Discover Artificial Intelligence

Research

Personalized learning path planning for higher education based on deep generative models and quantum machine learning: a multimodal learning analysis method integrating transformer, adversarial training and quantum state classification

Changzhi Sun[1] · Shijie Huang[1] · Banghui Sun[1] · Shiwei Chu[2]

Received: 26 November 2024 / Accepted: 18 March 2025
Published online: 26 March 2025

Abstract
In the field of personalized learning path planning for higher education, traditional methods lack in-depth analysis of students' dynamically changing learning status and interests, resulting in insufficient personalization, and focus on the analysis of a single data type, making it difficult to integrate data of different modalities. This paper proposes a multimodal learning analysis method for personalized learning path planning in higher education, addressing the limitations of traditional methods that do not account for dynamic changes in students' learning behaviors and interests. The approach integrates Transformer models, adversarial training, and quantum state classification to analyze multimodal data, including text, audio, and video, to capture learning patterns. The Transformer model uses a self-attention mechanism to generate personalized learning paths based on integrated data. Adversarial training is applied to simulate abnormal data and enhance the model's robustness to various learning scenarios. Quantum state classification improves data processing efficiency, addressing challenges in handling high-dimensional multimodal data. Experimental results show that the Transformer model achieves stable accuracy of 0.95 and recall of 0.91 for personalized learning path generation. The adversarial training method reduces the loss value to around 0.05, while the introduction of quantum state classification reduces processing time to 56 s in the 18th round, a 31-s improvement. These results confirm the effectiveness of the proposed method in enhancing the accuracy, robustness, and computational efficiency of personalized learning path generation in higher education.

Keywords Learning path planning · Deep generative model · Transformer model · Adversarial training · Quantum state classification

1 Introduction

The development of artificial intelligence and machine learning technology has made personalized learning path planning more and more widely used in higher education. Analyzing multimodal data such as text, audio and video can better understand students' learning behavior [1, 2]. Using deep generative models, it can more accurately extract valuable features from large-scale data [3, 4], providing students with a more personalized learning path [5, 6].

✉ Shiwei Chu, P23111026@stu.ahu.edu.cn; Changzhi Sun, bzxyscz@bzuu.edu.cn; Shijie Huang, 2022020003@bzuu.edu.cn; Banghui Sun, 2022020007@bzuu.edu.cn | [1]Electronics and Information Engineering, Bozhou University, Bozhou 236800, Anhui, China. [2]School of Electronic Information Engineering, Anhui University, Hefei 230039, China.

The traditional personalized learning path planning system relies on static student data [7, 8], ignoring the dynamic changes in students' learning process. This problem also causes these methods to be unable to accurately capture students' interest changes, changes in learning status, and their potential needs during the learning process [9, 10], which to some extent limits the accuracy and adaptability of personalized recommendations [11]. Traditional methods generally rely on single-modal data analysis, most of which only use text data and lack the ability to integrate and process multimodal data such as text, audio, and video [12, 13]. This problem prevents the system from fully understanding students' learning behaviors and diverse needs, affecting the effectiveness and personalization level of path planning [14, 15]. The above problems also lead to other problems. Traditional models face problems such as low computational efficiency and slow processing speed when processing high-dimensional and multi-type data [16, 17]. Especially when processing large amounts of student data and multimodal data, traditional methods cannot generate efficient learning path recommendations in real time [18], lack adaptability to abnormal data, and are prone to bias and misjudgment [19, 20]. Traditional methods have solved the needs of personalized learning path planning to a certain extent, but they still have some limitations. New technologies and methods need to be introduced to improve the intelligence level of the recommendation system and enhance its dynamic adaptability, data processing capabilities, and multimodal data integration capabilities. This is also the main background reason for this paper to propose a multimodal learning analysis method based on deep generative models and quantum machine learning.

The research contributions of this article include:

(1) This paper explores a personalized learning path planning approach that combines deep generative models with quantum machine learning techniques, hoping to address the limitations of existing methods in the field of higher education. By integrating the Transformer model, adversarial training, and quantum state classification technology, it breaks through the bottleneck of traditional methods and improves the accuracy and real-time performance of personalized learning path planning.
(2) In this process, the paper used the self-attention mechanism in the Transformer model to conduct in-depth analysis of students' multimodal data, accurately obtain students' learning status, interest changes, and behavior patterns, and generate personalized learning paths. In response to the challenge of abnormal data, adversarial training is introduced to generate adversarial samples to enhance the adaptability of the system. The quantum state classification method in quantum machine learning is combined to address the computational challenges of high-dimensional features in multimodal data processing, and quantum computing is used to accelerate data processing. Traditional methods have poor adaptability, delayed updates, and difficulty in real-time adjustment of recommended paths in response to students' dynamic learning status and interest changes.
(3) The method proposed in this paper can capture subtle changes in student behavior through the Transformer model self-attention mechanism and quantum state classification technology, generate efficient and personalized learning paths in real time, and significantly improve dynamic response capabilities and recommendation accuracy. These innovative steps are used to improve the accuracy of personalized learning path planning, generate the best learning path in real time in the dynamic environment of students' learning needs and behavior changes, realize efficient and intelligent personalized education recommendation, and promote the development of higher education to the intelligent and personalized direction.

The structure of this paper is organized as follows:

Section 1 serves as the Introduction, which delineates the background, significance, and limitations of current methodologies in personalized learning path planning. It also introduces the proposed multimodal learning analysis method grounded in deep generative models and quantum machine learning.

Section 2 presents the Related Work, offering a comprehensive review of existing recommendation algorithms, deep learning frameworks, and the application of Transformer models, adversarial training techniques, and quantum state classification within the context of personalized learning.

Section 3 elaborates on the Proposed Methodology. This section encompasses an overview of personalized learning path planning processes, integration and feature extraction from multimodal data sources, generation of personalized paths utilizing Transformer models, adaptive optimization through adversarial training strategies, and efficient data processing via quantum state classification.

Section 4 assesses the effectiveness of the proposed methodology by examining various aspects such as the quality of generated personalized learning paths, implications arising from adversarial training practices, and efficiency metrics associated with data processing using quantum state classification.

Finally, Sect. 5 concludes this paper by summarizing its contributions while emphasizing avenues for future research aimed at addressing identified limitations and further enhancing personalized learning systems.

2 Related work

In the field of personalized learning path planning, recommendation algorithms based on collaborative filtering are widely used in the field of education, which can infer personalized paths based on students' historical learning behaviors and similar user data [21, 22]. This type of algorithm has certain advantages in improving recommendation efficiency, but it relies on students' static historical data, ignores the dynamic changes in learning interests and behaviors, and makes the recommendations inaccurate. For instance, collaborative filtering may struggle to adapt to sudden changes in a student's learning preferences, resulting in suboptimal learning paths [23]. In order to further improve the applicability of path recommendations, some researchers have introduced deep learning models such as convolutional neural networks and long short-term memory networks into this field [24, 25]. This type of deep learning model has achieved good results in static learning state analysis due to its powerful feature extraction capabilities [26], but it is still difficult to cope with students' dynamically changing learning needs. For example, CNNs can effectively process text data to identify key learning concepts, while LSTMs can capture temporal dependencies in students' learning behaviors. However, these models still face challenges in handling dynamic and multimodal data, limiting their ability to provide comprehensive and adaptive learning paths [27]. The integration of multimodal data, including text, audio, and video, has been identified as a critical factor in improving the accuracy and personalization of learning path recommendations [14]. Several studies have explored the use of multimodal data to capture students' learning behaviors more comprehensively. For instance, some researchers have used text data to analyze students' written assignments and audio data to assess classroom interactions [28]. However, most existing methods focus on single-modal data analysis and lack the ability to integrate and process multimodal data effectively [11]. This limitation prevents the system from fully understanding students' learning behaviors and diverse needs, affecting the effectiveness and personalization level of path planning.

Some studies have shown that the attention mechanism of the Transformer model can realize the integrated analysis of multimodal data in personalized recommendation, and better improve the ability to respond to students' dynamic learning status [29, 30]. Transformer's self-attention mechanism allows the model to identify important feature information from large-scale multimodal data, so it shows good applicability when processing multimodal data such as text, audio and video [31, 32]. Adversarial training, as an important means in deep learning, uses the generation of adversarial samples to enhance the model's adaptability to abnormal data and different learning scenarios, and improve the model's robustness in dealing with complex data environments [33, 34]. Quantum state classification is gradually being applied to multimodal data processing due to its ability to efficiently process high-dimensional data [35, 36]. The rapid classification of quantum states can effectively solve the problem of speed limitations of traditional deep learning models in multimodal high-dimensional data processing. When these methods are applied alone, they cannot fully improve the accuracy and real-time performance of personalized learning paths.

However, despite these advancements, several literature gaps remain:

1. Integration of Multimodal Data: While there is a growing body of research on using multimodal data for personalized learning, few studies have effectively integrated these data types to capture the dynamic learning behaviors and interests of students. Most existing methods focus on single-modal data analysis, limiting the comprehensive understanding of students' learning needs and behaviors.
2. Dynamic Adaptation: Traditional methods often rely on static student data and fail to account for the dynamic changes in students' learning processes. This limitation results in inaccurate and less adaptive personalized learning paths, which do not fully address the evolving interests and learning states of students.
3. Robustness and Generalization: Existing models face challenges in handling high-dimensional and multi-type data, leading to low computational efficiency and slow processing speeds. Additionally, these models lack robustness in dealing with abnormal data and different learning scenarios, resulting in biased or incorrect recommendations.
4. Quantum Machine Learning: Although quantum machine learning has shown potential in handling high-dimensional data, its application in personalized learning path planning is still in its infancy. The integration of quantum state classification with deep generative models and adversarial training has not been thoroughly explored, leaving significant room for improvement in computational efficiency and real-time performance.

5. Interpretability and Transparency: Many advanced models, such as the Transformer, are often considered "black boxes" due to their complex structures. This lack of interpretability makes it difficult to understand the decision-making processes of these models, which is crucial for their application in educational settings where transparency and trust are essential.
6. Fairness and Bias: With the increasing use of data-driven personalized recommendation systems, concerns about fairness, transparency, and bias have emerged. Existing models may exacerbate achievement gaps between students, particularly when recommendations rely heavily on historical learning data, potentially leading to biased or unfair outcomes.

Addressing these gaps requires further research and development to enhance the integration of multimodal data, improve dynamic adaptation, increase robustness and generalization, explore the potential of quantum machine learning, and ensure interpretability, fairness, and transparency in personalized learning path planning systems.

3 Method

Figure 1 is the overall process of personalized learning path planning, which collects students' multimodal data in order to fully understand their learning behaviors and interest characteristics, and provides input for subsequent analysis based on these data. Transformer-based personalized path generation, inputs multimodal features into the Transformer model, uses its self-attention mechanism to analyze students' dynamic learning status, and generates personalized learning paths. The Transformer model can effectively capture important features in multimodal data and achieve more accurate and personalized recommendations. Adversarial training improves the model's adaptability to abnormal learning behaviors and data noise. The introduction of adversarial training to generate adversarial samples simulates possible abnormal scenarios, improves the reliability and adaptability of personalized path planning, and quantum state classification improves computational efficiency. After generating the personalized path, quantum state classification is used to efficiently classify the path data. Quantum state classification can quickly process high-dimensional data and effectively improve the processing efficiency of the model in multimodal data.

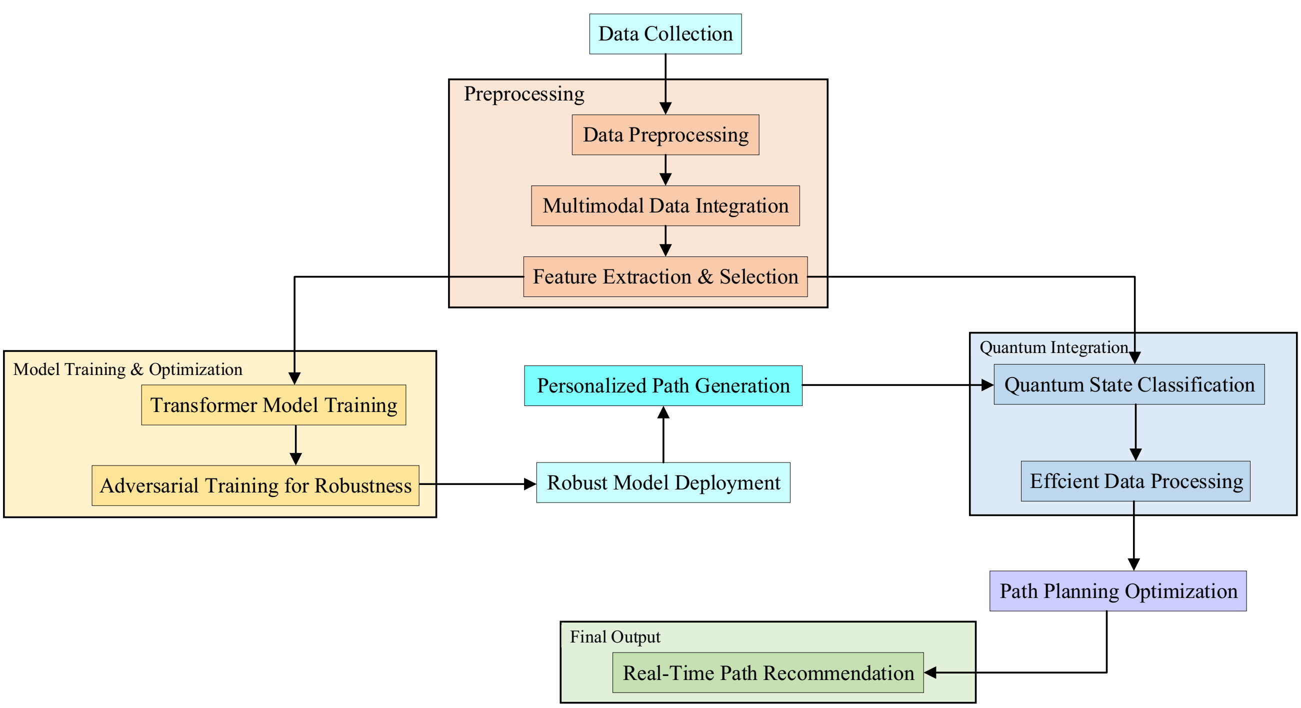

Fig. 1 Personalized learning path planning process

3.1 Multimodal data integration and feature extraction

3.1.1 Preprocessing and feature encoding of multimodal data

This study collects data from the Learning Through' learning management system, including student text, audio, video data, which can be used to analyze subsequent students' learning behavior and generate personalized learning paths. Text data contains 50,000 student learning records, including after-class notes and answer feedback, accounting for 50% of the overall data; the audio data covers 10,000 recordings, mainly for classroom interaction and Q & A records, accounting for 30% of the overall data; the video data includes 2,000 classroom teaching segments and student self-study process recording, accounting for 20%". In the third paragraph of 3.1.2, in order to ensure the efficient integration of different modal features, a multi-modal alignment mechanism based on the unified embedded space is studied and designed. These heterogeneous features are mapped to a shared embedded space through linear transformation to ensure that the features of different modes have the same dimension and semantic distribution. In the embedded space, the self-attention mechanism is further used to excavate the interaction relationship between the modes. The self-attention mechanism according to the importance of modal features dynamically adjusts the weight, emphasizes the capture of specific modal information, and weakens the interference of irrelevant features. This mechanism can realize the depth fusion of modal features to generate high-quality joint representations to meet the needs of multimodal data joint analysis and personalized learning path recommendation. Multi-modal data integration and feature extraction of students' text, audio and video data are conducted to obtain the multi-dimensional behavior characteristics and interest characteristics of students in the learning process. The three types of modal data are preprocessed, the text data is transformed into embedded vector representation, the audio data is used for feature extraction, the video data obtains visual features through image frame processing [37]. In the text data is processed, the original text is transformed into dense vector by word embedding method, and the BERT (Bidirectional Encoder Representations from Transformers) model is used to facilitate the subsequent modal integration. Audio data uses short-time Fourier transform (STFT) [38, 39] to obtain the frequency and time domain features of the audio signal, and then MFCC (Mel Frequency Cepstral Coefficients) is used to extract its low-order features to obtain the voice information and rhythm features of the audio data. The video data processing steps include segmenting and extracting video frames, and using the convolutional neural network ResNet (Residual Network) [40, 41] to extract visual feature vectors for each frame, focusing on extracting data such as students' facial expressions and gestures during the learning process to supplement emotion and attention-related information.

After obtaining the preliminary features of different modal data, the features of each modality are further encoded based on the autoencoder; the conditional autoencoder is used to map the multimodal data to a unified feature space to eliminate the distribution differences between the modalities. The feature vectors of text, audio, and video are $\mathbf{x}_t$, $\mathbf{x}_a$, and $\mathbf{x}_v$ respectively. The encoder E of the conditional autoencoder is used to encode the features of each modality into a unified potential representation $\mathbf{z}$, that is:

$$\mathbf{z} = E\left(\mathbf{x}_t|c_t\right) = E\left(\mathbf{x}_a|c_a\right) = E\left(\mathbf{x}_v|c_v\right) \tag{1}$$

c_t, c_a, and c_v are modal feature conditional labels, which are used to ensure that modality-specific semantic information is retained during the encoding process. After normalization in a unified feature space, the imbalance of feature weights caused by scale differences between modalities is avoided, laying the foundation for subsequent personalized path recommendations.

3.1.2 Feature fusion

The representation effect of multimodal features needs to be further improved. The study uses the variational autoencoder (VAE) in the deep generative model [42, 43] to achieve feature fusion and feature reconstruction. For the normalized multimodal data, it is input into the VAE encoder, and the mean μ and standard deviation σ of the latent variables are encoded to generate the latent feature distribution. The formula is as follows:

$$\mathbf{z} \sim \mathcal{N}\left(\mu, \sigma^2\right) \tag{2}$$

By using this probabilistic feature representation, VAE can effectively suppress noise while retaining feature information, improving the robustness and fusion of features in multimodal space.

During the fusion process, VAE not only maintains the feature information of each modality, but also introduces the correlation between modalities, and uses distribution learning in the shared latent space to achieve information complementarity between modalities. Information from different modalities is processed uniformly through the distribution of a shared latent space, avoiding the one-sidedness of single modality information. Text features can be combined with visual information from videos, and emotional information from audio can be supplemented with learning behavior data, providing a more comprehensive representation of students' learning status and interest changes. To ensure efficient fusion of different modal features, we designed a multimodal feature alignment mechanism based on a unified embedding space. These heterogeneous features are mapped to a shared embedded space through linear transformation to ensure that the features of different modes have the same dimension and semantic distribution. In the embedded space, the self-attention mechanism is further used to excavate the interactive relationship between the modes. The self-attention mechanism according to the importance of modal features dynamically adjusts the weight, emphasizes the capture of specific modal information, and weakens the interference of irrelevant features. This mechanism can realize the depth fusion of modal features, so as to generate high-quality joint representations to meet the needs of multimodal data joint analysis and personalized learning path recommendation. In the feature reconstruction stage, the latent variables are input into the VAE decoder to regenerate the features of each modality. The reconstructed features are represented as $\hat{\mathbf{x}}$, and the high-fidelity reconstruction of the features is achieved by maximizing the data likelihood $p\left(\hat{\mathbf{x}}|\mathbf{z}\right)$.

The loss function of VAE consists of two parts: reconstruction error and KL divergence (Kullback–Leibler Divergence):

$$\mathcal{L}_{VAE} = \mathrm{E}_{q(\mathbf{z}|\mathbf{x})}\left[\log p(\mathbf{x}|\mathbf{z})\right] - D_{KL}[q(\mathbf{z}|\mathbf{x})|\,|p(\mathbf{z})] \tag{3}$$

$\mathrm{E}_{q(\mathbf{z}|\mathbf{x})}$ is the reconstruction error and D_{KL} is the KL divergence, which are used to control the smoothness of the distribution of latent variables.

Figure 2 shows the whole process of multimodal data integration. It includes preprocessing, feature extraction, embedding generation, feature standardization and multimodal fusion of text, audio and video data. After word segmentation and cleaning, the text data is passed through the BERT encoder to extract semantic features; after the audio data is processed by noise reduction and frame segmentation, the frequency and time domain features are extracted using STFT and MFCC; the video data is processed by ResNet to extract visual feature vectors.

3.2 Personalized path generation based on transformer model

Based on the integration of multimodal data, the Transformer model is used to analyze students' learning behavior data and generate personalized learning paths through the self-attention mechanism. The Transformer model can capture the dynamic changes of students' learning interests and states, thereby generating real-time adaptive path recommendations and improving personalization.

3.2.1 Learning behavior analysis based on transformer model

After integrating multimodal data, this paper uses the Transformer model to conduct in-depth analysis of students' learning behavior data and generate personalized learning paths. The integrated text, audio, and video features are transformed into embedding vectors of the same dimension through linear mapping, and each modal feature vector is represented as an item in the time series. These vectors are input into the encoder part of the Transformer model. The model uses the self-attention mechanism to model the students' historical learning behavior data, and obtains the students' interest changes, learning status, and behavior patterns during the learning process. In each self-attention layer, for the input feature sequence $\mathbf{X} = \left[\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n\right]$, the model is used to calculate the attention weight between each pair of features:

$$\mathbf{A} = \mathrm{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{g_k}}\right) \tag{4}$$

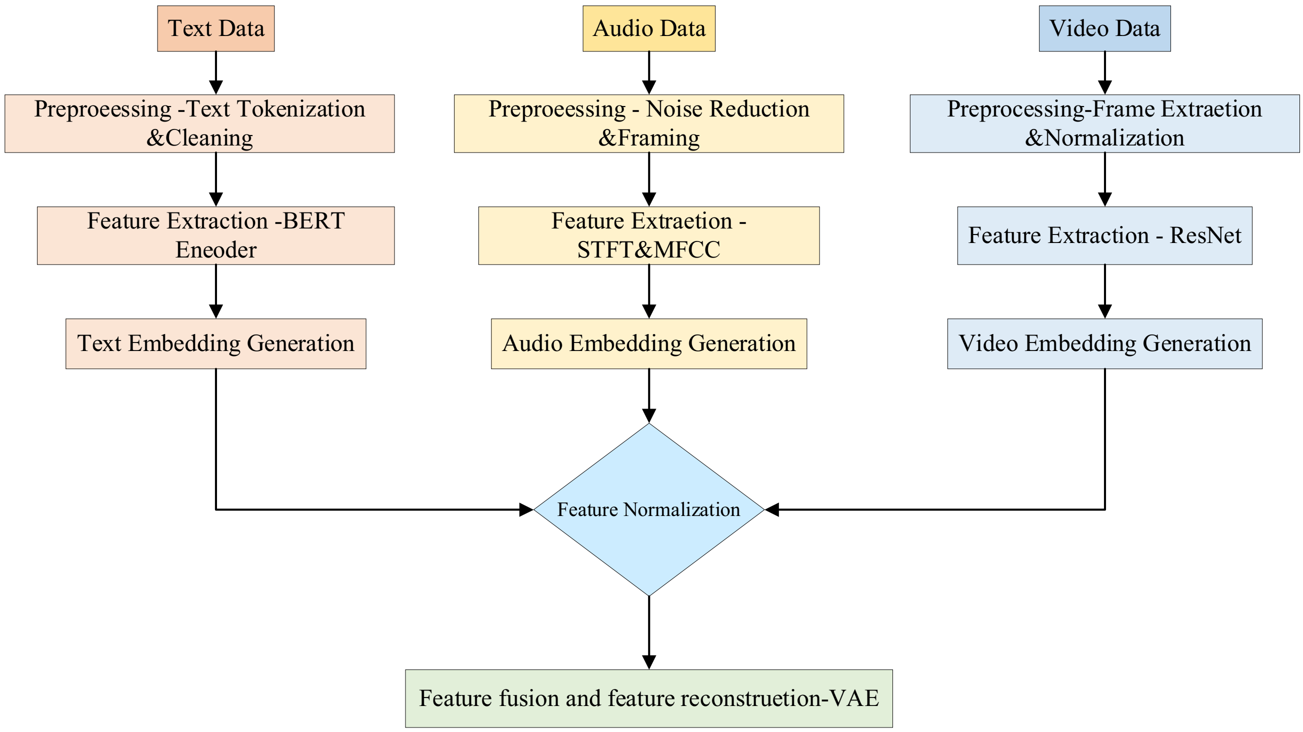

Fig. 2 Multimodal data integration flow chart

$\mathbf{Q}$ and $\mathbf{K}$ are the query vector and key vector generated by the input feature vector, $\mathbf{A}$ is the attention matrix, $\sqrt{g_k}$ is the normalization factor, and g_k is the feature dimension. This step calculates the importance of each feature vector in a given context and updates the output feature representation based on the weight.

The configuration of the Transformer model includes 6-layer encoder and 6-layer decoder, each layer adopts 8-head self-attention mechanism, the hidden layer dimension is set to 512, and the hidden layer dimension of the feedforward network is 2048, using the ReLU activation function. This structure was chosen based on the advantages of Transformer in handling long sequence dependencies and multiple coding decoding layers to effectively capture dynamic changes in students' learning behavior. Super parameter selection reason mainly considering the computing resources and model balance, deep network layers and larger hidden layer dimension helps to improve the expression ability of the model, but may bring high computing overhead, 8 head since the attention mechanism can capture more characteristic interaction information, suitable for processing the complexity of multimodal data.

The key factors affecting the model performance include learning rate, batch size and training rounds. Too high learning rate may lead to gradient explosion, while too low will lead to slow convergence; batch size affects the stability of gradient estimation; training rounds determines whether the model can effectively learn all modes. These hyperparameters were adjusted by cross-validation to ensure that the model achieves an optimal balance between accuracy and efficiency.

In this process, the Transformer model can automatically identify and focus on the most representative information in the student's learning process, especially in the dynamic changes in learning interests, fluctuations in learning status, and changes in learning behavior patterns. These analyses help generate accurate personalized learning paths for students and respond to changes in student needs in real time. By using the self-attention mechanism, the model can effectively integrate the temporal characteristics of multimodal data. It not only relies on static learning history, but also can obtain the short-term fluctuations and long-term evolution trends of students' interests.

In education, traditional Transformer models are usually regarded as a" black box " due to their self-attention mechanism and multi-level structure, which makes it difficult to intuitively understand it in the decision-making process. To address this problem, this article explores ways to improve the interpretability of the model, drawing on the mathematically interpretable white-box Transformer model CRATE (Contextualized Recurrent Attention Transformer Encoder) proposed by Ma Yi's team. CRATE introduces a transparent decision-making mechanism based on the attention mechanism. While ensuring the performance of the Transformer, it can provide a clear mathematical basis to explain each decision-making process. Using the visualized attention weights, researchers can trace how the model generates personalized paths based on students' learning history, interests, and status. CRATE also introduces the module of interpretability to clarify the decision basis of each layer, to help decision makers understand the reasoning process of the model, and to enhance the transparency of the model. This paper introduces a CRATE interpretability framework in the application of Transformer model. When selecting the recommended content, the model can clearly show which behavior or knowledge points influence the current recommended path, CRATE provides a traceable "interpretation module" in each decision node to help users understand how the model evaluates and processes the input data, and ensure the transparency and operability of the model. This interpretability framework provides higher trust and controllability for model application in the educational field.

3.2.2 Optimization and adaptability of generating personalized learning paths

In the process of generating personalized learning paths, the Transformer model can make path recommendations based on students' current learning status and use the self-attention mechanism to predict future learning needs, thereby improving the adaptability and personalization of the path to a certain extent. Based on the output feature representation of the model, $\mathbf{H} = \left[\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_n\right]$ is put into the multi-head self-attention mechanism for further processing to obtain useful information in different subspaces:

$$\mathbf{H}_{\text{out}} = \text{Concat}\left(\text{head}_1, \text{head}_2, ..., \text{head}_i\right)\mathbf{W}_O \tag{5}$$

In the formula, head_i is the attention output of the i th head, $\mathbf{W}_O$ is the output weight matrix, and $\mathbf{H}_{\text{out}}$ is the multi-head attention output. At this stage, the Transformer model can analyze the student's learning status from multiple angles, allowing the model to comprehensively consider the student's long-term interests, short-term needs, and learning progress, and generate a learning path that best meets the student's personalized needs.

This paper also introduces position encoding and time perception mechanisms to improve the model's adaptability to dynamic changes in learning behavior and better adapt to the time series characteristics of students' learning process. Position encoding adds information related to its position in the time series to each input feature vector, enabling the model to understand the relationship between the time order in the data:

$$\mathbf{PE}(t, 2i) = \sin\left(\frac{t}{10000^{2i/d}}\right) \tag{6}$$

$$\mathbf{PE}(t, 2i+1) = \cos\left(\frac{t}{10000^{2i/d}}\right) \tag{7}$$

t is the current time step, i is the feature dimension index, and d is the total dimension of the model. The introduction of position encoding allows the Transformer model to more accurately obtain the law of student behavior changes over time when processing students' dynamic learning status, and provide time-sensitive dynamic analysis for personalized path recommendation. This mechanism takes into account long-term effects and sustainability to a certain extent, especially the generation of personalized learning paths and real-time adaptability to students' dynamic learning status. In the long run, this continuous adaptability may have a positive impact on students' long-term learning outcomes, especially in cultivating students' autonomous learning and lifelong learning abilities.

Using the above method, the Transformer model can generate paths based on students' current learning data, respond to changes that may occur in the learning process in real time, and provide flexible and personalized learning path recommendations. The key advantage of this process is that the model can comprehensively consider students' historical data, current status and future trends, and provide real-time adaptive learning paths in a dynamically changing learning environment, greatly improving the accuracy and practicality of personalized learning paths.

Table 1 shows sample data of some students' learning behaviors, including text, audio, and video features related to their behaviors. Each student's behavior is tracked at multiple timestamps, and their corresponding feature values are provided. Taking the behavior of student 001 as an example, alternating between learning and resting, it has different feature values in different modes. This data is critical to understanding how students' behaviors and learning preferences evolve over time, which can be used to leverage the Transformer model to recommend personalized learning paths based on their real-time activities and preferences.

Table 2 shows the key steps of self-attention calculation when the Transformer model processes student behavior data. Different feature pairs of each student are calculated through the query vector, key vector and value vector, and

Table 1 Example table of some student behavior data and characteristics

Student ID	Timestamp	Text features	Audio features	Video features	Behavior type (study or rest)
001	2024/11/1 8:00	[0.1, 0.5, 0.3]	[0.4, 0.3, 0.2]	[0.2, 0.6, 0.1]	Study
001	2024/11/1 8:30	[0.2, 0.4, 0.5]	[0.5, 0.4, 0.3]	[0.3, 0.7, 0.4]	Rest
001	2024/11/1 9:00	[0.1, 0.6, 0.4]	[0.3, 0.2, 0.1]	[0.2, 0.5, 0.3]	Study
002	2024/11/1 8:00	[0.3, 0.6, 0.1]	[0.6, 0.2, 0.1]	[0.1, 0.4, 0.2]	Study
002	2024/11/1 8:30	[0.2, 0.5, 0.4]	[0.5, 0.4, 0.3]	[0.3, 0.6, 0.2]	Study
002	2024/11/1 9:00	[0.3, 0.7, 0.2]	[0.6, 0.5, 0.4]	[0.2, 0.6, 0.3]	Study
003	2024/11/1 8:00	[0.4, 0.3, 0.2]	[0.5, 0.3, 0.2]	[0.3, 0.4, 0.1]	Study
003	2024/11/1 8:30	[0.1, 0.5, 0.3]	[0.4, 0.2, 0.1]	[0.2, 0.5, 0.2]	Rest
003	2024/11/1 9:00	[0.2, 0.6, 0.3]	[0.5, 0.3, 0.2]	[0.3, 0.4, 0.2]	Rest

the attention weight is further fused with the value vector to generate the output feature. In this way, the model can dynamically adjust the learning path based on the relationship between features. Each row in the table represents the calculation process of student behavior data at a specific time point, reflecting how the model integrates multimodal information and uses the self-attention mechanism to capture the correlation between different features, thereby generating personalized learning path recommendations.

3.3 Adaptive optimization of adversarial training

This paper introduces adversarial training methods in the training process to improve the adaptability of the learning path planning system in the face of uncertain data and changes in student behavior, and uses the generation of adversarial samples to strengthen the model. The process of generating adversarial samples is based on the fast gradient symbol method to generate input data with slight perturbations to test the stability of the model in different learning scenarios. Given the label $\mathbf{Y}$ corresponding to the real input data $\mathbf{X}$, the gradient $\nabla_{\mathbf{X}}\mathcal{L}$ of the loss function relative to the input value is calculated to generate perturbation data $\mathbf{X}_{adv}$:

$$\mathbf{X}_{adv} = \mathbf{X} + \in \cdot \mathrm{sign}\left(\nabla_{\mathbf{X}}\mathcal{L}(\theta, \mathbf{X}, \mathbf{Y})\right) \tag{8}$$

ϵ is a hyperparameter that controls the size of the perturbation, and sign represents the sign function of the gradient direction. The generated adversarial sample $\mathbf{X}_{adv}$ contains a small amount of noise compared to the original data, simulating possible abnormal data. The addition of this adversarial data makes the model robust to potential deviations in different learning scenarios during training.

The selection criteria are based on the amplitude of sample disturbance, the sensitivity of classification boundaries, and the impact on model performance. By controlling the perturbation size ϵ to ensure that the generated confrontation samples still have labels consistent with the original sample. The paper selects the $\epsilon \in [0.01, 0.1]$, strike a balance between disturbance and sample authenticity; preferentially select samples near the classification boundary for perturbation.

Table 2 Transformer self-attention calculation example data

Student ID	Query vector	Key vector	Value vector	Attention weights	Output features H_{out}
001	[0.1, 0.3, 0.1]	[0.4, 0.3, 0.2]	[0.2, 0.6, 0.1]	[0.4, 0.2, 0.2]	[0.3, 0.5, 0.4]
001	[0.1, 0.3, 0.5]	[0.2, 0.7, 0.5]	[0.4, 0.4, 0.1]	[0.5, 0.4, 0.1]	[0.4, 0.6, 0.5]
001	[0.2, 0.5, 0.4]	[0.2, 0.6, 0.3]	[0.1, 0.4, 0.2]	[0.6, 0.4, 0.3]	[0.4, 0.5, 0.3]
002	[0.1, 0.4, 0.2]	[0.6, 0.5, 0.4]	[0.2, 0.6, 0.3]	[0.4, 0.4, 0.2]	[0.4, 0.1, 0.4]
002	[0.1, 0.3, 0.2]	[0.1, 0.5, 0.3]	[0.5, 0.3, 0.2]	[0.6, 0.4, 0.3]	[0.5, 0.6, 0.4]
002	[0.2, 0.4, 0.3]	[0.4, 0.2, 0.1]	[0.2, 0.5, 0.2]	[0.5, 0.3, 0.2]	[0.3, 0.5, 0.4]
003	[0.1, 0.4, 0.2]	[0.6, 0.4, 0.3]	[0.4, 0.4, 0.2]	[0.2, 0.6, 0.3]	[0.3, 0.5, 0.3]
003	[0.1, 0.3, 0.2]	[0.1, 0.5, 0.3]	[0.5, 0.3, 0.2]	[0.6, 0.4, 0.3]	[0.5, 0.6, 0.4]
003	[0.2, 0.4, 0.3]	[0.4, 0.2, 0.1]	[0.2, 0.5, 0.2]	[0.5, 0.3, 0.2]	[0.3, 0.5, 0.3]

Because these samples have the greatest impact on the model's decision, we can evaluate the effectiveness of using test adversarial samples to mutate the model output, and retain those samples that can significantly weaken the model's classification accuracy in the training process. This criterion ensures that the generated confrontation samples can effectively expose the vulnerable points of the model. At the same time, not too many uncontrollable factors will be introduced to affect the overall optimization of the model. The effectiveness of the sample is evaluated by comparing the performance of the model on the original test set and the test set, if the performance of the model on the test set is obviously reduced, it proves that the sample can expose the weaknesses of the model, it shows that the generation strategy of the sample is reasonable, and the training process is effective. The impact of confrontation training on model performance is mainly reflected in two aspects. On the one hand, it enhances the robustness of the model to noise and abnormal data, allowing the model to better adapt to multiple changes in the learning scenario; on the other hand, it may lead to the increase of training time because of the high complexity of generating confrontation samples and training time. Although the adversarial training may slightly reduce the adaptation speed of the model to simple samples, it significantly improves the performance of complex samples, which can optimize the stability and generalization ability of the overall performance.

In each training iteration, the model is optimized using real data and trained using adversarial samples. This measure enables the model to still output accurate prediction results when faced with such abnormal data, improves the generalization ability of the model, reduces the instability of path planning caused by abnormal data, and enhances the adaptability of the model in a variety of complex learning scenarios.

In order to further improve the effect of adversarial training, this paper adopts a mixed training strategy of adversarial samples and real samples, that is, adding a combination of real samples and adversarial samples in each training to optimize the performance of the model under normal and abnormal data. The adversarial sample set is represented as $\mathbf{X}_a$ and the real sample set is $\mathbf{X}_{real}$. The total loss function can be expressed as:

$$\mathcal{L}_{total} = \alpha \cdot \mathcal{L}\left(\theta, \mathbf{X}_{real}, \mathbf{Y}\right) + (1-\alpha) \cdot \mathcal{L}\left(\theta, \mathbf{X}_a, \mathbf{Y}\right) \tag{9}$$

α is a hyperparameter that controls weights and adjusts the model's emphasis on real samples and adversarial samples. By adjusting the value of α, the model's performance on different types of data can be balanced, allowing the model to maintain high accuracy in the case of regular learning data and show strong robustness in abnormal data environments.

Randomness is introduced in the adversarial sample generation stage to make the model more stable in more complex scenarios. Each time an adversarial sample is generated, a small random perturbation is added within the range of ϵ values to make the adversarial sample different each time it is trained, to avoid overfitting the model to a specific perturbation pattern, and to enhance the model's adaptability to different abnormal data. This random perturbation strategy utilizes the characteristics of dispersed adversarial samples, allowing the model to adaptively adjust when facing diverse abnormal data, and is not easily restricted by a specific perturbation pattern, thus achieving more robust learning path planning to a certain extent.

The influence of adversarial training on model performance is mainly reflected in two aspects: on the one hand, it enhances the robustness of the model to noise and abnormal data, and enables the model to better adapt to multiple changes in learning scenarios; on the other hand, it may lead to the increase of training time due to the high complexity of generating adversarial samples and training time. Although the adversarial training may slightly reduce the adaptation speed of the model to simple samples, it can significantly improve the performance of complex samples, which can optimize the stability and generalization ability of the overall performance.

In the above process, this paper introduces adversarial training to simulate abnormal data changes that may occur in the actual learning process, and gradually improves the robustness of the model. The model after adversarial training can still provide stable and accurate personalized path planning when the learning behavior data undergoes uncertain changes, providing a reliable guarantee for the personalized adaptation of the student learning process.

In solving the security problems caused by confrontation samples, multi-layer security protection mechanism is introduced in the model design. Combined with the combination of generated confrontation network and confrontation training, the generated confrontation samples are trained together with normal samples, so that the model can not only perform well on conventional data, but also resist malicious attacks. The feature transformation adversarial detection algorithm can be used to monitor the abnormal behavior of input data in real time, alarm the model and update the model.

3.4 Efficient data processing of quantum state classification

In order to better deal with the high-dimensional problem in multimodal data processing, this paper adopts quantum state encoding and classification in quantum machine learning methods, and uses quantum states to efficiently represent and classify feature vectors. The feature vectors in the multimodal data set are mapped to quantum states, and the superposition and entanglement characteristics of quantum states are used to significantly compress the storage and computing requirements of high-dimensional data. For a given input feature vector $\mathbf{x}$, it is encoded into quantum state $|\psi\rangle$:

$$\left|\psi\right\rangle = \frac{1}{\| \mathbf{x} \|} \sum\nolimits_{i=1}^{n} x_i \left| i \right\rangle \tag{10}$$

$|i\rangle$ is the standard ground state, and $\| \mathbf{x} \|$ is the norm of the feature vector. Using this mapping, the original high-dimensional features can be represented in the form of linear expansion in the quantum state, allowing subsequent calculations to be completed in the quantum state space, avoiding the overhead of performing item-by-item operations on high-dimensional data.

This paper uses quantum measurement operations to perform feature classification and clustering. After the quantum state $|\psi\rangle$ is transformed by a series of quantum gates, the classification result of the quantum state is obtained through measurement. The target quantum state is $|\phi\rangle$. Using the similarity information in the measurement results, the inner product $|\langle\phi|\psi\rangle|^2$ of $|\psi\rangle$ and $|\phi\rangle$ can be calculated, and the category of the data point can be determined in this way. The quantum measurement process is extremely efficient in multi-feature clustering and classification, reducing the complexity caused by the increase in dimensions in traditional computing.

A path recommendation algorithm based on quantum classification is also designed and optimized to meet the real-time requirements of the personalized path planning system. The quantum classification algorithm improves the data classification speed through quantum coherence and interference phenomena, realizes rapid clustering of features in high-dimensional space, and shortens the calculation time. In quantum classification, the quantum support vector machine algorithm (QSVM) is used to perform quantum feature mapping on each input quantum state and embed it into the quantum Hilbert space. For a given quantum state pair $\left(|\psi_i\rangle, |\psi_j\rangle\right)$, its kernel function can be expressed as:

$$K\left(\psi_i, \psi_j\right) = \left|\langle \psi_i \middle| \psi_j \rangle\right|^2 \tag{11}$$

This quantum kernel function directly calculates the similarity of two data points on the quantum state, achieving rapid data classification in high-dimensional space. The improvement of quantum state classification mainly comes from the characteristics of quantum computing, which uses the characteristics of quantum superposition and quantum entanglement to achieve data parallelization and feature mapping capabilities. This feature enables quantum state classification to perform feature clustering and classification more efficiently when processing complex multimodal data, especially when facing the curse of dimensionality and nonlinear patterns, showing obvious computational advantages.

During the training phase, QSVM determines the classification surface of various quantum states by maximizing the boundary distance of the quantum state, ensuring accurate classification under different personalized path requirements. In practical applications, this paper uses the results of quantum state classification to calculate personalized path recommendations in real time. After calculating the quantum inner product of each student's characteristic state and the path characteristic state, the paths with higher similarity are recommended first, realizing efficient path planning based on quantum classification. The quantum classification process has high parallelism in feature classification. The use of quantum state encoding reduces storage space and makes the calculation process of path planning more efficient. Quantum state classification is used to achieve efficient processing of multimodal data, especially in high-dimensional feature space, which can achieve the effect of real-time classification and path planning.

Table 3 shows the encoding and mapping methods of multimodal data in quantum computing. Different quantum state encoding methods are used depending on the data type. Text data uses quantum state superposition encoding to map its features to quantum space, while audio data uses quantum state quantization and compression encoding, which is suitable for processing higher-dimensional data. The processing of video data requires more complex quantum image feature extraction and encoding methods, which can adapt to higher data dimensions. In terms of computational complexity, the processing of text data is relatively simple, with a computational complexity of $O(n)$, while the computational complexity of audio and video data is $O\left(n^2\right)$ and $O(n\log n)$ respectively, and the data requires more computing resources.

Table 3 Quantum state encoding and mapping method

Data type	Feature dimension	Quantum state mapping method	Computational complexity
Text data	1076	Quantum state superposition encoding	O(n)
Audio data	504	Quantum state quantization and compression encoding	$O(n^2)$
Video data	2432	Quantum state image feature extraction and encoding	O(nlogn)

Table 4 shows the key processing steps of quantum classification, each of which involves specific data types and quantum operation methods. The data mapping step maps the student's behavioral feature vector to a quantum state vector, using quantum state superposition coding to improve data representation efficiency. The feature transformation operation processes the learning interest and state data through quantum gate transformation and coherence operation to obtain a new quantum state. These quantum states perform quantum inner product calculations through classification operations, evaluate the similarity between different data, determine the category to which the students belong, and calculate the similarity score; based on the students' categories and path characteristics, the quantum path recommendation mechanism uses quantum measurement and classification to generate personalized learning path recommendations.

4 Method effect evaluation

4.1 Personalized learning path generation effect

After a series of processing, the feature data can reflect the students' interest fluctuations and behavioral states during the learning process. It can also be used to analyze students' learning behavior and generate personalized learning paths. The feature data is divided into training sets and test sets in a ratio of 8:2 for subsequent experiments.

Students' learning interest data is calculated based on their participation and concentration in the learning content. The frequency with which students access certain learning resources or complete tasks within a specific time period reflects their interest in the content. The generation of learning paths is achieved through the Transformer model. The model uses input students' learning behavior data to learn students' interest patterns and adjusts position encoding based on these patterns. The model's self-attention mechanism can obtain the dynamic changes of students' interests during the learning process and recommend paths that adapt to the current learning status.

Figure 3 shows the changes in positional encoding and dynamic learning state. The horizontal axis is the time step of learning, reflecting the learning progress of students at different stages, and the vertical axis represents the relative level changes of students' positional encoding and interest; the learning state increases in the early stage of learning, indicating that students' interest in the learning content gradually deepens. As the time step increases, the learning state fluctuates, indicating that students experience a decline in interest or fatigue in the later stages of learning. The adjusted value of the position encoding reflects the ability of the Transformer model to adjust the learning path according to the learning state, and the changes in the position encoding value at different time steps show a certain regularity. As students' interests change, the adjustment of positional encoding values helps the model effectively obtain the relevance of learning content at different learning stages; these changes show that the Transformer model not only increases the recommendation of learning content when students' interest is high, but also appropriately adjusts the learning path when their interest is low, helping students better cope with different learning states, thereby improving the personalized adaptability of learning.

Accuracy and recall are the main indicators to evaluate the effect of personalized learning pathway generation. During the evaluation process, the Transformer-based recommendation model of this study was compared with the traditional collaborative filtering (CF) and long short-term memory (LSTM)-based recommendation models, using the same training set and evaluation framework to ensure the fairness of the comparison and the comparability of the results. Collaborative filtering is a classic recommendation method, which is widely used in the traditional personalized recommendation field. It uses user behavior or item similarity to realize recommendation. As a benchmark model, it is simple and easy to implement, and can effectively reflect the improvement of complexity and accuracy. LSTM is widely used in behavior prediction and recommendation system, and it is also the model used to generate the most learning paths. Its memory

Table 4 Schematic diagram of quantum classification and clustering algorithm process

Step	Input data type	Quantum operation method	Output result
Data mapping	Student Behavior Feature Vector	Quantum state superposition mapping	Quantum state vector
Feature transformation	Learning Interest and State Data	Quantum gate transformation and coherence operation	Transformed quantum state
Classification operation	Quantum states	Quantum inner product calculation and similarity measurement	Category label and similarity score
Path recommendation	Student category and path features	Quantum measurement and classification	Personalized learning path recommendation

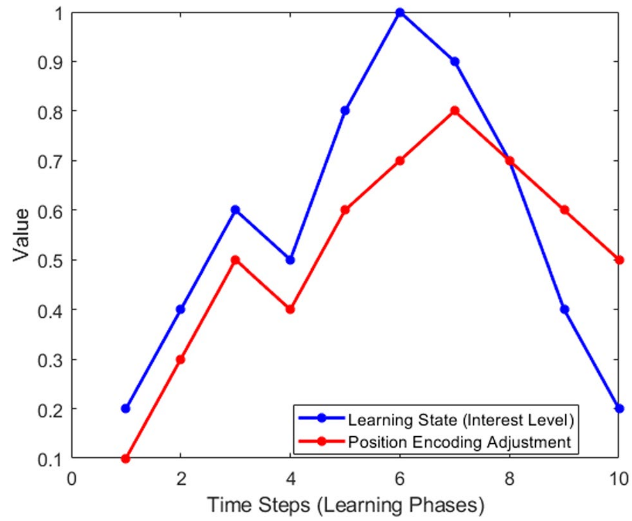

Fig. 3 Positional encoding and dynamic learning state

unit can learn long-term dependencies and help to capture the time evolution characteristics of students' behavior. For each model, the paper calculates its recommendation accuracy and recall for different student groups. When calculating the accuracy, it takes the ratio of the number of learning paths successfully recommended by the model to the total number of recommendations. When calculating the recall, the paper takes the ratio of the number of correct paths recommended by the model to the actual number of paths required by students, to reflect the responsiveness of the model to changes in the interests of different students.

Figure 4 shows the changes in accuracy and recall of the Transformer model, traditional collaborative filtering and LSTM model in 20 iterations. The Transformer model in Figure a shows volatility in the early stage, but eventually stabilizes, with an accuracy rate of 0.95, which can gradually adapt to changes in students' interests and status. The accuracy of the collaborative filtering model was relatively stable, remaining at 0.88 in the 17th round. The accuracy of the LSTM model also showed a gradual increase, and finally stabilized at 0.93, showing that its accuracy in obtaining students' learning behaviors has gradually improved, but it is lower than the Transformer model. Figure b shows the performance of the three models in terms of recall rate. The recall rate of the Transformer gradually increased from 0.78 to stabilize at 0.91, reflecting good recommendation diversity and adaptability. The recall rate of the collaborative filtering model fluctuates slightly and finally stabilizes at 0.85, indicating that its effect on diversity recommendation is relatively limited. The recall rate of the LSTM model is also stable at 0.91, which also shows its advantages in obtaining student needs, but its overall performance is not as good as the Transformer model in this paper.

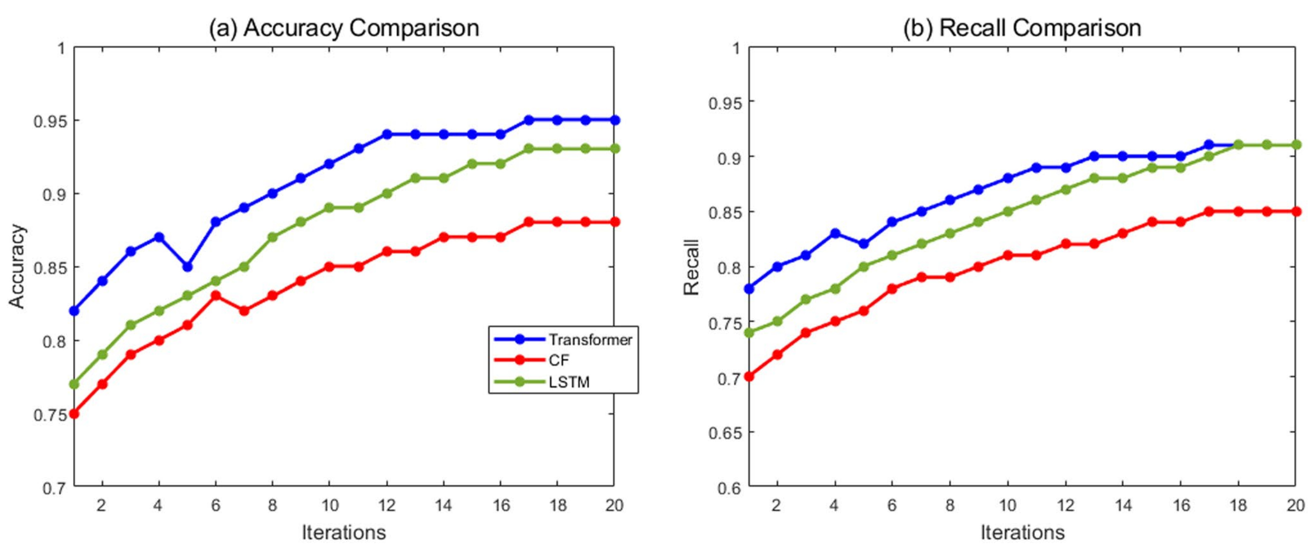

Fig. 4 Comparison of accuracy and recall of different recommendation models

4.2 Effect of adversarial training

The effectiveness of adversarial training is evaluated by comparing the loss functions of conventional training and adversarial training. In conventional training, only real student behavior data is used; in adversarial training, in addition to real data, adversarial samples generated by adversarial methods are also added to the training data.

Figure 5 shows the optimization process of the loss function under different weight parameters α, which helps analyze the adaptability improvement effect of the model under diversified data conditions. The horizontal axis represents the weight parameter α, ranging from 0 to 1, which is used to balance the relative weights of the original loss and the adversarial loss; the vertical axis represents the loss value, which includes the original loss, adversarial loss, and the weighted total loss of the two. The original loss decreases as α increases, the adversarial loss increases as α increases, and the total loss is a weighted combination of the two. The trend of the total loss shows that the appropriate selection of α value can balance the performance of the model on normal data and adversarial samples. Comparing the total loss values under different α s and finding the optimal parameters for improving the model's adaptability in complex data environments helps ensure the stability of the model in various learning scenarios.

Figure 6 has two sub-graphs, a and b, which are the comparison of loss functions and test set accuracy of conventional training and adversarial training, respectively. The horizontal axis is 100 iterations. The loss value of conventional training is lower in the early stage, decreases slowly, and the optimization speed is limited; the loss value of adversarial training decreases more rapidly and lasts longer, and the loss value in the later stage of iteration is significantly lower than that of conventional training, fluctuating around 0.05, indicating that adversarial training is more adaptable to abnormal data. In the test set accuracy graph, the accuracy of conventional training is higher than that of adversarial training at the beginning, but gradually stagnates at a low level in the later stage. The accuracy of adversarial training continues to improve, and in the later stage it is significantly higher than that of conventional training, fluctuating around 0.95, showing its adaptability to complex scenarios. Adversarial training is superior to conventional training in adaptability and robustness, which helps to improve the stability and accuracy of the system in complex environments.

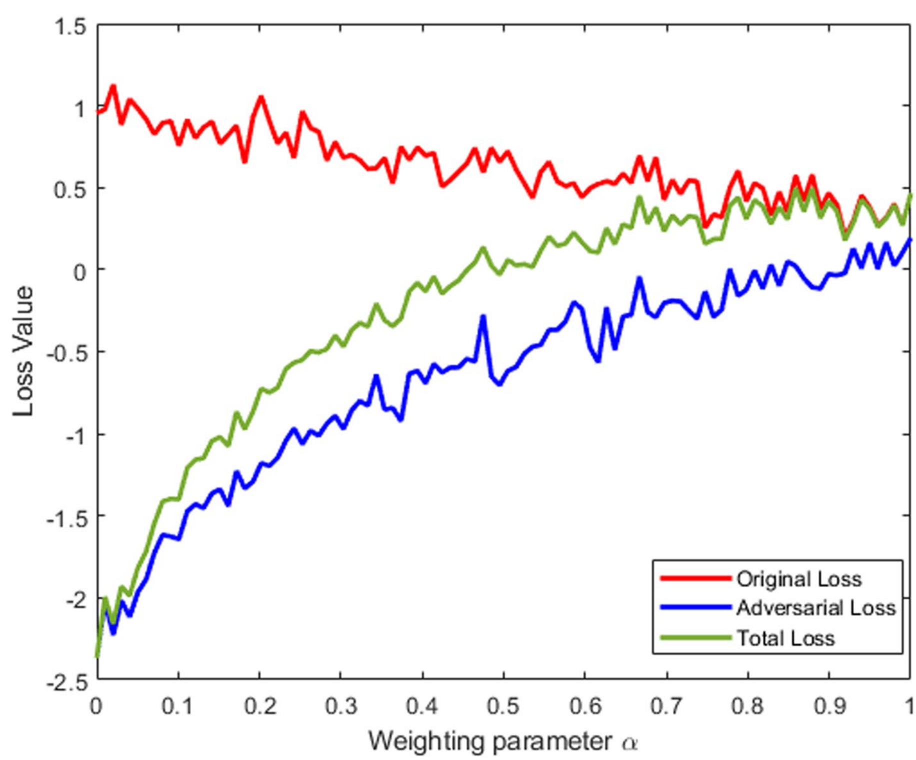

Fig. 5 Schematic diagram of loss function optimization

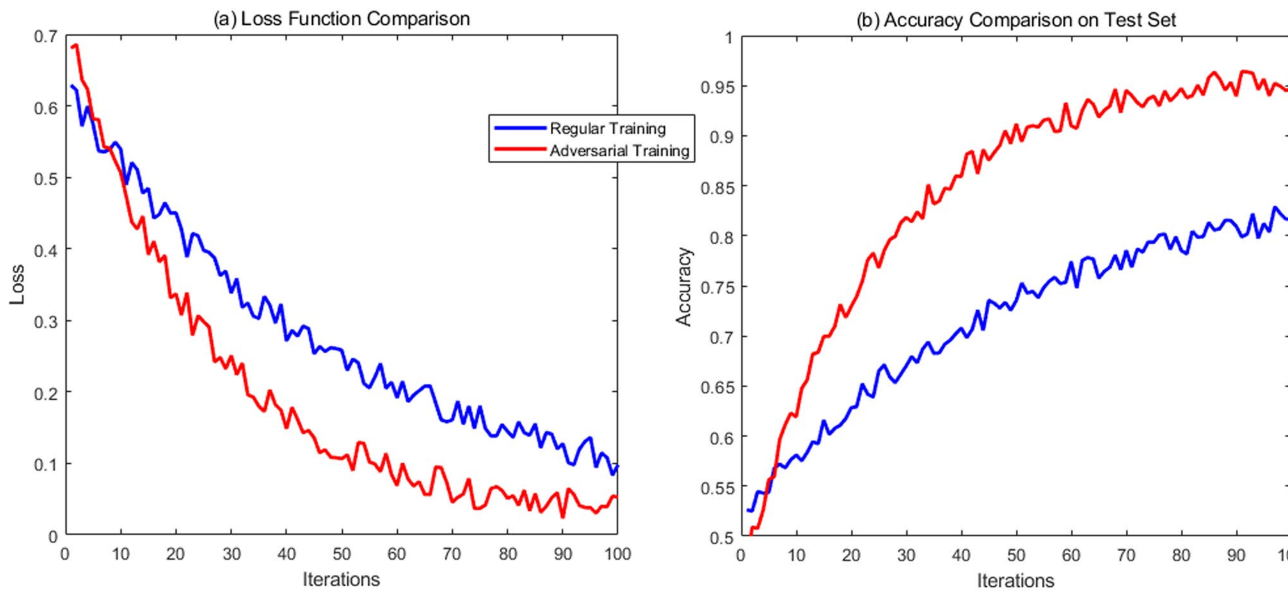

Fig. 6 Comparison of adversarial training and conventional training

4.3 Data processing efficiency of quantum state classification

When evaluating data processing efficiency, this paper uses processing time as the main indicator and compares the computing time before and after the introduction of quantum state classification in the Transformer model. The training and prediction time of the Transformer model in multimodal data processing before the introduction of quantum state classification was measured, and then the training and prediction time combined with the quantum state classification method was calculated. The same training set was processed, the time difference between the two was compared, and the acceleration effect of quantum state classification in data processing was evaluated.

Figure 7 shows the comparison of processing time between the traditional Transformer model and the one after the introduction of quantum state classification. In Figure a, the processing time of the Transformer model before the introduction of quantum state classification gradually decreases with the number of iterations, but the change is small, and the overall processing time is long, reaching 87 s in the 18th iteration. In Figure b, the model after the introduction of quantum state classification shows significant time optimization, and the processing time drops rapidly in the early stage, and finally drops to 56 s in the 18th iteration. Quantum computing accelerates data processing, especially in later iterations. Quantum state classification effectively improves the computational efficiency of the Transformer model, especially in large-scale data processing.

5 Conclusions

This paper introduces a personalized learning path recommendation method based on the Transformer model and quantum state classification, combined with adversarial training to optimize robustness, and solves the limitations of traditional methods in dynamic learning scenarios. The Transformer model extracts key features from multimodal data through the self-attention mechanism, and integrates information from different modalities through model integration and fusion strategies to improve the learning effect of the model. The interactive mechanism enables the model to handle the complex relationships between different data sources, optimize feature combinations, and more accurately capture the dynamic learning state of students. By generating samples to simulate potential abnormal situations, the robustness and stability of the model are improved, the system is guaranteed to operate effectively under data uncertainty, and the quantum state classification method is introduced into quantum computing. Quantum feature processing can greatly accelerate the data processing process, and cluster and classify high-dimensional

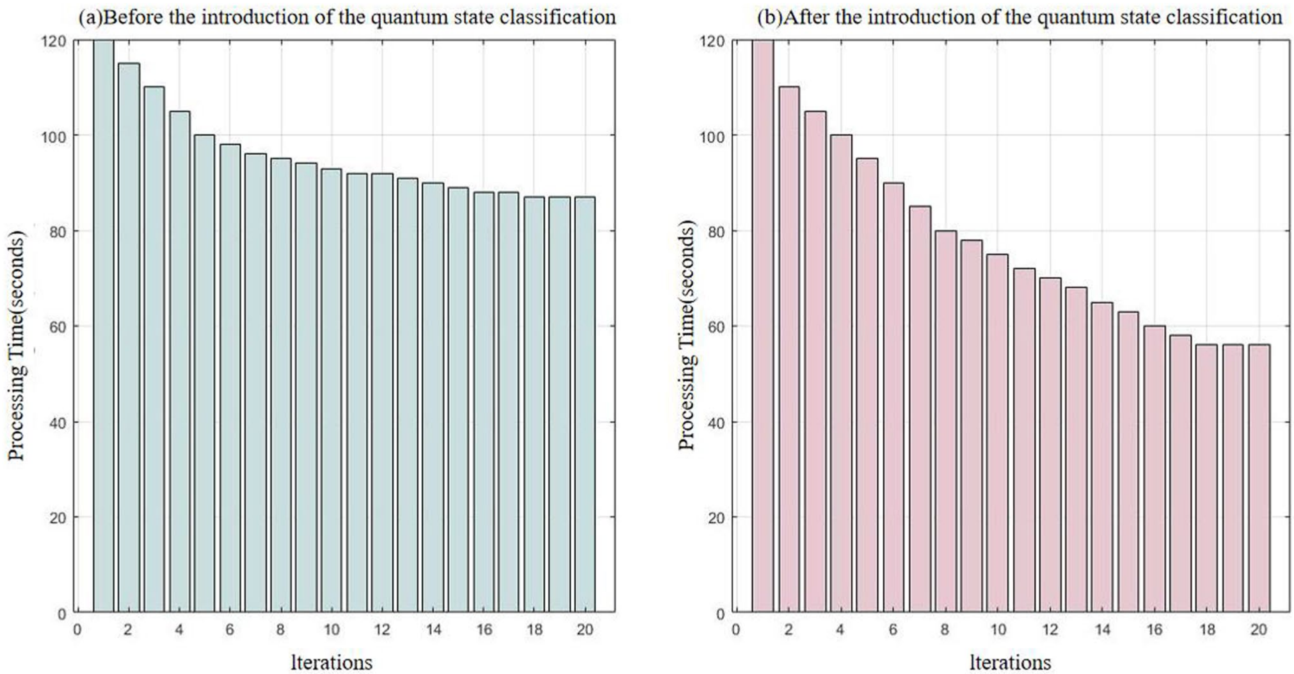

Fig. 7 Comparison of computing time before and after the introduction of quantum state classification in the Transformer model

data, further improving the computational efficiency and real-time performance of the personalized recommendation system. Experimental results show that the Transformer-based recommendation model outperforms traditional collaborative filtering and long short-term memory network models in terms of accuracy and recall. In terms of adaptability, adversarial training is used to simulate possible abnormal data, which enhances the model's adaptability to different learning scenarios and makes the recommendation model more stable and reliable when facing uncertain data. In terms of computational efficiency, the introduction of quantum state classification methods has significantly improved the data processing efficiency of the Transformer model, especially in the later iterations, the processing time is significantly reduced, meeting the real-time requirements. This study introduces Transformer model and quantum state classification method, which realizes the efficient generation and recommendation of personalized learning path under multimodal data, and improves the accuracy and processing efficiency of the system. Studies that still have some limitations. Quantum state classification introduces significant computational overhead and requires complex quantum circuit design, which is challenging given current hardware limitations. Integrating multimodal data is difficult due to heterogeneity and alignment issues, and scaling this integration is resource-intensive. The Transformer model's complexity and lack of interpretability make training challenging and limit transparency. Data privacy and security are also concerns, with sensitive student data needing robust protection and compliance with regulations like GDPR. Future research can expand the size of the dataset, improve the fairness and transparency of the model in more practical scenarios, and explore the improvement of quantum computing hardware and its potential for large-scale deployment in educational scenarios, so as to promote further development and innovation of personalized learning systems.

Author contributions Changzhi Sun conceived and designed the study, performed the experiments, and analyzed the data. Shijie Huang and Banghui Sun contributed to the collection and preprocessing of the data. Shiwei Chu wrote and revised the manuscript. All authors discussed the results and implications of the research and contributed to the final version of the manuscript.

Funding This work was supported by Quality Enhancement Project for New Specialties in Electronic Information Engineering (2022xjzlts027), Bozhou College Anhui Zhongke Zhibo Technology Development Co., Ltd. cooperative practice education base (2023XJXM003) Reform and Practice of DSP Application Technology Curriculum Based on OBE-CDIO Concept (2022XJXM064).

Data availability The data are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate This study was approved by Bozhou College and informed consent was obtained from all participants in accordance with the ethical standards of the University. For participants under 16 years of age, we obtained consent from their parents or guardians. Participants were informed of their rights and a sample consent form was provided upon request.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

1. Tapalova O, Zhiyenbayeva N. Artificial intelligence in education: AIEd for personalised learning pathways. Electron J e-Learn. 2022;20(5):639–53.
2. Alamri H, Lowell V, Watson W, Watson SL. Using personalized learning as an instructional approach to motivate learners in online higher education: learner self-determination and intrinsic motivation. J Res Technol Educ. 2020;52(3):322–52.
3. Peng H, Ma S, Spector JM. Personalized adaptive learning: an emerging pedagogical approach enabled by a smart learning environment. Smart Learn Environ. 2019;6(1):1–14.
4. Alamri HA, Watson S, Watson W. Learning technology models that support personalization within blended learning environments in higher education. TechTrends. 2021;65(1):62–78.
5. Bishop PA, Downes JM, Netcoh S, Farber K, DeMink-Carthew J, Brown T, Mark R. Teacher roles in personalized learning environments. Elem Sch J. 2020;121(2):311–36.
6. Castro R. Blended learning in higher education: trends and capabilities. Educ Inf Technol. 2019;24(4):2523–46.
7. Bonfield CA, Salter M, Longmuir A, Benson M, Adachi C. Transformation or evolution?: Education 4.0, teaching and learning in the digital age. High Educ Pedagog. 2020;5(1):223–46.
8. Pataranutaporn P, Danry V, Leong J, Punpongsanon P, Novy D, Maes P, Sra M. AI-generated characters for supporting personalized learning and well-being. Nat Mach Intell. 2021;3(12):1013–22.
9. Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education–where are the educators? Int J Educ Technol High Educ. 2019;16(1):1–27.
10. Jaiswal A, Arun CJ. Potential of Artificial Intelligence for transformation of the education system in India. Int J Educ Dev Using Inf Commun Technol. 2021;17(1):142–58.
11. Ifenthaler D, Yau JYK. Utilising learning analytics to support study success in higher education: a systematic review. Educ Tech Res Dev. 2020;68(4):1961–90.
12. Anthonysamy L, Koo AC, Hew SH. Self-regulated learning strategies in higher education: fostering digital literacy for sustainable lifelong learning. Educ Inf Technol. 2020;25(4):2393–414.
13. Williamson B, Bayne S, Shay S. The datafication of teaching in Higher Education: critical issues and perspectives. Teach High Educ. 2020;25(4):351–65.
14. Chen X, Zou D, Xie H, Cheng G, Liu C. Two decades of artificial intelligence in education. Educ Technol Soc. 2022;25(1):28–47.
15. Jones KM. Learning analytics and higher education: a proposed model for establishing informed consent mechanisms to promote student privacy and autonomy. Int J Educ Technol High Educ. 2019;16(1):1–22.
16. Rapanta C, Botturi L, Goodyear P, Guàrdia L, Koole M. Balancing technology, pedagogy and the new normal: post-pandemic challenges for higher education. Postdigit Sci Educ. 2021;3(3):715–42.
17. Farrokhnia M, Banihashem SK, Noroozi O, Wals A. A SWOT analysis of ChatGPT: implications for educational practice and research. Innov Educ Teach Int. 2024;61(3):460–74.
18. Zhao Y, Watterston J. The changes we need: education post COVID-19. J Educ Change. 2021;22(1):3–12.
19. Patle BK, Pandey A, Parhi DRK, Jagadeesh AJDT. A review: On path planning strategies for navigation of mobile robot. Def Technol. 2019;15(4):582–606.
20. Rudolph J, Tan S, Tan S. ChatGPT: bullshit spewer or the end of traditional assessments in higher education? J Appl Learn Teach. 2023;6(1):342–63.
21. Mu M, Yuan M. Research on a personalized learning path recommendation system based on cognitive graph with a cognitive graph. Interact Learn Environ. 2024;32(8):4237–55.
22. Chen W, Shen Z, Pan Y, Tan K, Wang C. Applying machine learning algorithm to optimize personalized education recommendation system. J Theory Pract Eng Sci. 2024;4(01):101–8.

23. Liu Z, Wang L, Li X, Pang S. A multi-attribute personalized recommendation method for manufacturing service composition with combining collaborative filtering and genetic algorithm. J Manuf Syst. 2021;58:348–64.
24. Zhang X, Liu S, Wang H. Personalized learning path recommendation for e-learning based on knowledge graph and graph convolutional network. Int J Software Eng Knowl Eng. 2023;33(01):109–31.
25. Li H, Gong R, Zhong Z, Xing L, Li X, Li H. Research on personalized learning path planning model based on knowledge network. Neural Comput Appl. 2023;35(12):8809–21.
26. Chen L, Wang Y, Miao Z, Mo Y, Feng M, Zhou Z, Wang H. Transformer-based imitative reinforcement learning for multirobot path planning. IEEE Trans Industr Inf. 2023;19(10):10233–43.
27. Zhou J, Li Y, Li Y, Wang Y. A novel hybrid deep learning IChOA-CNN-LSTM model for multimodal sentiment analysis during the COVID-19 pandemic. Sci Rep. 2024;14(1):1–16.
28. Mu S, Cui M, Huang X. Multimodal data fusion in learning analytics: a systematic review. Sensors. 2020;20(23):6856.
29. Heston TF, Khun C. Prompt engineering in medical education. Int Med Educ. 2023;2(3):198–205.
30. Alqahtani T, Badreldin HA, Alrashed M, Alshaya AI, Alghamdi SS, Bin Saleh K, Albekairy AM. The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research. Res Soc Admi Pharm. 2023;19(8):1236–42.
31. Alshahrani A. The impact of ChatGPT on blended learning: current trends and future research directions. Int J Data Netw Sci. 2023;7(4):2029–40.
32. Xiong L, Chen Y, Peng Y, Ghadi YY. Improving robot-assisted virtual teaching using transformers, GANs, and computer vision. J Organ End User Comput (JOEUC). 2024;36(1):1–32.
33. Singh R, Ren J, Lin X. A review of deep reinforcement learning algorithms for mobile robot path planning. Vehicles. 2023;5(4):1423–51.
34. Zhang S, Wijesinghe A, Ding Z. Rme-gan: a learning framework for radio map estimation based on conditional generative adversarial network. IEEE Internet Things J. 2023;10(20):18016–27.
35. Kavitha SS, Kaulgud N. Quantum machine learning for support vector machine classification. Evol Intel. 2024;17(2):819–28.
36. Gharehchopogh FS. Quantum-inspired metaheuristic algorithms: comprehensive survey and classification. Artif Intell Rev. 2023;56(6):5479–543.
37. Bilal M, Ali G, Iqbal MW, Anwar M, Malik MSA, Kadir RA. Auto-Prep: efficient and robust automated data preprocessing pipeline. IEEE Access. 2022. https://doi.org/10.1109/ACCESS.2022.3198662.
38. Cai H, Jiang Q, Li L, Suter BW. Analysis of adaptive short-time Fourier transform-based synchrosqueezing transform. Anal Appl. 2021;19(01):71–105.
39. Li M, Liu Y, Zhi S, Wang T, Chu F. Short-time Fourier transform using odd symmetric window function. J Dyn Monit Diagn. 2022;1(1):37–45.
40. Wen L, Li X, Gao L. A transfer convolutional neural network for fault diagnosis based on ResNet-50. Neural Comput Appl. 2020;32(10):6111–24.
41. Bharati S, Podder P, Mondal MRH, Prasath VS. CO-ResNet: optimized ResNet model for COVID-19 diagnosis from X-ray images. Int J Hybrid Intell Syst. 2021;17(1–2):71–85.
42. Kingma DP, Welling M. An introduction to variational autoencoders. Found Trends® Mach Learn. 2019;12(4):307–92.
43. Xu W, Tan Y. Semisupervised text classification by variational autoencoder. IEEE Trans Neural Netw Learn Syst. 2019;31(1):295–308.