Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 20 No. 16 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijim.v20i16.62750

PAPER

Adaptive Teaching Path Modeling and Dynamic Intervention Mechanisms for Higher Education Driven by Interactive Mobile Technologies

Jianjiang Wang, Yangyang Cui, Lingling Sun (✉)

Hebei Minzu Normal University, Chengde, China

sunlingling0378@hbun.edu.cn

ABSTRACT

The proliferation of mobile intelligent terminals and ubiquitous networks has established mobile learning as a core modality within the digital–intelligent higher education ecosystem. Conventional adaptive teaching systems, which rely on centralized cloud architectures, suffer from limited contextual awareness dimensionality, delayed intervention responses, and pronounced privacy risks—deficiencies that render them ill-suited for the fragmented and dynamically fluctuating contexts characteristic of mobile learning. To address these challenges, a three-tier adaptive teaching intervention framework was constructed, integrating on-device perception, edge-side inference, and cloud-based optimization. A gated adaptive multi-modal context fusion mechanism was designed to integrate multiple contextual dimensions and cognitive states into a unified joint representation. A lightweight efficiency prediction network was trained via knowledge distillation, enabling real-time local path fine-tuning on mobile terminals through constrained optimization, while raw data remained entirely on the device. An edge-empowered end–edge–cloud collaborative intervention strategy was proposed, wherein structured pruning facilitated low-latency inference at the edge layer, and a heterogeneity-aware personalized federated learning algorithm was introduced to perform global distributed optimization. A dual-timescale closed-loop feedback mechanism was established to synchronously enable instantaneous policy iteration and long-term knowledge system refinement. This study achieves a deep integration of mobile terminal sensing capabilities with adaptive teaching logic, reconciling real-time adaptability, global optimization, and privacy preservation, thereby providing technical support for the scalable deployment of digital–intelligent learning ecosystems in higher education.

KEYWORDS

mobile adaptive learning, multi-modal context awareness, edge computing, federated learning, dynamic learning path intervention, end–edge–cloud collaboration

Wang, J., Cui, Y., Sun, L. (2026). Adaptive Teaching Path Modeling and Dynamic Intervention Mechanisms for Higher Education Driven by Interactive Mobile Technologies. *International Journal of Interactive Mobile Technologies (iJIM)*, 20(16), pp. 118–133. https://doi.org/10.3991/ijim.v20i16.62750

Article submitted 2026-05-22. Revision uploaded 2026-06-20. Final acceptance 2026-06-21.

1 INTRODUCTION

Within the ongoing digital–intelligent transformation of higher education, mobile intelligent terminals—enabled by ubiquitous networking and multi-modal sensing capabilities [1, 2]—have become the primary vehicles supporting ubiquitous and fragmented learning. With the continued advancement of 5G universal coverage and on-device artificial intelligence computing power [3], a technical foundation has been established for adaptive learning grounded in real-time contextual awareness. Conventional adaptive teaching systems, which predominantly adopt centralized cloud architectures [4, 5], rely heavily on remote computation for path planning and intervention decision-making [6]. Such systems are constrained not only by network fluctuations that induce elevated response latency [7], but also by limited contextual awareness dimensionality, which precludes effective adaptation to the complex and dynamically changing scenarios inherent in mobile learning. The construction of a full-chain dynamic intervention system, achieved through deep integration of mobile computing with educational pedagogy, constitutes a critical pathway for resolving current bottlenecks in mobile learning experience, and offers a reusable technical paradigm for the scalable deployment of digital–intelligent learning ecosystems in higher education.

Contemporary research on mobile adaptive learning remains constrained by four core limitations. At the contextual awareness level, existing path modeling predominantly relies on learners' cognitive states as the primary basis [8], while underutilizing the multi-dimensional contextual information retrievable via mobile terminals—including physical environment, device status, and interaction behaviors. Moreover, statically weighted feature fusion logic proves inadequate in the presence of sensor signal loss and noise perturbations [9], resulting in notable deficiencies in system robustness. At the architectural level, pure cloud-based solutions are precluded by transmission latency constraints from meeting millisecond-level path adjustment requirements [10], whereas pure on-device solutions are limited by computational capacity in supporting global complex path optimization [11]. Consequently, a layered collaborative architecture that balances real-time responsiveness with optimization depth has yet to be established. At the privacy and optimization level, centralized training paradigms necessitate the upload of learners' raw behavioral data, which conflicts with regulatory compliance requirements for personal information protection [12]; existing federated learning approaches predominantly employ generic aggregation strategies [13] without fully accounting for heterogeneity across different learning populations, thereby degrading the adaptation accuracy of global models in local contexts. At the feedback mechanism level, most existing studies adopt single-timescale optimization logics, which fail to simultaneously accommodate both real-time contextual adaptation in mobile scenarios and the dual objective of long-term learning efficacy improvement.

To address the aforementioned research limitations, a three-tier collaborative adaptive teaching path dynamic intervention framework is constructed in this work, integrating on-device perception, edge-side inference, and cloud-based optimization. The core academic contributions can be summarized in three aspects. First, a gated adaptive multi-modal context fusion and local path fine-tuning mechanism is proposed, through which a joint context–cognitive state representation space is established. A lightweight prediction network is trained via knowledge distillation,

enabling millisecond-level adaptive path adjustment locally on mobile terminals, while raw data remain entirely on devices to ensure privacy preservation. Second, an edge computing-empowered end–edge–cloud collaborative intervention architecture is designed, and a heterogeneity-aware personalized federated learning algorithm is introduced. Global optimization is accomplished solely through model gradient transmission, while simultaneously adapting to population differences across different edge nodes, thereby balancing global commonality with local specificity. Third, a dual-timescale closed-loop feedback mechanism is constructed, through which synergistic coordination is achieved between second-level policy iteration at the edge layer and long-period knowledge system optimization at the cloud layer. The effectiveness of the proposed solution is validated through multiple sets of controlled experiments across the dimensions of technical performance, learning efficacy, and privacy overhead.

2 OVERALL ARCHITECTURE OF THE END–EDGE–CLOUD COLLABORATIVE ADAPTIVE TEACHING INTERVENTION

A three-tier collaborative overall architecture for adaptive teaching path dynamic intervention is constructed in this work, integrating on-device perception, edge-side inference, and cloud-based optimization, through which full-chain adaptive regulation is achieved following the closed-loop logic of perception, modeling, decision-making, intervention, and feedback. The on-device perception layer is deployed on learners' terminal devices, undertaking the functions of multi-modal contextual data acquisition, local feature extraction, and path fine-tuning, while raw learning and sensor data remain entirely on terminals, with only anonymized feature vectors being output for participation in upper-layer computation. The edge-side inference layer is deployed at campus edge computing nodes, receiving anonymized features uploaded from multiple terminals, performing low-latency group intervention decision-making and local model iteration, and supporting cross-individual coordination and local–global information scheduling. The cloud-based optimization layer is deployed at cloud data centers, maintaining the global knowledge graph and global path planning model, aggregating model gradients uploaded from each edge node to accomplish global optimization, and simultaneously advancing the long-period evolutionary update of the knowledge system. A bidirectional data flow pathway is formed within the architecture: the forward data flow sequentially transmits anonymized features and model gradients from bottom to top, while the backward data flow distributes optimized model parameters and global path strategies from top to bottom. Two core technical mechanisms achieve deep synergy within the architecture: the data flow is fully connected from terminal acquisition to cloud optimization in a layer-by-layer manner; decision granularity progresses hierarchically from individual millisecond-level adaptation to local collaborative intervention and then to global long-period optimization; privacy protection forms full-chain coverage from raw data retention on terminals to gradient transmission at the transport layer. A unified architectural foundation is thereby provided for the detailed exposition of the two subsequent core technologies. Figure 1 presents the overall architecture and operational mechanism diagram of the end–edge–cloud collaborative adaptive teaching dynamic intervention.

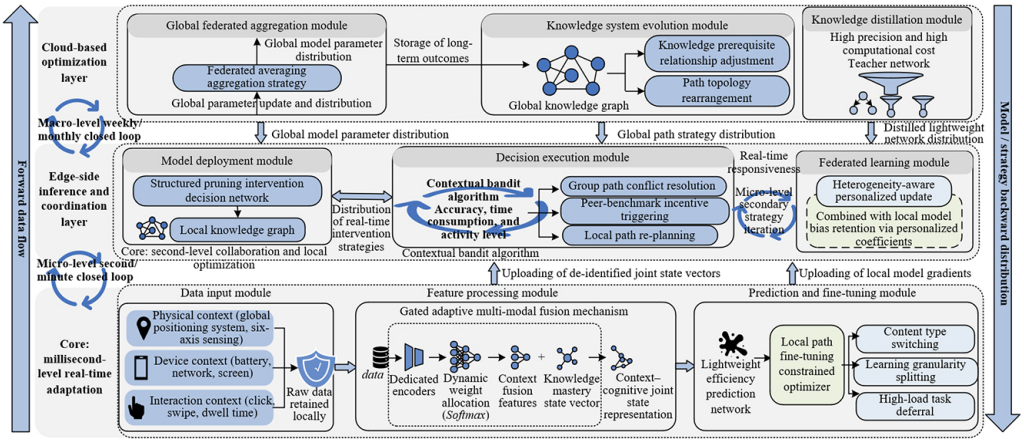

Fig. 1. Overall architecture and operational mechanism diagram of the end–edge–cloud collaborative adaptive teaching dynamic intervention

3 REAL-TIME ADAPTIVE PATH FINE-TUNING MECHANISM DRIVEN BY MOBILE MULTI-MODAL CONTEXTUAL AWARENESS

Multi-dimensional learning contextual data is synchronously acquirable by mobile terminals. The raw contextual data vector at time step t is defined as $S_t = [S_t^{phys}, S_t^{dev}, S_t^{inter}]$, where S_t^{phys} corresponds to the physical context dimension, encompassing global positioning system coordinates, timestamps, and six-axis motion sensing data; S_t^{dev} corresponds to the device context dimension, encompassing remaining battery level, network type, and screen hardware parameters; and S_t^{inter} corresponds to the interaction context dimension, encompassing four categories of behavioral statistics: click frequency, swipe velocity, page dwell time, and input interval. To address the robustness deficiency of statically weighted fusion caused by sensor signal fluctuations and frequent data missing in mobile scenarios, a gated adaptive multi-modal fusion mechanism is adopted to generate a unified contextual representation. Raw features of each modality are first mapped into a common feature space via a dedicated encoder network ϕ_m, and then a fused feature z_t is obtained through weighted summation with dynamic confidence weights, computed as:

$$Z_t = \sum_{m \in \{phys, dev, inter\}} \alpha_m^t \cdot \phi_m(S_t^m) \tag{1}$$

where, the modality confidence weight α_m^t is obtained by a lightweight multi-layer perceptron MLP_m that outputs an intermediate variable β_m^t based on the current modality data, which is then normalized via a SoftMax function:

$$\alpha_m^t = \frac{exp(\beta_m^t)}{\sum_k exp(\beta_k^t)}, \quad \beta_m^t = \mathrm{MLP}_m(S_t^m) \tag{2}$$

The weight values are dynamically adjusted according to the quality of single-modal data. Modalities with signal missing or higher noise levels automatically receive lower contribution proportions, thereby enabling adaptive accommodation of sensing data characteristics across different mobile scenarios.

Building upon the contextual feature representation, the learner's knowledge mastery state is introduced to fully characterize the learning state. The knowledge mastery vector at time t is defined as $k_t \in \mathrm{R}^D$, where D represents the total number of knowledge concepts in the target discipline, and its values are updated at each time step through dynamic assessment via embedded quizzes, with the update rule specified as:

$$k_{t+1} = k_t + \eta \cdot \nabla L(response_t, k_t) \tag{3}$$

where, L denotes the knowledge tracing loss function, $response_t$ denotes the learner's quiz response at the current time step, and η denotes the learning rate for state updating. The fused contextual feature z_t and the knowledge mastery vector k_t are concatenated along the feature dimension to construct a context–cognitive joint state representation:

$$h_t = [z_t; k_t] \in \mathrm{R}^{d_z + D} \tag{4}$$

where, d_z denotes the dimensionality of the contextual fused features. This joint representation simultaneously encompasses both the learner's external environmental constraints and internal cognitive level, thereby enabling more accurate reflection of the actual learning state boundaries in mobile scenarios and providing more comprehensive state input for subsequent path decision-making. Based on the joint state representation, a lightweight prediction network is constructed to evaluate the degree of suitability of different learning activities under the current state. The expected learning efficiency coefficient $\hat{e}_{t,a}$ for learning activity a at time t is defined, with a value range of [0,1], serving to quantify the learning efficiency gain or discount effect of the current context on a specific learning activity. This is computed as:

$$\hat{e}_{t,a} = f_\theta(h_t, emb_a) \tag{5}$$

where, f_θ denotes the parameterized prediction network, and emb_a denotes the learnable embedding vector corresponding to learning activity a, with different content types of learning activities corresponding to different embedding representations. Considering the computational and power constraints of mobile terminals, the prediction network is trained via a knowledge distillation strategy: a high-precision teacher network is deployed at the cloud for offline training, after which knowledge is transferred to the on-device student network, whose parameter count is substantially compressed, through a distillation loss. The single-step inference latency of the finally deployed on-device network is controllable within 10 milliseconds, with the entire computation process completed locally on the terminal, requiring no reliance on cloud computing support.

When the expected efficiency coefficient of the current learning activity falls below a preset threshold corresponding to its type, the mobile terminal automatically triggers the path fine-tuning process. Path fine-tuning is formulated as a constrained combinatorial optimization problem, with the optimization objective being the maximization of the product of the expected efficiency of candidate learning activities and the learning gain, where the learning gain $R(\tilde{a}, k_t)$ is derived from knowledge graph reasoning, corresponding to the potential knowledge mastery improvement obtainable upon completion of candidate activity $\tilde{a}$ under the

current knowledge state. The formal formulation of the optimization problem is expressed as:

$$\max_{\tilde{a}} \quad \hat{e}_{t,\tilde{a}} \cdot R(\tilde{a}, k_t) \tag{6}$$

$$s.t.(\tilde{a}) \le T_{remain}, (\tilde{a}) \le B_{available} \tag{7}$$

where, the constraints respectively specify that the duration of the candidate activity does not exceed the current remaining available learning time T_{remain}, and that the bandwidth requirement of the activity does not exceed the current available network bandwidth $B_{available}$. The optimization solution outputs three categories of fine-tuning strategies: content type switching, learning granularity splitting, and high-load task deferral. The entire solution and decision execution process is completed locally on the mobile terminal, with no raw sensing or behavioral data uploaded to the cloud, thereby achieving learner privacy protection at the data source while ensuring real-time responsiveness of path adjustment.

4 EDGE COMPUTING-EMPOWERED END–CLOUD COLLABORATIVE INTERVENTION STRATEGY AND FEDERATED OPTIMIZATION MECHANISM

Edge computing nodes serve as the inference hub between terminals and the cloud, undertaking group intervention decision-making functions within local regions and forming a hierarchically complementary decision-making system with on-device local fine-tuning. The mobile terminal is responsible for handling real-time contextual adaptation for individual learners, while the edge layer focuses on cross-learner collaborative decisions and complex scenarios that depend on local–global information, including group learning path conflict resolution, peer-benchmark-based incentive strategy triggering, and path re-planning tasks requiring local knowledge graph support. For the *i*-th learner, the output of the edge intervention decision model is expressed as:

$$d_t^{(i)} = g_\psi(\tilde{h}_t^{(i)}, K_t^{edge}) \tag{8}$$

where, $\tilde{h}_t^{(i)}$ denotes the anonymized joint state vector uploaded from the mobile terminal, K_t^{edge} denotes the local knowledge graph state maintained by the edge node, and g_ψ denotes the parameterized intervention decision network. To satisfy the low-latency requirements in high-concurrency scenarios, the edge-side model is compressed via a structured pruning strategy, through which redundant network channels and fully connected layer parameters are pruned, with single-sample inference latency controlled within 50 milliseconds under the constraint that accuracy loss remains below 2%. Compared with pure cloud architectures that require full wide-area network transmission and remote inference, edge deployment shortens the intervention response pathway to within the campus local area network, reducing end-to-end latency by order of magnitude, thereby better accommodating the real-time requirements of mobile learning scenarios.

The cloud undertakes the core functions of global model aggregation and knowledge system evolution, with a federated learning framework adopted to achieve

distributed collaborative optimization across multiple edge nodes, transmitting only model gradients throughout the process without involving raw learning data. The base aggregation adopts a federated averaging strategy, with the update rule for global model parameters in the $(r+1)$-th round specified as:

$$\Psi^{(r+1)} = \Psi^{(r)} - \eta \sum_{j=1}^{M} \frac{n_j}{n} \nabla L_j(\Psi^{(r)}) \tag{9}$$

where M denotes the total number of edge nodes, n_j denotes the number of learners covered by the j-th edge node, n denotes the total number of all learners, η denotes the global learning rate, and ∇L_j denotes the local gradient of the j-th node. Given that the learner populations covered by different edge nodes exhibit significant heterogeneity in knowledge foundations and learning behavioral patterns, direct averaging aggregation would lead to degraded adaptation accuracy of the global model in local contexts. Therefore, a personalized federated update mechanism is introduced, with the local model update rule for each edge node specified as:

$$\Psi_j^{(r+1)} = \Psi^{(r+1)} + \lambda \cdot (\psi_j^{(r)} - \Psi^{(r)}) \tag{10}$$

where, λ denotes the personalized coefficient, used to control the retention proportion of local model bias, and $\psi_j^{(r)}$ denotes the local model parameters of the j-th node prior to the update. The personalized coefficient is dynamically adjusted according to the Kullback–Leibler divergence between the local data distribution of the node and the global data distribution; a larger distribution difference yields a higher coefficient value, thereby adapting to the learning characteristics of different populations while preserving globally shared knowledge. Concurrently, the global knowledge graph is maintained at the cloud. When the cumulative gradient updates uploaded from each edge node reach a preset threshold, adjustments to knowledge concept prerequisite relationships, reordering of learning path topology, and relabeling of learning resources are triggered, enabling continuous evolution of the knowledge system.

The entire intervention system achieves continuous iterative optimization of strategies through a dual-timescale closed-loop feedback mechanism, simultaneously accommodating real-time responsiveness and long-term learning efficacy. The micro-level closed loop operates at the edge layer, with a temporal granularity ranging from seconds to minutes, and employs a contextual bandit algorithm to accomplish real-time updating of intervention strategies, with the parameter update form specified as:

$$\psi \leftarrow \psi + \gamma \cdot \delta_t \cdot \nabla_\psi \log \pi_\psi \left(d_t \mid \tilde{h}_t \right) \tag{11}$$

where, π_ψ denotes the policy function of the intervention strategy, γ denotes the policy gradient learning rate, and δ_t denotes the immediate reward signal, calculated from the differences in short-term behavioral metrics—including task accuracy, completion duration, and interaction activity level—before and after the intervention. Following each intervention execution, the reward signal is immediately fed back to the edge model, enabling single-step policy updates and ensuring rapid adaptation of intervention strategies to dynamic contexts. The macro-level closed loop operates at the cloud layer, with a temporal granularity ranging from

weeks to months. Long-term learning efficacy data reported from each edge node are aggregated at the cloud, including metrics such as periodic assessment scores, course completion rates, and learning persistence. Systemic learning bottlenecks are identified at the global level, triggering deep optimization of the global knowledge graph reconstruction and the global path planning model. The dual closed loops are complementary in temporal scale: the micro-level closed loop ensures immediate adaptation accuracy in dynamic contexts, while the macro-level closed loop governs the achievement of long-term learning objectives, preventing short-term optimization from deviating from long-term cultivation goals and avoiding global optimization lagging behind real-time contextual changes.

5 EXPERIMENTAL DESIGN AND RESULT ANALYSIS

5.1 Experimental setup

A real-world anonymized dataset from a mobile learning platform at a domestic higher education institution was adopted for the experiments, encompassing 12-week course learning records from 216 undergraduate learners, including three categories of contextual data—physical sensing, device status, and interaction behavior—as well as embedded quiz response data from corresponding knowledge modules. The publicly available knowledge tracing dataset ASSISTments2017 was additionally introduced as a supplement to validate the generalization performance of the cognitive state modeling. All data were subjected to identity anonymization in compliance with institutional ethics review requirements and personal information protection regulations.

The mobile-side testing device was equipped with a Snapdragon 8 Gen 2 processor and 12 GB memory, running the Android 14 operating system. The edge node was configured as an edge server with dual Intel Xeon Silver 4314 processors and 64 GB memory, deployed at a campus 5G multi-access edge computing node, with a one-way network latency of approximately 5 ms to the mobile terminal. The cloud was configured as a cloud server instance with 8 cores and 16 GB specification, with a one-way wide-area network latency of approximately 30 ms to the edge node. All experiments were repeated five times, with the average values taken as the final results.

5.2 Performance validation of mobile-side contextual awareness and path fine-tuning

The prediction accuracy, response speed, and scenario robustness of the gated adaptive multi-modal fusion method and the path fine-tuning algorithm were validated in this experiment. Three categories of comparison schemes were included: a statically weighted fusion scheme, in which fixed weights were assigned to the three modality features followed by linear fusion; a deep concatenation fusion scheme, in which the three modality features were directly concatenated and then mapped through a multi-layer perceptron for output; and the proposed gated adaptive fusion scheme. Two typical mobile scenarios were respectively configured for robustness testing: single-modality missing and 20% gaussian noise.

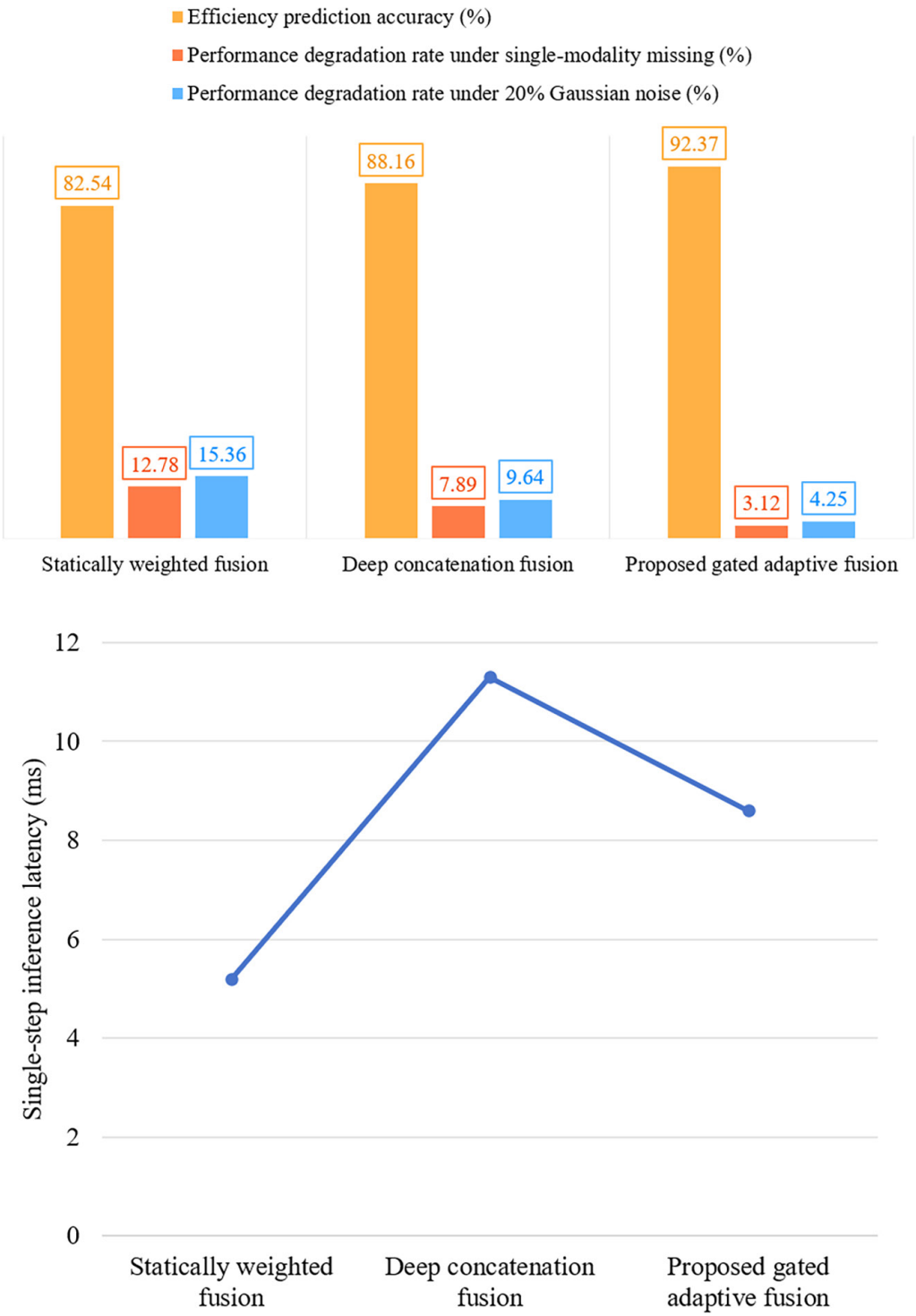

Fig. 2. Performance comparison of mobile-side multi-modal fusion and path fine-tuning

As shown in Figure 2, the proposed method significantly outperformed the two baseline schemes in efficiency prediction accuracy, achieving an improvement of 9.83 percentage points over statically weighted fusion and 4.21 percentage points over deep concatenation fusion. The core advantage was attributed to the adaptive accommodation of different modality data qualities by the dynamic confidence weights. In terms of inference latency, the proposed method achieved a single-step

inference time of 8.6 ms, satisfying the real-time design target of within 10 ms on mobile terminals. Although slightly higher than that of the structurally simplest statically weighted scheme, this latency was far lower than that of the deep concatenation fusion scheme, demonstrating a reasonable balance between accuracy and speed. In robustness testing, under sensor data missing and noise interference scenarios, the performance degradation rate of the proposed method remained below 5% in all cases, far lower than those of the two baseline schemes. This validated the adaptability of the dynamic weighting mechanism in unstable sensing environments, rendering it more suitable for the complex and dynamically changing characteristics of real-world mobile learning scenarios.

5.3 Comparative experiment on edge-side intervention inference performance

The inference efficiency, concurrent processing capability, and accuracy loss control effectiveness of the pruned and compressed edge model were validated in this experiment. Four categories of comparison schemes were included: a pure cloud inference scheme, in which the entire intervention model was deployed at the cloud with computation performed via wide-area network transmission; an unpruned full model, in which the complete-parameter intervention model was directly deployed at the edge node; a conventional edge lightweight scheme, in which a lightweight intervention model constructed with depthwise separable convolutions was adopted; and the proposed structured pruning scheme.

Table 1. Performance comparison of edge-side intervention inference

Method	Single-Sample Inference Latency (ms)	100-Concurrent Throughput (req/s)	Decision Accuracy Loss Rate (%)	Edge-Side Memory Usage (MB)
Pure Cloud Inference	287.3	42.6	0	–
Unpruned Full Model	62.4	128.3	0	186.2
Conventional Edge Lightweight Scheme	38.7	215.4	3.87	92.5
Proposed Structured Pruning Scheme	41.2	247.1	1.72	78.3

As shown in Table 1, the end-to-end latency of the pure cloud scheme reached 287.3 ms, far higher than that of all edge-deployed schemes, rendering it incapable of meeting the real-time intervention requirements of mobile learning scenarios. The proposed pruning scheme achieved a single-sample inference latency of 41.2 ms, meeting the design target of within 50 ms. Compared with the unpruned full model, latency was reduced by 34.0%, 100-concurrent throughput was improved by 92.6%, and edge-side memory footprint was reduced by 57.9%, while the decision accuracy loss was only 1.72%. Thus, inference efficiency and concurrent processing capacity were substantially improved while accuracy loss was maintained within a controllable range. Compared with the conventional lightweight scheme, the proposed scheme reduced accuracy loss by 2.15 percentage points under comparable latency and throughput, demonstrating the technical advantage of structured pruning in preserving core decision-making features.

5.4 Experiment on federated learning convergence and personalization effectiveness

The convergence efficiency and local adaptation performance of the heterogeneity-aware personalized federated learning algorithm in scenarios with population heterogeneity were validated in this experiment. Five edge nodes were configured, and medium-to-high heterogeneity learning population scenarios were constructed by adjusting the data distribution of each node. Four categories of comparison schemes were included: local training only, in which each edge node trained exclusively with local data without participating in federated aggregation; the classical federated averaging algorithm; the federated proximal algorithm, a federated optimization algorithm with a proximal term constraint; and the proposed personalized federated learning algorithm.

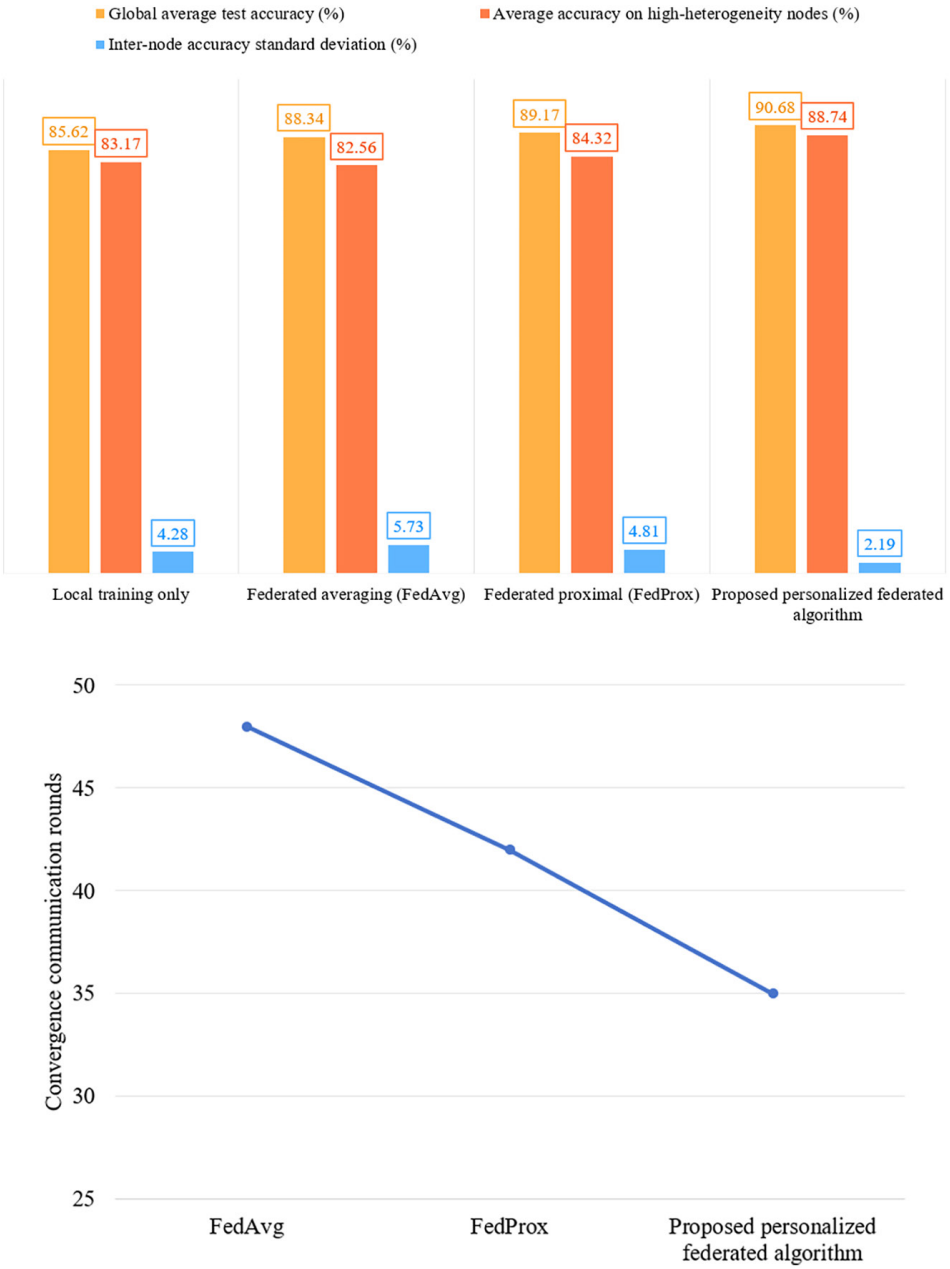

Fig. 3. Performance comparison of different federated learning algorithms

As shown in Figure 3, the proposed algorithm required only 35 communication rounds to achieve convergence, representing a reduction of 27.1% and 16.7% in communication rounds compared with federated averaging and federated proximal, respectively, demonstrating superior convergence efficiency. In terms of global average test accuracy, the proposed algorithm achieved 90.68%, outperforming the three baseline schemes, thereby validating the effectiveness of global knowledge aggregation. On high-heterogeneity nodes, the average accuracy of the proposed algorithm reached 88.74%, significantly higher than the 82.56% of federated averaging and the 84.32% of federated proximal, while the inter-node accuracy standard deviation was only 2.19%, far lower than those of the other federated schemes. The experimental results indicate that the personalized coefficient dynamically adjusted based on distribution differences effectively accommodates the learner population characteristics of different nodes, improving local adaptation accuracy while preserving globally shared knowledge, and alleviating the global model performance bias caused by data heterogeneity.

5.5 Comprehensive evaluation of learning efficacy and user experience

The practical educational application value of the proposed approach was validated in this experiment within real-world higher education teaching scenarios. A total of 180 undergraduate learners enrolled in three public foundational courses at a higher education institution were selected and randomly divided into three groups of 60 learners each, with an eight-week controlled teaching experiment conducted. The three groups respectively adopted: a conventional mobile learning system without contextual awareness, a traditional cloud-based adaptive learning system, and the proposed end–edge–cloud collaborative adaptive learning system. Standardized knowledge level tests were administered before and after the experiment, while learning behavioral data and course satisfaction questionnaire results were also collected.

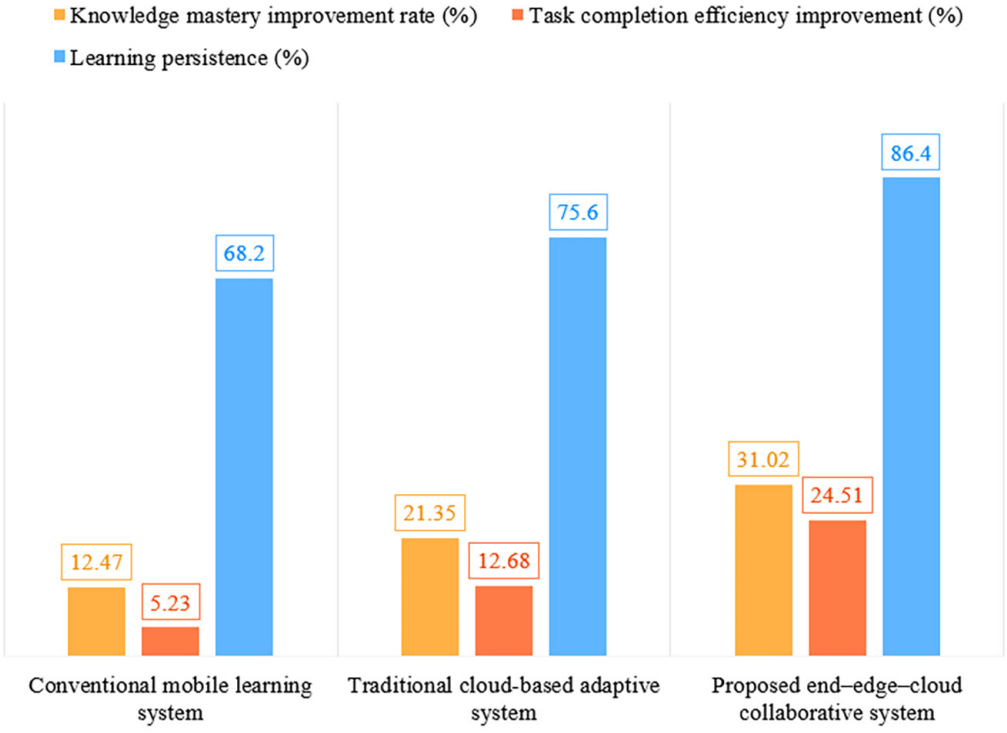

Fig. 4. *(Continued)*

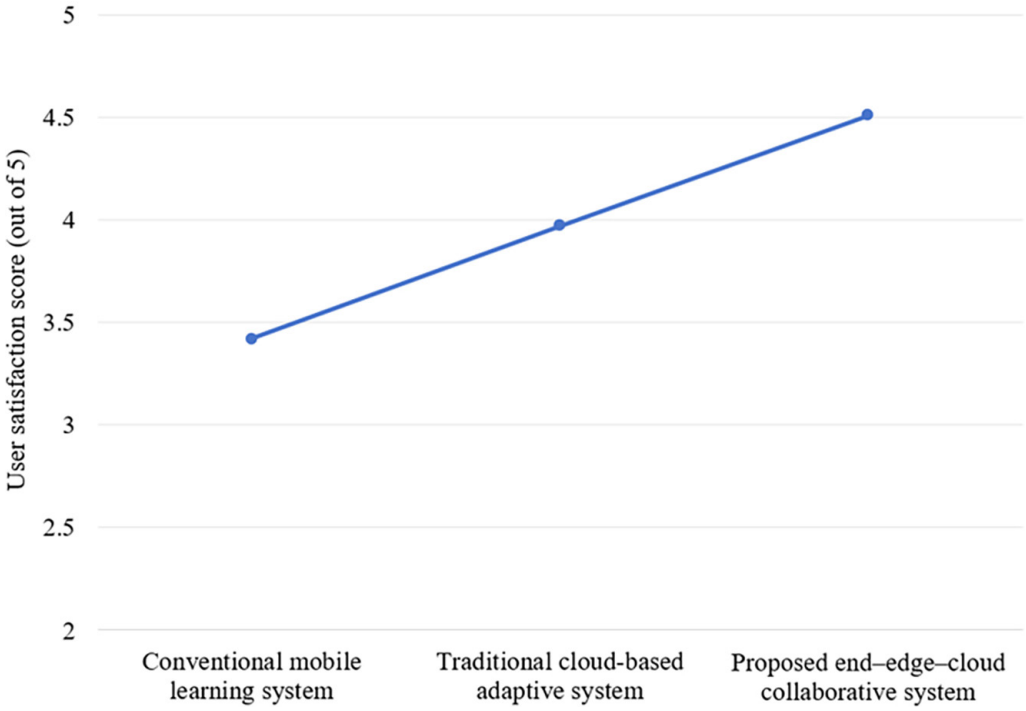

Fig. 4. Comparison of learning efficacy and user experience across different systems

As shown in Figure 4, the teaching experiment results demonstrated that the proposed system significantly outperformed the two control systems across all learning efficacy metrics. The knowledge mastery improvement rate reached 31.02%, representing an increase of 18.55 percentage points over the conventional mobile learning system and 9.67 percentage points over the traditional adaptive system, thereby validating the positive promotive effect of context–cognitive joint modeling and dynamic path adjustment on learning outcomes. The task completion efficiency improvement reached 24.51%, which was primarily attributed to the reduction of inefficient learning time by millisecond-level path fine-tuning, better accommodating the fragmented mobile learning scenarios. Learning persistence and user satisfaction were also significantly higher than those of the control groups, indicating that real-time contextual adaptation and low-latency intervention effectively enhanced the fluency and user acceptance of mobile learning.

5.6 Privacy protection efficacy and system overhead analysis

The privacy protection level and terminal resource overhead performance of the proposed approach were validated in this experiment. Three categories of comparison schemes were included: a centralized cloud-based system, in which all raw behavioral data of learners were uploaded to the cloud for processing; a conventional federated learning system, in which an end–cloud architecture with federated averaging was adopted and only model gradients were uploaded; and the proposed end–edge–cloud collaborative system. The privacy leakage risk level was comprehensively assessed based on data sensitivity and transmission volume.

Table 2. Comparison of privacy overhead and resource occupation across different systems

System Scheme	Average Daily Uploaded Data Volume per User (KB)	Terminal Energy Consumption per Unit Time (mAh/h)	Terminal Memory Usage (MB)	Privacy Leakage Risk Level
Centralized Cloud-Based System	1284.7	11.2	42.6	High
Conventional Federated Learning System	168.3	9.7	68.4	Medium
Proposed End–Edge–Cloud Collaborative System	27.9	7.8	56.2	Low

As shown in Table 2, the average daily uploaded data volume per user of the proposed system was only 27.9 KB, representing a reduction of 97.8% compared with the centralized cloud-based system and 83.4% compared with the conventional federated learning system. This advantage was attributed to the fact that raw data were retained entirely on terminals, with only anonymized low-dimensional feature statistics and model gradients being uploaded, thereby reducing the risk of privacy leakage at the data source and achieving an overall low-risk privacy protection level. In terms of terminal resource overhead, the energy consumption per unit time of the proposed system was 7.8 mAh/h, and the terminal memory footprint was 56.2 MB, both within the acceptable range of mainstream mobile devices, imposing no significant burden on daily terminal usage and demonstrating feasibility for large-scale deployment.

6 CONCLUSION

To address the issues of high response latency, limited contextual awareness dimensionality, and inadequate privacy protection mechanisms in conventional adaptive teaching systems within mobile learning scenarios, a three-tier collaborative adaptive teaching path dynamic intervention framework was constructed in this work, integrating on-device perception, edge-side inference, and cloud-based optimization. A gated adaptive multi-modal context fusion method was proposed, through which a joint context–cognitive state representation space was established. Combined with a lightweight prediction network trained via knowledge distillation, real-time millisecond-level learning path fine-tuning was achieved locally on mobile terminals. A heterogeneity-aware personalized federated learning algorithm and a dual-timescale closed-loop feedback mechanism were designed, enabling distributed optimization of the global model and long-period iteration of the knowledge system without raw data leaving the terminals. Multiple controlled experiments demonstrated that the proposed approach outperformed existing state-of-the-art technical solutions across dimensions including inference efficiency, prediction accuracy, learning efficacy, and privacy overhead, effectively accommodating the fragmented and dynamically changing characteristics of mobile learning in higher education. This study transforms the multi-modal sensing capabilities and on-device computing potential of mobile terminals into core driving factors for adaptive teaching, providing a systematic technical pathway for the deep integration of mobile computing technologies with higher education, and offering theoretical and methodological support for the large-scale deployment of digital–intelligent learning ecosystems.

7 ACKNOWLEDGEMENT

This paper was funded by the Research Project on Education and Teaching Reform in 2025 of the National Ethnic Affairs Commission of the People's Republic of China Titled with "Investigation on Immersive Integrating the Consciousness of Chinese National Community into the Basic Principles of Marxism in the Context of Digital Intelligence Enablement" (Grant No.: 2025-GMJ-105).

8 REFERENCES

[1] T. Zhao, Z. Ma, X. Sun, Q. Yan, and D. Wang, "Dynamic prediction and optimization of energy consumption in mining equipment using mobile computing platforms," *International Journal of Interactive Mobile Technologies*, vol. 19, no. 10, pp. 236–250, 2025. https://doi.org/10.3991/ijim.v19i10.55837

[2] M. Pham Ngoc, T. Ngo Duy, H. Huynh Duc, and K. Tran Anh, "A proposed CNN model for audio recognition on embedded device," *International Journal of Interactive Mobile Technologies*, vol. 18, no. 8, pp. 116–126, 2024. https://doi.org/10.3991/ijim.v18i08.45917

[3] Y. Siriwardhana, P. Porambage, M. Liyanage, and M. Ylianttila, "A survey on mobile augmented reality with 5G mobile edge computing: Architectures, applications, and technical aspects," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 1160–1192, 2021. https://doi.org/10.1109/COMST.2021.3061981

[4] A. R. Nepomuceno, E. L. Domínguez, S. D. Isidro, M. A. M. Nieto, A. Meneses-Viveros, and J. de la Calleja, "Software architectures for adaptive mobile learning systems: A systematic literature review," *Applied Sciences*, vol. 14, no. 11, p. 4540, 2024. https://doi.org/10.3390/app14114540

[5] I. Gligorea, M. Cioca, R. Oancea, A. Gorski, H. Gorski, and P. Tudorache, "Adaptive learning using artificial intelligence in e-learning: A literature review," *Education Sciences*, vol. 13, no. 12, p. 1216, 2023. https://doi.org/10.3390/educsci13121216

[6] N. S. Raj and V. G. Renumol, "An improved adaptive learning path recommendation model driven by real-time learning analytics," *Journal of Computers in Education*, vol. 11, no. 1, pp. 121–148, 2022. https://doi.org/10.1007/s40692-022-00250-y

[7] X. Zheng, W. Zhang, C. Hu, L. Zhu, and C. Zhang, "Cloud-edge-end collaborative inference in mobile networks: Challenges and solutions," *IEEE Network*, vol. 39, no. 4, pp. 90–96, 2025. https://doi.org/10.1109/MNET.2025.3533581

[8] G. Abdelrahman, Q. Wang, and B. P. Nunes, "Knowledge tracing: A survey," *ACM Computing Surveys*, vol. 55, no. 11, pp. 1–37, 2023. https://doi.org/10.1145/3569576

[9] C. P. Gumbheer, K. K. Khedo, and A. Bungaleea, "Personalized and adaptive context-aware mobile learning: Review, challenges and future directions," *Education and Information Technologies*, vol. 27, no. 6, pp. 7491–7517, 2022. https://doi.org/10.1007/s10639-022-10942-8

[10] S. Deng, H. Zhao, W. Fang, J. Yin, S. Dustdar, and A. Y. Zomaya, "Edge intelligence: The confluence of edge computing and artificial intelligence," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7457–7469, 2020. https://doi.org/10.1109/JIOT.2020.2984887

[11] K. Hoffpauir *et al.*, "A survey on edge intelligence and lightweight machine learning support for future applications and services," *Journal of Data and Information Quality*, vol. 15, no. 2, pp. 1–30, 2023. https://doi.org/10.1145/3581759

[12] Q. Li *et al.*, "A survey on federated learning systems: Vision, hype and reality for data privacy and protection," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 4, pp. 3347–3366, 2021. https://doi.org/10.1109/TKDE.2021.3124599

[13] H. Zhu, J. Xu, S. Liu, and Y. Jin, "Federated learning on non-IID data: A survey," *Neurocomputing*, vol. 465, pp. 371–390, 2021. https://doi.org/10.1016/j.neucom.2021.07.098

9 AUTHORS

Jianjiang Wang has completed Doctor of Philosophy, works at College of Marxism, Hebei Minzu Normal University, and is interested in the research of ideological and political education, philosophy (E-mail: wangjianjiang0091@hbun.edu.cn).

Yangyang Cui has a Master of Laws. He works at College of Marxism, Hebei Minzu Normal University. He is interested in the research of ideological and political education (E-mail: Cuiyangyang_1015@163.com).

Lingling Sun has completed Master of Management and is working at the Library, Hebei Minzu Normal University. His area of interest include research of information retrieval and knowledge service (E-mail: sunlingling0378@hbun.edu.cn).