means	0.535	0.531	0.394	0.250
PCA+k-means	0.535	0.542	0.394	0.299
AE+k-means	0.652	0.640	0.581	0.448

Table 4
Comparison of the results of each clustering algorithm CH score

	Knowledge Skills	Learning Attitude	Course Interest	Total Score
Fuzzy C Clustering	172.12	158.88	153.69	63.54
Hierarchical clustering	199.96	155.79	149.40	65.438
K-means	230.15	184.12	166.73	71.09
PCA+k-means	230.16	202.04	166.74	131.62
AE+K-means	3628.15	1240.04	476.76	94.29

According to the experimental results, the silhouette coefficient of the k-means clustering algorithm based on auto-encoder outperforms both the conventional clustering method and the k-means algorithm based on PCA data reconstruction by an average of 20%, the average CH score is 2-15 times higher, and the average DBI value is more than 20% lower. And it is especially obvious on the knowledge skill dataset. It shows that the feature extraction and data reconstruction algorithm of auto-encoder is more effective than

PCA algorithm. The data are closer inside the class and further apart from one another following data reconstruction based on auto-encoder. Because AE reconstructs the data, the output layer of the model learns the nonlinear relationship between different features and increases the accuracy of clustering. In addition, it reduces the sensitivity of K-means clustering algorithm to outliers, improves the robustness of K-means clustering algorithm, makes the students ‘ability score more accurate.

Table 5
Comparison of the results of DBI values of each clustering algorithm

	Knowledge Skills	Learning Attitude	Course Interest	Total Score
Fuzzy C Clustering	0.733	0.854	0.976	1.810
Hierarchical clustering	0.660	0.845	0.914	1.383
K-means	0.686	0.781	0.876	1.203
PCA+K-means	0.687	0.758	0.876	1.047
AE+K-means	0.457	0.475	0.607	0.757

3.8. Radar chart of multidimensional capacity

We complete the evaluation and grading of students’ multidimensional learning abilities by dimensionality reduction and k-means clustering. Then we make radar charts of the multidimensional learning ability of the students under this course. The radar chart of a student’s capacity for multidimensional learning is shown in Fig. 6.

Fig. 6. Li’s learning ability radar chart

The student outperforms other students in the course in terms of homework completion, exam performance, and other areas, as shown by the multidimensional ability radar chart. He is positive but doesn’t show enough interest in the course. He needs to engage more in the course's relevant discussions. Teachers and students can receive these radar maps for use as a resource in their ensuing research.

4. Summary and future outlook

The purpose of this study is to create an algorithm for assessing students' course ability for online education platforms. The ability of students to learn in a particular course can be assessed through long-term data analysis of their performance. However, there are too many dimensions of data regarding students' learning behaviour. This study assesses students' abilities using a clustering algorithm based on auto-encoder in order to extract the intrinsic connection between students' performance and improve clustering efficacy. The experiment shows that the auto-encoder-based K-means clustering algorithm outperforms other traditional clustering algorithms in terms of performance.

Future analysis of this study may take into account the following three points:

1. There are only a few types of data available, and only the learning dimensions collected through the online learning platform are used to analyze the students' abilities. We can therefore consider including data from scientific research competitions, awards, and student extracurricular activities for a fuller analysis.

2. Only 152 pieces of data were collected for this research, which confirms the efficiency of the proposed approach in a small sample. The data amount can be increased to tens of thousands in the future, and three or four rounds of clustering can be carried out to raise the ability evaluation score to a range of 0.0 to 100.0.

3. Despite being superior to traditional clustering algorithms, the k-means algorithm based on auto-encoder dimension reduction is still not effective enough at clustering the total score. The more recent autoencoder-based deep clustering algorithm may be used in further research to assess students' comprehensive abilities.

Author Statement

The authors declare that there is no conflict of interest.

Acknowledgements

This research was funded by the National Natural Science Foundation of China (61877017), the Natural Science Foundation of Hunan Province (2022JJ50002, 2021JJ50049).

ORCID

Junfeng Man https://orcid.org/0000-0002-9867-178X

Rongke Zeng https://orcid.org/0009-0006-6758-3767

Xiangyang He https://orcid.org/0009-0000-7220-6102

Hua Jiang https://orcid.org/0009-0008-1072-1908

References

Chen, C. (2018). *Design and implementation of college students' professional ability evaluation system*. University of Jinan, China.

Chen, Z. (2015). Evaluation technology of students' academic level based on radar chart. *Chinese Journal of ICT in Education*, 22 ,76–78.

Chen, Z., Jiang, H., Hao, D., & Man, J. (2022). X-ray fluorescence spectral inspection of plastic express packaging bags based on k-means and within-cluster sum of squared errors. *Laser & Optoelectronics Progress, 59*(11), 489–495. https://doi.org/10.3788/LOP202259.1130001

Dong, Y., Wang, Y., Dong, Y., & Deng, Y. (2023). Survey of online learning resource recommendation. *Journal of Computer Applications, 43*(6), 1655–1663. https://doi.org/10.11772/j.issn.1001-9081.2022091335

Du, J., Guo, L., & Wang, X. (2010). Comprehensive evaluation of primary and secondary school teachers' educational technology ability based on radar chart. *Information Technology Education in Primary and Secondary Schools, 2010*(2), 71–73.

Guleria, H. V., Luqmani, A. M., Kothari, H. D., Phukan, P., Patil, S., Pareek, P., Kotecha, K., Abraham, A., & Gabralla, L. A. (2023). Enhancing the breast histopathology image analysis for cancer detection using variational autoencoder. *International Journal of Environmental Research and Public Health*, *20*(5): 4244. https://doi.org/10.3390/ijerph20054244

Hu, H., Liu, J., Zhang, X., & Fang, M. (2023). An effective and adaptable K-means algorithm for big data cluster analysis. *Pattern Recognition*, *139*: 109404. https://doi.org/10.1016/j.patcog.2023.109404

Ji, M., Cao, C., Zhang, J., & Zhang, H. (2017). Research on student ability evaluation based on student performance in the context of big data. *Statistics and Management, 9*(5), 98–100. https://doi.org/10.16722/j.issn.1674-537X.2017.05.034

Li, J., Hao, Y., & Yang, Y. (2022). A learner class profile method based on improved DBSCAN feature clustering. *Computer & Digital Engineering, 22*(4), 703–708. https://doi.org/10.3969/j.issn.1672-9722.2022.04.004

Li, P., Pei, Y., & Li, J. (2023). A comprehensive survey on design and application of autoencoder in deep learning. *Applied Soft Computing*, *138*: 110176. https://doi.org/10.1016/j.asoc.2023.110176

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (pp. 281–297).

Mo, S., Gulanbaier, T., & Mairidan, W. (2022). Research on one-class classification algorithm for mixture and high dimensional intrusion detection data. *Computer Applications and Software, 23*(4), 300–310. https://doi.org/10.3969/j.issn.1000-386x.2022.04.047

Peng, L., Wu, Q., Li, S., Zhou, X., & Xiao, C. (2020). Research on comprehensive evaluation and classification of students based on k-means clustering and SVM algorithms. *Digital Technology & Application, 38*(10), 88–91. https://doi.org/10.19695/j.cnki.cn12-1369.2020.10.29

Qu, C., Zhu, R., Qu, N., Cao, L., Lv, H., & Hu, K. (2019). Identification and correction methods of grid operation abnormal data based on spark framework. *Science Technology and Engineering, 19*(25), 211–219. https://doi.org/10.3969/j.issn.1671-1815.2019.25.031

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65.

https://doi.org/10.1016/0377-0427(87)90125-7

Song, X., & He, L. (2022). Online course recommendation model based on enhanced auto-encoder. *Computer Systems & Applications, 31*(3), 288–293. https://doi.org/10.15888/j.cnki.csa.008356

Tao W., Qian Y., Zhang Y., Ma H., Leng H., & Ma M. (2022). Survey of deep clustering algorithm based on autoencoder. *Computer Engineering and Applications*, *18*, 16–25.

Wang, Z., Li, J., & Song, Y. (2020). Improved k-means algorithm based on distance and weight. *Computer Engineering and Applications, 56*(23), 87–94. https://doi.org/10.3778/j.issn.1002-8331.2009-0103

Wu, J., Shi, L., Yang, L., Niu, X., Li, Y., Cui, X., Tsai, S.-B., & Zhang, Y. (2021). User value identification based on improved RFM model and *K* -means++ algorithm for complex data analysis. *Wireless Communications and Mobile Computing*~~, *2021*(1): 9982484~~. https://doi.org/10.1155/2021/9982484

Xie, Y., Dai, X., Huang, X., Li, J., & Liu, S. (2021). Research on quantitative evaluation of students' comprehensive quality based on academic achievements. *Higher Education of Sciences, 21*(1), 115–123. https://doi.org/10.3969/j.issn.1000-4076.2021.01.020

Xiong, H., Song, Y., Zhang, P., & Liu, L. (2019). Personalized question recommendation based on autoencoder and two-step collaborative filtering. *Computer Science, 2019*(2), 172–177.

Yang, H., Zhao, X., & Wang, L. (2023). Review of data normalization methods. *Computer Engineering and Applications, 59*(3), 13–22. https://doi.org/10.3778/j.issn.1002-8331.2207-0179

Yang, J., Hu, S., Wang, Q., & Fong, S. (2021). Discriminable multi-label attribute selection for pre-course student performance prediction. *Entropy*, *23*(10): 1252. https://doi.org/10.3390/e23101252

Yang, J., & Zhao, C. (2019). Survey on k-means clustering algorithm. *Computer Engineering and Applications*, *55*(23), 7–14. https://doi.org/10.3778/j.issn.1002-8331.1908-0347

Yuan, G. D., He, M., Ma, Z. Y., Z., Zhang, W. S., & L, W. (2023). Multiagent following multileader algorithm based on k-means clustering. *Journal of System Simulation, 35*(3), 616–622. https://doi.org/10.16182/j.issn1004731x.joss.21-1158

Zhao, Z., Chang, X., Fan, Y., Xu, S., & Fan, M. (2021). Clustering analysis of power load curve based on auto-encoder. *Science Technology and Engineering, 21*(32), 13737–13743. https://doi.org/10.3969/j.issn.1671-1815.2021.32.017

Zheng, Q., Dong, B., Qian, B., Tian, F., Wei, B., Zhang, W., & Liu, J. (2019). The state of the art and future tendency of smart education. *Journal of Computer Research and Development, 51*(1), 209–224. https://doi.org/10.7544/issn1000-1239.2019.20180758