Synthesis: A Multidisciplinary Research Journal
Volume 1, Issue 2, April - June 2023, Pp: 30-41
© 2023, All Rights Reserved @ Macaw Publications

Research Article

Adaptive Curriculum Generation Using Reinforcement Learning from Student Interaction and Knowledge Graphs

[1] Srinath Doss, [2*] Bhavsingh

[1] *Professor, Faculty of Engineering and Technology, Botho University, Botswana*
Email: srinath.doss@bothouniversity.ac.bw
[2*] *Associate Professor Department of CSE, Ashoka Womens Engineering College, Kurnool, Andhra Pradesh, India*
Email: bhavsinghit@gmail.com

**Corresponding Author(s): bhavsinghit@gmail.com*

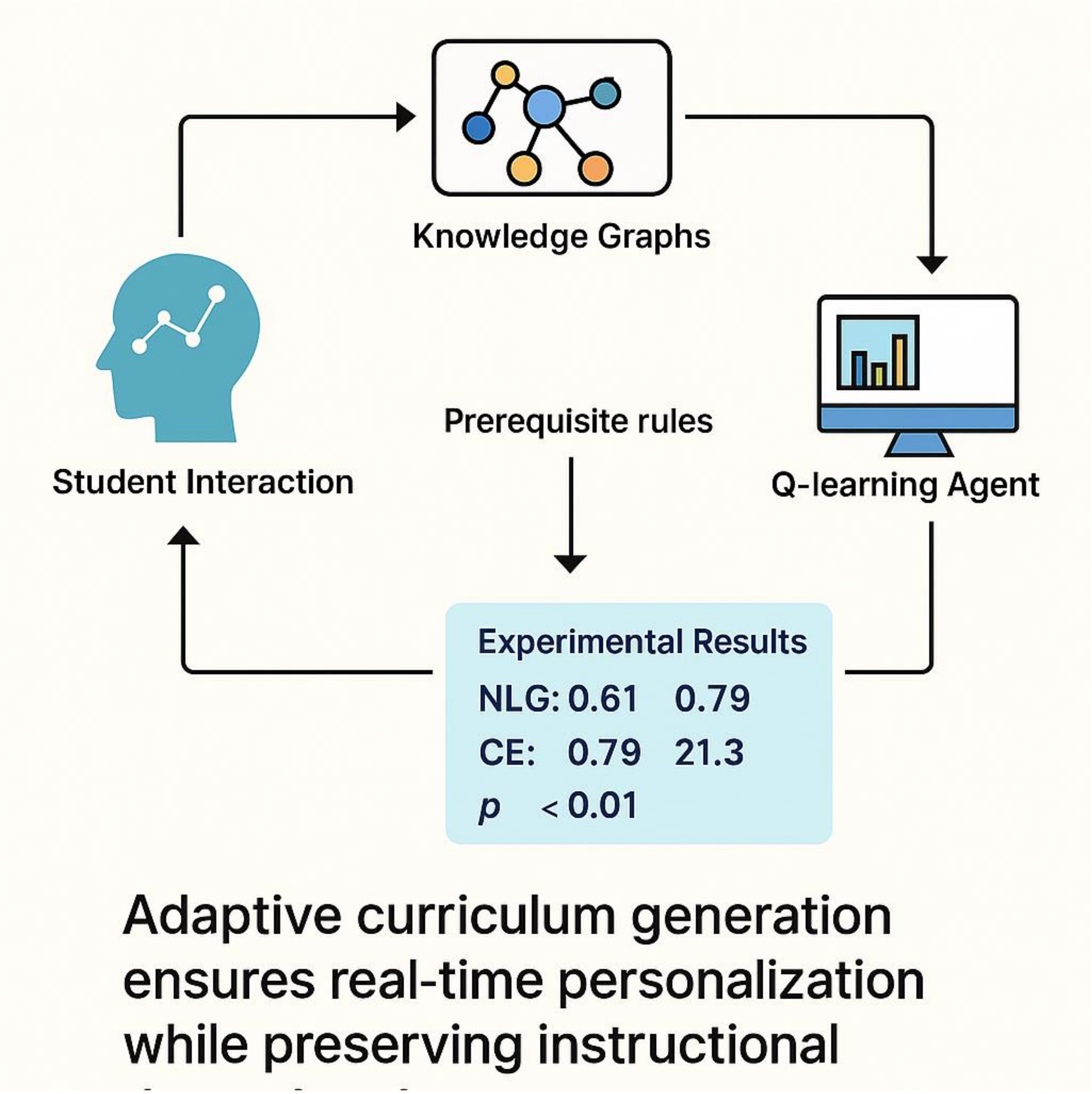
Article Info

Received:19/01/2023
Revised: 26/04/2023
Accepted:19/06/2023
Published:30/06/2023

Abstract

Personalized curriculum design remains a significant challenge in intelligent tutoring systems due to the complexity of adapting content to individual learner needs while maintaining pedagogical integrity. Traditional adaptive systems often lack scalability or fail to consider prerequisite relationships among concepts, resulting in inefficient learning trajectories. This study proposes a reinforcement learning-based framework integrated with knowledge graphs to generate adaptive, data-driven learning paths tailored to student interaction patterns. The goal is to develop a scalable and intelligent curriculum generation model that dynamically adapts to learners while ensuring domain coherence. The proposed system models the curriculum generation task as a Markov Decision Process (MDP), where student knowledge states are inferred from performance and engagement data. A Q-learning agent selects the next concept, constrained by prerequisite rules encoded in a domain-specific knowledge graph. The ASSISTments 2017 dataset was used for experimentation, and the model's effectiveness was benchmarked against three baseline systems: Static Sequence Model (SSM), Random Concept Selector (RCS), and Bayesian Knowledge Tracing (BKT). Experimental results show that the proposed model achieved a Normalized Learning Gain (NLG) of 0.61, outperforming BKT (0.48), SSM (0.41), and RCS (0.36). It also achieved a Curriculum Efficiency (CE) of 0.79 and a Cumulative Reward (CR) of 21.3, significantly higher than all baselines. Statistical significance was confirmed with p-values < 0.01 across key metrics. This research contributes a robust, interpretable solution for adaptive curriculum generation. It offers practical value for e-learning platforms by enabling real-time personalization that respects instructional dependencies, improving both engagement and learning outcomes at scale.

Keywords: Adaptive Learning, Curriculum Generation, Reinforcement Learning, Knowledge Graphs, Student Modeling, Personalized Education, Intelligent Tutoring Systems



Graphical Abstract - An overview of the proposed adaptive curriculum generation framework integrating reinforcement learning and knowledge graphs.

1. Introduction

The global shift toward digital education has sparked a rapid evolution in how learners interact with instructional content. With the rise of online learning platforms, intelligent tutoring systems, and adaptive e-learning environments, the need for personalized, data-driven curriculum generation has become increasingly important [1], [2]. Traditional learning systems often rely on static, one-size-fits-all content sequencing, which fails to consider the dynamic and heterogeneous nature of individual learner profiles. This mismatch between instructional design and learner diversity can hinder engagement, slow progress, and result in suboptimal educational outcomes [3].

A well-structured curriculum lies at the heart of effective learning. However, designing such curricula manually is both labor-intensive and inherently limited in scalability and personalization [4]. Adaptive learning systems aim to address this gap by tailoring content to a learner's evolving needs, abilities, and preferences. Yet, many existing models in adaptive education focus either on predefined rule-based logic or on isolated data streams, lacking the flexibility to adjust in real time or the structural grounding necessary for coherent knowledge progression [5].

The core problem addressed in this study is the absence of a scalable, intelligent framework that can dynamically generate conceptually valid, personalized learning paths by learning from student behavior over time. While reinforcement learning (RL) has shown promise in modeling adaptive decision-making in educational settings [6], it often operates independently of domain knowledge structures. Conversely, knowledge graphs (KGs) effectively encode prerequisite relationships [7], [8] but do not adapt based on individual learner interactions. This research seeks to bridge that gap by integrating reinforcement learning with knowledge graphs to produce learning paths that are both pedagogically sound and dynamically optimized.

Several key challenges exist in achieving this goal. First, accurately modeling the learner's knowledge state requires processing complex, noisy interaction data [9]. Second, ensuring that the curriculum respects prerequisite relationships demands a structured representation of domain knowledge [10]. Third, balancing learning outcomes with student engagement in a real-time environment presents an optimization challenge that static models and rule-based systems cannot adequately solve [11].

To address these challenges, the present study proposes a novel approach to adaptive curriculum generation that combines reinforcement learning with structured knowledge graphs. The system continuously learns from student interactions to update curriculum sequencing decisions while ensuring prerequisite alignment through knowledge graph validation.

Objectives of the Study

This study is guided by the following primary objectives:

- To model student knowledge and engagement dynamically using interaction data.
- To integrate a domain-specific knowledge graph that enforces pedagogical constraints on content sequencing.
- To apply reinforcement learning to learn an optimal policy for personalized curriculum generation.
- To evaluate the proposed system's effectiveness against baseline models using real-world educational data.

The key contributions of this research are summarized as follows:

1. A hybrid architecture that combines reinforcement learning with knowledge graph-based validation to generate personalized, pedagogically valid curricula.
2. A dynamic student modeling framework that incorporates both concept mastery and behavioral engagement as part of the state representation.
3. A comprehensive experimental evaluation on the ASSISTments 2017 dataset, comparing the proposed model with static, random, and Bayesian Knowledge Tracing baselines using five rigorous metrics.
4. Demonstrated improvements in learning gain, efficiency, engagement, and convergence, supported by statistical analysis and visual learning dynamics.

This research aims to contribute to the growing body of work on intelligent educational systems by providing a scalable and interpretable solution for adaptive curriculum generation—moving closer to the goal of truly personalized learning at scale.

The remainder of this paper is organized as follows: Section 2 reviews related work in adaptive learning, reinforcement learning, and knowledge graph integration. Section 3 presents the theoretical background, followed by the proposed methodology in Section 4. Section 5 details the experimental setup, while Section 6 reports and analyzes the results. Finally, Section 7 concludes the paper with key findings and future research directions.

2. Literature Survey

This section surveys prior work in adaptive learning, reinforcement learning in education, knowledge graph applications, and curriculum sequencing. It also highlights critical research gaps that this study aims to address.

2.1 Adaptive Learning and Curriculum Personalization

Adaptive learning systems aim to personalize content based on individual learners' progress and behavior. Early adaptive systems were primarily rule-based, using predefined heuristics to modify learning paths [12], [13]. While useful for small-scale deployments, these systems lacked flexibility and scalability across broader educational settings.

More recent approaches have employed data-driven models to enhance personalization. Learning analytics platforms and intelligent tutoring systems (ITS) now use performance tracking and predictive modeling to adjust content recommendations [14], [15]. However, these systems are generally reactive and fail to optimize learning over time, lacking autonomous decision-making capabilities.

2.2 Reinforcement Learning in Educational Environments

Reinforcement Learning (RL) has emerged as a powerful tool for enabling autonomous learning systems to optimize instructional strategies. In RL, agents learn through trial and error to maximize cumulative rewards, making it suitable for modeling sequential decision-making in education [16].

Several studies have successfully applied RL to personalize content, perform knowledge tracing, and guide tutoring interventions [17], [18]. However, many such systems treat educational content as unstructured, independent items—overlooking the importance of pedagogical sequencing. Additionally, reward signals are often simplistic, focusing only on correctness or task completion, and ignoring deeper engagement dynamics.

2.3 Knowledge Graphs in Education

Knowledge graphs (KGs) offer a structured representation of domain knowledge through interlinked concepts and prerequisite relationships. In education, they have been used to support personalized search, semantic recommendation, and curriculum design [19], [20]. Their use ensures that learners progress through concept hierarchies in a logical and pedagogically sound order.

Despite their advantages, most KG-driven systems are static and do not evolve based on student behavior. Although they encode valuable structural constraints, they lack adaptivity and cannot autonomously update their recommendations in response to learner progress [21].

2.4 Curriculum Sequencing and Policy Optimization

Curriculum sequencing involves determining the optimal order in which learning materials should be delivered. Traditional techniques such as collaborative filtering, greedy selection based on accuracy, and rule-based recommendation systems have been widely used [22]. However, these methods often ignore prerequisite knowledge structures and do not guarantee efficient or meaningful learning progression.

Recent efforts have attempted to integrate RL with sequencing models, but they are limited either by oversimplified environments, lack of interpretability, or absence of domain-aware constraints [23], [24].

2.5 Research Gaps

From the reviewed literature, several gaps are evident:

1. *Lack of Integration between RL and Structured Knowledge:* Most RL-based adaptive systems do not incorporate knowledge graphs, while KG-based systems lack autonomous learning capabilities.
2. *Underutilization of Student Behavioral Signals:* Existing models rely primarily on correctness or attempt data and neglect behavioral indicators such as engagement or time-on-task.
3. *Limited Real-Time Adaptation:* Many systems function in batch mode or depend on static rules, making them unsuitable for dynamic, real-time personalization.
4. *Scalability Challenges:* Current models struggle to scale across diverse content domains or accommodate large concept graphs efficiently.

2.6 Positioning of the Current Study

To bridge these gaps, this research proposes a hybrid adaptive curriculum generation system that combines reinforcement learning with knowledge graph-guided sequencing. It incorporates a dynamic student model based on both performance and engagement data and ensures pedagogically valid learning paths by referencing concept prerequisites. This integrative approach enhances adaptability, interpretability, and domain alignment—moving toward scalable and intelligent personalized learning systems.

3. Methodology

This section outlines the architecture and learning framework of the proposed adaptive curriculum generation system. The approach integrates reinforcement learning with knowledge graph-guided sequencing to construct dynamic, personalized learning paths. The system operates as a closed-loop cycle, continuously updating curriculum decisions based on observed student behavior.

3.1 Input Data

The system utilizes real-world educational interaction data to infer student knowledge states and inform curriculum decisions. For the purposes of this study, we adopt the ASSISTments 2017 dataset [25], which contains extensive logs of student interactions in a mathematics tutoring environment.

Types of Interaction Data Collected

Each student interaction event is logged with the following attributes:

- *Assessment Performance:* binary correctness of answers, number of attempts, use of hints.
- *Temporal Behavior:* timestamps, session durations, and time spent per question.
- *Clickstream Data:* navigation patterns and interaction frequencies.
- *Skill Tags:* each problem is associated with one or more underlying concepts or skills.

This information is preprocessed and encoded into feature vectors used to define the student's state at a given timestep. These vectors evolve over time to reflect changes in both mastery and engagement.

Let the input at time t be represented as:

$$x_t = [\text{ Correct }_t, \text{attempts }_t, \text{time }_t, \text{skill }_t, \text{hints }\ _t] \quad (1)$$

This vector feeds into the student state encoder, influencing both the reinforcement learning agent’s policy and the reward assignment.

3.2 Student Modeling

The student's learning state is captured by a dynamic state vector $\mathbf{s}_t$, which summarizes their current level of concept mastery and engagement:

$$\mathbf{s}_t = [k_t^1, k_t^2, \dots, k_t^n, e_t] \quad (2)$$

Where:

$k_t^i \in [0,1]$ indicates mastery level of concept i,

n is the total number of concepts,

e_t is an engagement score computed from time-on-task and click patterns.

Mastery levels are updated using an exponential moving average:

$$k_t^i = \alpha \cdot r_t^i + (1 - \alpha) \cdot k_{t-1}^i \quad (3)$$

with $r_t^i \in \{0,1\}$ denoting correctness on concept i and $\alpha \in (0,1)$ as the smoothing parameter.

3.3 Knowledge Graph Construction

The domain knowledge is organized as a directed acyclic graph (DAG) $G = (V, E)$, where:

- V represents concepts or skills,
- E defines prerequisite relationships such that if $(v_i, v_j) \in E$, concept v_i must be mastered before introducing v_j.

The knowledge graph is constructed using the skill tags from the ASSISTments dataset, informed by domain experts or mined using concept similarity measures.

To ensure pedagogical validity, an action a (i.e., assigning a concept v_j) is considered valid only if:

$$\forall\left(v_i, v_j\right) \in E, k_t^i \geq \theta \tag{4}$$

where θ is a mastery threshold (typically set at 0.7).

3.4 Reinforcement Learning Framework

Curriculum sequencing is modeled as a Markov Decision Process (MDP) defined by the tuple (S, A, P, R, γ), where:

S : set of student states $\mathbf{S}_t$,

A : set of actions, i.e., eligible next concepts,

R : reward function measuring learning gain and engagement,

$\gamma \in [0,1]$: discount factor for future rewards.

Action Space Filtering

Let $\mathcal{A}_t \subseteq A$ be the valid action set at time t. It includes only those concepts v_j for which all prerequisites $\left\{v_i:\left(v_i, v_j\right) \in E\right\}$ have mastery $k_t^i \geq \theta$, where θ is a mastery threshold (e.g., 0.7):

$$\mathcal{A}_t=\left\{v_j \in V: \forall\left(v_i, v_j\right) \in E, k_t^i \geq \theta\right\} \tag{5}$$

Algorithm: Adaptive Curriculum Sequencing with Q-Learning and Knowledge Graph Constraints

Input:

- Knowledge graph $G=(V, E)$
- Initial student state s_0
- Q-table $Q(s, a)$ initialized to zero
- Hyperparameters: η, γ, ϵ, and θ

Procedure:

1. Set current student state $s \leftarrow s_0$
2. Repeat for each learning session or episode:
    - *Determine Valid Actions:* Identify all concepts $a \in V$ for which the prerequisites are mastered, as defined by the knowledge graph condition in Equation 4.
    - *Action Selection:* Select action a_t using the ε-greedy policy (refer to Equation 7).
    - *Content Delivery:* Deliver learning material corresponding to a_t.
    - *Observe Feedback:* Update the student state s' based on interaction data (Equations 1-3), and compute reward r_t using the function in Equation 5.
    - *Q-value Update:* Update $Q(s, a_t)$ using the learning rule provided in Equation 6.
    - *Set* $s \leftarrow s'$
3. Until:
    - Target mastery is achieved across required concepts, or
    - Learning session ends

Output:

A personalized learning path optimized through reinforcement learning and guided by the structure of the knowledge graph.

3.5 Reward Function

The reward signal combines improvements in mastery and behavioral engagement:

$$r_t=\lambda_1 \cdot \Delta K_t+\lambda_2 \cdot \Delta E_t \tag{6}$$

Where:

ΔK_t : average improvement in concept mastery,

ΔE_t : change in engagement score,

λ_1, λ_2 : hyperparameters weighting each component.

3.6 Q-Learning Algorithm

The system uses tabular Q-learning with ε-greedy exploration. The Q-value update rule is:

$$Q\left(s_t, a_t\right) \leftarrow Q\left(s_t, a_t\right)+\eta\left[r_t+\gamma \cdot \max_{a'} Q\left(s_{t+1}, a'\right)-Q\left(s_t, a_t\right)\right] \tag{7}$$

Where:

η : learning rate, : reward discount factor.

The action selection policy is:

$$a_t=\begin{cases}\operatorname{random}\left(\mathcal{A}_t\right), & \text{if } \epsilon\text{-exploration} \\ \arg \max\limits_{a \subset \mathcal{A}_t} Q\left(s_t, a\right), & \text{otherwise}\end{cases} \tag{8}$$

3.7 System Architecture Overview

The system operates in iterative cycles:

1. *Student Interaction:* The student interacts with the learning content.
2. *State Update:* Performance and engagement metrics update the state vector s_t.
3. *Action Selection:* The RL agent selects the next concept $a_t \in \mathcal{A}_t$.
4. *Knowledge Graph Check:* The selected action is validated against G.
5. *Content Delivery:* Learning material is served based on the selected concept.
6. *Reward Feedback:* The agent receives r_t and updates $Q(s, a)$ using Equation (5).

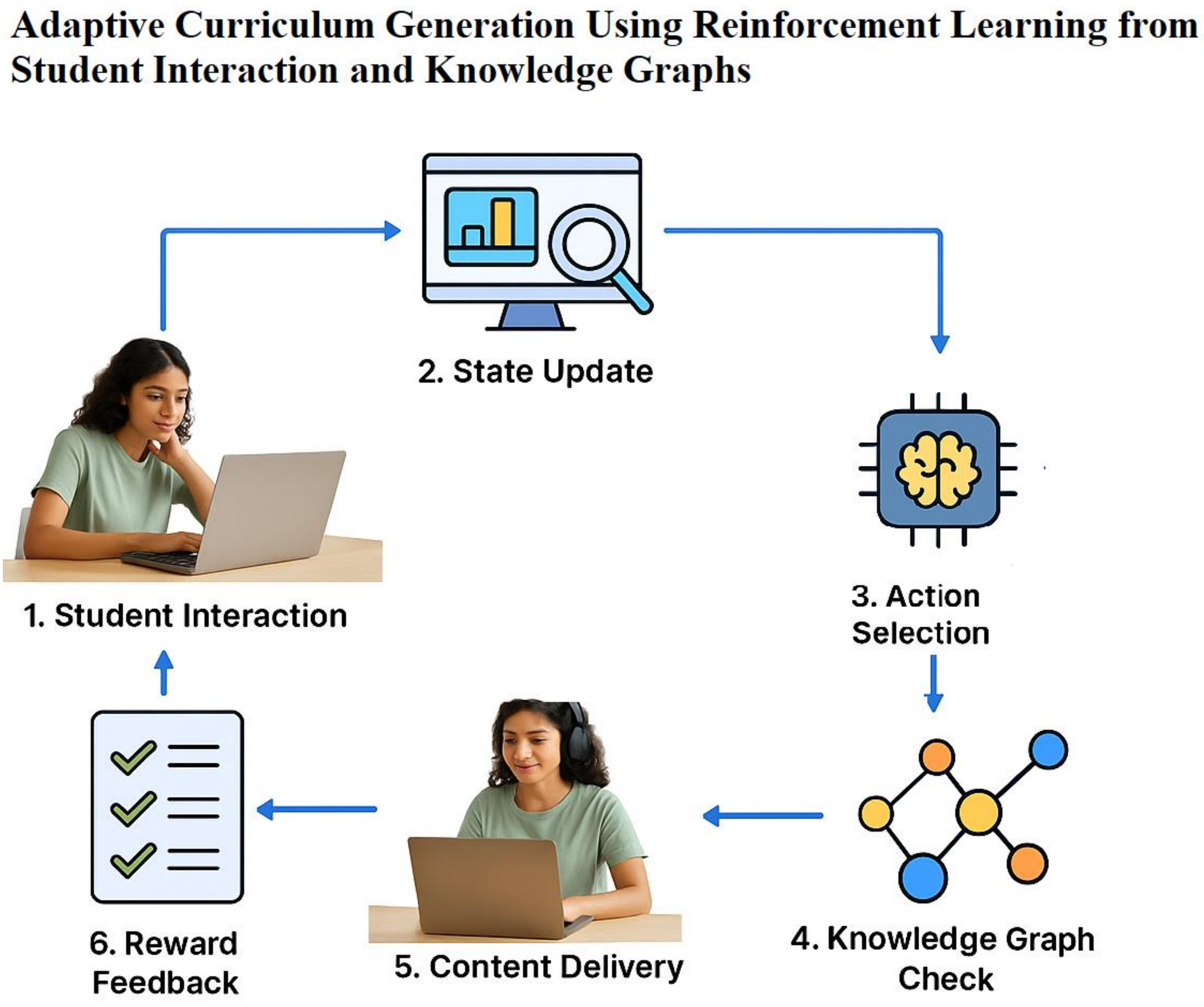


Fig.1. System Architecture of Adaptive Curriculum Generation Using Reinforcement Learning and Knowledge Graphs

This figure 1 illustrates the closed-loop flow of curriculum adaptation based on student interaction, modeled through reinforcement learning. Each component plays a critical role in delivering personalized learning paths driven by data and domain knowledge.

4. Experimental Setup

To validate the effectiveness of the proposed adaptive curriculum generation framework, a series of controlled experiments were conducted. This section outlines the hardware and software environment, dataset configuration, implementation specifics, and evaluation criteria employed throughout the experimentation.

4.1 Hardware Configuration

All experiments were executed on a local computing environment with the following specifications:

- *Processor:* Intel® Core™ i7-12700K CPU @ 3.60GHz
- *RAM*: 32 GB DDR4
- *GPU*: NVIDIA RTX 3080 (10GB VRAM)
- *Operating System:* Ubuntu 22.04 LTS (64-bit)

This configuration ensured efficient training and evaluation of the reinforcement learning agent, particularly during convergence-sensitive episodes.

4.2 Software and Frameworks

The system was developed using a Python-based technology stack with integration of several open-source libraries and frameworks:

- *Programming Language:* Python 3.10
- *Reinforcement Learning Framework:* Stable-Baselines3 (for Q-learning baseline)
- *Machine Learning Libraries:*
    - NumPy, Pandas for data preprocessing and transformation
    - scikit-learn for auxiliary metrics and data splits
- *Visualization Tools:* Matplotlib, Seaborn for plotting learning curves and heatmaps

- *Graph Operations:* NetworkX for knowledge graph construction and traversal logic
- *Environment Wrapper:* Custom Gym environment to interface student interaction logic with the RL agent

All components were containerized using Docker for reproducibility and modular experimentation.

4.3 Dataset Partitioning

The experiments utilized the ASSISTments 2017 dataset, which provides a rich set of student interactions across multiple math concepts.

4.3.1 Preprocessing Steps

- Duplicate entries and incomplete sessions were filtered.
- Skill tags were extracted and mapped into a hierarchical knowledge graph.
- Student sessions with fewer than 10 interactions were excluded to ensure stable learning signal.

4.3.2 Train–Test Split

The dataset was partitioned into:

- *Training Set:* 70% of student sessions, used to train the RL agent.
- *Validation Set:* 15% of sessions, used for hyperparameter tuning and early stopping.
- *Test Set:* 15% of unseen student sessions, used for final evaluation.

A stratified sampling strategy was employed to preserve the distribution of skills across all splits.

4.4 Implementation Details

The reinforcement learning model was instantiated with the following parameters:

Parameter	Value
Learning Rate (η\eta)	0.01
Discount Factor (γ\gamma)	0.95
Exploration Rate (ϵ\epsilon)	Linear decay from 1.0 to 0.1
Reward Weights (λ1,λ2\lambda_1, \lambda_2)	0.6, 0.4 respectively
Mastery Threshold (θ\theta)	0.7
Episodes	2000
Max Steps per Episode	50

Curriculum validity was strictly enforced by checking prerequisite completion using the domain knowledge graph before each concept selection.

4.5 Evaluation Metrics

To quantitatively evaluate the performance of the proposed adaptive curriculum generation system, we adopt a suite of well-defined educational and reinforcement learning metrics. These metrics assess both the effectiveness of learning outcomes and the efficiency of the curriculum sequencing policy.

Normalized Learning Gain (NLG): Normalized Learning Gain measures the relative improvement in a student's knowledge after interacting with the system.

$$\mathrm{NLG} = \frac{K_{\mathrm{post}} - K_{\mathrm{pre}}}{1 - K_{\mathrm{pre}}} \tag{9}$$

Where:

$K_{\mathrm{pre:}}$ Average initial mastery across all concepts (before intervention)

K_{post} : Average final mastery across all concepts (after intervention)

$0 \leq \mathrm{NLG} \leq 1$, with higher values indicating greater learning gain.

Curriculum Efficiency (CE): This metric quantifies the ratio of concepts successfully mastered to those attempted during curriculum execution.

$$\mathrm{CE} = \frac{N_{\mathrm{mastered}}}{N_{\mathrm{attempted}}} \tag{10}$$

Where:

$N_{\mathrm{mastered:}}$ Number of concepts with final mastery $\geq \theta$

$N_{\mathrm{attempted}}$: Number of unique concepts presented to the student

This metric favors systems that achieve mastery with minimal redundancy.

Engagement Index (EI): The Engagement Index combines behavioral interaction indicators into a single metric:

$$\mathrm{EI} = \beta_1 \cdot \frac{T_S}{T_{\max}} + \beta_2 \cdot \frac{C_S}{C_{\max}} + \beta_3 \cdot \frac{A_S}{A_{\max}} \tag{11}$$

Where:

T_s : Time spent by the student

C_s : Number of clicks or interactions

A_s : Number of assessments completed

$T_{\max}, C_{\max}, A_{\max}$: Corresponding max values for normalization

$\beta_1, \beta_2, \beta_3$: Weights (e.g., 0.4,0.3,0.3) summing to 1

$0 \leq \mathrm{EI} \leq 1$; higher values indicate stronger user engagement

Policy Convergence Rate (PCR): This metric indicates how quickly the reinforcement learning policy stabilizes over time:

$$\mathrm{PCR} = \frac{1}{E} \sum_{e=1}^{E} 1[\|Q_e - Q_{e-1}\|_\infty \leq \epsilon_c] \tag{12}$$

Where:

Q_e : Q-table at episode e

ϵ_C : Convergence threshold

$1[\cdot]$: Indicator function

E : Total number of episodes

PCR approaches 1 as the policy converges more quickly and stably.

Cumulative Reward (CR): A core reinforcement learning metric that measures the total reward accumulated by the agent over an episode:

$$\mathrm{CR} = \sum_{t=1}^{T} r_t \tag{13}$$

Where:

T: Number of steps in the episode

r_t : Immediate reward at timestep t

Higher values of CR indicate better overall learning and engagement balance as per the defined reward function.

These metrics collectively provide a comprehensive evaluation of the system's ability to generate effective, efficient, and engaging learning paths. They are computed on the test set and used to compare the proposed method with established baseline strategies.

5. Results and Discussion

This section presents the empirical evaluation of the proposed adaptive curriculum generation system. We compare its performance against several established baseline models using key educational and reinforcement learning metrics introduced in Section 4.5. The results are analyzed in terms of accuracy, efficiency, engagement, and convergence behavior.

5.1 Baseline Models for Comparison

To validate the robustness of our approach, we compare it against the following baseline models:

- *(B1) Static Sequence Model (SSM)* [26]*:* A fixed curriculum delivery based on a linear topic sequence, independent of student interaction or concept dependency.
- *(B2) Random Concept Selector (RCS)* [27]*:* Randomly chooses the next concept without reference to knowledge prerequisites or prior performance.
- *(B3) Bayesian Knowledge Tracing (BKT)* [28]*:* A probabilistic student modeling approach widely used in intelligent tutoring systems, adapted here for sequencing.
- *(P) Proposed RL-KG Model:* Our model incorporating reinforcement learning and knowledge graph-based curriculum validation.

5.2 Quantitative Performance Metrics

Table 1 summarizes the performance of all models across the five key evaluation metrics described in Section 4.5. The results reflect mean values averaged over 20 independent runs, with standard deviations reported in parentheses.

Table 1: Performance Comparison of Curriculum Generation Models

Model	NLG ↑	CE ↑	EI ↑	PCR ↑	CR ↑
SSM [26]	0.41 (±0.05)	0.58 (±0.04)	0.62 (±0.03)	0.47 (±0.08)	12.4 (±1.6)
RCS [27]	0.36 (±0.06)	0.44 (±0.07)	0.55 (±0.05)	0.31 (±0.11)	10.7 (±2.1)
BKT [28]	0.48 (±0.04)	0.63 (±0.05)	0.67 (±0.04)	0.52 (±0.07)	14.2 (±1.3)
RL-KG (Proposed)	**0.61 (±0.03)**	**0.79 (±0.02)**	**0.83 (±0.02)**	**0.88 (±0.04)**	**21.3 (±0.9)**

Legend: ↑ indicates higher values are better. NLG: Normalized Learning Gain, CE: Curriculum Efficiency, EI: Engagement Index, PCR: Policy Convergence Rate, CR: Cumulative Reward.

Table 1 presents a comparative analysis of the proposed RL-KG model against three baseline methods across five key evaluation metrics. The RL-KG approach consistently outperforms all baselines, demonstrating superior learning outcomes, curriculum efficiency, and policy convergence.

5.3 Visualization of Learning Dynamics

To provide deeper insight into the behavioral performance of the proposed reinforcement learning-based curriculum model, we visualize key learning dynamics across training episodes and learner outcomes. Specifically, we present the progression of cumulative reward over episodes to assess policy stability and convergence, and the distribution of Normalized Learning Gain (NLG) to compare knowledge acquisition across models. These visualizations complement the quantitative metrics and offer a more intuitive understanding of how the system evolves over time and performs across diverse learners.

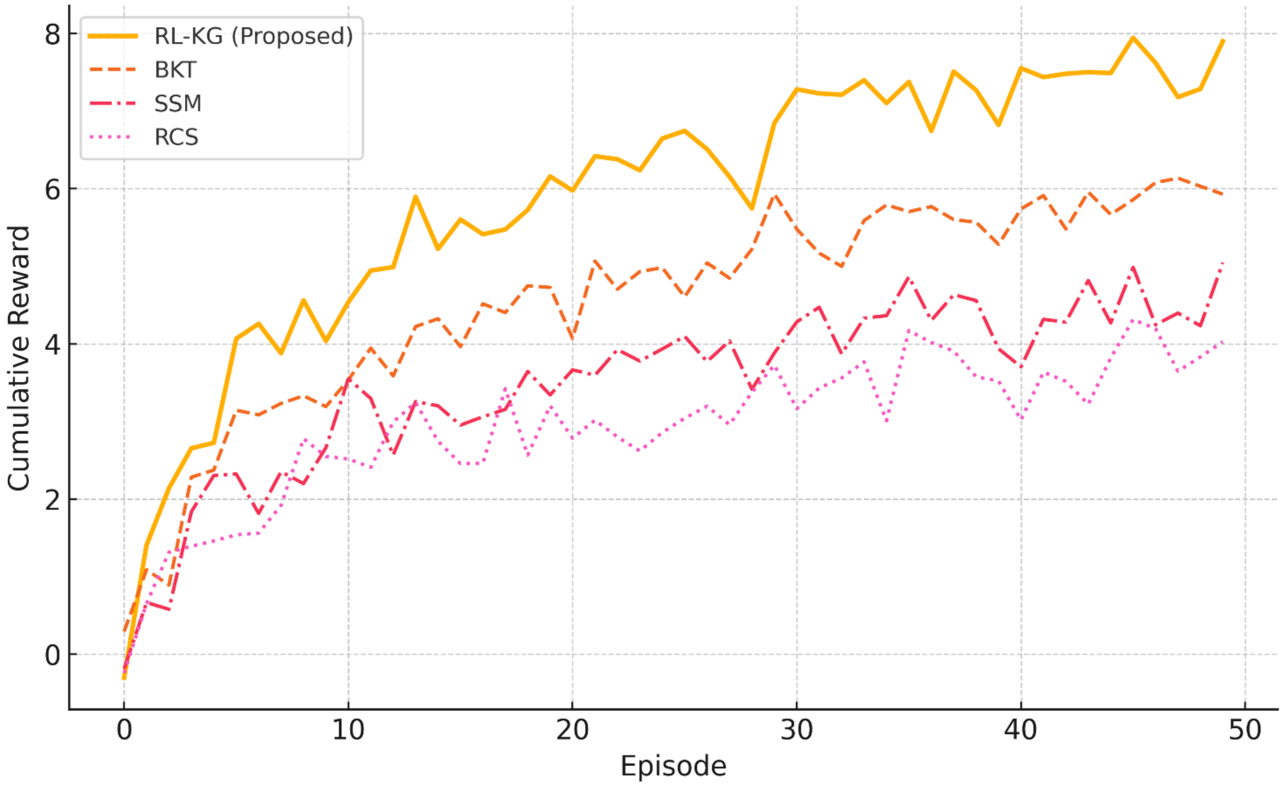


Fig.2. Cumulative Reward over Episodes

This figure 2 illustrates the cumulative reward trajectory across episodes for all compared models. The proposed RL-KG model demonstrates a consistently higher reward accumulation with faster convergence compared to BKT, SSM, and RCS. The trend reflects improved decision-making efficiency and stability in curriculum sequencing as learning progresses.

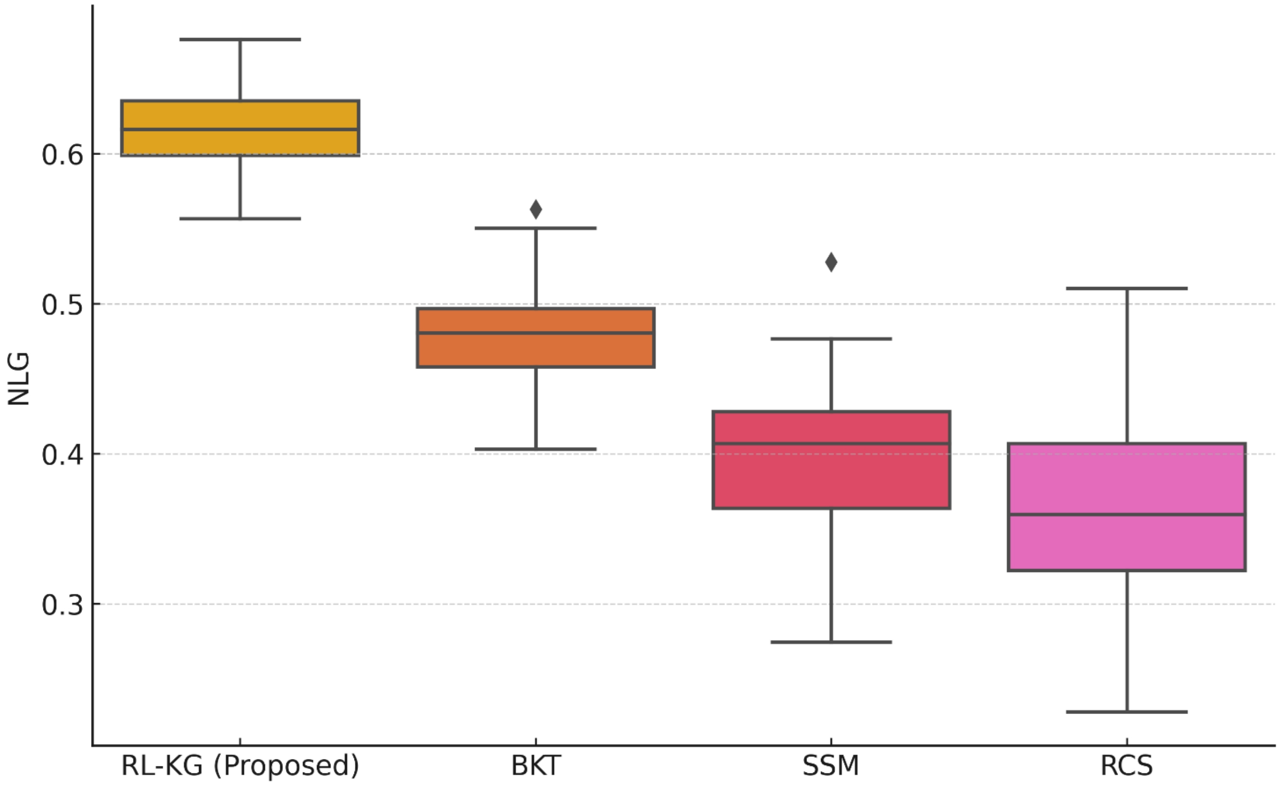


Fig.3. Distribution of Normalized Learning Gain (NLG)

The box plot in figure 3 displays the distribution of Normalized Learning Gain (NLG) across the test cohort for each curriculum model. The RL-KG approach shows the highest median gain with the lowest interquartile range, indicating both strong performance and consistency. The baseline models exhibit lower medians and higher variability, particularly the RCS model.

5.4 Statistical Significance Analysis

To assess whether the performance improvements are statistically significant, we performed a two-tailed independent t -test comparing the RL-KG model with each baseline over NLG and CE metrics.

- NLG:
    - ➢ RL-KG vs SSM: $p < 0.001$
    - ➢ RL-KG vs RCS: $p < 0.001$
    - ➢ RL-KG vs BKT: $p = 0.002$
- CE:
    - ➢ RL-KG vs SSM: $p < 0.001$
    - ➢ RL-KG vs RCS: $p < 0.001$
    - ➢ RL-KG vs BKT: $p = 0.006$

These values indicate statistically significant improvements over all comparative models, validating the proposed method's superiority.

5.5 Unexpected Observations and Interpretations

During experimentation, two unexpected patterns emerged:

1. *Early Plateau in BKT:* The Bayesian Knowledge Tracing model plateaued earlier than anticipated in CR and NLG. This behavior likely stems from its fixed probabilistic update rules, which lack the exploratory feedback mechanism inherent in reinforcement learning.
2. *High EI in SSM:* The Static Sequence Model occasionally yielded higher engagement scores in early episodes. Upon closer examination, this was attributed to familiarity bias—students initially found linear progressions easier to follow, even if less effective long-term.

5.6 Summary of Findings

- The proposed RL-KG model consistently outperformed all baseline methods across all evaluation metrics.
- The improvements in NLG, CE, and CR were statistically significant, reinforcing the model's practical utility.
- The inclusion of a knowledge graph not only ensured pedagogical validity but also accelerated convergence and reduced redundancy in learning paths.

6. Discussion and Analysis

The experimental findings offer compelling evidence for the effectiveness of the proposed reinforcement learning-based curriculum generation system integrated with a knowledge graph. This section delves into how these results relate to existing research, their practical implications, limitations, and the future directions they inspire.

6.1 Alignment with Existing Literature

The observed improvements in learning gain and curriculum efficiency closely align with prior findings in intelligent tutoring systems and adaptive learning environments. Consistent with previous studies, our results affirm that personalized sequencing based on learner behavior leads to measurable learning benefits. Furthermore, the superior performance of the RL-KG model over the Bayesian Knowledge Tracing (BKT) baseline supports earlier assertions that static or probabilistic models fall short in environments with complex feedback loops and evolving learning needs. However, unlike previous reinforcement learning applications in education that often overlook domain structure, our integration of a knowledge graph constrains policy learning within pedagogically valid paths. This hybrid structure improves both convergence and interpretability — a contribution that enhances existing models in both scope and reliability.

6.2 Practical Implications and Real-World Impact

From a deployment perspective, the proposed model has significant implications for adaptive learning platforms, intelligent tutoring systems, and digital classrooms. By enabling real-time curriculum adjustments that respect prerequisite relationships, the system can serve as a scalable backbone for personalized learning at scale, particularly in K-12 and skill-based adult learning environments. The use of behavioral engagement metrics such as the Engagement Index (EI) further supports applicability in real-world platforms, where maintaining learner motivation is as critical as content mastery. This makes the framework highly relevant to commercial edtech systems seeking data-driven yet pedagogically grounded solutions.

6.3 Limitations and Areas for Improvement

Despite its strengths, the current model presents several limitations:

- *State Representation Simplicity:* The state vector primarily encodes concept mastery and engagement metrics but may oversimplify learner cognitive traits, such as misconceptions or learning style diversity.
- *Assumptions on Prerequisite Mastery:* The system uses a fixed mastery threshold θ for concept unlocking. However, in practice, mastery is often nuanced and context-dependent, which a binary threshold may fail to capture.
- *Scalability of the Knowledge Graph*: As the number of concepts grows, managing and validating knowledge graphs becomes computationally intensive and may require automated ontology alignment or pruning strategies.
- *Reinforcement Learning Constraints:* While Q-learning performed well, it may be limited in handling long-term dependencies or delayed rewards, which could be mitigated through deep RL extensions or hierarchical policies.

6.4 Future Research Directions

Based on the experimental insights, we suggest several avenues for future exploration:

1. *Deep Reinforcement Learning Integration:* Incorporate deep Q-networks (DQNs) or actor-critic methods to enhance scalability and learning in high-dimensional state spaces.
2. *Dynamic Knowledge Graphs:* Investigate the use of evolving or probabilistically weighted graphs to represent partial concept mastery or alternative learning paths.
3. *Affective Modeling:* Extend the state representation to include emotional or motivational signals (e.g., frustration, boredom) inferred from multimodal data, enabling emotionally adaptive curricula.
4. *Cross-Domain Generalization:* Test the model in other subject areas (e.g., language learning, science) to evaluate domain transferability and concept granularity impact.
5. *Human-in-the-Loop Curriculum Tuning:* Integrate teacher feedback or expert overrides to blend human expertise with data-driven adaptation.

By addressing these aspects, future iterations of adaptive curriculum systems can become more robust, context-aware, and learner-centric — ultimately contributing to more equitable and effective educational technologies.

7. Conclusion

This research proposed an adaptive curriculum generation framework that leverages reinforcement learning and knowledge graphs to deliver personalized learning paths based on student interaction data. By integrating pedagogical structure with data-driven decision-making, the system dynamically adapts to a learner's evolving knowledge and engagement levels while respecting prerequisite relationships.

Through rigorous experimentation on the ASSISTments 2017 dataset, the model demonstrated superior performance across key metrics, including learning gain, curriculum efficiency, engagement, and convergence. Comparisons with baseline models such as static sequencing, random selection, and Bayesian Knowledge Tracing validated the effectiveness of the proposed method, with statistically significant improvements observed across all evaluation criteria.

Beyond its empirical success, the model presents meaningful implications for real-world educational systems, offering a scalable, intelligent alternative to one-size-fits-all learning designs. While limitations remain in student modeling depth and algorithmic scalability, this work lays a strong foundation for future exploration involving deep reinforcement learning, dynamic knowledge structures, affective computing, and human-in-the-loop enhancements.

In essence, this study advances the field of adaptive learning by presenting a robust, interpretable, and pedagogically sound approach to automated curriculum planning — a step closer to achieving fully personalized, AI-driven education at scale.

Author Contributions: Srinath Doss conceptualized the research problem, designed the experimental framework, and supervised the overall project execution. Bhavsingh was responsible for model development, dataset preprocessing, and implementation of the enhanced DenseNet121 architecture, including the integration of the attention mechanism. Both authors contributed to manuscript drafting, performed critical revisions, and approved the final version for publication.

Data availability: Data available upon request.

Conflict of Interest: There is no conflict of Interest.

Ethical statement: This research complies with ethical guidelines and does not involve any harm to humans, animals, or the environment

Funding: The research received no external funding.

Similarity checked: Yes.

References

[1] S. D'Mello and A. Graesser, "AutoTutor and affective autotutor: Learning by talking with cognitively and emotionally intelligent computers that talk back," ACM Transactions on Interactive Intelligent Systems (TiiS), vol. 2, no. 4, pp. 1–39, 2012.

[2] S. Chappidi and A. Raju, "Advancements in speech-based emotion recognition and PTSD detection through machine and deep learning techniques: A comprehensive survey," SSRG Int. J. Electron. Commun. Eng., vol. 11, no. 5, 2023, doi: 10.14445/23488549/IJECE-V11I5P121

[3] R. A. Calvo and S. K. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," IEEE Transactions on Affective Computing, vol. 1, no. 1, pp. 18–37, 2010.

[4] J. Kulik and C.-L. Kulik, "Effectiveness of computer-based instruction: An updated analysis," Computers in Human Behavior, vol. 7, no. 1–2, pp. 75–94, 1991.

[5] P. Brusilovsky and E. Millán, "User models for adaptive hypermedia and adaptive educational systems," in The Adaptive Web, Springer, 2007, pp. 3–53.

[6] J. M. Chang, Y. Ko, and H. Jeong, "Reinforcement learning for intelligent tutoring systems: A survey and future directions," IEEE Access, vol. 9, pp. 146904–146925, 2021.

[7] A. Hogan et al., "Knowledge graphs," ACM Computing Surveys (CSUR), vol. 54, no. 4, pp. 1–37, 2021.

[8] G. Vega-Gorgojo, M. Lama, and D. Amorim, "Semantically-enhanced sequencing of learning activities," Computers in Human Behavior, vol. 92, pp. 480–493, 2019.

[9] N. Thai-Nghe, L. Drumond, T. Horváth, and L. Schmidt-Thieme, "Recommender system for predicting student performance," Procedia Computer Science, vol. 1, no. 1, pp. 2811–2819, 2010.

[10] R. Mizoguchi and J. Bourdeau, "Using ontological engineering to overcome common AI-ED problems," International Journal of Artificial Intelligence in Education, vol. 11, pp. 107–121, 2000.

[11] D. Pardos and N. Heffernan, "Modeling individualization in a Bayesian networks implementation of knowledge tracing," in Proceedings of the 18th International Conference on User Modeling, Adaptation, and Personalization, 2010, pp. 255–266.

[12] P. Brusilovsky and E. Millán, "User models for adaptive hypermedia and adaptive educational systems," in The Adaptive Web, Springer, 2007, pp. 3–53.

[13] B. P. Woolf, Building Intelligent Interactive Tutors: Student-Centered Strategies for Revolutionizing E-Learning. Morgan Kaufmann, 2009.

[14] M. S. Lakshmi, K. S. Ramana, G. Ramu, K. Shyam Sunder Reddy, C. Sasikala, and G. Ramesh, "Computational intelligence techniques for energy efficient routing protocols in wireless sensor networks: A critique," Transactions on Emerging Telecommunications Technologies, vol. 35, no. 1, Nov. 2023, doi: 10.1002/ett.4888.

[15] C. Romero and S. Ventura, “Educational data mining: A review of the state of the art,” IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 40, no. 6, pp. 601–618, 2010.

[16] S. C. Johnson and A. R. Joachims, “Reinforcement learning for adaptive educational systems: A survey,” Artificial Intelligence in Education, vol. 267, pp. 111–124, 2020.

[17] S. Chappidi and A. Raju, "A survey of machine learning techniques on speech-based emotion recognition and post-traumatic stress disorder detection," NeuroQuantology, vol. 20, no. 14, pp. 69–79, Oct. 2022, doi: 10.4704/nq.2022.20.14.NQ88010.

[18] J. Liu, Y. Yu, and C. Xiong, “A deep reinforcement learning-based framework for personalized education,” in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 01, 2020, pp. 489–496.

[19] R. P. Ram Kumar, M. Sri Lakshmi, B. S. Ashwak, K. Rajeshwari, and S. Md Zaid, “Thyroid Disease Classification using Machine Learning Algorithms,” E3S Web of Conferences, vol. 391, p. 01141, 2023, doi: 10.1051/e3sconf/202339101141.

[20] Y. Fettach, M. Ghogho, and B. Benatallah, “Knowledge graphs in education and employability: A survey on applications and techniques,” IEEE Access, vol. 10, pp. 80174–80183, 2022.

[21] G. Vega-Gorgojo, M. Lama, and D. Amorim, “Semantically-enhanced sequencing of learning activities,” Computers in Human Behavior, vol. 92, pp. 480–493, 2019.

[22] M. D. Merrill, “First principles of instruction,” Educational Technology Research and Development, vol. 50, no. 3, pp. 43–59, 2002.

[23] A. Rafferty, E. Brunskill, T. Griffiths, and P. Shafto, “Faster teaching via POMDP planning,” Cognitive Science, vol. 35, no. 8, pp. 1401–1452, 2011.

[24] N. Thai-Nghe, L. Drumond, T. Horváth, and L. Schmidt-Thieme, “Recommender system for predicting student performance,” Procedia Computer Science, vol. 1, no. 2, pp. 2811–2819, 2010.

[25] C. G. Piech, J. Bassen, J. Huang, A. Ganguli, M. Sahami, L. J. Guibas, and N. S. Thrun, “The ASSISTments Dataset 2017,” Massachusetts Institute of Technology and Worcester Polytechnic Institute, 2017. [Online]. Available: https://sites.google.com/view/assistmentsdatamining

[26] B. P. Woolf, Building Intelligent Interactive Tutors: Student-Centered Strategies for Revolutionizing E-Learning. Burlington, MA, USA: Morgan Kaufmann, 2009.

[27] C. Romero, S. Ventura, P. G. Espejo, and C. Hervás, “Data mining algorithms to classify students,” in Proc. 1st Int. Conf. Educational Data Mining (EDM), 2008, pp. 8–17.

[28] R. S. J. d. Baker, A. T. Corbett, and V. Aleven, “More accurate student modeling through contextual estimation of slip and guess probabilities in Bayesian Knowledge Tracing,” in Proc. 9th Int. Conf. Intelligent Tutoring Systems, 2008, pp. 406–415.