International Scientific Technical and Economic Research **(ISSN: 2959-1309)**
Vol. 4, No. 2, (June 2026), pp. 166-185
Journal homepage: https://www.istaer.online

RESEARCH ARTICLE **DOI: 10.71451/ISTAER2620**

Real Time Adaptive Teaching Digital Human System Based on Large Language Model

Wumeng Yang*

AI Science and Technology Department, Beijing SRT Education & Technology Co., Ltd., Beijing, China

Abstract: With the diversification and personalization of educational needs, the traditional teaching model is facing many challenges, especially in meeting students' personalized learning needs and providing real-time feedback. This study proposes a real-time adaptive teaching digital human system based on a large language model, which aims to improve the quality of education and student engagement through intelligent technology. The system obtains students' learning status in real time through a variety of data acquisition devices (such as learning behavior tracking, question-answering records, and speech recognition) and uses the large language model to generate personalized teaching content and feedback. The system calculates a comprehensive learning status score by evaluating students' learning progress, answer accuracy, participation, and learning time, thereby dynamically adjusting the teaching content and learning path. Experimental results show that with the system, students' knowledge mastery rate increases by 15%, understanding depth by 17%, and learning interest by 20%. The system's response time is reduced from 5 seconds (in traditional systems) to 1.5 seconds, and its processing capacity is increased by 2.5 times, supporting more concurrent users. The successful implementation of the system provides a new solution for personalized education and has broad application prospects, especially in online education and distance learning.

Keywords: Large language model; Real time adaptive; Teaching digital people; Individualized education; System optimization

How to Cite: Yang, W. (2026). Real Time Adaptive Teaching Digital Human System Based on Large Language Model. *International Scientific Technical and Economic Research, 4*(2), 166–185. **https://doi.org/10.71451/ISTAER2620**

Article history: Received: 27 Jan 2026; **Revised:** 01 Mar 2026; **Accepted:** 19 Apr 2026; **Published:** 26 Apr 2026

1. INTRODUCTION

1.1 Research background

The current education model faces many challenges, particularly in the personalization of teaching content and the evaluation of teaching effectiveness. The traditional education model often suffers from problems such as teacher-student imbalance, monolithic teaching methods, and inflexible adjustment of teaching content, making it difficult to meet the unique needs of

* **Corresponding author:** Wumeng Yang, AI Science and Technology Department, Beijing SRT Education & Technology Co., Ltd., Beijing, China. Email: yangwumeng0209@163.com

each student [1]. With the diversification and personalization of educational needs, the limitations of the traditional education model are gradually exposed [2],[3]. Therefore, a more flexible and intelligent educational approach is needed to meet students' learning needs, provide personalized content and real-time feedback, and improve learning outcomes and student engagement.

In this context, the rise of large language models (such as the GPT series) provides significant potential for innovation in education [4],[5],[6],[7]. With their powerful natural language processing capabilities, large language models can understand and generate text in multiple languages, providing strong technical support for personalized generation of educational content, real-time analysis of students' learning behavior, and adaptive learning path recommendations. These technologies can help educators better understand students' learning status, respond in a timely manner, and provide targeted teaching content and feedback, thereby improving educational quality [8],[9],[10].

At the same time, as a new educational tool, teaching digital human technology has been gradually applied in many fields, such as virtual teachers and teaching assistants [11],[12]. A teaching digital human can not only interact with students through speech recognition, speech synthesis, and other technologies but also provide real-time feedback based on students' learning status and emotional changes, thereby enhancing immersion and engagement in the learning process. With the continuous development of computer vision, deep learning, natural language processing, and other technologies, teaching digital humans possess strong emotional recognition, language understanding, and personalized service capabilities, providing new possibilities for educational innovation.

1.2 Research purpose and significance

The main goal of this study is to propose a real-time adaptive teaching digital human system based on a large language model. The system can dynamically adjust teaching content according to students' learning behavior and emotional state and interact with students in real time through a virtual teaching digital human, thereby improving learning outcomes and engagement [13],[14],[15].[16],[17],[18]. Compared with the traditional teaching mode, this system can better meet students' personalized learning needs, adjust learning progress in real time, provide personalized feedback, and promote students' in-depth learning.

The innovative significance of the system lies in its combination of the powerful language generation capabilities of the large language model with the emotional interaction capabilities of the virtual teaching digital human, forming a new educational model [19],[20],[21]. This model can optimize teaching content and learning paths in real time in a data-driven manner, allowing each student to learn at their most suitable pace, thus greatly improving learning efficiency and outcomes [22],[23]. In addition, the system's real-time feedback mechanism and personalized recommendation function provide a new solution for personalized teaching in education, holding important theoretical value and practical significance.

With the continuous progress of information technology and artificial intelligence technology, the real-time adaptive teaching digital human system based on large language model has broad application prospects [24],[25],[26]. The system can not only play a role in traditional classroom education, but also be widely used in online education, distance education, vocational training and other fields, providing strong technical support for the popularization and fairness of education.

1.3 Overview of paper structure

The structure of this paper is as follows: Section 2 introduces the overall architecture and design of the real-time adaptive teaching digital human system based on a large language model, including its various modules and functions. Section 3 focuses on the application method and algorithm improvement of the large language model within the system, detailing the process of model selection, fine-tuning, and optimization. Section 4 deeply analyzes the technical details of system implementation, especially the design and implementation of the virtual teaching digital human, as well as system performance optimization and scalability. Section 5 presents the experimental design and verification, including experiment settings, evaluation metrics, and analysis of experimental results. Finally, Section 6 summarizes the research work and discusses future research directions.

In terms of research methods, this study adopts a combination of system design, algorithm optimization, and experimental verification. By constructing an experimental platform, we verify the effectiveness of the real-time adaptive teaching digital human system based on a large language model and analyze its performance in different teaching environments through comparative experiments, thereby providing a theoretical basis and technical support for the intelligent development of education.

2. SYSTEM ARCHITECTURE AND DESIGN

This system adopts a framework based on a large language model, aiming to provide a real-time adaptive teaching environment and deliver a personalized, interactive educational experience through teaching digital humans. The system architecture consists of three core modules: the large language model module, the real-time adaptive mechanism module, and the teaching digital human module. Through the organic combination of these three modules, the system can monitor students' learning status in real time, dynamically adjust teaching content and optimize learning path according to students' feedback.

First, in the overall system architecture, the large language model module serves as the core computing unit, responsible for generating teaching content and feedback based on the input learning content and student responses. The basic framework of the large language model is based on modern pre training models, such as GPT-4 or BERT, and uses fine-tuning technology to transform the general language model into a special model in the field of education [27],[28],[29],[30]. Under this framework, the system can dynamically generate and adjust teaching content according to students' knowledge, learning interest and real-time feedback. For example, the model uses formulas to represent the generation process of teaching content:

$$C(t) = f\big(K(t), I(t), F(t)\big) \tag{1}$$

Where $C(t)$ is the teaching content generated at time t, $K(t)$ is the students' knowledge at that time, $I(t)$ is the students' immediate feedback, and $F(t)$ is the instant feedback from students' interaction. Through this design, the system can adjust the content in real time to ensure the personalization and accuracy of the content.

Second, the integration of the interactive learning platform and the digital human enables students to interact with the system more naturally. As the system's virtual avatar, the teaching digital human possesses natural language processing, emotion recognition, and speech synthesis capabilities. It can interact with students through multimodal means such as voice and images, thereby enhancing immersion and learning interest during the teaching process. In teaching, digital humans not only act as knowledge imparters but also provide timely incentives and guidance based on students' learning status, forming an effective learning feedback loop [31],[32]. The design of this part realizes a highly integrated interactive experience. Students input through voice or text, while digital people feedback through voice or text output.

To ensure efficient operation and real-time performance, the system includes several collaborative modules. The user management module records and analyzes students' historical learning data, ensuring personalized learning records and progress tracking for each student. The data feedback module supports the real-time adaptive mechanism by continuously collecting students' learning behaviors and responses, such as correct answer rates, learning time, and knowledge points. The teaching content generation module generates content matching students' needs based on the large language model and interacts with data from other modules to ensure accuracy and relevance. All these modules exchange and coordinate data through a background cooperation mechanism, ensuring that the system can be adjusted and optimized in real time based on students' learning feedback [33],[34].

Next, the real-time adaptive mechanism is one of the core components of the architecture design. The adaptive algorithm framework dynamically adjusts teaching content and learning paths by obtaining students' learning progress and feedback in real time. The key of algorithm design is the accurate evaluation and prediction of students' state. We can describe the update process of student status through the formula:

$$\hat{S}(t) = g\big(S(t), F(t)\big) \tag{2}$$

Where $\hat{S}(t)$ is the system's prediction of the student's state at time t, $S(t)$ represents the student's current state, and $F(t)$ is the student's immediate feedback. Through this framework, the system can automatically evaluate students' mastery at each learning stage, and adjust the follow-up teaching content according to the evaluation results. Specifically, when the system recognizes that students have difficulties in some knowledge points, it will automatically adjust the difficulty of the content and provide more tips or further explanations.

In addition, the real-time update and feedback mechanism enables the system to continuously track students' learning status and adjust learning paths in real time when necessary. Each feedback of students in the learning process will be used as a new input, affecting the subsequent teaching strategies. This mechanism can be expressed by the following formula:

$$C(t+1) = \phi\left(C(t), \hat{S}(t), R(t)\right) \tag{3}$$

Where $C(t+1)$ represents the teaching content updated at time $t+1$, $C(t)$ is the teaching content at time t, $\hat{S}(t)$ is the student state prediction at time t, and $R(t)$ is the immediate feedback of students. This feedback mechanism ensures the individuation and adaptability of students' learning process.

Finally, the design and implementation of the teaching digital human constitute another important part of the system. In the design of the virtual avatar, the teaching digital human uses high-quality 3D modeling and animation technology to simulate real teaching scenes and facial expressions, thereby enhancing students' learning immersion through the expressions and actions of the virtual character. In the aspect of interaction design, teaching digital people can understand students' oral problems through speech recognition technology, provide natural language feedback through speech synthesis technology, support image recognition and generation technology, and further improve interaction and learning effect through visual feedback. The implementation of multimodal input-output design enables the teaching digital human to have richer and more diverse interactions with students, providing more comprehensive learning support.

3. LARGE LANGUAGE MODEL METHODOLOGY AND ALGORITHM IMPROVEMENT

In the real-time adaptive teaching digital human system based on a large language model, model selection and optimization are central to system performance. First, it is important to choose a large language model suitable for the educational context. Compared with traditional natural language processing tasks, the requirements in educational contexts are more complex and dynamic, requiring the model to effectively handle a large amount of subject knowledge and make dynamic adjustments based on students' feedback. In this context, GPT-4 and BERT have become the first choice due to their powerful generation and understanding capabilities. GPT-4 can generate fluent and natural text and has strong ability of context understanding, while BERT performs well in comprehension tasks, especially in blank filling and sentence classification tasks. For the specific needs of the educational context, we combine the two models to ensure content diversity and naturalness, as well as deep knowledge understanding.

After model selection, domain-specific fine-tuning is a key step to improve model performance. In practical applications, the field of education has a unique semantic structure and terminology, where general language models often perform poorly. Therefore, we need to fine-tune the selected large language model and retrain it on specialized educational datasets so that the model can better understand and generate subject-related content. Specifically, we adjusted the model's loss function, introduced knowledge graphs and textbook data from education, and enhanced the model's reasoning ability in specific disciplinary contexts [35],[36],[37]. The loss function during fine-tuning is defined as:

$$\mathcal{L} = \sum_{i=1}^{N} | y_i - \hat{y}_i | + \lambda \sum_{j=1}^{M} | \mathrm{W}_j - \mathrm{W}_j^* | \tag{4}$$

Where y_i and $\hat{y}_i$ are the true label and the predicted value model's predicted value for i th sample, W_j are the parameters of the model, W_j^* is the parameter adjusted by incorporating educational knowledge in the process of fine-tuning, λ is the regularization coefficient, which is used to balance the matching of task-specific optimization and knowledge map.

Next, for the algorithm improvement of teaching content generation, we generate customized content by combining with the adaptive learning schedule. Based on students' mastery of each knowledge point, the system dynamically generates content that matches their current learning level. In this process, the system adjusts the difficulty and depth of the generated content by evaluating the students' learning progress. Assuming that $P_i(t)$ represents the state of students' learning progress at time t, the output $C(t)$ of content generation can be expressed by the following formula:

$$C(t) = f\big(P_i(t), F(t), G(t)\big) \tag{5}$$

Among them, $P_i(t)$ is the students' mastery of the ith knowledge point at time t, $F(t)$ is the students' immediate feedback, and $G(t)$ is the generation strategy of the current teaching objectives and content. In this way, the model can adjust the depth, breadth and form of the content in real time, so as to achieve personalized teaching content generation.

In addition, to address students' comprehension difficulties during learning, the system can provide real-time correction and hints based on students' feedback. When students have comprehension deviations or learning difficulties at a certain knowledge point, the system automatically identifies these and provides corresponding error correction suggestions through the model. This process involves the immediate recognition of students' errors and the generation of corrective text through the model to help students overcome learning obstacles. By formula:

$$E(t) = h\big(F(t), C(t)\big) \tag{6}$$

Where $E(t)$ is the error correction information generated by the system according to the students' feedback, $F(t)$ is the students' real-time feedback, and $C(t)$ is the current teaching content. The model can be updated in time and generate appropriate error correction prompts after each interaction.

For the improvement of the adaptive learning path planning algorithm, the system accurately plans students' learning paths by analyzing their learning behavior. By modeling students' historical learning data, the system can predict potential learning bottlenecks and adjust subsequent learning content and paths based on each student's learning situation. The optimization of the learning path planning model relies on a combination of reinforcement learning and collaborative filtering algorithms, particularly using reinforcement learning to optimize behavioral strategies and thus improve student learning outcomes. The reward function of reinforcement learning is designed as follows:

$$R(s_t, a_t) = \gamma \cdot V(s_{t+1}) - V(s_t) \tag{7}$$

Where $R(s_t, a_t)$ is the reward after taking action a_t in state s_t, $V(s_t)$ and $V(s_{t+1})$ represent the value estimation of the current state and the next state respectively, and γ is the discount factor. Through this reward function, the model can optimize behavior strategies according to students' learning progress, so that each learning path can maximize students' learning benefits.

The algorithm improvement for multimodal learning support enables the system not only to support text and voice input but also to enhance interactivity and adaptability through cross-modal data processing. By combining speech recognition and text generation technologies, the model enables students to communicate naturally with the system via speech and enhances visual interaction through image processing technology. The key to cross-modal data processing is effectively integrating different types of data to provide richer feedback and teaching content. Assuming that $\mathrm{X}_{\mathrm{audio}}$ and $\mathrm{X}_{\mathrm{text}}$ are the features of voice and text input respectively, the system completes data fusion by generating a cross modal joint representation of $\mathrm{X}_{\mathrm{joint}}$:

$$\mathrm{X}_{\mathrm{joint}} = \alpha \cdot \mathrm{X}_{\mathrm{audio}} + (1 - \alpha) \cdot \mathrm{X}_{\mathrm{text}} \tag{8}$$

Where α is the parameter to adjust the weight of voice and text input. Through this method, the system can intelligently select the most suitable teaching feedback mode based on the student's input type, further improving system interactivity and adaptability.

Through the improvement of the above algorithm, our system not only realizes the adaptation and personalization in the generation of teaching content, but also greatly improves the teaching effect and interactive experience by optimizing the learning path and cross modal data processing technology.

4. SYSTEM IMPLEMENTATION AND KEY TECHNOLOGIES

In the implementation of a real-time adaptive teaching digital human system based on a large language model, the selection of the technical framework and platform, the application of key technologies, and system performance optimization are all crucial aspects. This paper will introduce the technology selection, implementation platform, system integration and deployment scheme of the system in detail, and deeply discuss the real-time feedback and adaptive mechanism of the system, the implementation details of virtual teaching digital human, and the performance optimization and scalability design of the system.

4.1 Technical framework and implementation platform

To ensure the efficiency, scalability, and stability of the system, we adopt a modern technology stack and high-performance frameworks. The technology selection of the system mainly involves programming language, framework, platform and other aspects.

First, the choice of programming language is key to system implementation. We chose Python as the main development language because it has rich deep learning frameworks (such as TensorFlow and PyTorch) and natural language processing libraries (such as Hugging Face Transformers), which can effectively support the training, fine-tuning and reasoning of large language models. At the same time, Python has high readability and flexibility, which is very suitable for rapid development and prototype verification.

Second, for framework selection, we use TensorFlow and PyTorch as the main deep learning frameworks. TensorFlow provides highly optimized computational graph support, which is suitable for the deployment of large-scale models. Due to its dynamic graph characteristics, PyTorch can carry out experiments and iterations quickly, which is convenient for model fine-tuning and experimental verification. In practical applications, TensorFlow is used for model reasoning and deployment, while PyTorch is mainly used for training and fine-tuning models.

For the platform, the system is deployed on cloud platforms such as AWS and Google Cloud. The selection of cloud platform can provide flexible computing resources and elastic expansion ability, especially the support of GPU and TPU, which can greatly improve the efficiency of training and reasoning. In terms of data storage, we use Amazon S3 and Google Cloud Storage to store a large number of training data and model parameters.

Figure 1 illustrates the overall architecture of the real-time adaptive teaching digital avatar system based on a large language model.

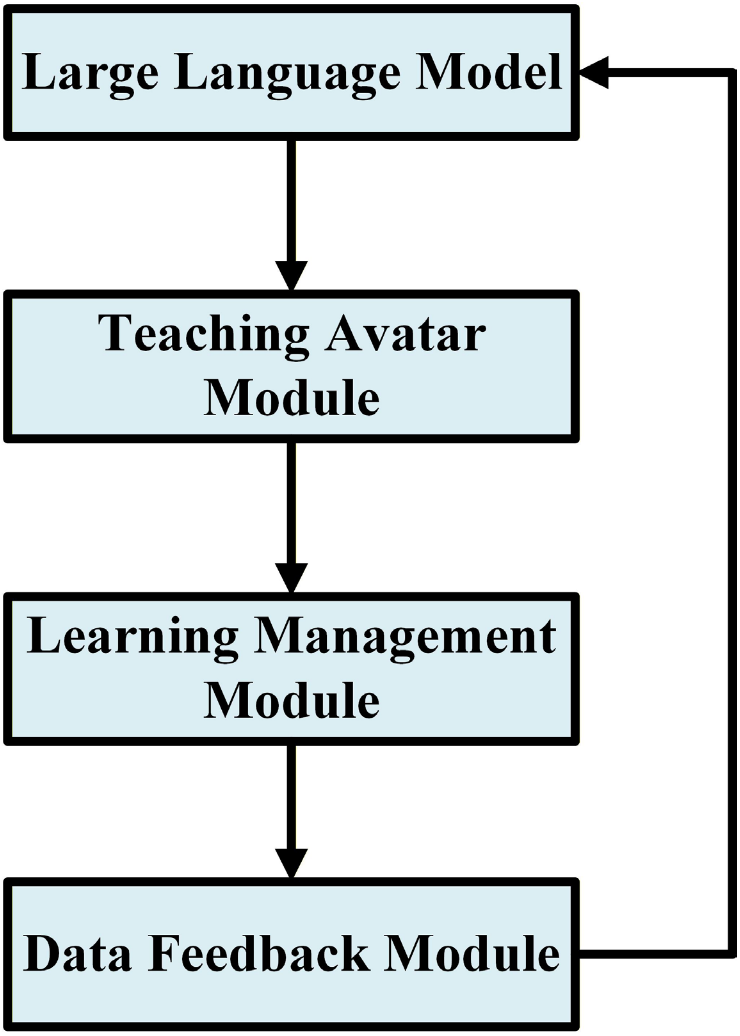

Figure 1. System architecture diagram

The system consists of four main modules: the Large Language Model, the Teaching Avatar Module, the Learning Management Module, and the Data Feedback Module. The Large Language Model serves as the core of the system, generating personalized teaching content based on the student's learning progress. The Teaching Avatar Module interacts with the students through a virtual avatar, providing real-time feedback and emotional expressions, enhancing the interactivity and engagement of the learning process. The Learning Management Module tracks the student's progress and assesses knowledge acquisition, dynamically adjusting the content based on these data. The Data Feedback Module collects student behavior and feedback, continuously optimizing the outputs of the large language model to ensure the content's adaptability and real-time response.

4.2 Implementation of real-time feedback and adaptive mechanism

One of the core components of this system is the real-time feedback and adaptive mechanism, which dynamically adjusts teaching content and progress based on the student's learning state. The real-time monitoring and evaluation technology of students' learning status obtains real-time learning information through various sensors and data acquisition devices (such as learning behavior tracking, question-answering records, and speech recognition). These information will be transmitted to the background system to generate students' learning status data through analysis and calculation.

Specifically, the system will dynamically evaluate students' learning status $S(t)$ according to the following formula:

$$S(t) = \alpha_1 \cdot \text{accuracy}(t) + \alpha_2 \cdot \text{engagement}(t) + \alpha_3 \cdot \text{time}_{\text{spent}(t)} \quad (9)$$

Among them, $\text{accuracy}(t)$ is the accuracy rate of students' answers at the current stage, $\text{engagement}(t)$ is students' participation (such as interaction frequency, learning depth), $\text{time}_{\text{spent}(t)}$ is the time students spend on the learning module, $\alpha_1, \alpha_2, \alpha_3$ are weight coefficients, which are adjusted based on different learning objectives.

Using this formula, the system comprehensively evaluates students' learning status and adjusts learning content based on the evaluation results. The technical implementation of dynamically adjusting the learning schedule and content depends on the adaptive learning algorithm of the model. When students master some knowledge points well, the system will automatically push more difficult content; On the contrary, when students have difficulty understanding a certain knowledge point, the system will provide more review materials or tips to ensure that students can obtain the most valuable learning content at the right time. Table 1 shows the assessment results of students' learning status, including learning progress, answer accuracy, participation, learning time and comprehensive learning status score.

Table 1. Assessment of students' learning status

Student ID	Learning progress (%)	Answer accuracy (%)	Participation (%)	Learning time (minutes)	Learning status (score)
001	80	90	75	50	0.85
002	60	70	60	40	0.72
003	40	60	50	30	0.65

The system can dynamically evaluate the learning status of each student and decide whether to adjust the learning content or progress. Student 001's learning status score is 0.85, indicating that his/her learning progress, accuracy and participation are relatively high, indicating that he/she has a good grasp of the learning content. The score of student 002 was 0.72. Although the learning progress and accuracy rate were reduced, it was still at a good level. In contrast, student 003's score is 0.65, indicating that he/she has some difficulties in the learning process and needs more teaching support and review materials.

4.3 Implementation details of virtual teaching digital human

The design and implementation of the virtual teaching digital human constitute another innovation of this system. Through speech synthesis and speech recognition technologies, the system enables digital humans to communicate with students in natural language. Speech recognition technology uses a deep learning model to accurately recognize students' speech input and convert it into text for processing. Speech synthesis technology uses the latest neural network models (such as Tacotron 2 and WaveGlow) to generate natural and fluent speech feedback.

The generation of character images and dynamic expression rendering is another key part of the virtual teaching digital human. We use high-quality 3D modeling technology to generate digital human image, and combine face recognition and dynamic expression synthesis technology, so that the expression of digital human can reflect students' emotions and learning status in real time. Expression generation is based on the following formula:

$$E(t) = f\big(\text{engagement}(t),\text{emotion}(t)\big) \tag{10}$$

Among them, $E(t)$ represents the expression of the digital person at time t, $\text{engagement}(t)$ is the participation of students, $\text{emotion}(t)$ is the emotional state generated according to students' voice or text input.

Figure 2 demonstrates the emotional expression changes of the virtual teaching avatar based on the student's emotional state.

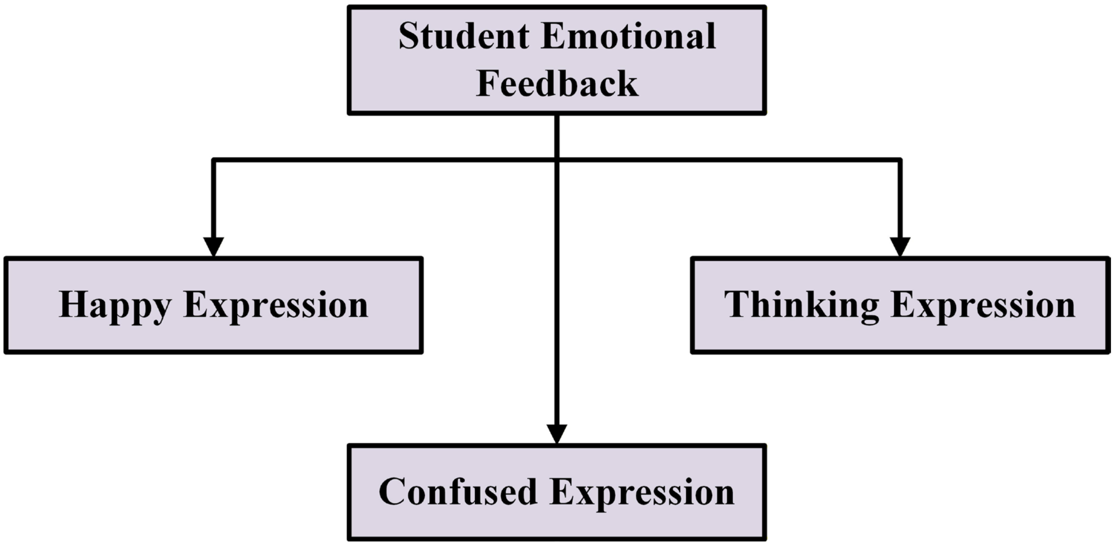

Figure 2. Virtual teaching avatar expression change diagram

As shown in the diagram, the avatar’s expressions change according to the student's emotions, such as smiling when the student feels happy or showing a thoughtful expression when the student is confused. This emotional feedback mechanism not only enhances the immersive learning experience but also provides emotional support to the student, improving their overall learning experience. By adjusting the avatar's expression according to the student's emotions, the system fosters a stronger emotional connection, thus increasing student motivation and engagement throughout the learning process.

4.4 Performance optimization and scalability design

As the system is continuously applied and the number of users increases, its performance and scalability become particularly important. For performance optimization, first, the system has greatly improved response speed and real-time performance by using asynchronous

https://doi.org/10.71451/ISTAER2620

processing and distributed computing technologies. In large-scale parallel computing, the system uses the elastic computing resources of the cloud platform to realize the automatic scheduling and load balancing of computing tasks. Specifically, the system optimizes the calculation load through the following formula:

$$\mathrm{Load}(t) = \sum_{i=1}^{N} \frac{\mathrm{task}_i}{\mathrm{available\_resources}_i} \tag{11}$$

Where $\mathrm{Load}(t)$ is the computing load at time t, task_i represents the ith computing task, $\mathrm{available\_resources}_i$ represents the available computing resources. Table 2 shows the comparison of the system before and after performance optimization, mainly including three performance indicators: response time, system throughput and maximum concurrency.

Table 2. Comparison of system performance before and after optimization

Before optimization	After optimization	Improvement range
Response time (seconds)	5.2	1.3
System throughput (times/second)	80	200
Maximum concurrent number	50	200

After optimization, the response speed and processing capacity of the system have been significantly improved. The response time is reduced from 5.2 seconds to 1.3 seconds, the system throughput was increased from 80 times/second to 200 times/second, and the maximum number of concurrent operations was also increased from 50 to 200. These improvements demonstrate that the adaptability of the system in high concurrency scenarios is greatly enhanced, and it can support more user access and interaction at the same time.

5. EXPERIMENTAL DESIGN AND VERIFICATION

To verify the effectiveness of the real-time adaptive teaching digital human system based on a large language model, this paper designs several experiments covering teaching effectiveness, system performance, and user experience. The experimental design involves detailed environmental settings, data set selection, comparative experimental design of baseline model, and quantitative analysis of different evaluation indicators.

5.1 Experimental setup

For experimental environment settings, we use two different setups: a local environment and a cloud platform environment. Specifically, the local environment is mainly used for small-scale experiments and preliminary tests, equipped with high-performance GPU (NVIDIA A100) for training and testing models; The cloud platform environment (AWS) provides flexible computing resources, supports larger scale experiments and parallel computing, and can simulate the actual scenarios of highly concurrent users.

Dataset selection: To evaluate system performance, we designed two core datasets: a student behavior dataset and a teaching content dataset. The student behavior data set contains the student's learning behavior information, such as each student's learning time, correct answer rate, interaction frequency, learning progress, etc. The teaching content data set includes teaching materials, exercises, and learning path information in different disciplines.

Comparative experiment design and baseline model selection: in the comparative experiment, we selected two baseline models for comparison. The first baseline model is the traditional teaching system, where students and teachers interact directly. The teaching content is static and cannot be adjusted based on students' learning progress and feedback. The second baseline model is the existing adaptive learning system, which can adjust the content according to students' learning state, but does not integrate the real-time virtual teaching digital human. Therefore, these two baseline models provide a reference framework for our system and help us verify the advantages of the large language model adaptive teaching digital human system.

Figure 3 illustrates the experimental setup and data flow process within the system.

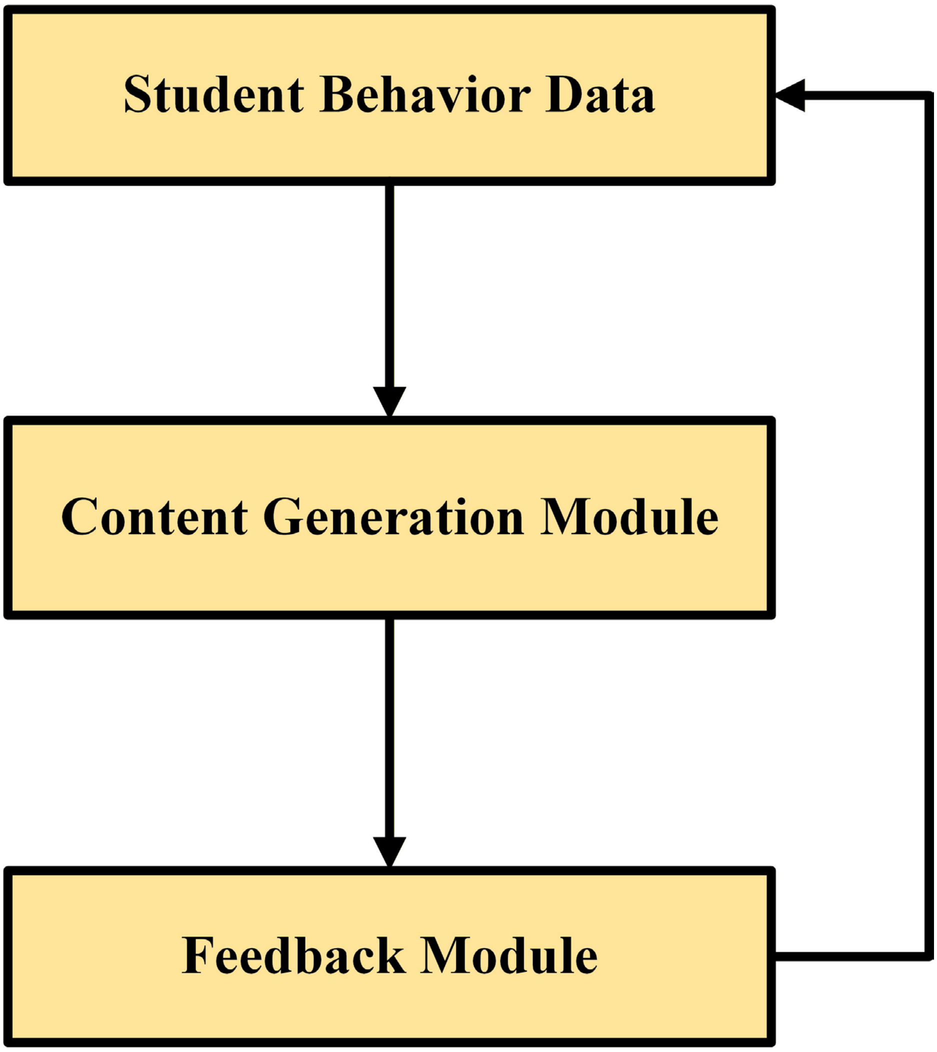

Figure 3. Experimental environment and data flow diagram

In the experimental environment, student behavior data is collected in real-time and transmitted to the Content Generation Module, which generates personalized learning content based on the student's feedback and progress. Simultaneously, the student’s learning data is sent to the Feedback Module, which analyzes the student's understanding depth and engagement, feeding the results back into the Content Generation Module to further refine and adapt the teaching content. This continuous data flow ensures that the system can dynamically adjust teaching strategies based on the student’s real-time performance, providing an adaptive and personalized learning experience.

5.2 Experimental evaluation index

To comprehensively evaluate system performance, we establish several experimental evaluation metrics, covering teaching effectiveness, system performance, and user experience.

Teaching effectiveness evaluation: We use students' knowledge mastery rate, depth of understanding, and learning interest as the main metrics. Specifically, the knowledge mastery rate is measured by students' correct answer rate at each knowledge point, the depth of understanding is measured by students' understanding of complex concepts, and the learning interest is evaluated by students' interaction frequency and learning duration.

The calculation formula of knowledge mastery rate is as follows:

$$\text{Mastery Rate}(t) = \frac{\text{Correct Answers}(t)}{\text{Total Questions}(t)} \tag{12}$$

Where $\text{Correct Answers}(t)$ is the number of correct answers given by students at time t, and $\text{Total Questions}(t)$ is the total number of questions completed by students at that time.

To ensure system efficiency and stability, we define three key performance metrics: response speed, system stability, and real-time performance. The response speed mainly measures the time required by the system from receiving student input to feedback output; The system stability is evaluated by monitoring the number of crashes and error rate of the system; Real time performance is evaluated by measuring the ability of the system to process requests in high concurrency scenarios. User experience assessment is quantified by interactivity, participation and satisfaction. Interactivity mainly assesses the fluency of students' interaction with the system, the sense of participation assesses students' participation in the learning process, and satisfaction is quantified through questionnaires and user feedback. Table 3 lists the experimental indexes and calculation methods used to evaluate the system performance. These indicators cover three aspects: teaching effect, system performance and user experience.

Table 3. Description of experimental evaluation metrics

Indicator name	Computing method	Explain
Knowledge mastery rate	$\frac{\text{Correct Answers}}{\text{Total Questions}}$	Measuring students' mastery of teaching content
Learning interest	Interaction frequency+Learning duration	Assess students' interest and focus through student engagement
Response speed	Total response time/number of requests	Average response time of the system to each request
System stability	Errors/total requests	Stability of the system under high load
User satisfaction	Based on the survey results	Students' overall satisfaction with teaching digital people

Knowledge mastery rate and learning interest are the key indicators to evaluate students' learning effect, which can reflect students' mastery and participation in the content. Response speed and system stability are important parameters to evaluate the technical performance of the system, which affect students' learning experience. As the final evaluation standard, user satisfaction reflects students' overall satisfaction with teaching digital people.

5.3 Experimental results and analysis

Figure 4 compares the traditional teaching system, the existing adaptive learning system, and the real-time adaptive teaching digital human system based on a large language model in terms of students' knowledge mastery rate, understanding depth, and interest.

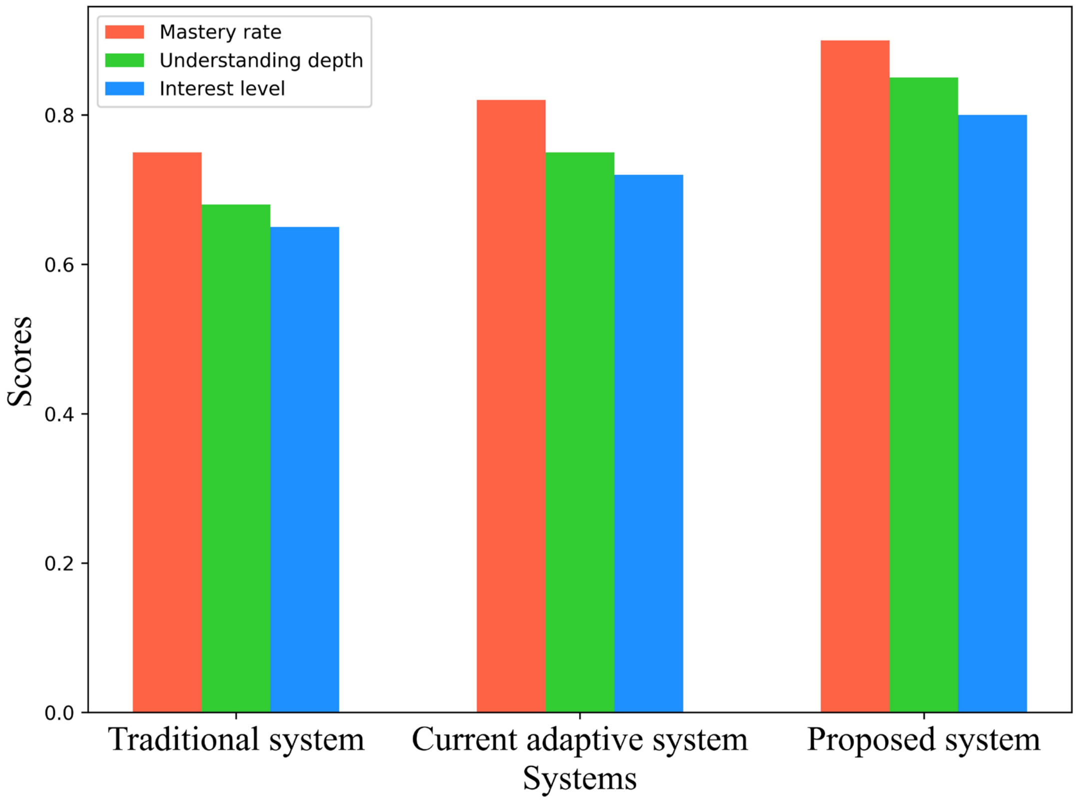

Figure 4. Teaching effect comparison

As shown in Figure 4, the three teaching approaches exhibit significant differences in these three metrics, especially in improving students' knowledge and learning interest. Our proposed system based on a large language model shows significant advantages. In the comparison of teaching effect, the system based on the large language model shows significant advantages in the knowledge mastery rate, understanding depth and interest dimension. The students' knowledge mastery rate of the traditional teaching system is 75%, and the existing adaptive system is increased to 82%, while this system is up to 90%, which shows the effectiveness of personalized learning path. In terms of the depth of understanding, the traditional system is 68%, the existing system is 75%, and the current system is 85%, which proves the advantages of real-time feedback and dynamic content adjustment. The interest dimension of students has also been significantly improved, from 65% of the traditional system to 72% of the existing system, and finally the system reaches 80%, which further demonstrates that the system effectively improves student engagement and learning interest through interactive feedback and emotional regulation.

Table 4 shows the comparison of response speed, system stability and real-time performance among traditional teaching system, existing adaptive learning system and system based on large language model.

Table 4. System performance evaluation

System version	Response speed (seconds)	System stability (crash rate)	Real time (request processing speed)
Traditional teaching system	5.0	0.05	Low
Existing adaptive learning system	3.2	0.02	In
System based on large language model	1.5	0.01	High

The performance of the system based on a large language model is significantly better than that of other systems in terms of response speed, system stability, and real-time performance. The significant reduction in response time enables the system to process student requests more quickly, reduce delays, and improve user experience. The improvement of system stability means that the system is more stable under high concurrency and is not prone to collapse or errors. The improvement of real-time performance ensures that the system can efficiently handle a large number of concurrent requests, and can still provide fast and smooth feedback when large-scale users visit.

6. CONCLUSION AND PROSPECT

6.1 Research summary

This paper proposes a real-time adaptive teaching digital human system based on a large language model and studies its design principles, core technologies, implementation scheme, and experimental verification in detail. The main contribution of this paper is the creative combination of a large language model with a virtual teaching digital human to build an adaptive teaching platform that can dynamically adjust content and progress based on students' learning states state. By introducing a real-time feedback mechanism, the system can provide customized learning content tailored to students' personalized needs, thereby improving learning outcomes and engagement. At the same time, this research also made positive improvements in system performance, using cloud computing resources and parallel computing technology, significantly improved the response speed and processing capacity of the system, and ensured the stability and fluency in high concurrency scenarios. In addition, the virtual teaching digital human of the system can enhance the interaction between students and learning content through speech synthesis and emotional expression, and improve the immersion and interest of learning.

Although the system has achieved some results in design and implementation, certain limitations remain. First, while the system can adjust teaching content based on students' learning progress, there is still room for improvement in its adaptability to highly personalized learning needs. Second, although the virtual teaching digital human can provide real-time feedback on students' emotional states, there are still limitations in emotional expression and interaction, especially regarding complex emotional changes and emotional understanding in cross-cultural contexts. In addition, the deployment and operation of the system require substantial computing resources, which may affect its popularity and applicability in resource-constrained settings. Finally, although we have adopted various evaluation metrics to test system effectiveness, further optimizing and adjusting the system for different educational contexts and student groups remains an urgent problem to be solved.

6.2 Future work outlook

In the future, with continuous technological progress and expanding application scenarios, the adaptive teaching digital human system based on a large language model will have ample room for optimization and upgrading. First, for further system optimization, we can enhance its ability to understand and generate content across different disciplines and complex knowledge points by improving training data quality and the generalization ability of the large language model. In addition, by strengthening integration with other emerging technologies such as deep learning and reinforcement learning, the system's intelligence level can be further improved. For example, deep learning can be used for more accurate recognition and generation of images and speech, while reinforcement learning can help the system better adapt to students' learning behavior and carry out dynamic strategy optimization in real-time feedback.

The combination of large language models with other emerging technologies holds broad prospects. The combination of deep learning and natural language processing can enhance the system's ability to understand students' emotions, solve complex problems, and generate diversified teaching content. With the introduction of reinforcement learning, the system can not only adjust the content according to the historical performance of students, but also continuously optimize the teaching strategy through autonomous learning and improve the intelligent level of the system. With the improvement of computing power and the continuous optimization of the model, the combination of these technologies in the future is expected to make the teaching system more flexible and efficient, and provide students with a more personalized learning experience.

In addition to its application in the field of education, the application potential of real-time adaptive teaching digital human system based on large language model in other fields also deserves attention. Especially in the field of medical treatment and psychological counseling, the introduction of virtual assistant can provide personalized health advice, psychological counseling and emotional support for patients. Through the real-time monitoring and analysis of patients' behavior and emotion, the system can adjust the feedback and suggestions in time to help patients with self-management and treatment. In addition, the system can also be applied to vocational training, customer service and other fields to provide customized learning and service experience by simulating real interactive scenes. With the continuous maturity of technology and the expansion of cross domain applications, the virtual assistant based on large language model will play an increasingly important role in many industries.

In summary, the real-time adaptive teaching digital human system based on a large language model demonstrates great potential in personalized education and interactive learning, providing new ideas for the innovation of future educational models. With the continuous evolution of technology, the system will continue to improve and expand, and has broad application prospects in education and other fields, which is worthy of further exploration and research.

Abbreviations

LLM, Large Language Model;
GPT, Generative Pre-trained Transformer;
BERT, Bidirectional Encoder Representations from Transformers;
NLP, Natural Language Processing;
CV, Computer Vision;
DL, Deep Learning;
ASR, Automatic Speech Recognition;
TTS, Text-to-Speech;
GPU, Graphics Processing Unit;
TPU, Tensor Processing Unit;
AWS, Amazon Web Services;
S3, Simple Storage Service;

API, Application Programming Interface;
CNN, Convolutional Neural Network;
RNN, Recurrent Neural Network;
ReLU, Rectified Linear Unit;
AdamW, Adaptive Moment Estimation with Weight Decay;
MSE, Mean Squared Error;
MAE, Mean Absolute Error.

Supplementary Material

Not applicable.

Appendix

Not applicable.

Ethics approval and consent to participate.

This study did not involve human participants, animal subjects, or any data requiring ethical approval. Therefore, ethics approval and consent to participate are not applicable.

Acknowledgements

The authors would like to thank the editors of this journal and all the anonymous reviewers who provided valuable comments on this work.

Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

Author contributions

All authors have read and agreed to the published version of the manuscript. The author's contributions are specified as follows: **W.Y.:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing – Original draft, Writing – Review & Editing, Visualization, Supervision, Project administration.

Funding information

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability

The data that support the findings of this study are available upon request from the corresponding authors, **W.Y.**

Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. It does not necessarily reflect the official policy or position of any affiliated institution, funder, agency, or that of the publisher. The authors are responsible for this article's results, findings, and content.

Declaration of AI and AI-assisted Technologies in the Writing Process

During the writing of this article, the author used DeepSeek for spelling and grammar checking. After using this tool, the author reviewed and edited the content as needed and assumes full responsibility for the final published content.

REFERENCES

[1] Martin, A. D. (2022). Teacher identity perspectives. In *Encyclopedia of teacher education* (pp. 1856-1860). Singapore: Springer Nature Singapore. DOI: https://doi.org/10.1007/978-981-13-1179-6_222-1

[2] Tetzlaff, L., Schmiedek, F., & Brod, G. (2021). Developing personalized education: A dynamic framework. *Educational psychology review*, *33*(3), 863-882. DOI: https://doi.org/10.1007/s10648-020-09570-w

[3] Dumont, H., & Ready, D. D. (2023). On the promise of personalized learning for educational equity. *Npj science of learning*, *8*(1), 26. DOI: https://doi.org/10.1038/s41539-023-00174-x

[4] Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, *103*, 102274. DOI: https://doi.org/10.35542/osf.io/5er8f

[5] Wang, S., Xu, T., Li, H., Zhang, C., Liang, J., Tang, J., ... & Wen, Q. (2026). Large language models for education: A survey and outlook. *IEEE Signal Processing Magazine*, *42*(6), 51-63. DOI: https://doi.org/10.48550/arXiv.2403.18105

[6] Alqahtani, T., Badreldin, H. A., Alrashed, M., Alshaya, A. I., Alghamdi, S. S., Bin Saleh, K., ... & Albekairy, A. M. (2023). The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research. *Research in social and administrative pharmacy*, *19*(8), 1236-1242. DOI: https://doi.org/10.1016/j.sapharm.2023.05.016

[7] Oprea, S. V., & Bâra, A. (2025). Transforming Education with Large Language Models. Trends, Themes and Untapped Potential. *IEEE Access*. DOI: https://doi.org/10.1109/access.2025.3570649

[8] Chen, E., Heritage, M., & Lee, J. (2024). Identifying and monitoring students' learning needs with technology. In *Transforming Data Into Knowledge* (pp. 309-332). Routledge. DOI: https://doi.org/10.1207/s15327671espr1003_6

[9] Gm, D., Goudar, R. H., Kulkarni, A. A., Rathod, V. N., & Hukkeri, G. S. (2024). A digital recommendation system for personalized learning to enhance online education: A review. *IEEe Access*, *12*, 34019-34041. DOI: https://doi.org/10.1109/access.2024.3369901

[10] Almufarreh, A., & Arshad, M. (2023). Promising emerging technologies for teaching and learning: Recent developments and future challenges. *Sustainability*, *15*(8), 6917. DOI: **https://doi.org/10.3390/su15086917**

[11] Audras, D., Zhao, A., Isgar, C., & Tang, Y. (2022). Virtual teaching assistants: A survey of a novel teaching technology. *International Journal of Chinese Education*, *11*(2), 2212585X221121674. DOI: **https://doi.org/10.1177/2212585x221121674**

[12] Zhou, Y., Xu, K., Yin, B., & Liu, N. (2024, June). Research on the application of digital humans in english oral teaching based on AI models. In *Proceedings of the 2024 9th International Conference on Distance Education and Learning* (pp. 49-56). DOI: **https://doi.org/10.1145/3675812.3675820**

[13] Aly, M. (2025). Revolutionizing online education: Advanced facial expression recognition for real-time student progress tracking via deep learning model. *Multimedia Tools and Applications*, *84*(13), 12575-12614. DOI: **https://doi.org/10.1007/s11042-024-19392-5**

[14] Vistorte, A. O. R., Deroncele-Acosta, A., Ayala, J. L. M., Barrasa, A., López-Granero, C., & Martí-González, M. (2024). Integrating artificial intelligence to assess emotions in learning environments: a systematic literature review. *Frontiers in psychology*, *15*, 1387089. DOI: **https://doi.org/10.3389/fpsyg.2024.1387089**

[15] Pabba, C., & Kumar, P. (2022). An intelligent system for monitoring students' engagement in large classroom teaching through facial expression recognition. *Expert Systems*, *39*(1), e12839. DOI: **https://doi.org/10.1111/exsy.12839**

[16] Dahleez, K. A., El-Saleh, A. A., Al Alawi, A. M., & Abdelmuniem Abdelfattah, F. (2021). Higher education student engagement in times of pandemic: the role of e-learning system usability and teacher behavior. *International Journal of Educational Management*, *35*(6), 1312-1329. DOI: **https://doi.org/10.1108/ijem-04-2021-0120**

[17] Kong, K., Isleem, H. F., Aluvalu, R., Tejani, G. G., & Metwally, A. S. M. (2025). Real-time cognitive and emotional state tracking in intelligent tutoring systems for enhanced learning outcomes. *Journal of Big Data*, *12*(1), 266. DOI: **https://doi.org/10.1186/s40537-025-01333-0**

[18] Zhou, Z. (2026). Artificial Intelligence and Student Engagement in Online Learning: A Literature Review. *American Journal of Distance Education*, *40*(1), 8-30. DOI: **https://doi.org/10.1080/08923647.2025.2594282**

[19] Song, Y., & Xiong, W. (2025). Large language model-driven 3D hyper-realistic interactive intelligent digital human system. *Sensors*, *25*(6), 1855. DOI: **https://doi.org/10.3390/s25061855**

[20] Llanes-Jurado, J., Gómez-Zaragozá, L., Minissi, M. E., Alcañiz, M., & Marín-Morales, J. (2024). Developing conversational virtual humans for social emotion elicitation based on large language models. *Expert Systems with Applications*, *246*, 123261. DOI: **https://doi.org/10.1016/j.eswa.2024.123261**

[21] Alqahtani, T., Badreldin, H. A., Alrashed, M., Alshaya, A. I., Alghamdi, S. S., Bin Saleh, K., ... & Albekairy, A. M. (2023). The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research. *Research in social and administrative pharmacy*, *19*(8), 1236-1242. DOI: **https://doi.org/10.1016/j.sapharm.2023.05.016**

[22] Song, C., Shin, S. Y., & Shin, K. S. (2024). Implementing the dynamic feedback-driven learning optimization framework: a machine learning approach to personalize educational pathways. *Applied Sciences*, *14*(2), 916. DOI: **https://doi.org/10.3390/app14020916**

[23] Ruan, S., & Lu, K. (2025). Adaptive deep reinforcement learning for personalized learning

pathways: A multimodal data-driven approach with real-time feedback optimization. *Computers and Education: Artificial Intelligence*, 100463. DOI: **https://doi.org/10.1016/j.caeai.2025.100463**

[24] Yu, M. (2024, September). Application of an Artificial Intelligence-based adaptive learning system to chinese language education in universities. In *Proceedings of the 2024 International Symposium on Artificial Intelligence for Education* (pp. 416-421). DOI: **https://doi.org/10.1145/3700297.3700368**

[25] Tian, X. (2024). Personalized translator training in the era of digital intelligence: Opportunities, challenges, and prospects. *Heliyon*, *10*(20). DOI: **https://doi.org/10.1016/j.heliyon.2024.e39354**

[26] Sajja, R., Sermet, Y., Cikmaz, M., Cwiertny, D., & Demir, I. (2024). Artificial intelligence-enabled intelligent assistant for personalized and adaptive learning in higher education. *Information*, *15*(10), 596. DOI: **https://doi.org/10.3390/info15100596**

[27] Raiaan, M. A. K., Mukta, M. S. H., Fatema, K., Fahad, N. M., Sakib, S., Mim, M. M. J., ... & Azam, S. (2024). A review on large language models: Architectures, applications, taxonomies, open issues and challenges. *IEEE access*, *12*, 26839-26874. DOI: **https://doi.org/10.1109/access.2024.3365742**

[28] Myers, D., Mohawesh, R., Chellaboina, V. I., Sathvik, A. L., Venkatesh, P., Ho, Y. H., ... & Jararweh, Y. (2024). Foundation and large language models: fundamentals, challenges, opportunities, and social impacts. *Cluster Computing*, *27*(1), 1-26. DOI: **https://doi.org/10.1007/s10586-023-04203-7**

[29] Bharathi Mohan, G., Prasanna Kumar, R., Vishal Krishh, P., Keerthinathan, A., Lavanya, G., Meghana, M. K. U., ... & Doss, S. (2024). An analysis of large language models: their impact and potential applications. *Knowledge and Information Systems*, *66*(9), 5047-5070. DOI: **https://doi.org/10.1007/s10115-024-02120-8**

[30] Annepaka, Y., & Pakray, P. (2025). Large language models: a survey of their development, capabilities, and applications. *Knowledge and Information Systems*, *67*(3), 2967-3022. DOI: **https://doi.org/10.1007/s10115-024-02310-4**

[31] Grigoriev, V. V., Loginov, E. L., & Loginova, V. E. (2022, May). Agent-based human interaction with the digital educational environment, based on a set of digital intelligent agents. In *Computer Applications for Management and Sustainable Development of Production and Industry (CMSD2021)* (Vol. 12251, pp. 120-125). SPIE. DOI: **https://doi.org/10.1117/12.2631456**

[32] Tan, S. C., Chan, C., Bielaczyc, K., Ma, L., Scardamalia, M., & Bereiter, C. (2021). Knowledge building: Aligning education with needs for knowledge creation in the digital age. *Educational Technology Research and Development*, *69*(4), 2243-2266. DOI: **https://doi.org/10.1007/s11423-020-09914-x**

[33] Yu, X., & Tian, Y. (2025). Enhancing Computer Education through IoT-Enabled Learning Environments Leveraging Mobile Edge Computing for Real-Time Feedback. *Systems and Soft Computing*, 200433. DOI: **https://doi.org/10.1016/j.sasc.2025.200433**

[34] Aly, M. (2025). Revolutionizing online education: Advanced facial expression recognition for real-time student progress tracking via deep learning model. *Multimedia Tools and Applications*, *84*(13), 12575-12614. DOI: **https://doi.org/10.1007/s11042-024-19392-5**

[35] Li, X., Henriksson, A., Duneld, M., Nouri, J., & Wu, Y. (2023). Evaluating embeddings from pre-trained language models and knowledge graphs for educational content recommendation. *Future Internet*, *16*(1), 12. DOI: **https://doi.org/10.3390/fi16010012**

[36] Yu, L. (2026). Research on the Construction and Reasoning of Educational Knowledge

Graph Based on Pre-trained Language Model. *Human-Centric Intelligent Systems*, 1-18. DOI: **https://doi.org/10.1007/s44230-026-00139-4**

[37] Li, Y., Liang, Y., Yang, R., Qiu, J., Zhang, C., & Zhang, X. (2024). CourseKG: an educational knowledge graph based on course information for precision teaching. *Applied Sciences*, *14*(7), 2710. DOI: **https://doi.org/10.3390/app14072710**