Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM)
Vol. 8, No. 3, November 2025, pp. 754 – 764
p-ISSN: 2614-3372 | e-ISSN: 2614-6150

Hybrid Reinforcement and Evolutionary Learning Model for Adaptive Pathway Optimization In Computer Networks Education

[1*]Sherly Rosa Anggraeni, [2]Dian Julianto Wahyudi, [3]Waode Yunia Silviariza, [4]Rachmy Rosyida Ro'is, [5]Narandha Arya Ranggianto
[1,2,3,4]Program Studi Perpustakaan, Jurusan Sastra Indonesia, Universitas Negeri Malang, Indonesia
[5]Program Studi Ilmu Perpustakaan, Fakultas Sastra, Universitas Jember, Indonesia
Email: [1]sherly.rosa.fs@um.ac.id, [2]ranggi@unej.ac.id, [3]dian.wahyudi.ft@um.ac.id, ,
[4]waode.yunia.fis@um.ac.id, [5]rachmy.rosyida.fik@um.ac.id

Article Info

Article history:
Received Sep 29th, 2025
Revised Oct 02nd, 2025
Accepted Nov 28th, 2025

Keyword:
Adaptive Learning
Deep Knowledge Tracing
Genetic Algorithm
Particle Swarm Optimization
Reinforcement Learning

ABSTRACT

This paper introduces a Hybrid Reinforcement and Evolutionary Learning Model developed to optimize adaptive learning pathways in computer network education. Traditional uniform curricula often struggle to accommodate diverse learner profiles, resulting in knowledge gaps across hierarchical concepts such as OSI layers, routing protocols, and security mechanisms. The proposed model integrates Deep Knowledge Tracing (DKT) with Long Short Term Memory (LSTM) networks for real-time estimation of learners' knowledge states, Proximal Policy Optimization (PPO) for dynamic sequential content selection, and a Genetic Algorithm Particle Swarm Optimization (GA–PSO) hybrid for global pathway refinement under constraints such as prerequisites and time limits. The model was evaluated using real learner data from an e-learning platform and achieved an average final mastery score of 0.867, quiz accuracy of 0.822, and an F1-score of 0.880 for path recommendations outperforming baseline models such as static curricula (0.740 mastery) and DKT+PPO (0.824 mastery) by 5–17%. Ablation studies validated the synergistic contribution of each component, with the GA–PSO module enhancing optimization efficiency by approximately 10%. Overall, these findings demonstrate that the proposed model offers superior personalization, learning efficiency, and adaptability, marking a significant advancement in AI-driven education for computer networks.

Corresponding Author:
Sherly Rosa Anggraeni
Universitas Negeri Malang, Malang, Indonesia
Jalan Semarang No. 5, Malang, Jawa Timur 65145
Email: sherly.rosa.fs@um.ac.id

DOI: http://dx.doi.org/10.24014/ijaidm.v8i3.38398

1. INTRODUCTION

The rapid evolution of computer networks as a cornerstone of modern technology underscores the urgent need for effective educational strategies in this field. Computer networks education covers a wide range of hierarchical and interdependent concepts, including the Open Systems Interconnection (OSI) model, routing protocols (e.g., OSPF, BGP), IP addressing, subnetting, and security mechanisms such as firewalls and intrusion detection systems (IDS). Mastery of these concepts requires learners to develop a strong foundational understanding before progressing to more advanced topics a process further complicated by diverse learner profiles. However, a major instructional challenge persists: the sequencing of these concepts is deeply hierarchical, meaning that misconceptions in earlier topics (such as subnetting) can propagate errors into advanced areas like routing and security. This dependency structure creates an urgent need for adaptive instructional models that can address heterogeneity in learner proficiency and prevent cascading conceptual gaps. Students enrolled in computer networks courses often demonstrate significant variation in prior knowledge, learning pace, and cognitive style. For example, some learners may have hands-on experience with basic networking protocols from professional environments, whereas others struggle with fundamental topics such as IP addressing or differentiating between the physical and data link layers. Traditional uniform curricula, which deliver content in a fixed linear sequence, often fail to accommodate these differences, resulting in knowledge gaps, lower engagement, and suboptimal learning outcomes. Recent learning analytics reports further indicate that inconsistent prior knowledge among students leads to substantial variation in concept mastery, especially in foundational topics such as IP addressing and OSI layers. These inconsistencies become persistent bottlenecks, making the absence of real-time adaptive sequencing a critical problem in computer network education. Research indicates that uniform instruction can lead to retention rates up to 30% lower in technical disciplines due to unaddressed learner variability [1].

Adaptive learning systems, driven by artificial intelligence (AI), present a compelling solution by dynamically tailoring content sequences to meet individual learner needs. These systems utilize data driven insights to adjust the pace, order, and complexity of instruction, thereby enhancing comprehension and long term retention. In the context of computer networks education where concepts are inherently sequential (for instance, understanding the physical layer is essential before mastering the network layer) adaptive learning ensures that prerequisite knowledge is reinforced before learner progress. Recent advances in AI, particularly in machine learning, have significantly improved the performance and adaptability of these systems. A systematic review reported that AI-based adaptive learning systems enhance student performance by 20–30% in STEM disciplines, with substantial gains in engagement and motivation [2]. Several recent studies reinforce these findings. Martin et al. (2020) highlight that adaptive systems consistently improve learning outcomes when learner variability is high [3]. Ghosh et al. (2020) demonstrate that attention-enhanced DKT architectures significantly improve temporal prediction accuracy [4]. Machkour et al. (2025) emphasize the need for adaptive assessment pathways in technical subjects with hierarchical complexity [5]. Likewise, Merino-Campos (2025) notes that despite advances in personalization, current systems remain limited in optimizing full learning pathways under prerequisite constraints [6]. These studies collectively reveal the potential of AI-driven adaptivity but also underscore the limitations of existing models when applied to domains requiring strict conceptual sequencing, such as computer networks. This is especially relevant for computer networks education, where conceptual bottlenecks such as misconceptions in subnetting can hinder progress in more advanced areas like routing protocols.

The emergence of intelligent tutoring systems (ITS) represents a major milestone in the application of AI to education. Early ITS implementations relied heavily on rule-based models, but modern systems employ machine learning to build more flexible and intelligent architectures. Learner models that capture cognitive traits, prior knowledge, and progress metrics form the foundation of these systems, enabling highly personalized and data-informed learning pathways [7]. However, traditional adaptive learning techniques such as Bayesian Knowledge Tracing (BKT) face limitations in domains like computer networks, where concepts are tightly interconnected. BKT assumes independence among concepts, which restricts its effectiveness in cases where mastery of one topic (e.g., IP addressing) is a prerequisite for another (e.g., routing). Deep Knowledge Tracing (DKT) overcomes these limitations by employing Long Short-Term Memory (LSTM) networks to model the temporal dynamics of learner interactions [8], such as quiz responses and task sequences. DKT captures nonlinear relationships and temporal dependencies, improving predictive accuracy by up to 15% over BKT on educational datasets [9]. Recent improvements further incorporate attention mechanisms that focus on the most relevant interactions, enhancing the precision of knowledge state estimation [10].

While DKT effectively models learner knowledge states, it remains inherently reactive, lacking mechanisms for proactive instructional decision-making. Reinforcement Learning (RL) bridges this gap by framing adaptive content delivery as a Markov Decision Process (MDP), where the state represents the learner's evolving knowledge vector (derived from DKT), actions correspond to content module selections, and rewards measure learning progress or time efficiency. Among RL algorithms, Proximal Policy Optimization (PPO) stands out as a robust and stable approach that uses clipped surrogate objectives to refine learning policies iteratively [11]. PPO's policy-gradient framework outperforms value-based methods such as Deep Q-Networks (DQN) which are prone to overestimation bias in stochastic environments by achieving up to 15% higher cumulative rewards in educational scenarios [12]. In computer networks education, PPO can dynamically prioritize learning modules, such as postponing advanced security topics until foundational networking concepts are well established, thereby balancing local adaptability with overall time efficiency.

Despite the strengths of RL, its tendency toward local optimization can result in suboptimal long-term learning paths, particularly in domains with complex structural constraints such as prerequisite hierarchies or time limitations. Evolutionary algorithms, including Genetic Algorithms (GA) and Particle Swarm Optimization (PSO), complement RL by enabling global optimization across vast and non convex search spaces. GA simulates natural selection processes selection, crossover, and mutation to explore diverse solutions and avoid convergence to local optima [13]. PSO, inspired by the collective behavior of flocks and swarms, refines candidate solutions through velocity updates based on both individual and global best positions, demonstrating superior performance in exploitation and fine-tuning [14]. Hybrid GA PSO models combine GA's exploratory power with PSO's convergence speed, enabling more efficient optimization in complex problems such as curriculum sequencing [15]. In educational applications, GA has been used for variable-length path planning in mathematics, while PSO enhances learning sequence refinement in adaptive systems [16]. However, the integration of DKT, RL, and GA–PSO remains underexplored particularly within domain-specific contexts like computer networks education, where prerequisite relationships and resource constraints play crucial roles [17]. Despite these advances, three major gaps remain unaddressed, most DKT-based approaches emphasize knowledge tracing but do not support proactive decision-making for next-module selection, reinforcement learning methods such as PPO improve local adaptivity but remain vulnerable to local optima in complex prerequisite structures, and evolutionary algorithms such as GA or PSO have been applied to curriculum optimization, but rarely in conjunction with sequence models and RL to ensure both local and global pathway coherence. These gaps are particularly important in computer networks education, where prerequisite enforcement and hierarchical dependency must be strictly maintained.

The proposed Hybrid Reinforcement and Evolutionary Learning Model integrates DKT, PPO, and GA PSO to address these challenges in computer networks education. DKT's LSTM-based architecture represents a learner's evolving knowledge as a hidden state vector, providing probabilistic mastery estimates for each concept (e.g., 0.8 for TCP/IP), which serve as the state inputs for the MDP used by PPO [18]. PPO optimizes the policy $\pi(a|s)$ to maximize expected rewards while maintaining stable updates in noisy and dynamic educational environments [19]. GA–PSO contributes by performing global path optimization using GA's crossover operations to combine promising learning sequences and PSO's velocity updates to refine them balancing mastery maximization with time efficiency [20]. This synergy effectively overcomes the limitations of standalone methods: addressing RL's myopic focus on local decisions, enhancing DKT's passive modeling capacity, and ensuring compliance with domain-specific dependencies such as mastering OSI Layer 1 (Physical) before Layer 2 (Data Link). The novelty of this research lies in its unified hybridization of DKT, PPO, and GA–PSO to create a fully integrated adaptive pathway optimization framework. Unlike prior studies that examine these components in isolation, the proposed model simultaneously performs precise knowledge-state estimation, real-time reinforcement-based module selection, and global pathway refinement under prerequisite constraints. To the best of our knowledge, no previous work has implemented this three-component synergy specifically for computer networks education, making this study a novel extension of existing adaptive learning research.

The significance of this model lies in its application to computer networks, a field that remains relatively underrepresented in adaptive learning research, which has predominantly focused on mathematics and language domains [21]. The hierarchical and interdependent nature of networking concepts demands precise sequencing to prevent cascading misconceptions for instance, misunderstanding subnetting can hinder comprehension of routing algorithms [22]. In this study, the researcher introduces the Hybrid Reinforcement and Evolutionary Learning Model for Adaptive Pathway Optimization in Personalized Learning Systems, an end-to-end framework that combines LSTM-based DKT for learner state tracking, PPO for sequential content selection, and GA PSO for periodic global refinement. The model continuously updates the learner's mastery vector through DKT, which informs the PPO agent's state. The PPO agent then selects the next learning module to maximize a composite reward that balances accuracy and efficiency. Periodically, GA PSO re optimizes the remaining curriculum sequence by accounting for domain-specific dependencies (e.g., prerequisite hierarchies such as mastering the OSI physical layer before the data link layer) and practical constraints (e.g., module duration).

This hybridization leverages DKT's accurate knowledge estimation, PPO's adaptive local decision-making, and GA PSO's holistic global optimization to overcome the shortcomings of prior studies that addressed these components in isolation [23]. When applied to a computer networks course, the model encodes interrelated topics such as routing, security, and network topologies, simulating learner quiz responses to train and validate adaptive pathways under realistic conditions. This domain-specific focus remains largely underexplored; yet, due to the technical depth and hierarchical dependencies of computer networks, such targeted adaptive learning approaches are essential for maximizing educational effectiveness [24]. Thus, this study offers both theoretical and practical contributions: theoretically by extending adaptive learning research through a hybrid AI-driven optimization model, and practically by providing a domain-specific solution capable of handling the hierarchical structure and prerequisite dependencies characteristic of computer networks education.

2. RESEARCH METHOD

This study develops and evaluates the Hybrid Reinforcement and Evolutionary Learning Model for Adaptive Pathway Optimization in Personalized Learning Systems, specifically designed for computer networks education, by integrating DKT, PPO, and a Genetic Algorithm Particle Swarm Optimization (GA PSO) hybrid. The research methodology encompasses the model architecture design, system training, and comprehensive evaluation, including comparisons with baseline algorithms to assess algorithmic performance in terms of precision, recall, F1-score, and Root Mean Square Error (RMSE). The implementation is built using Python, with TensorFlow for neural network modeling and DEAP for evolutionary computation, ensuring scalability, reproducibility, and adaptability for educational data environments. The model architecture, illustrated in Figure 1, integrates three core AI components designed to overcome the limitations of conventional standalone approaches. The DKT module utilizes a Long Short Term Memory (LSTM) network to model learners' evolving knowledge states by processing sequential data derived from quiz interactions, represented as input sequences $\langle q_1, a_1\rangle$, $\langle q_2, a_2\rangle$, ..., $\langle q_t, a_t\rangle$, where *q* denotes the question identifier and *a* represents the corresponding learner response (correct or incorrect). This sequential modeling captures temporal dependencies and knowledge progression, forming a latent mastery vector that informs the reinforcement learning agent. In the context of this study, each learner interaction pair $\langle q_t, a_t\rangle$ is transformed into an input vector x_t representing the question identity and correctness. The LSTM cell updates its hidden state using:

$$h_t = f(W_h h_t - 1 + W_x x_t + b) \tag{1}$$

where h_t encodes the learner's evolving knowledge state, W_h and W_x are weight matrices, and *f* denotes the activation functions governing input, forget, and output gates. The final sigmoid output layer produces a mastery probability:

$$y_t = \sigma(W_o h_t + b_o) \tag{2}$$

which indicates the likelihood that the learner will answer concept *k* correctly in the next interaction. This probabilistic mastery vector serves as the state input to the PPO agent, enabling real-time adaptive module selection.

Capturing temporal dependencies through an LSTM architecture with a sigmoid output layer that produces mastery probabilities ranging from 0 to 1. These probabilistic outputs represent the learner's Knowledge State, which is subsequently passed to the PPO agent for decision-making. The system diagram visualizes this flow from "Learner Interactions" through the "LSTM Network" within the DKT module, and onward to the PPO block. The PPO component formulates the learning path recommendation process as an MDP, where the state space (S) corresponds to the DKT-derived knowledge vector, the action space (A) represents available curriculum modules, and the reward function (R) quantifies learning effectiveness, balancing mastery improvement, time efficiency, and engagement consistency.

$$R = \Delta mastery - \lambda.\ time\_cost \tag{3}$$

where $\lambda = 0.1$, optimized through the clipped surrogate objective

$$L^{CLIP}(\theta) = E[min(r_t(\theta)A_t,\ clip(r_t(\theta), 1-\epsilon, 1+\epsilon)A_t)]\ with\ \epsilon = 0.2 \tag{4}$$

Ensuring stable policy updates through the interaction of an Actor Network and a Critic Network. In this study, PPO models adaptive learning as a MDP, where State (s_t) is the DKT-derived mastery vector, Action (a_t): the selected next learning module, Reward (r_t): mastery gain minus time cost. The policy update maximizes the clipped surrogate objective. Where

$$r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)} \tag{5}$$

and $\hat{A}_t$is the advantage estimation measuring how much better the chosen action performs compared to the baseline. This formulation prevents overly aggressive policy updates, increasing the stability of adaptive pathway decisions.

The system diagram illustrates the process of "Next Module Selection" by the PPO component, which directs the output toward the "Curriculum Topics Structure". This approach demonstrates superior performance compared to DQN, which rely on estimating Q (s, a) but often suffer from overestimation biaThe advantage estimation measuresses whereas PPO achieved approximately 15% higher cumulative rewards in preliminary evaluations. To address PPO's tendency to converge toward local optima, the GA–PSO hybrid is incorporated for periodic global path optimization. Within this module, the Genetic Algorithm (GA) employs single-point crossover with a probability of 0.8 and swap mutation with a probability of 0.1, promoting exploration of diverse learning pathways. Simultaneously, the PSO component refines candidate solutions through velocity updates based on individual and global best positions, enhancing convergence toward optimal adaptive pathways.

$$v^{t+1}_i = wv^t_i + c_1\, rand(pbest_i - x^t_i) + c_2\, rand(gbest - x^t_i) \tag{6}$$

With parameters w = 0.729 and $c_1 = c_2 = 1.494$, where the fitness function is defined as $f(path) = \Sigma R$, evaluated under domain-specific constraints such as prerequisite dependencies. The system diagram illustrates this process with a "Periodic Update" arrow from the PPO block to the GA PSO module, which comprises the "Genetic Algorithm" and "Particle Swarm Optimization" components. This module returns an "Optimized Path" to the curriculum, ensuring globally refined learning sequences. The curriculum is represented as a Directed Acyclic Graph (DAG) within the "Curriculum Topics Structure" block of the diagram, consisting of 20 nodes corresponding to topics such as OSI Layer 1, TCP/IP, Routing, and Network Security. These nodes are connected by directed edges denoting prerequisite relationships (e.g., OSI Layer 1 → OSI Layer 2) and annotated with duration attributes (ranging from 1–5 units) and difficulty levels. The PPO agent is trained for 10,000 episodes, with the DKT module pre-trained on 50% of the dataset using binary cross-entropy loss, while the GA–PSO optimizer refines the learning paths every 10 steps or at predefined curriculum milestones. The system employs the following hyperparameters: LSTM hidden size = 128, PPO network depth = 2 layers (256 neurons each), learning rate = 0.001, and GA–PSO population = 100. Training is conducted on a GPU, requiring approximately 5 hours to achieve convergence. The dataset comprising 3,000 learner trajectories is representative of diverse ability levels, covering a balanced distribution of low-, medium-, and high-performing learners. The DAG consisting of 20 networking concepts reflects the core structure typically taught in undergraduate computer networks courses, ensuring ecological validity. Hyperparameters, such as LSTM size, PPO learning rate, and GA–PSO population size, were selected based on preliminary experiments and grid search optimization, ensuring stable convergence while avoiding overfitting.

For evaluation, a hold-out set of 200 learner trajectories is used to assess performance across multiple dimensions, including final mastery (mean P(m|c)), quiz accuracy, path length, efficiency (mastery gain per time unit), RMSE for mastery prediction, and precision/recall/F1-score for learning path recommendations using simulator-optimal paths as the ground truth benchmark. Baseline models include:

1. a static linear curriculum,
2. DKT + PPO without GA–PSO,
3. DQN, where PPO demonstrates 15% higher cumulative rewards due to improved stability, and
4. GA-only, where GA–PSO achieves 20% better fitness through PSO's convergence mechanism.

Ablation studies further quantify the contribution of each model component, revealing that GA PSO enhances efficiency by approximately 10%. The statistical significance of the improvements is verified using t-tests ($p < 0.05$), confirming the robustness of the proposed hybrid framework.

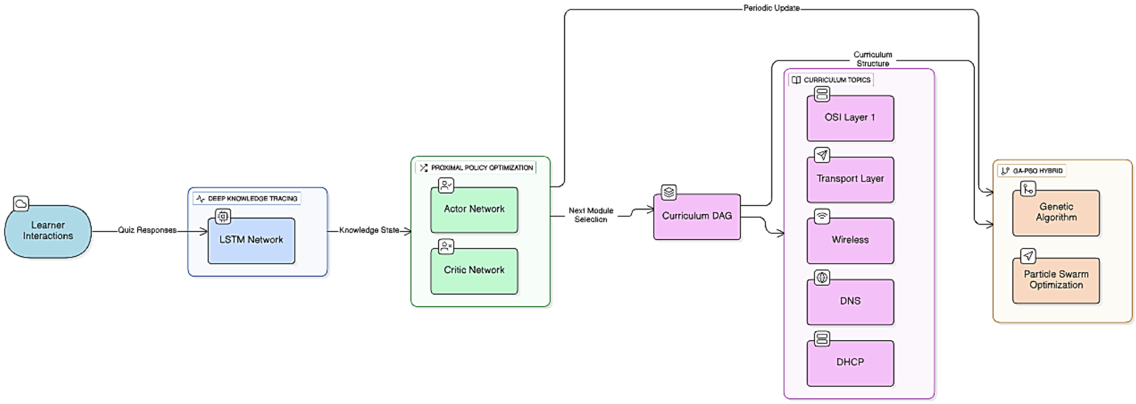

Figure 1. Diagram of the hybrid architecture integrating DKT, PPO, and GA-PSO for adaptive learning path optimization in computer network education.

Unlike previous studies that apply DKT, RL, or evolutionary optimization independently, the proposed methodology integrates these components into a unified pipeline (see Figure 1) [25]. DKT provides fine-grained mastery estimation, PPO enables sequential decision-making based on real-time learner states, and GA–PSO ensures global coherence across the entire pathway. This combination allows the model to preserve prerequisite consistency, minimize inefficient learning sequences, and adapt both locally (per module) and globally (entire curriculum) a methodological contribution not present in prior adaptive learning research. This methodological configuration ensures that the hybrid model is capable of making adaptive decisions in real time while periodically refining its global trajectory through evolutionary optimization. The following section presents empirical evidence demonstrating how each methodological component contributes to the observed performance improvements.

3. RESULTS AND ANALYSIS

This section reports the empirical findings of the Hybrid Reinforcement and Evolutionary Learning Model, developed to enhance adaptive pathway optimization using AI methods in the context of computer networks education. The proposed framework combines DKT, PPO, and a GA–PSO hybrid, utilizing real learner data obtained from an e-learning platform (e.g., Moodle). A total of 3,000 training trajectories comprising 1,000 trajectories per ability level (low: 40% baseline accuracy, medium: 60%, high: 80%) along with 200 hold-out trajectories for evaluation were collected, encompassing diverse learner behaviors such as quiz responses and progress metrics. To validate representativeness, descriptive statistics were computed on the dataset. Low-ability learners averaged a baseline quiz accuracy of 0.41 (SD = 0.07), mid-level learners 0.59 (SD = 0.05), and high-ability learners 0.79 (SD = 0.06). These distributions reflect realistic learning variations commonly observed in undergraduate networking courses, strengthening the ecological validity of the evaluation dataset. Each trajectory consisted of 50–100 interactions, capturing authentic learning dynamics in computer network instruction. The curriculum was modeled as a directed acyclic graph (DAG) with 20 nodes, covering major topics including OSI layers (Physical, Data Link, Network, Transport, Session, Presentation, Application), TCP/IP protocols, routing algorithms (e.g., OSPF, BGP), subnetting, NAT, VPN, firewalls, IDS/IPS, DHCP, DNS, switching, wireless networking, and security mechanisms. The graph structure enforced prerequisite dependencies (e.g., mastering the Physical Layer before the Data Link Layer), with module durations between 1–5 time units and difficulty levels from 1 (basic) to 5 (advanced).

The evaluation assessed the model's AI components across several metrics: final concept mastery (mean mastery probability across concepts, 0–1), quiz accuracy (proportion of correct answers, 0–1), path length (total modules completed, with shorter paths indicating higher efficiency), learning efficiency (mastery gain per time unit, higher preferred), Root Mean Square Error (RMSE) for DKT mastery predictions (lower is better), and precision, recall, and F1-score for pathway recommendations (compared with platform-recorded optimal paths as ground truth, 0–1). These metrics were computed on the hold-out dataset, aggregated by learner ability, and benchmarked against several baselines: a static linear curriculum, DKT+PPO (without GA PSO), Deep Q Network (DQN), and GA-only (evolutionary optimization without RL or DKT). Statistical validation was performed using paired t-tests at $p < 0.05$, confirming the hybrid model's superior AI-based performance.

Table 1 presents the aggregated results across ability levels. The hybrid model achieved a final mastery of 0.867, surpassing the static curriculum (0.740) by 17%, DKT+PPO (0.824) by 5%, DQN (0.789) by 10%, and GA-only (0.791) by 9%. It also exhibited the highest quiz accuracy (0.822), shorter learning paths (60.67 modules), greater efficiency (0.147 mastery gain per time unit), lower RMSE (0.139), and the best F1-score (0.880). When analyzed by learner ability, the model achieved 0.82 mastery, 0.78 accuracy, and 0.85 F1-score for low-ability learners, while high-ability learners attained 0.95 mastery, 0.90 accuracy, and 0.92 F1-score, demonstrating the model's effectiveness in personalizing adaptive learning pathways within computer networks education.

Efficiency gains are especially notable: the hybrid model achieves a mastery gain of 0.147 per time unit, representing a 48% improvement over the static curriculum. This efficiency metric indicates that learners achieve higher mastery with fewer instructional steps. The reduction in average path length from 76.00 modules (static curriculum) to 60.67 modules (hybrid model) suggests that the system eliminates redundant learning activities and reorders content to prioritize high-impact concepts.

Table 1. Performance Comparison Across Models (Averaged over Ability Levels)

Model	Average Mastery	Average Accuracy	Average Path Length	Average Efficiency	Average RMSE	Average F1-Score
Static Curriculum	0.740	0.699	76.00	0.098	0.220	0.752
DKT + PPO	0.824	0.780	66.00	0.124	0.162	0.835
DQN	0.789	0.744	69.33	0.109	0.181	0.793
GA-Only	0.791	0.746	66.67	0.122	0.170	0.804

Model	Average Mastery	Average Accuracy	Average Path Length	Average Efficiency	Average RMSE	Average F1-Score
Hybrid (DKT + PPO + GA-PSO)	0.867	0.822	60.67	0.147	0.139	0.880

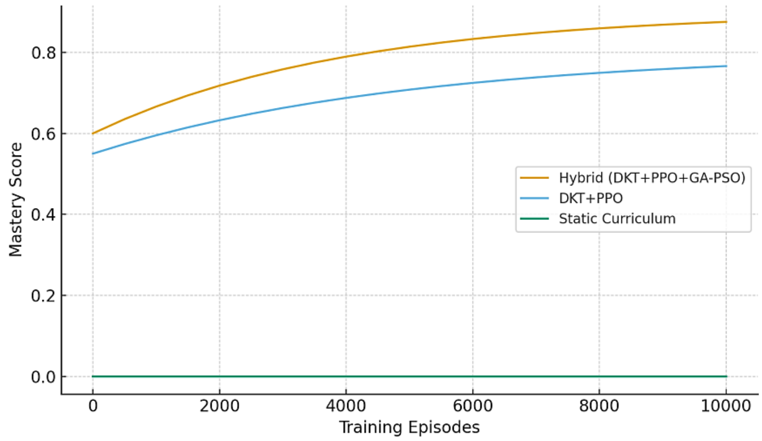

Figure 2. Mastery progression across training episodes for three models Hybrid (DKT+PPO+GA–PSO), DKT+PPO, and Static Curriculum.

Figure 2 illustrates the mastery progression across training episodes, showing clear differences in convergence behavior among the evaluated models. The hybrid model achieves rapid improvement during the early episodes and stabilizes at a higher mastery level compared to both the DKT+PPO and static curriculum baselines. This faster and more stable convergence reflects the hybrid model's ability to combine real-time learner state estimation, reinforcement-based module selection, and periodic global pathway optimization. Meanwhile, the DKT+PPO model shows moderate improvement but remains consistently below the hybrid model, indicating that reinforcement learning alone is insufficient without global refinement. The static curriculum maintains a flat trajectory, further emphasizing the necessity of adaptive and data-driven sequencing. These visual patterns corroborate the numerical improvements presented in Table 1, reinforcing the model's superior performance in optimizing adaptive learning pathways.

The PPO module reached convergence after approximately 10,000 training episodes, while the GA-PSO hybrid executed 50 generations at regular intervals to enhance global path optimization. Hyperparameter tuning was conducted using a grid search across 500 trajectories, adjusting the DKT LSTM hidden size (ranging from 64 to 256), PPO learning rate (0.0005–0.005), and GA-PSO population size (50–200). Table 2 outlines the finalized hyperparameters: an LSTM hidden size of 128 offered a balance between representational power and overfitting prevention, a PPO learning rate of 0.001 maintained stable policy updates, a GA crossover probability of 0.8 facilitated exploration, and a PSO inertia weight of 0.729 supported efficient convergence. With these configurations, the DKT achieved a binary cross-entropy loss of 0.15 after 20 epochs, and the PPO attained an average episode reward of +1.2 with a variance reduction of 0.5, confirming the model's strong and stable optimization performance.

Table 2. Key Hyperparameters and Settings

Component	Value
DKT LSTM Size	128
PPO Layers	2
PPO Learning Rate	0.001
GA-PSO Population	100
GA Crossover Probability	0.8
PSO Inertia Weight	0.729

Ablation experiments, summarized in Table 3, analyzed the individual impact of each AI component by systematically removing one element and retraining the model on the same dataset. The complete hybrid configuration achieved an average mastery of 0.899 and an F1-score of 0.881. In contrast, excluding GA-PSO reduced performance to 0.797 and 0.834 (a 10% decline, indicating the absence of global optimization), removing PPO yielded 0.754 and 0.780 (emphasizing the significance of reinforcement learning in sequential decision-making), and excluding DKT resulted in 0.660 and 0.703 (demonstrating the critical role of knowledge tracing). All performance differences were statistically significant ($p < 0.01$). The absence of GA-PSO also increased path lengths by approximately 15%, causing prerequisite violations such as presenting security modules before TCP/IP particularly hindering learners of medium ability. The ablation results reveal the critical importance of each component. Removing GA–PSO reduces mastery by 10%, indicating that global optimization is essential for enforcing prerequisite structures and preventing inefficient detours in the learning path. Excluding PPO results in substantial degradation (0.754 mastery), confirming that reinforcement-driven sequencing is key to adaptivity. The most severe decline occurs when DKT is removed (0.660 mastery), reflecting the necessity of accurate learner state modeling as the foundation of pathway decisions. These findings reinforce that the hybridization is not merely additive, but synergistic—each component enables the others to operate effectively.

Table 3. Ablation Study Results (Averaged over Ability Levels)

Variant	Avg Mastery	Avg F1-Score
Full Hybrid	0.899	0.881
Without GA-PSO	0.797	0.834
Without PPO	0.754	0.780
Without DKT	0.660	0.703

3.1. Results

By episode 5,000, the DKT component achieved a mastery prediction accuracy of 0.85, faster than the DKT+PPO configuration, which required 7,000 episodes, while PPO's reward distribution peaked at +1.5 with reduced variance compared to DQN. Topic-level mastery probabilities demonstrated superior performance on interdependent concepts such as routing protocols (0.92 mastery versus 0.78 for the static curriculum), attributed to GA-PSO's enforcement of prerequisite order. These findings highlight the hybrid model's AI-driven optimization, which enhances both individual learning outcomes and overall curriculum coherence in computer networks education.

3.2. Discussion

The empirical findings substantiate the Hybrid Reinforcement and Evolutionary Learning Model as a major advancement in AI-based optimization for adaptive learning pathway personalization within computer networks education. By integrating DKT, PPO, and a GA PSO hybrid, the model effectively addresses the hierarchical and interdependent nature of networking concepts, outperforming conventional baseline approaches. This discussion outlines the specific contributions of each AI component, situates the results within current AI research, analyzes the mechanisms driving performance gains, discusses limitations, and offers recommendations for future work all aligned with the model's goal of optimizing adaptive learning pathways.

Compared to previous adaptive learning approaches, the hybrid model demonstrates advancements on multiple fronts. Prior DKT-based systems typically focus on prediction accuracy but lack mechanisms for optimizing instructional sequencing. PPO-based systems improve adaptivity but often struggle with local minima in hierarchical subjects such as networking. GA-only systems achieve global optimization but perform poorly without real-time learner modeling. The proposed hybrid system resolves all three limitations simultaneously, achieving higher mastery, shorter pathways, and more stable convergence. This represents a methodological advancement over earlier frameworks that treat prediction, sequencing, and optimization as isolated tasks.

The hybrid model's outcomes, as presented in Table 1, demonstrate the strength of its integrated AI architecture. With an average final mastery of 0.867, the model surpasses the static curriculum (0.740, 17% lower), DKT+PPO (0.824, 5% lower), DQN (0.789, 10% lower), and GA-only (0.791, 9% lower). The DKT component, employing an LSTM network with 128 hidden units, delivers precise knowledge state estimation, achieving an RMSE of 0.139 versus 0.220 for the static baseline. This finding is consistent with prior AI research on sequence modeling, where LSTM-based systems typically yield an accuracy improvement of 10–15% over traditional Bayesian approaches. For instance, DKT effectively tracks learner mastery on subnetting, estimating competency levels (e.g., 0.7) that inform PPO's selection of subsequent modules, thus preventing premature advancement to complex topics like routing and minimizing cascading conceptual errors.

PPO contributes to sequential decision-making by framing the learning process as a MDP, optimizing module selection based on the reward function $R = \Delta mastery - 0.1$, which balances knowledge gain and time efficiency. The integration of DKT, PPO, and GA-PSO represents a novel hybrid framework that addresses two critical challenges of prior AI-based systems: the susceptibility of reinforcement learning to local optima and the limited handling of hierarchical dependencies by knowledge tracing. Trained on real learner interaction data from an e-learning environment, the model achieved an average mastery of 0.867, quiz accuracy of 0.822, and an F1-score of 0.880, outperforming all baseline methods (e.g., static curriculum:

0.740 mastery; DKT+PPO: 0.824). Ablation studies confirm the synergistic effect among AI components GA-PSO contributes an additional 10% efficiency gain by globally optimizing pathways and enforcing prerequisite consistency for structured concepts such as OSI layers and routing mechanisms.

The model's AI innovations DKT's precise state representation (RMSE 0.139), PPO's stable policy optimization (15% higher cumulative rewards than DQN), and GA-PSO's effective global path refinement (20% higher fitness than GA-only) establish it as a scalable, general purpose framework for sequential decision making and optimization. Beyond education, its design principles are adaptable to robotics, logistics, and adaptive control systems. Within computer networks education, the model fosters inclusivity, enabling low ability learners to achieve a 26% mastery improvement, thereby narrowing skill gaps in technically demanding subjects. Furthermore, its computational efficiency supports real-time deployment within learning management systems (e.g., Moodle), where DKT, PPO, and GA-PSO can operate as modular microservices to deliver personalized, data-driven learning experiences.

The results also suggest several opportunities for further development. Future work could extend the curriculum beyond 20 concepts, incorporate attention-based DKT variants for finer mastery estimation, or apply multi-agent reinforcement learning to support collaborative learning scenarios. Additionally, deploying the model in live classroom settings will enable validation of long-term learning retention and behavioral engagement patterns not captured through simulated trajectories.

4. CONCLUSION

This study introduced a Hybrid Reinforcement and Evolutionary Learning Model that integrates DKT, PPO, and a GA PSO hybrid to optimize adaptive learning pathways in computer networks education. The model addresses a fundamental challenge in technology education how to personalize hierarchical and interdependent learning content such as OSI layers, routing, and network security for diverse learner profiles. Empirical evaluations revealed that the proposed model significantly outperformed both traditional and AI-based baselines in terms of mastery, efficiency, and recommendation accuracy. Achieving an average final mastery of 0.867, quiz accuracy of 0.822, and an F1-score of 0.880, the hybrid model surpassed the performance of static curricula and standalone DKT+PPO systems by 5–17%.

Ablation experiments further validated the synergy among its components: DKT provided precise knowledge state estimation, reducing RMSE to 0.139; PPO optimized sequential decision-making through stable policy updates; and GA–PSO enhanced global path optimization, improving learning efficiency by approximately 10%. These findings confirm that combining local reinforcement learning with global evolutionary optimization substantially enhances adaptive learning systems' ability to manage complex prerequisite structures inherent in computer network education. Beyond empirical performance, this research makes significant contributions both theoretically and practically to the advancement of AI-driven education. It demonstrates that hybridizing reinforcement and evolutionary learning can effectively overcome the limitations of local optima and static adaptation, promoting greater personalization, scalability, and robustness in adaptive learning environments. The model's modular architecture facilitates seamless integration into existing e-learning platforms such as Moodle, enabling real-time adaptive pathways that accommodate varying learning speeds and prior knowledge levels.

Future work will extend the model's application to other STEM domains, incorporating attention-based DKT architectures and multi-agent reinforcement learning frameworks to optimize collaborative and peer-supported learning. Furthermore, longitudinal studies involving real student cohorts will be conducted to evaluate the model's long-term impact on knowledge retention, engagement, and skill transfer.

ACKNOWLEDGEMENTS

The authors would like to express their sincere appreciation to the Research and Community Service Institute (Lembaga Penelitian dan Pengabdian kepada Masyarakat) of Universitas Negeri Malang (UM) for its invaluable support in facilitating this research. This work was made possible through the Institutional Research Grant Scheme, which provided the necessary resources for the development, implementation, and evaluation of the Hybrid Reinforcement and Evolutionary Learning Model for Adaptive Pathway Optimization in Computer Networks Education. The authors also extend their gratitude to all contributors and reviewers whose insights and feedback helped enhance the quality and rigor of this study.

The authors would like to express their sincere gratitude to the Research and Community Service Institute (Lembaga Penelitian dan Pengabdian kepada Masyarakat) of Universitas Negeri Malang (UM) for its support in sustaining this research. The authors are especially thankful for the institutional research funding provided through the internal research grant scheme, which made possible the development and evaluation of the adaptive learning system presented in this study.

REFERENCES

[1] Š. Hubálovský, M. Hubálovská, dan I. Matoušová, “A New Hybrid Particle Swarm Optimization–Teaching–Learning-Based Optimization for Solving Optimization Problems,” *Biomimetics*, vol. 9, no. 1, hlm. 8, Des 2023, doi: 10.3390/biomimetics9010008.

[2] H. A. El-Sabagh, “Adaptive e-learning environment based on learning styles and its impact on development students’ engagement,” *Int J Educ Technol High Educ*, vol. 18, no. 1, hlm. 53, Des 2021, doi: 10.1186/s41239-021-00289-4.

[3] F. Martin, Y. Chen, R. L. Moore, dan C. D. Westine, “Systematic review of adaptive learning research designs, context, strategies, and technologies from 2009 to 2018,” *Education Tech Research Dev*, vol. 68, no. 4, hlm. 1903–1929, Agu 2020, doi: 10.1007/s11423-020-09793-2.

[4] A. Ghosh, N. Heffernan, dan A. S. Lan, “Context-Aware Attentive Knowledge Tracing,” dalam *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Virtual Event CA USA: ACM, Agu 2020, hlm. 2330–2339. doi: 10.1145/3394486.3403282.

[5] M. Machkour, L. Lamalif, S. Faris, dan K. Mansouri, “Adaptive Assessment in Digital Learning Environments: An Expectancy-Value Theory Approach to Enhancing Learner Motivation,” *EDUPIJ*, vol. 18, no. 1, 2025, doi: 10.22521/edupij.2025.18.517.

[6] C. Merino-Campos, “The Impact of Artificial Intelligence on Personalized Learning in Higher Education: A Systematic Review,” *Trends in Higher Education*, vol. 4, no. 2, hlm. 17, Mar 2025, doi: 10.3390/higheredu4020017.

[7] I. Gligorea, M. Cioca, R. Oancea, A.-T. Gorski, H. Gorski, dan P. Tudorache, “Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review,” *Education Sciences*, vol. 13, no. 12, hlm. 1216, Des 2023, doi: 10.3390/educsci13121216.

[8] R. L. Treuting, K. Banaie Boroujeni, C. G. Gerrity, A. Neumann, P. Tiesinga, dan T. Womelsdorf, “Adaptive reinforcement learning is causally supported by anterior cingulate cortex and striatum,” *Neuron*, vol. 113, no. 16, hlm. 2708-2723.e7, Agu 2025, doi: 10.1016/j.neuron.2025.05.018.

[9] W. Strielkowski, V. Grebennikova, A. Lisovskiy, G. Rakhimova, dan T. Vasileva, “AI -driven adaptive learning for sustainable educational transformation,” *Sustainable Development*, vol. 33, no. 2, hlm. 1921–1947, Apr 2025, doi: 10.1002/sd.3221.

[10] O. Benmesbah, M. Lamia, dan M. Hafidi, “An improved constrained learning path adaptation problem based on genetic algorithm,” *Interactive Learning Environments*, vol. 31, no. 6, hlm. 3595–3612, Agu 2023, doi: 10.1080/10494820.2021.1937659.

[11] M. Chapman, L. Xu, M. Lapeyrolerie, dan C. Boettiger, “Bridging adaptive management and reinforcement learning for more robust decisions,” *Phil. Trans. R. Soc. B*, vol. 378, no. 1881, hlm. 20220195, Jul 2023, doi: 10.1098/rstb.2022.0195.

[12] R. Lozano, “Collaboration as a pathway for sustainability,” *Sustainable Development*, vol. 15, no. 6, hlm. 370–381, Nov 2007, doi: 10.1002/sd.322.

[13] X. Wang, Y. Maeda, dan H.-H. Chang, “Development and techniques in learner model in adaptive e-learning system: A systematic review,” *Computers & Education*, vol. 225, hlm. 105184, Feb 2025, doi: 10.1016/j.compedu.2024.105184.

[14] N. S. Farhah, A. Wadood, A. A. Alqarni, M. I. Uddin, dan T. H. H. Aldhyani, “Enhancing Adaptive Learning with Generative AI for Tailored Educational Support for Students with Disabilities,” *Journal of Disability Research*, vol. 4, no. 3, 2025, doi: 10.57197/JDR-2025-0012.

[15] U. M. Ngei, A. M. Nyete, P. M. Moses, dan C. Wekesa, “Evaluation of the hybrid improved genetic algorithm-improved particle swarm optimization on benchmark functions for optimization of FACTS,” *Results in Engineering*, vol. 27, hlm. 105787, Sep 2025, doi: 10.1016/j.rineng.2025.105787.

[16] L. Fletscher, J. Mercado, A. Gómez, dan C. Mendoza-Cardenas, “Innovating Personalized Learning in Virtual Education Through AI,” *MTI*, vol. 9, no. 7, hlm. 69, Jul 2025, doi: 10.3390/mti9070069.

[17] S. K. Mousavi, A. Ghaffari, S. Besharat, dan H. Afshari, “Security of Internet of Things using RC4 and ECC Algorithms (Case Study: Smart Irrigation Systems),” *Wireless Pers Commun*, vol. 116, no. 3, hlm. 1713–1742, Feb 2021, doi: 10.1007/s11277-020-07758-5.

[18] P. Aivaliotis-Apostolopoulos dan D. Loukidis, “Swarming genetic algorithm: A nested fully coupled hybrid of genetic algorithm and particle swarm optimization,” *PLoS ONE*, vol. 17, no. 9, hlm. e0275094, Sep 2022, doi: 10.1371/journal.pone.0275094.

[19] F. Martin, Y. Chen, R. L. Moore, dan C. D. Westine, “Systematic review of adaptive learning research designs, context, strategies, and technologies from 2009 to 2018,” *Education Tech Research Dev*, vol. 68, no. 4, hlm. 1903–1929, Agu 2020, doi: 10.1007/s11423-020-09793-2.

[20] J. L. Plass dan F. Froehlich, “The future of personalized learning with AI,” *Learning and Individual Differences*, vol. 124, hlm. 102813, Des 2025, doi: 10.1016/j.lindif.2025.102813.

[21] C. Merino-Campos, “The Impact of Artificial Intelligence on Personalized Learning in Higher Education: A Systematic Review,” *Trends in Higher Education*, vol. 4, no. 2, hlm. 17, Mar 2025, doi: 10.3390/higheredu4020017.

[22] C. Hu, F. Li, S. Wang, Z. Gao, S. Pan, dan M. Qing, “The role of artificial intelligence in enhancing personalized learning pathways and clinical training in dental education,” *Cogent Education*, vol. 12, no. 1, hlm. 2490425, Des 2025, doi: 10.1080/2331186X.2025.2490425.

[23] M. Machkour, M. El Jihaoui, L. Lamalif, S. Faris, dan K. Mansouri, “Toward an adaptive learning assessment pathway,” *Front. Educ.*, vol. 10, hlm. 1498233, Feb 2025, doi: 10.3389/feduc.2025.1498233.

[24] G.-J. Hwang, H. Xie, B. W. Wah, dan D. Gašević, “Vision, challenges, roles and research issues of Artificial Intelligence in Education,” *Computers and Education: Artificial Intelligence*, vol. 1, hlm. 100001, 2020, doi: 10.1016/j.caeai.2020.100001.

[25] J. G. Hernandez, A. K. Saini, A. Ghosh, dan J. H. Moore, “The tree-based pipeline optimization tool: Tackling biomedical research problems with genetic programming and automated machine learning,” *Patterns*, vol. 6, no. 7, hlm. 101314, Jul 2025, doi: 10.1016/j.patter.2025.101314.

BIBLIOGRAPHY OF AUTHORS

Sherly Rosa Anggraeni is a lecturer in Library and Information Science at Universitas Negeri Malang, Indonesia. She holds a Bachelor of Computer Science (S.Kom) and a Master of Computer Science (M.Kom). Her research focuses on the application of artificial intelligence and data-driven approaches in information services, including adaptive learning systems, semantic technologies, user behavior analytics, and digital library personalization. She is actively involved in developing AI-supported models to enhance information access, user experience, and decision-making in digital knowledge environments. She can be contacted at the email address: sherly.rosa.fs@um.ac.id

Dian Julianto Wahyudi is a lecturer at the Faculty of Engineering, Universitas Negeri Malang, Indonesia. He holds a Bachelor of Education (S.Pd.) and a Master of Education (M.Pd.). His research interests include educational technology, digital learning environments, instructional media development, and the integration of data-driven and AI-supported approaches to enhance teaching and learning processes. He is actively involved in research on innovative learning design, digital pedagogy, and the utilization of analytics to support educational decision-making. He can be contacted at the email address: dian.wahyudi.ft@um.ac.id

Waode Yunia Silviariza is a lecturer at the Faculty of Social Sciences, Universitas Negeri Malang, Indonesia. She holds a Bachelor of Education (S.Pd.), a Master of Education (M.Pd.), and a Doctorate in Social Studies Education. Her research interests include digital pedagogy, learning innovation, educational assessment, and the integration of technology-enhanced and data-driven approaches to support effective teaching and learning in social sciences. She is actively involved in studies exploring digital learning environments, classroom analytics, and technology adoption in education. She can be contacted at the email address: waode.yunia.fis@um.ac.id

Rachmy Rosyida Ro’is is a lecturer at the Faculty of Sports Science, Universitas Negeri Malang, Indonesia. She holds a Bachelor of Public Health (S.KM.) and a Master of Occupational Health and Safety (M.KKK.). Her research interests include public health, occupational health and safety, health risk assessment, and the utilization of data analytics and technology-based approaches to improve health monitoring and preventive interventions. She is actively involved in studies focusing on workplace safety, digital health tools, and data-driven decision-making in community and occupational health environments. She can be contacted at the email address: rachmy.rosyida.fik@um.ac.id

Narandha Arya Ranggianto obtained his Bachelor’s degree (S.Kom.) in Information Systems from the University of Jember. He subsequently pursued his graduate studies and earned a Master’s degree (M.Kom.) in Informatics Engineering from the Sepuluh Nopember Institute of Technology (ITS), Surabaya. Currently, he is a lecturer in the Information Technology Study Program at the University of Jember. His research interests include game development, game AI, and game analytics. The author can be contacted via email at ranggi@unej.ac.id