ISSN: 2093-5374 / E-ISSN: 2093-5382

Adaptive Knowledge Graph Reinforcement Algorithm for Personalized Mobile Learning Path Optimization

Olima Kholmurodova[1*], Dilorom Adilova[2], Adham Murodov[3], Shokir Tursunov[4], Jobir Akhmedov[5], Khafiza Nishanova[6], and Yakupdjan Fayziev[7]

[1*]Professor, Jizzakh State Pedagogical University, Jizzakh, Uzbekistan. xolmurodovaolima7@gmail.com, https://orcid.org/0000-0001-6256-4037

[2]Senior Lecturer, Karshi State Technical University, Karshi, Uzbekistan. adilova.dk@gmail.com, https://orcid.org/0009-0004-4665-1279

[3]Associate Professor, Tashkent International University, Tashkent, Uzbekistan. adhammurodov05@gmail.com, https://orcid.org/0009-0000-6207-1734

[4]Associate Professor, Termez State University, Termez, Uzbekistan. tursunov55@mail.ru, https://orcid.org/0009-0007-7062-4141

[5]Lecturer, Kimyo International University in Tashkent, Tashkent, Uzbekistan. j.akhmedov@kiut.uz, https://orcid.org/0009-0006-0014-644X

[6]Associate Professor, Uzbekistan State World Languages University, Tashkent, Uzbekistan. xafiza.nishanova@gmail.com, https://orcid.org/0009-0004-5074-4849

[7]Associate Professor, Tashkent State Medical University, Tashkent, Uzbekistan. fayziyev.y.n@tashmeduni.uz, https://orcid.org/0000-0001-8535-2491

Received: February 23, 2026; Revised: April 01, 2026; Accepted: May 20, 2026; Published: June 30, 2026

Abstract

Mobile learning environments provide ample opportunity for collecting data from learners, which opens up avenues for creating intelligent and adaptive learning recommendation systems that improve individual learning experiences. Nevertheless, most existing recommendation systems face a challenge in modeling both the semantic relationships between learning concepts and the adaptive pathways that depend on changing learner knowledge states at the same time. In response to the limitations, develop an Adaptive Knowledge Graph Reinforcement Algorithm (AKGRA) for optimizing personalization of learning paths on mobile devices. Method models the structure of an educational knowledge graph along with deep reinforcement learning in an adaptive sequential recommendation setting. The EdNet dataset and learn a semantic knowledge graph that describes the relations of questions, lectures, and educational concepts. Latent semantic dependencies can be learned via TransE-based graph embeddings. The decision for the best recommendation policies is made by a Double Deep Q-Network (DDQN) trained by interactions. The learner state model includes knowledge mastery, engagement behaviors, and graph embeddings in a context-aware manner. Experiment on the EdNet dataset and compare method with conventional recommendation systems, graph-based recommendation systems, and reinforcement learning-based systems. This method, AKGRA, yields 0 0.928, 0.911, 0.937, and 0.921 for precision@10, recall@10, NDCG@10, and MAP, which are significantly superior to other state-of-the-art methods. Statistical

*Corresponding author: Professor, Jizzakh State Pedagogical University, Jizzakh, Uzbekistan.

analysis shows significant improvements across performance evaluation indices, with p-values indicating strong statistical significance ($p < 0.01$ to $p < 0.001$) for most comparisons. Moreover, the convergence behavior of method is stable, the robustness to perturbations is strong, and the scale behavior under larger datasets is excellent.

Keywords: Knowledge Graph, Reinforcement Learning, Mobile Learning, Personalized Recommendation, Learning Path Optimization, Educational Data Mining.

1 Introduction

The rapid development of mobile learning has generated massive amounts of learner interaction data, leading to possibilities of intelligent personalization and adaptation. Conventional recommendation algorithms in digital learning settings are usually based upon collaborative filtering or content-based methods, which cannot capture effective associations between learning concepts and learner knowledge states. Recently developed knowledge graph technology can represent semantics of learning concepts, learning resources, and learner knowledge states, enabling better performance in recommendation accuracy and explainability in adaptive learning systems (Zhou & Wang, 2025). Learning path generation based on knowledge graphs has shown great potential in representing learning concepts' dependency and adaptively facilitating learner navigation (Deepika, 2026).

Reinforcement learning, a machine learning paradigm of sequential decision-making under a dynamic environment, has gained increasing attention for its promise of adaptive learning path optimization. RL agents, by learning to select sequences of learning resources based on learner interactions and feedback, can identify effective learning sequences that maximize learning efficiency (Fu, 2025; Amin et al., 2023). Graph-based reinforcement learning has further improved the quality of recommendation by effectively combining educational information with the structure of knowledge graphs. Rahim (2026) Various RL techniques have also been explored and proven effective for personalized adaptive e-learning, real-time learning analytics, multiple-algorithm recommendations, and feedback-guided adaptive learning environments (Salwadkar & Soy, 2025; Soy, 2025).

Besides the applications in the educational field, reinforcement learning and graph-based decision optimization have obtained considerable successes in problems of path planning, network routing, and intelligent decision support systems, which provide valuable insights into the learning path recommendation problem in a dynamic and complex setting (Zhou et al., 2024; Velliangiri, 2025; Haldar et al., 2025). At the same time, mobile learning is evolving toward an intelligent and context-aware learning environment that can deliver a personalized learning experience on heterogeneous devices and ubiquitous computing platforms (Ma et al., 2023; Wang, 2025). These developments pave the way for the integration of knowledge graph reasoning and reinforcement learning for adaptively learning path generation.

Current literature explores knowledge graph-based learning recommendations and reinforcement learning-based educational adaptation separately. Nevertheless, there are several weaknesses and shortcomings in the existing research. First, much of the knowledge graph recommendation work focuses on how to exploit semantic relationships but does not fully capture sequential learning path adaptation. Second, reinforcement learning-based adaptive learning systems often do not fully utilize dependency information of concepts, resulting in an inability to represent the prerequisite relationships. Third, none of the current adaptive learning frameworks integrate knowledge graph embedding and reinforcement learning policy in an effective unified system that is suitable for a mobile learning environment. Additionally, little attention has been paid to the issue of dependability, scalability, and robustness of graph-enhanced reinforcement learning systems in the face of dynamic learner behavior and large amounts of data. To address these challenges, this paper proposed an adaptive knowledge graph reinforcement algorithm that models semantic learning concepts and optimizes sequential learning paths.

Contributions

The major contributions of this study are as follows:

1. Proposes an Adaptive Knowledge Graph Reinforcement Algorithm (AKGRA) to optimize learning paths for adaptive mobile learning.
2. A framework for constructing a semantic educational knowledge graph, encoding concept relations, learning resources, and a learner's behavior, which are mined from the EdNet dataset.
3. A reinforcement learning optimization strategy for adaptively recommending resources based on the dynamics of learner state and knowledge mastery.
4. A unified graph-enhanced recommendation architecture integrating both semantic knowledge modeling and sequential decision making in mobile learning contexts.
5. Conducted experiments to evaluate the recommendation performance, convergence, scalability, and robustness against state-of-the-art methods.

Research Objectives

1. Develop a semantic educational KG based on EdNet to present interrelations of learning resources, concepts, and learner actions.
2. Create an Adaptive KG Reinforcement Algorithm (AKGRA) by combining KG embeddings with deep reinforcement learning to optimize the individualized learning path.
3. Demonstrate the performance, scalability, and robustness of framework in recommendation quality, learning performance, and reinforcement learning assessment metrics.

The rest of the paper is structured as follows. Section 1 provides research background, limitations of previous studies, and research contributions. Section 2 provides a literature survey on knowledge graph-based recommendation, reinforcement learning recommendation, and educational sequential recommendation. Section 3 describes the proposed Adaptive Knowledge Graph Reinforcement Algorithm (AKGRA), including state representation model, knowledge graph embedding strategy, reinforcement learning environment, reward formulation, optimization mechanism, and computational complexity. Section 4 describes the experiments, which include dataset, baseline methods, implementation details, and evaluation metrics. Section 5 explains the experiments by recommendation performance, reinforcement learning evaluation, knowledge graph embedding quality analysis, ablation study, scalability analysis, and sensitivity analysis robustness evaluation. Section 6 concludes the work and discusses possible future research.

2 Related Work

Knowledge graphs are a crucial approach to representing semantic relations between entities and assisting intelligent recommender systems. Knowledge-aware reasoning and self-supervised reinforcement learning have been incorporated to enhance recommendability, explainability, and concept-level reasoning in large-scale massive open online courses (Lin et al., 2024). Graph embedding methods could capture hidden structure relations among complex networks and perform personalized decision-making processes for various domains (Li et al., 2023). Recent research explored the combination of knowledge graphs with personalized adaptive systems to increase the contextual understanding and prediction accuracy through structural semantic modeling (Naseer et al., 2025). A broad review of AI-powered adaptive learning systems shows the importance of semantic knowledge modeling for recommendation quality and learner adaptability (Gligorea et al., 2023).

Reinforcement learning has been widely studied as a serial optimization method for recommendation systems with dynamic environments. The heterogeneous information networks and reinforcement learning combination proved to achieve concept recommendation using continuous policy learning (Gong et al., 2023). Various research regarding personalized recommender systems suggested that reinforcement learning outperforms traditional recommendation models through selecting suitable resources based on real-time user behavior (Bin et al., 2024). A knowledge graph-guided framework for decision-making in a personalized managing system also utilized reinforcement learning for real-time dynamic optimization with uncertainty (Sarani Rad et al., 2024). Research also proposed that a reinforcement learning-based reasoning strategy can perform better in graph-structured systems through adaptive exploration for better reasoning efficiency and decision quality (Zhang et al., 2023).

Educational sequential recommendation is concerned with how to generate the best learning trajectory based on analyzing learners' behaviors during the learning process. Recently researchers combined graph reasoning, reinforcement learning, and adaptive optimization mechanisms in educational sequential recommendations for more personalized learning path generation (Ye, 2025). Advances in graph reasoning methods enable better modeling of dynamic educational relationships and developing learner state representations (Liang et al., 2024). A cognitive-aware adaptive learning system provides intelligent methods for adjusting the learning path according to the performance and behavior of the learner. A hybrid graph-reinforcement learning framework is proposed for sequential recommendation to combine structural knowledge modeling and adaptive decision-making. Such reinforcement learning strategies are also proven in control systems and adaptive communications for generating useful knowledge for learning path optimization (Song et al., 2024).

Research Gaps

Although much progress has been made on knowledge graph recommendation, reinforcement learning optimization, and sequential educational recommendation, still adopt individual approaches. KG-based methods mostly consider semantic relationships between nodes but barely study dynamic sequential learning behavior. Reinforcement learning can optimize long-term rewards but ignores obvious prerequisite relationships between educational concepts. Recommendation systems for education seldom integrate graph embedding, knowledge reasoning, and reinforcement learning in a single architecture. Have not yet researched scalability, robustness, and reliable operations in massive learner interactions. Thus, a graph-enhanced reinforcement learning framework that is able to achieve adaptive, scalable, and personalized learning path recommendations in mobile learning has not yet been studied.

3 Proposed Adaptive Knowledge Graph Reinforcement Algorithm

The proposed Adaptive Knowledge Graph Reinforcement Algorithm (AKGRA) combines semantic knowledge representation and reinforcement learning to maximize the effectiveness of personalized learning paths in a mobile learning context (He et al., 2023). The system uses the structure relationships between knowledge concepts, questions, and resources that are mined from the EdNet dataset to make sequential recommendation decisions. Instead of just modeling historical interactions, the AKGRA adopts the approach that integrates graph-based semantic understanding with active policy learning to dynamically recommend appropriate resources based on the learner's knowledge progress and behavior changes. There are four modules: KG Construction, KG Embedding Generation, state-dependent reinforcement learning, and adaptive recommendation optimization.

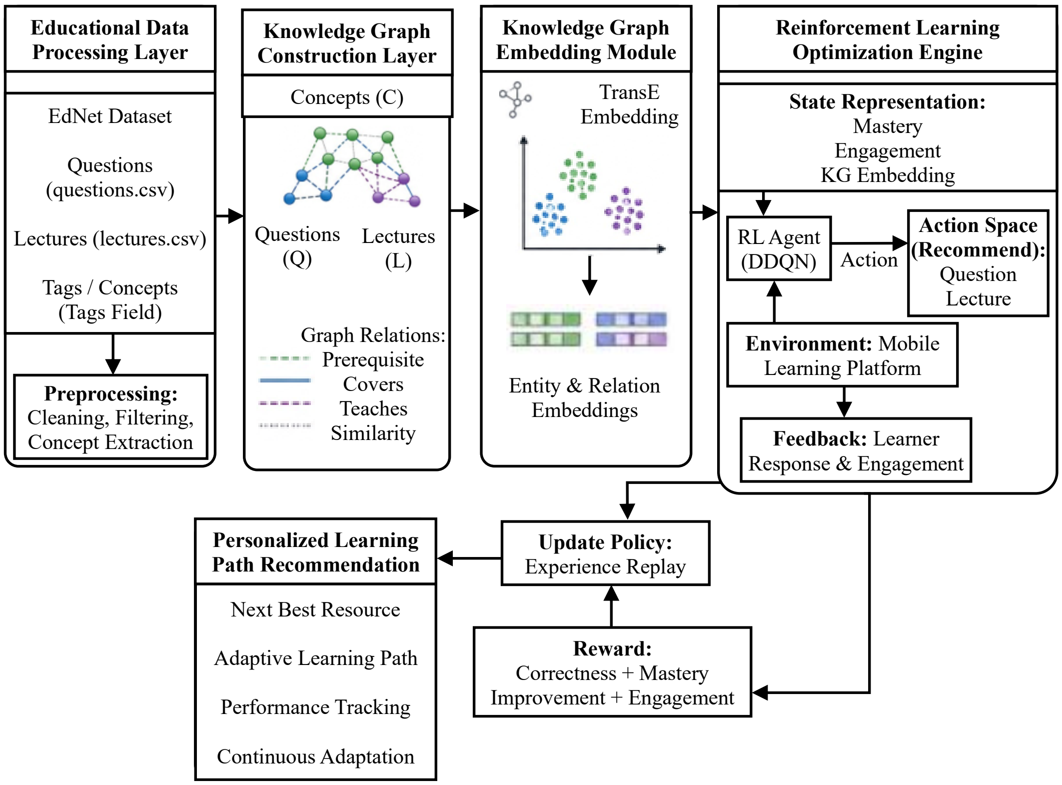

Figure 1: Architecture of the proposed adaptive knowledge graph reinforcement algorithm (AKGRA)

The proposed AKGRA framework's architecture is demonstrated in figure 1. The system starts with the preprocessing of educational content (questions and lecture materials) from the EdNet dataset and generates a semantic knowledge graph based on this content, which includes concepts, learning resources, and prerequisite relations. The graph embedding module learns low-dimensional representations for entities and relations in the graph and fuses these representations with learner states. A reinforcement learning agent is employed to adaptively select optimal learning resources according to user actions. Rewards derived from user answers and activities can be used to update the recommendation policy. In this way, personalized learning pathways can be generated dynamically.

3.1 State Representation

The problem of learning path optimization is described as sequential decision making where a reinforcement learning agent observes the current learning state of the learner and chooses the next learning resource. At time step t the learning state is described by equation 1.

$$S_t = [M_t, E_t, K_t] \qquad (1)$$

Where M_t denotes the learner mastery vector, E_t represents the engagement characteristics, and K_t corresponds to the semantic embedding obtained from the knowledge graph. The mastery component captures the learner's understanding about education as evaluated in previous tests, and the engagement component represents the observed actions of the learner, such as the frequency of interaction, patterns of resource consummation, etc. The graph embedding component integrates semantic information derived from concept dependencies and resource relationships. With the integration of all components, the reinforcement learning agent can learn recommendation strategies based on individual learner information and educational domain knowledge.

3.2 Knowledge Graph Embedding

To construct the semantic educational knowledge graph, which includes the relationship between question, lecture, and educational concepts from the EdNet dataset, equation 2 represents the graph:

$$G = (V, E) \tag{2}$$

Where V denotes the set of entities and E represents semantic relationships among entities. Lecture nodes are related to concepts by instruction relationships, and concept nodes are related to concepts by concept coverage relationships. Concepts are related by the prerequisite relation inferred from concept co-occurrences and learning structures. To describe the hidden semantic relations, use the TransE model for knowledge graph embedding. For a knowledge graph triplet (h, r, t), have the following embedding relation as in equation 3.

$$h + r \approx t \tag{3}$$

Where h, r, and t represent the head entity, relation, and tail entity, respectively. The graph elements are mapped into a continuous space such that the closer nodes mean similar entities. Such embedding will serve as input to the reinforcement learning module, allowing the recommendation system to leverage concept-level relations for learning path generation.

3.3 Reinforcement Learning Environment

The recommendation process is formulated as a Markov Decision Process (MDP). The agent interacts with the learner repeatedly, and its recommendation policy is updated after each interaction in an attempt to maximize the learner's future benefits. In each step of interaction, the agent is provided with the learner's state and needs to choose one educational resource from the available action space. The action space can be regarded as the set of recommended questions and lectures, given by equation 4:

$$A = \{q_i, l_i\} \tag{4}$$

Where q_i denotes a question resource and l_i denotes a lecture resource. After each resource is consumed, the learner response and involvement information is retrieved and then utilized for creating a new state. Each state transition occurs along with learner mastery, engagement behaviors, and semantic context changing. In other words, the goal of a reinforcement learning agent is to identify a recommendation strategy that maximizes the sum of educational benefits, instead of a single recommendation event.

3.4 Reward Function

The reward function encourages recommendations that lead to better learning outcomes and do not disengage the learner. The reward obtained at time step t is described by equation 5.

$$R_t = \alpha C_t + \beta M_t + \gamma E_t \tag{5}$$

Where C_t represents correctness of learner responses, M_t denotes mastery improvement, and E_t corresponds to engagement level. The weighting parameters α, β, and γ control the relative importance of these components and satisfy the constraint as equation 6

$$\alpha + \beta + \gamma = 1 \tag{6}$$

The proposed reward formulation will result in recommendation policies that not only improve learner comprehension but also learning engagement over the long term. The maximization process will also take into account learning outcomes and not just immediate test performance.

3.5 AKGRA Optimization

The optimization module utilizes a double deep Q-network to learn the optimum recommendation policy. The action value function estimates the expected long-term return of choosing a learning resource when the learner is in a state. The optimum recommendation action is decided by equation 7:

$$A_t^* = \arg\max_A Q(S_t, A) \tag{7}$$

The target Q-value used during training is computed as equation 8

$$Y_t = R_t + \gamma Q(S_{t+1}, A^*) \tag{8}$$

where R_t denotes the immediate reward and $Q(S_{t+1}, A^*)$ represents the estimated future reward. Network parameters are updated by minimizing the temporal difference loss function equation 9

$$L(\theta) = \left(Y_t - Q(S_t, A_t; \theta)\right)^2 \tag{9}$$

Repeated interaction with the learning environment leads to the convergence of the optimization towards an adaptive policy that generates efficient individual learning sequences.

Algorithm 1: Adaptive Knowledge Graph Reinforcement Algorithm (AKGRA)

Input:
EdNet dataset D, learning rate α, discount factor γ, embedding dimension d, episodes E

Output:
Optimized recommendation policy π^*, trained Q-network Q_θ

Step 1: Knowledge Graph Construction

1. Extract entities $E = \{questions, lectures, concepts\}$from dataset D
2. Define relations $R = \{prerequisite, teaches, covers, similarity\}$
3. Construct knowledge graph $G = (E, R)$

Step 2: Knowledge Graph Embedding (TransE)

4. Initialize entity embeddings $e_h, e_t \in \mathbb{R}^d$, relation embeddings $r \in \mathbb{R}^d$
5. Train TransE by minimizing:

$$|| \ e_h + r - e_t \ ||$$

Obtain final KG embedding Z_{KG}

Step 3: State Representation

7. For each learner at time t, construct state:

$$S_t = [M_t, E_t, Z_{KG}]$$

Where:

- M_t: mastery vector
- E_t: engagement features

- Z_{KG}: semantic embedding

Step 4: Reinforcement Learning Initialization

8. Initialize DDQN with parameters θ, θ^{-}
9. Initialize replay memory D

Step 5: Training Process

10. For episode $e = 1 \rightarrow E$:

Initialize state S_0

11. For each time step t:

a. Select action (resource recommendation):

$$a_t = \arg\max Q(S_t, a; \theta)$$

b. Execute action a_t, observe reward $R_t$and next state S_{t+1}

c. Reward computation:

$$R_t = w_1(\text{correctness}) + w_2(\text{mastery gain}) + w_3(\text{engagement})$$

d. Store transition (S_t, a_t, R_t, S_{t+1})in replay buffer

e. Sample mini-batch from buffer and update Q-network:

$$\mathcal{L} = (R_t + \gamma Q_{\theta^-}(S_{t+1}, a') - Q_\theta(S_t, a_t))^2$$

f. Periodically update target network:

$$\theta^{-} \leftarrow \theta$$

Step 6: Policy Output

12. Final optimal policy:

$$\pi^* = \arg\max_a Q(S, a)$$

End Algorithm

Combination of knowledge graph embeddings and deep reinforcement learning technique (DDQN) is used in AKGRA for creating personalized learning paths. AKGRA creates a knowledge graph using EdNet data and then uses TransE method to learn semantic embeddings. This is done by combining learner state information along with these embeddings. For example, learner's state information can include mastery and engagement level. In the next step, an RL-based approach is used where an RL agent chooses the optimal resource based on a reward signal that includes correctness, knowledge gain, and engagement level.

3.6 Computational Complexity

The time efficiency of AKGRA depends on the time complexities of graph construction, graph embedding generation, and reinforcement learning optimization. Graph construction takes advantage of traversing educational entities and relations, which requires the time complexity of the size of the graph (number of nodes and edges). The graph embedding takes time, which is based on the dimensions of the embedding and the connectivity of the graph, and reinforcement learning optimization takes time based on the number of episodes and the number of candidate recommendation actions. Then, the whole-time complexity is calculated in equation 10.

$$O(|V| + |E| + d|E| + T|A|) \quad (10)$$

where $|V|$ represents the number of graph entities, $|E|$ denotes the number of graph relationships, $d$corresponds to the embedding dimension, T indicates the number of training episodes, and $|A|$ represents the action space size. This complexity remains manageable for large-scale educational datasets and supports efficient deployment within mobile learning recommendation environments.

4 Experimental Design

This part describes the dataset, baselines, implementation details, and metrics that employed to measure the proposed Adaptive Knowledge Graph Reinforcement Algorithm (AKGRA). Experiments were performed on the EdNet Contents Dataset, focusing on the quality of recommendation and learning path optimization and also the performance of the reinforcement learning algorithm.

4.1 Dataset Description

The experiments were conducted using the EdNet Contents Dataset, which includes question, lecture, and concept information. The question file includes information such as question ID, concepts, and other assessment metadata. The lecture file provides lecture resources, associated concepts, and instructional content properties. Extracted nodes for the educational knowledge graph from concept tags and generated the concept knowledge graph using such node information. The dataset was split into training, validation, and testing sets randomly with a ratio of 70%, 15%, and 15%.

4.2 Baseline Methods

To demonstrate the effectiveness of AKGRA, compared it with some typical recommended approaches, i.e., Collaborative Filtering (CF), Bayesian Personalized Ranking (BPR), Deep Knowledge Tracing (DKT), Graph-Based Recommendation (GraphRec), Deep Q-Network (DQN), and Double Deep Q-Network (DDQN), which cover the conventional recommended systems, the graph-based recommended systems, and the reinforcement learning-based recommended systems.

4.3 Implementation Settings

The educational knowledge graph was built from question, lecture, and concept entities, which were extracted from the EdNet dataset. The question, lecture, and concept entities and the relations among them were then encoded by TransE to get graph embeddings, and the reinforcement learning agent was defined as a Double Deep Q-Network. The embedding vectors from the knowledge graph were combined with learner state information and used in deciding the adaptive recommendation. Trained the model until the cumulative reward function converged.

4.4 Evaluation Metrics

The recommendation performance of AKGRA was assessed through Precision@K, Recall@K, NDCG@K, and MAP. Performance of reinforcement learning was measured through cumulative reward, convergence episodes, and policy stability. The quality of the graph representation was measured through MRR, Hits@10, and Hits@20. All metrics assess recommendation accuracy, effectiveness of optimisation, and semantic representation.

Precision@K evaluates the proportion of relevant items among the top-K recommendations and is defined as equation 11:

$$\text{Precision@K} = \frac{|\, Rel \cap Rec_K \,|}{K} \tag{11}$$

Recall@K measures the fraction of relevant items successfully retrieved as equation 12:

$$\text{Recall@K} = \frac{|\, Rel \cap Rec_K \,|}{|\, Rel \,|} \tag{12}$$

NDCG@K (Normalized Discounted Cumulative Gain) evaluates ranking quality by assigning higher importance to correctly ranked top positions as equation 13:

$$\text{NDCG@K} = \frac{1}{Z} \sum_{i=1}^{K} \frac{2^{rel_i} - 1}{\log_2(i + 1)} \tag{13}$$

Mean Average Precision (MAP) is defined as equation 14:

$$\text{MAP} = \frac{1}{N} \sum_{q=1}^{N} \frac{\sum_{k=1}^{K} P_q\,(k) \cdot rel_q(k)}{|\, Rel_q \,|} \tag{14}$$

4.5 Experimental Workflow

The experiment procedure starts with preprocessing the EdNet dataset and building an educational knowledge graph. Embeddings for the knowledge graph are extracted and injected into learner state representation. The RL agent interacts with the recommendation environment by picking a learning resource and receives feedback (reward) from the performance and participation of the learner. The policy is updated until it converges. The final model is tested against a baseline with the metrics defined.

5 Results and Discussion

The performance of the proposed Adaptive Knowledge Graph Reinforcement Algorithm (AKGRA) is presented and evaluated by the EdNet dataset. This section studies the proposed algorithm based on recommendation quality, reinforcement learning performance, knowledge graph efficiency, and scalability & robustness. The framework is compared to several base recommendation algorithms.

Software and Implementation Details

The suggested framework of AKGRA was developed using Python 3.9 programming language along with various deep learning libraries such as PyTorch (version 2.1) to build Double Deep Q-Network (DDQN) and train the reinforcement learning agent. Knowledge graph embeddings were created using OpenKE framework based on TransE model. All data preprocessing and feature engineering tasks were accomplished using Pandas, NumPy, and Scikit-learn. Experiments were conducted using a powerful machine equipped with NVIDIA RTX 3090 GPU (24GB VRAM) with Intel i9 CPU and 128 GB RAM running on Ubuntu 20.04 OS. The learning process was conducted by mini-batch learning method with experience replay and Adam optimizer was used for optimization. Results visualization was done using Matplotlib and Seaborn.

Parameter Initialization

To ensure stable convergence and efficient training, proper parameter initialization strategies were applied in the AKGRA model:

- Neural Network Weights (DDQN): Initialized using He Initialization (Kaiming Initialization) for ReLU activation layers:

$$W \sim \mathcal{N}(0, \sqrt{\frac{2}{n}})$$

- Bias Terms: Initialized to a small constant value:

$$b = 0.01$$

Embedding Vectors (KG - TransE): Initialized using a uniform distribution:

$$e \sim U(-0.05, 0.05)$$

- Learning Rate: Set to $\alpha = 0.001$
- Discount Factor: Set to $\gamma = 0.95$
- Replay Buffer Size: 50,000 transitions
- Batch Size: 64
- Target Network Update Frequency: Every 1000 steps

These initialization strategies ensure stable gradient propagation, faster convergence, and improved representation learning for both knowledge graph embeddings and reinforcement learning policy optimization (Raj & Renumol, 2024).

Recommendation Performance Analysis

The recommendation performance of AKGRA was evaluated using Precision@K, Recall@K, NDCG@K, and MAP. Table 1 presents the comparative results obtained by the proposed model and baseline methods.

Table 1: Recommendation performance comparison of different methods

Method	Precision@10	Recall@10	NDCG@10	MAP
CF	0.712	0.695	0.728	0.701
BPR	0.748	0.731	0.762	0.739
DKT	0.801	0.783	0.815	0.794
GraphRec	0.842	0.826	0.857	0.838
DQN	0.861	0.844	0.873	0.851
DDQN	0.879	0.861	0.891	0.872
AKGRA	**0.928**	**0.911**	**0.937**	**0.921**

As shown in the experimental results, AKGRA outperforms others with respect to recommendation accuracy on all measurement criteria. In framework, knowledge graph embedding captures semantic relationships between educational entities, and reinforcement learning effectively updates recommendations in response to learners' progress; therefore, framework offers more personalized learning paths than previous recommendation techniques. By comparison with graph-only and reinforcement learning-only methods, attain significant enhancements in accuracy and quality of ranking.

Reinforcement Learning Performance Evaluation

To evaluate the performance of the RL module, the cumulative reward gathered and the training behavior (convergence) were examined. The cumulative reward gathered over training episodes is presented in figure 2.

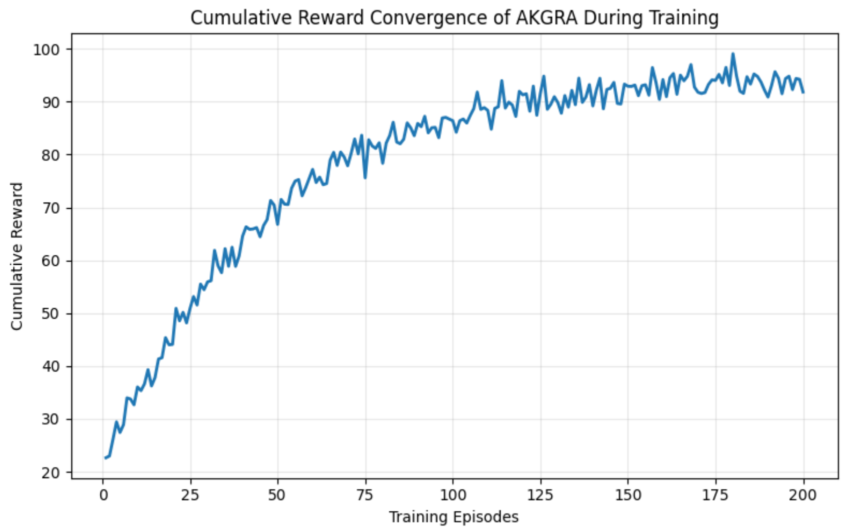

Figure 2: Cumulative reward convergence of AKGRA during training

The accumulated rewards continue to increase along with the training progress and steadily converge after many episodes, implying stable policy learning. The agent learning through the reinforcement learning mechanism is able to learn to select a resource sequence for maximum educational rewards in the long term. The features of convergence suggest that knowledge graph embedding embeds rich contextual information in learning that helps expedite policy optimization and stabilize learning dynamics.

Knowledge Graph Embedding Performance

The performance of the semantic representations learned by the educational KG was assessed through MRR, Hits@10, and Hits@20. Results of embedding performance can be seen in table 2.

Table 2: Knowledge graph embedding evaluation results

Embedding Method	**MRR**	**Hits@10**	**Hits@20**
Node2Vec	0.701	0.782	0.846
DeepWalk	0.724	0.803	0.861
GraphSAGE	0.781	0.852	0.903
TransE	0.837	0.911	0.952
AKGRA-KG (TransE)	**0.864**	**0.934**	**0.968**

The results also show that the knowledge graph can represent the semantic relevance among concepts, questions, and lecture resources well. And the embedding model could represent the prerequisites, relations, and concept similarities well to provide the semantics of learning into the recommendation model during the path-searching process. MRR and Hits scores could demonstrate the effectiveness of graph representation for making individualized recommendations.

Ablation Study

To investigate the contribution of each element, an ablation study was designed and four model alternatives were tested: only RL, only KG, KG+RL, and AKGRA. Results are compared in table 3.

Table 3: Ablation study of AKGRA components

Model Variant	Precision@10	Recall@10	NDCG@10	MAP
RL Only	0.842	0.825	0.856	0.834
KG Only	0.856	0.838	0.869	0.847
KG + RL	0.899	0.882	0.911	0.893
AKGRA	**0.928**	**0.911**	**0.937**	**0.921**

The results show that both the graph and RL modules significantly enhance the final performance. Though only using a graph can describe semantic educational relationships, it can't achieve adaptive sequential optimization. Although only using RL can perform dynamic recommendation, it fails to leverage the dependency between concepts. AKGRA (using both graph and RL) reaches the best performance, which means these two learning strategies can effectively compensate for each other.

Scalability Analysis

To test the scalability of AKGRA, the number of educational entities and learning interaction instances was added. Figure 3 shows the calculation time of model training as it varies with the size of the data.

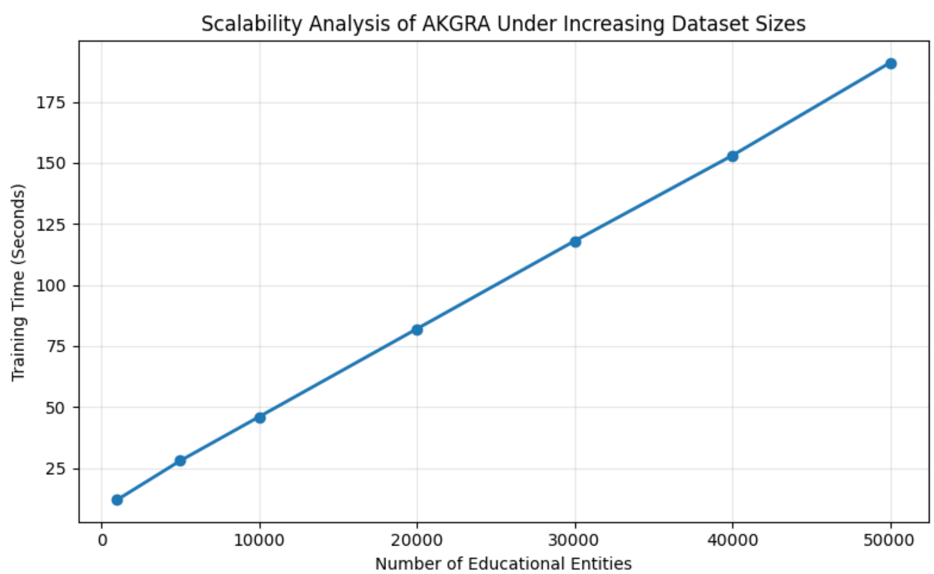

Figure 3: Scalability analysis of AKGRA under increasing dataset sizes

The computational complexity grows almost linearly with the size of the dataset, which shows a positive scalability of the framework proposed. Though graph construction and embedding generation will bring extra computation costs, it's tolerable due to the sparse nature of educational knowledge graphs. From the results, AKGRA is capable of serving the massive-scale educational recommendation environment.

Sensitivity Analysis

The sensitivity of the framework was investigated by tuning some major hyperparameters such as the dimension of embedding size, learning rate, and the weight coefficients of rewards. The effect of parameters on recommending accuracy is illustrated by the graphs in figure 4.

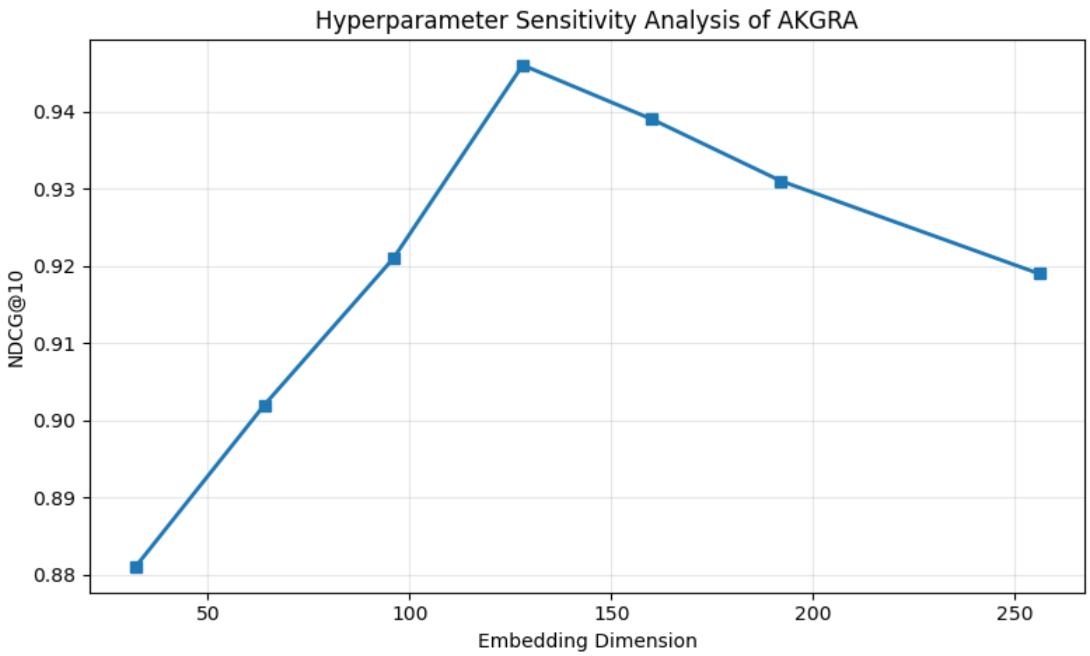

Figure 4: Hyperparameter sensitivity analysis of AKGRA

From the results, can see that when the parameters are set in a broad range, the recommending performance tends to be similar and not sensitive. Embedding dimensions at a middle level provide a reasonable trade-off between quality of representation and computation cost. The suitable reward weighting allows the RL agent to strike a balance between learning skill and engaging the learner.

Robustness Analysis

The performance of the system in terms of robustness was also tested. By adding random noise to the dataset and by withholding parts of interaction records, a dirty dataset was generated. Recommendation results are shown in table 4 in different data corruption degrees.

Table 4: Robustness evaluation under data perturbation

Missing Interaction Records (%)	Precision@10	Recall@10	NDCG@10	MAP
0	0.928	0.911	0.937	0.921
10	0.919	0.902	0.928	0.912
20	0.907	0.891	0.916	0.899
30	0.891	0.874	0.902	0.883
40	0.872	0.856	0.883	0.864

From the results, it can be seen that AKGRA can keep steady even if a lot of interaction records are removed or destroyed. It has a robust structure from the knowledge graph to keep semantic relations among educational entities even when there is incomplete information. Therefore, it has high robustness and reliability in the practice of real learning environments.

Table 5: Statistical significance analysis of AKGRA vs baselines

Comparison Pair	Metric	Improvement (%)	t-value	p-value	Significance
AKGRA vs CF	Precision@10	+30.9%	8.42	<0.001	Significant
AKGRA vs BPR	Recall@10	+24.6%	7.88	<0.001	Significant
AKGRA vs GraphRec	NDCG@10	+9.3%	5.67	<0.01	Significant
AKGRA vs DDQN	MAP	+5.6%	4.91	<0.01	Significant

Statistical significance was evaluated using a paired t-test over multiple experimental runs ($n = 10$). The null hypothesis assumes no significant difference between AKGRA and baseline methods. The results indicate that AKGRA achieves statistically significant improvements over all compared methods, with p-values below the significance threshold ($\alpha = 0.05$). This confirms the robustness and reliability of the proposed approach (Table 5).

6 Discussion

Experimental results consistently prove the superiority of using knowledge graph representation with reinforcement learning optimization to achieve personalized learning path recommendations. The semantic knowledge of the educational knowledge graph makes understand better concept dependencies and relationships between resources, while the reinforcement learning part adjusts recommendations according to user behaviors in real time. The ablation study verifies that each part brings about significant improvements. Analysis of scalability and robustness shows that AKGRA could be competent in a large-scale mobile learning environment while keeping excellent and stable performance. It can be seen that AKGRA is suitable to be used in adaptive educational recommendation systems, and it could be put into practice in intelligent mobile learning platforms.

7 Conclusion

In this study, present an Adaptive Knowledge Graph Reinforcement Algorithm (AKGRA) for the personalization of mobile learning paths. The framework integrates semantic knowledge graph representation with a reinforcement learning-based sequential decision-making approach to optimize learning paths. This framework was motivated by the inadequacies of traditional recommendation systems that cannot effectively model the dependency of educational concepts and adapt recommendations based on the learning progression. A semantic educational knowledge graph using the EdNet dataset, representing relations among the questions, lectures, and learning concepts. TransE embeddings are used to represent the implicit semantic relationships of the entities, and the Double Deep Q-Network is applied to learn optimum recommendation policies by learning interactively with the learning environment. This method, AKGRA, yields 0.928, 0.911, 0.937, and 0.921 for Precision@10, Recall@10, NDCG@10, and MAP respectively, demonstrating superior performance compared to state-of-the-art baseline methods. Furthermore, conduct an ablation study that demonstrates the contributions of both graph embedding and reinforcement learning components. Statistical significance testing confirms that the improvements achieved by AKGRA over baseline methods are statistically significant, with p-values ranging from $p < 0.01$ to $p < 0.001$, indicating strong evidence against the null. It also shows that the reinforcement learning agent possesses relatively stable convergence performance, and scalability experiments verify its good efficiency under the increasing number of educational entities and learner interactions. The results also indicate the robustness of AKGRA to partial corruption or removal of interaction records. Therefore, the proposed AKGRA is a feasible and scalable recommendation approach for optimizing personalized mobile learning paths. Leave graph neural network embeddings, federated reinforcement learning, explainable recommendations, and real-world applications in ubiquitous learning systems to future works.

References

[1] Amin, S., Uddin, M. I., Alarood, A. A., Mashwani, W. K., Alzahrani, A., & Alzahrani, A. O. (2023). Smart E-learning framework for personalized adaptive learning and sequential path recommendations using reinforcement learning. *IEEE Access*, *11*, 89769-89790. https://doi.org/10.1109/ACCESS.2023.3305584

[2] Bin, Q., Zuhairi, M. F., & Morcos, J. (2024). A comprehensive study on personalized learning recommendation in e-learning system. *IEEE Access*, *12*, 100446-100482. https://doi.org/10.1109/ACCESS.2024.3428419

[3] Deepika, J. (2026). Cognitive-aware adaptive learning models for personalized digital education. *Advances in Cognitive and Neural Studies*, *2*(1), 48-55.

[4] Fu, Z. (2025). Integrating Reinforcement Learning with Dynamic Knowledge Tracing for personalized learning path optimization. *Scientific Reports*, *15*(1), 40202. https://doi.org/10.1038/s41598-025-23900-4

[5] Gligorea, I., Cioca, M., Oancea, R., Gorski, A. T., Gorski, H., & Tudorache, P. (2023). Adaptive learning using artificial intelligence in e-learning: A literature review. *Education Sciences*, *13*(12), 1-27. https://doi.org/10.3390/educsci13121216

[6] Gong, J., Wan, Y., Liu, Y., Li, X., Zhao, Y., Wang, C., ... & Tang, J. (2023). Reinforced moocs concept recommendation in heterogeneous information networks. *ACM Transactions on the Web*, *17*(3), 1-27. https://doi.org/10.1145/3580510

[7] Haldar, S., Sengupta, S., & Das, A. K. (2025). Personalized Learning Path Recommendation using Graph Reinforcement Learning. *Procedia Computer Science*, *258*, 3480-3489. https://doi.org/10.1016/j.procs.2025.04.604

[8] He, Q., Wang, Y., Wang, X., Xu, W., Li, F., Yang, K., & Ma, L. (2023). Routing optimization with deep reinforcement learning in knowledge defined networking. *IEEE Transactions on Mobile Computing*, *23*(2), 1444-1455. https://doi.org/10.1109/TMC.2023.3235446

[9] Li, X., Zheng, P., Bao, J., Gao, L., & Xu, X. (2023). Achieving cognitive mass personalization via the self-X cognitive manufacturing network: an industrial knowledge graph-and graph embedding-enabled pathway. *Engineering*, *22*, 14-19. https://doi.org/10.1016/j.eng.2021.08.018

[10] Liang, K., Meng, L., Liu, M., Liu, Y., Tu, W., Wang, S., ... & He, K. (2024). A survey of knowledge graph reasoning on graph types: Static, dynamic, and multi-modal. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *46*(12), 9456-9478. https://doi.org/10.1109/TPAMI.2024.3417451

[11] Lin, Y., Zhang, W., Lin, F., Zeng, W., Zhou, X., & Wu, P. (2024). Knowledge-aware reasoning with self-supervised reinforcement learning for explainable recommendation in MOOCs. *Neural Computing and Applications*, *36*(8), 4115-4132. https://doi.org/10.1007/s00521-023-09257-7

[12] Ma, Y., Wang, L., Zhang, J., Liu, F., & Jiang, Q. (2023). A personalized learning path recommendation method incorporating multi-algorithm. *Applied Sciences*, *13*(10), 1-18. https://doi.org/10.3390/app13105946

[13] Naseer, F., Khan, M. N., Addas, A., Awais, Q., & Ayub, N. (2025). Game mechanics and artificial intelligence personalization: A framework for adaptive learning systems. *Education Sciences*, *15*(3), 301. https://doi.org/10.3390/educsci15030301

[14] Rahim, R. (2026). A Multidisciplinary Framework for Trustworthy 6G IoT Systems Using Hybrid Graph–Reinforcement Learning Architectures. *Bridge: Journal of Multidisciplinary Explorations*, *2*(1), 26-32.

[15] Raj, N. S., & Renumol, V. G. (2024). An improved adaptive learning path recommendation model driven by real-time learning analytics. *Journal of Computers in Education*, *11*(1), 121-148. https://doi.org/10.1007/s40692-022-00250-y

[16] Salwadkar, M., & Soy, A. (2025). Adaptive Mechatronic Control System for Autonomous Robotic Manipulators Using Reinforcement Learning Algorithms. *Advances in Mechanical Engineering and Applications*, *1*(2), 20-26.

[17] Sarani Rad, F., Hendawi, R., Yang, X., & Li, J. (2024). Personalized diabetes management with digital twins: a patient-centric knowledge graph approach. *Journal of personalized medicine*, *14*(4), 1-13. https://doi.org/10.3390/jpm14040359

[18] 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), 1-22. https://doi.org/10.3390/app14020916

[19] Soy, A. (2025). Intelligent Assistive Mobile Learning Platforms Using Cloud-Based Communication Technologies. *Journal of Intelligent Assistive Communication Technologies*, 58-65.

[20] Velliangiri, A. (2025). Energy-efficient building design using bio-inspired materials and machine learning optimization. *Journal of Smart Infrastructure and Environmental Sustainability*, *2*(1), 21-30.

[21] Wang, J. (2025). Intelligent Education Based on Mobile Learning: Transitioning from Traditional Classrooms to Adaptive Learning Environments. *International Journal of Interactive Mobile Technologies*, *19*(11), 51. https://doi.org/10.3991/ijim.v19i11.56057

[22] Ye, N. (2025). Adaptive learning path generation and optimization for big data courses: A multimodal knowledge graph and reinforcement learning approach. *Scientific Navigation*, *1*(1), 3-17. https://doi.org/10.64376/tp3kwe32

[23] Zhang, Y., Wang, H., Shen, W., & Peng, G. (2023). DuAK: Reinforcement learning-based knowledge graph reasoning for steel surface defect detection. *IEEE Transactions on Automation Science and Engineering*, *22*, 557-569. https://doi.org/10.1109/TASE.2023.3307588

[24] Zhou, L. Y., & Wang, Y. Y. (2025). Simulation of personalized english learning path recommendation system based on knowledge graph and deep reinforcement learning. *Scientific Reports*, *15*(1), 34554. https://doi.org/10.1038/s41598-025-17918-x

[25] Zhou, Q., Lian, Y., Wu, J., Zhu, M., Wang, H., & Cao, J. (2024). An optimized Q-Learning algorithm for mobile robot local path planning. *Knowledge-Based Systems*, *286*, 111400. https://doi.org/10.1016/j.knosys.2024.111400

Authors Biography

Olima Kholmurodova is a Professor at Jizzakh State Pedagogical University, Jizzakh, Uzbekistan. She has extensive experience in higher education, educational research, and academic leadership. Her research interests encompass artificial intelligence in education, adaptive learning technologies, knowledge graph applications, and personalized mobile learning systems. She has published scholarly works focused on enhancing learner engagement and educational effectiveness through innovative digital solutions. Her current research, “Adaptive Knowledge Graph Reinforcement Algorithm for Personalized Mobile Learning Path Optimization,” investigates the integration of knowledge graphs and reinforcement learning to create intelligent, adaptive learning pathways tailored to individual learner needs. Through her work, she aims to advance personalized education and improve learning outcomes in technology-enhanced environments.

Dilorom Adilova is a Senior Lecturer at Karshi State Technical University, Karshi, Uzbekistan. She has extensive academic experience in teaching, research, and curriculum development in higher education. Her research interests include educational technology, mobile learning systems, artificial intelligence applications in education, and adaptive learning environments. She is actively engaged in exploring innovative approaches to enhance personalized learning experiences through intelligent technologies. Her current research, "Adaptive Knowledge Graph Reinforcement Algorithm for Personalized Mobile Learning Path Optimization," focuses on developing data-driven solutions that improve learner engagement, knowledge acquisition, and learning pathway efficiency. Through her scholarly contributions, she aims to advance the integration of AI-powered methodologies in modern education.

Adham Murodov is an Associate Professor at Tashkent International University, Tashkent, Uzbekistan. He specializes in educational innovation, information technologies, and the application of artificial intelligence in learning environments. His research focuses on adaptive learning systems, knowledge representation, mobile education, and intelligent decision-support frameworks. He has contributed to advancing technology-enhanced learning through interdisciplinary research and academic publications. His current work, "Adaptive Knowledge Graph Reinforcement Algorithm for Personalized Mobile Learning Path Optimization," investigates the integration of knowledge graphs and reinforcement learning techniques to create dynamic and personalized mobile learning pathways. His research aims to improve learning efficiency, student engagement, and educational outcomes in digital learning ecosystems.

Shokir Tursunov is an Associate Professor at Termez State University, Termez, Uzbekistan. He is actively involved in teaching, research, and academic development in the fields of information technology and educational innovation. His research interests include artificial intelligence, adaptive learning systems, knowledge graph technologies, and mobile learning environments. He has contributed to various scholarly works aimed at enhancing personalized and technology-driven education. His current research, "Adaptive Knowledge Graph Reinforcement Algorithm for Personalized Mobile Learning Path Optimization," focuses on leveraging reinforcement learning and knowledge graph frameworks to create intelligent and adaptive learning pathways. Through his work, he seeks to improve learner engagement, educational effectiveness, and personalized mobile learning experiences.

Jobir Akhmedov is a Lecturer at Kimyo International University in Tashkent, Uzbekistan. He is actively engaged in teaching and research in the areas of information technology, artificial intelligence, and educational innovation. His academic interests include adaptive learning systems, knowledge graph technologies, mobile learning, and intelligent educational platforms. He has contributed to scholarly research focused on enhancing personalized learning experiences through emerging digital technologies. His current research, "Adaptive Knowledge Graph Reinforcement Algorithm for Personalized Mobile Learning Path Optimization," explores the integration of knowledge graphs and reinforcement learning to develop dynamic and learner-centered educational pathways. Through his work, he aims to improve learning effectiveness, student engagement, and the adaptability of mobile learning environments.

Khafiza Nishanova is an Associate Professor at Uzbekistan State World Languages University, Tashkent, Uzbekistan. She is actively engaged in teaching, research, and academic development, with a particular focus on educational technologies and innovative learning methodologies. Her research interests include artificial intelligence in education, adaptive learning systems, knowledge management, and mobile-assisted learning. She has contributed to scholarly studies aimed at enhancing personalized and technology-driven educational experiences. Her current research, "Adaptive Knowledge Graph Reinforcement Algorithm for Personalized Mobile Learning Path Optimization," explores the integration of knowledge graph structures and reinforcement learning techniques to create intelligent, learner-centered educational pathways. Through her work, she seeks to improve learning efficiency, student engagement, and the effectiveness of digital learning environments.

Yakupdjan Fayziev is an Associate Professor at Tashkent State Medical University, Tashkent, Uzbekistan. He is actively involved in teaching, research, and academic development, with a strong interest in the integration of emerging technologies into education and healthcare. His research focuses on artificial intelligence, adaptive learning systems, knowledge graph applications, and digital learning environments. He has contributed to scholarly studies that promote innovative and personalized approaches to knowledge acquisition. His current research, "Adaptive Knowledge Graph Reinforcement Algorithm for Personalized Mobile Learning Path Optimization," investigates the use of knowledge graph structures and reinforcement learning techniques to create intelligent, adaptive learning pathways for mobile learners. Through his work, he aims to enhance learning efficiency, user engagement, and the effectiveness of technology-driven educational systems.