Διεθνές Συνέδριο για την Ανοικτή & εξ Αποστάσεως Εκπαίδευση

Τόμ. 13, Αρ. 5 (2026)

ICODL2025

Reinforcement Learning: Educational Approaches in Distance Education

Athanasios Sypsas, Dimitris Kalles

doi: [10.12681/icodl.8631](https://doi.org/10.12681/icodl.8631)

Reinforcement Learning: Educational Approaches in Distance Education

Ενισχυτική Μάθηση: Εκπαιδευτικές Προσεγγίσεις στην Εξ Αποστάσεως Εκπαίδευση

Athanasios Sypsas
Adjunct Faculty Member
Hellenic Open University
athanasios.sypsas@ac.eap.gr

Dimitris Kalles
Professor
Hellenic Open University
kalles@eap.gr

Περίληψη

Η εξ αποστάσεως εκπαίδευση έχει εξελιχθεί σημαντικά, αξιοποιώντας πλέον διαδραστικές ψηφιακές πλατφόρμες· ωστόσο, εξακολουθεί να αντιμετωπίζει προκλήσεις, όπως ο περιορισμένος βαθμός εξατομίκευσης, η χαμηλή ενεργός εμπλοκή των εκπαιδευομένων και η ανεπαρκής παροχή υποστήριξης σε πραγματικό χρόνο. Η Ενισχυτική Μάθηση (Reinforcement Learning, RL), ως μέρος της Τεχνητής Νοημοσύνης (AI), προσφέρει δυνατότητες αντιμετώπισης αυτών των προβλημάτων, μοντελοποιώντας τη μάθηση ως μια διαδικασία λήψης αποφάσεων με ανατροφοδότηση. Σε αντίθεση με τις μεθόδους επιβλεπόμενης και μη επιβλεπόμενης μάθησης, οι πράκτορες RL προσαρμόζονται δυναμικά λαμβάνοντας ανταμοιβές ή ποινές, γεγονός που τους καθιστά κατάλληλους για προσομοιωμένα περιβάλλοντα προσανατολισμένα στον εκπαιδευόμενο. Αυτή η μελέτη ανασκοπεί τις εφαρμογές της RL στην εξ αποστάσεως εκπαίδευση, καλύπτοντας διάφορους τομείς. Τονίζει τον ρόλο της RL στη βελτιστοποίηση των μαθησιακών διαδρομών, στην εξατομίκευση της διδασκαλίας μέσω προφίλ μαθητών και στην παροχή προσαρμοστικών προτάσεων με αλγορίθμους όπως ο Q-learning. Με την ενσωμάτωση της RL σε υβριδικές προσεγγίσεις AI, η εξ αποστάσεως εκπαίδευση μπορεί να γίνει πιο ευέλικτη, ελκυστική και αποτελεσματική, προσφέροντας κλιμακούμενες λύσεις για διαφορετικούς μαθητές.

Λέξεις-Κλειδιά

μάθηση μέσω ενίσχυσης, εξ αποστάσεως εκπαίδευση, προσαρμοστικά μαθησιακά συστήματα, εξατομικευμένη ανατροφοδότηση

Abstract

Distance education has progressed from correspondence courses to interactive digital platforms, yet challenges such as limited personalization, low engagement, and insufficient real-time support persist. Reinforcement Learning (RL), as part of Artificial Intelligence (AI), offers unique potential to address these issues by modeling learning as a feedback-driven, sequential decision-making process. Unlike supervised and unsupervised methods, RL agents adapt dynamically by receiving rewards or penalties, making them well-suited for simulation-based, learner-centered environments. This study reviews applications of RL in distance education, covering various domains. It highlights the role of RL in optimizing learning paths, personalizing instruction through learner profiles, and enabling adaptive recommendations with algorithms like Q-learning. By integrating RL with hybrid AI approaches, distance education can become more responsive, engaging, and effective, offering scalable solutions for diverse learners.

Keywords

reinforcement learning, distance education, adaptive learning systems, personalized feedback

Introduction

Distance education has evolved from text-based correspondence models to highly interactive online platforms. In various domains, online learning has shown strong outcomes in terms of knowledge acquisition (Sypsas, Paxinou, Zafeiropoulos & Kalles, 2024). Yet, distance learners face barriers such as limited personalization, reduced engagement, and insufficient real-time support.

Artificial Intelligence (AI) has increasingly been applied in education to address persistent challenges, with intelligent tutoring systems, chatbots, and adaptive testing frameworks exemplifying its role in guiding and supporting learners (Dong et al., 2022). While supervised and unsupervised learning dominate many conventional machine learning (ML) applications, reinforcement learning (RL) is distinguished by its capacity to train agents through direct interaction with the environment (Alqahtani & Rajaraman, 2021; Wiering & van Otterlo, 2012). Unlike supervised learning, which depends on labeled data provided by an external supervisor, and unsupervised learning, which uncovers latent structures in unlabeled datasets, RL emphasizes interaction and trial-and-error learning, enabling agents to refine their behavior in dynamic environments. As such, RL is recognized as a distinct third paradigm of ML, oriented toward the maximization of long-term cumulative rewards (Sutton & Barto, 1998). Within educational contexts, RL appears particularly well-suited to distance education, as it conceptualizes learning as an interactive feedback loop, allowing systems to dynamically adapt instruction and learning materials (Muniasamy & Alasiry, 2020). Moreover, RL's reliance on iterative trial-and-error processes makes it highly applicable to educational scenarios where sequential decision-making and continuous feedback are central to effective learning (El Gourari et al., 2021; Fahad Mon et al., 2023; Thomaz & Breazeal, 2006). Figure 1, below, illustrates the relationship between AI, machine learning, and reinforcement learning.

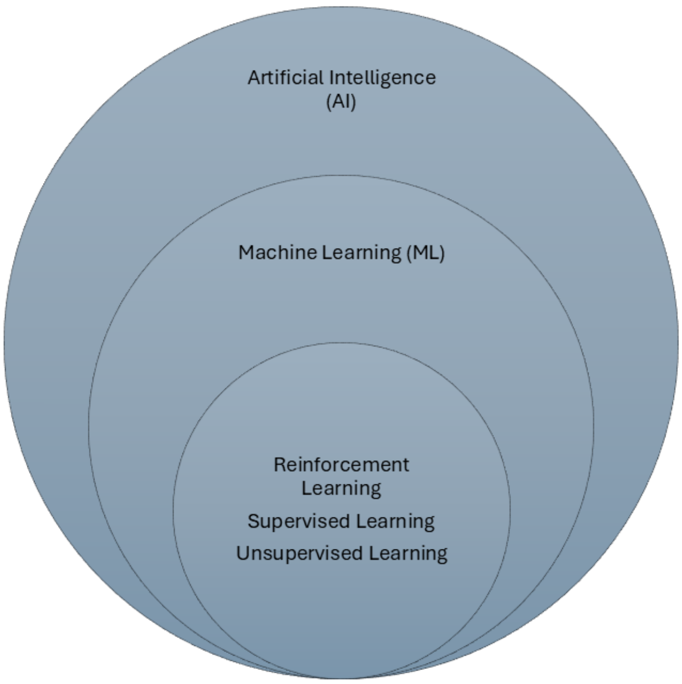

Figure 1: *The relationship and connections between AI, machine learning, and reinforcement learning.*

Given its strength in representing sequential decision-making processes and dynamically responding to learner behavior, RL has gained prominence as a framework for intelligent technologies in distance education (Riedmann, Schaper, & Lugrin, 2025). These capabilities are particularly valuable in virtual learning environments, where effective distance education depends on personalized, adaptive, and responsive instructional strategies. The present study therefore aims to: a) review recent advances in the application of RL within distance and virtual education; b) examine concrete use cases across various fields; and c) analyze current trends, implementation strategies, and the challenges and opportunities RL poses for shaping learner-centered distance education systems.

This study seeks to broaden existing perspectives by examining the role of Reinforcement Learning (RL) in distance education. It further considers the integration of hybrid AI techniques, such as Q-learning for adaptive sequencing, which enhance the adaptability and effectiveness of RL-driven educational systems.

To explore the transformative potential of RL in distance education environments, this paper is structured as follows. Section 2 describes the applied methodology. Section 3 explores RL's role in distance education, focusing on its use in learning path optimization, adaptive recommendations through Q-learning, and the development of learner profiles. Section 5 discusses future directions and key challenges, including issues of interpretability, ethical implications, and data efficiency. The final section provides concluding remarks, addresses limitations, and highlights prospects for future advancements.

Methodology

This section presents our scoping review of the literature on RL in distance education. The search was applied in ACM DL, Web of Science and Scopus databases, using the string: "Reinforcement Learning" AND ("distance education" OR "e-learning"), with no time restrictions to capture the full trajectory of RL in distance education. Screening followed three stages: (1) title and abstract review, (2) duplicate removal, and (3) full-text assessment for alignment with the study scope.

Inclusion criteria required studies to explicitly apply RL in distance education and e-learning approaches, across all educational levels. Both empirical studies and reviews were considered. Excluded were non-English papers, studies outside education (e.g., robotics, finance), and works mentioning RL in distance education only peripherally. During full-text screening, only studies implementing RL as a core methodological or pedagogical element were retained. From an initial 646 retrieved articles, 15 studies met the criteria and form the basis of this review.

Reinforcement Learning in Distance Education

RL has emerged as a transformative approach in distance education, offering personalized and adaptive learning experiences (Jing et al., 2023). By enabling systems to tailor content and feedback based on individual learner behaviors, RL enhances student engagement and learning outcomes (Zafeiropoulos, 2021). El Gourari et al. (2021) demonstrated the application of Deep Q-Networks (DQNs) to model remote practical work, allowing virtual agents to guide students through complex tasks by analyzing their interactions and providing real-time feedback. This approach not only supports individualized learning paths but also fosters a deeper understanding of the subject matter. Furthermore, a systematic literature review by Hayajneh et al. (2023) highlighted the potential of RL in education, identifying various techniques such as the Markov Decision Process and deep RL networks, and emphasizing the importance of adaptive learning environments in enhancing educational outcomes. These advancements underscore the growing significance of RL in reshaping distance education to meet the diverse needs of learners.

Learning Path Optimization

One of the most significant applications of reinforcement learning (RL) in distance education is learning path optimization. In RL-based instructional systems, agents are rewarded for selecting tasks and learning objects that improve the learner's experience and penalized when their choices result in disengagement or ineffective outcomes (Li, Xu, Zhang & Chang, 2021). This reward structure closely mirrors the pedagogical principle of scaffolding, where learners are guided toward increasingly complex skills in a way that maximizes comprehension while minimizing frustration. For example, in systems with hierarchical skill structures, RL agents can recommend content sequences tailored to individual learners' prior knowledge and performance, ensuring progression through concepts at a personalized pace (El Gourari et al., 2021). Such optimization is particularly important in distance education, where learners often study without direct supervision. Adaptive RL systems can replace static course sequencing with dynamic strategies that adjust in real time to the learner's performance (Cai, Zhang & Dai, 2021). This includes advancing a learner when mastery is demonstrated or revisiting earlier material when gaps are detected. By framing the learning process as a sequential decision-making problem, RL provides a more flexible instructional design that accommodates both remedial support and enrichment for advanced learners. As highlighted by Chakraborty et al. (2022), this approach represents a significant improvement in overrule-based systems, which cannot easily respond to unexpected variations in learner behavior. In another approach, a knowledge tracing approach was designed to model learners' knowledge progress over time, enabling more accurate prediction of their learning states (Cai, Zhang, and Dai, 2021). Building on this model, a learning path recommendation algorithm was developed, integrating knowledge tracing with reinforcement learning to deliver personalized instructional sequences.

Q-Learning for Adaptive Recommendations

Q-learning, one of the most widely applied RL algorithms, has gained attention in distance education for its ability to map learner states to optimal instructional actions. Unlike more complex deep reinforcement learning approaches, Q-learning maintains interpretability, making it easier to integrate into practical e-learning systems. Learners are represented as states in the system, while educational activities (e.g., quizzes, readings, videos) are modeled as actions. The Q-learning algorithm iteratively learns the expected rewards of each action in each state, thus building a policy for recommending the best next activity (Chakraborty et al., 2022). In a study (Tan, Han, Ye & Chen, 2020), Q-learning is applied to design adaptive recommendation systems that guide learners through personalized learning paths based on their behavior and proficiency data. By modeling each learner's state and selecting actions that maximize cumulative rewards, the algorithm recommends the most efficient sequence of activities. In another study, Q-learning was used to adapt an e-learning system, based on learner behavior (Shi, 2025). Then it adjusts recommendations in real time to enhance learning efficacy. User feedback, knowledge recall, and interest levels are used as rewards for the decision. To enhance accuracy, Q-learning is often combined with machine learning methods that build deep learner profiles. Clustering techniques such as k-means are used to identify groups of learners with similar characteristics, while linear regression predicts performance outcomes. Supervised models refine these predictions further, creating a robust representation of learner needs. Based on these profiles, Q-learning agents can recommend personalized learning trajectories, thereby maximizing cumulative rewards for engagement and achievement. This integration of reinforcement learning with other AI techniques reflects a trend toward hybridized systems that balance interpretability, adaptability, and scalability (El Gourari et al., 2021).

In distance education, Q-learning-based recommendations are particularly effective in large-scale online classrooms where teacher oversight is limited. For instance, in programming courses, Q-learning systems have been shown to guide learners through problem sets in a way that reduces dropout rates and ensures exposure to progressively challenging concepts (Riedmann, Schaper & Lugrin (2025). Another study in distance education proposed an RL based smart e-learning framework with Markov Decision Process (MDP) that has the potential to enhance the learning experience for each learner by providing them with a personalized and effective learning path (Amin et al., 2023). Combining MDP and Q-learning for Sequential Path Recommendation (SPR) the learning development is more achievable. This is because the MDP allows for adjusting the recommendations method to find new activities and learning paths based on the learners' feedback on recommendations results.

By systematically mapping states to actions, Q-learning ensures that learners do not stagnate but instead follow adaptive, efficient, and personalized educational trajectories.

Inferring and Adjusting to Learner Profiles in Distance Education

In distance education, where learners engage across diverse modalities such as online discussions, quizzes, interactive assignments, and virtual labs, the ability to unify data from multiple channels is particularly valuable. A central challenge of adaptive e-learning systems is the ability to recommend individualized learning scenarios that align with the specific needs of each learner (Chatzimparmpas et al., 2023). The effectiveness of such systems depends on constructing and continuously updating a deep learner profile that captures cognitive, behavioral, and contextual characteristics. However, many existing adaptive e-learning systems place limited emphasis on the adequacy of real learner profiles or their timely updates in relation to learning path recommendations (El Gourari et al., 2021; Riad et al., 2023).

To address this limitation, an adaptive e-learning framework has been proposed that combines machine learning and reinforcement learning (Riad et al., 2023). It comprises three modules: data preprocessing, deep learner profile creation, and learning path recommendation. Learner profiles are built using k-means clustering and linear regression, while Q-learning tailors adaptive learning paths to each learner.

El Gourari et al. (2021) highlight the role of RL agents in dynamically interpreting these profiles to adjust learning experiences. Learners who face difficulties can be supported through timely interventions such as hints or supplementary resources, while advanced learners can be challenged with more complex tasks to maintain engagement. Riad et al. (2023) proposed an intelligent adaptive e-learning system that leverages deep learner profiles to recommend the most suitable learning objects. The system employs machine learning and reinforcement learning algorithms to generate a tailored list of learning resources aligned with each learner's profile.

Findings indicate that incorporating deep learner profiles into the recommendation process significantly improves the quality of learning by ensuring that the suggested learning paths reflect the evolving characteristics of each learner. This highlights the importance of hybrid AI frameworks in distance education, where personalization must adapt dynamically to learners' ongoing progress and challenges (Chakraborty et al., 2022).

Future Trends and Challenges

Future developments in RL for distance education are expected to be shaped by several emerging trends and unresolved challenges. A major research direction involves improving interpretability and transparency of RL models, ensuring that educators and learners can understand and trust algorithmic decisions. Closely linked to this is the need for ethical and equitable frameworks, where issues of bias, inclusivity, and learner data protection are prioritized.

Another challenge is data efficiency: RL models often require large volumes of interaction data, which is difficult to obtain in educational contexts where learner time and effort are limited. Future work will focus on lightweight, data-efficient algorithms that balance personalization with feasibility. Hybrid approaches, combining RL with supervised, unsupervised, and deep learning techniques, are also likely to become more prominent, particularly for constructing robust learner profiles and delivering adaptive recommendations.

Finally, scalability and generalizability remain key challenges. Many current applications are small-scale prototypes or discipline-specific pilots. For RL-driven systems to be impactful, they must operate reliably across diverse disciplines, cultures, and institutional contexts. Meeting these challenges requires interdisciplinary collaboration, drawing from education, computer science, psychology, and ethics, to ensure RL serves pedagogical goals rather than purely technical optimization.

Conclusions and Limitations

Reinforcement Learning holds transformative potential for distance education by enabling adaptive, personalized, and scalable learning systems. Its strength lies in modeling the sequential, feedback-driven nature of learning, which allows instructional strategies to be dynamically tailored to individual learners. Evidence from case studies and systematic reviews demonstrates its capacity to optimize learning paths, enhance engagement, and deliver real-time feedback.

Nevertheless, current implementations are limited in scale and scope, often constrained to experimental or prototype environments. Broader validation across disciplines and learner populations is essential. Equally important are the challenges of complexity, interpretability, and resource requirements, which hinder adoption in real-world educational systems.

The scope of this study is not exhaustive, as the diversity of distance education contexts, instructional models, and reinforcement learning (RL) approaches cannot be fully represented within a single review. Many of the applications discussed remain at the level of prototypes or early-stage deployments, limiting the extent to which findings can be generalized across institutions, disciplines, or learner populations. These limitations, however, highlight important avenues for future inquiry. Longitudinal and cross-domain investigations are needed to assess the long-term effectiveness of RL-enhanced systems in supporting learner retention, achievement, and engagement in distance learning. Future research should also focus on hybrid frameworks that combine RL with data-efficient algorithms, multimodal learner modeling, and interpretable agent policies to improve both scalability and transparency.

To fully realize RL's promise, future research must emphasize ethical, transparent, and learner-centered design. By bridging innovation with pedagogy, RL can redefine distance education, making it more engaging, inclusive, and effective, while supporting lifelong learning in an increasingly digital world.

References

Alqahtani, A., & Rajaraman, V. (2021). A review of generative adversarial networks and their applications in education. *Journal of Computer Science Education, 14*(2), 45–60.

Amin, S., Uddin, M. I., Alarood, A. A., Mashwani, W. K., Alzahrani, A., & Alzahrani, A. O. (2023). Smart e-learning framework for personalized adaptive learning and sequential path recommendations using reinforcement learning. *IEEE Access, 11*, 89769–89790.

Cai, D., Zhang, Y., & Dai, B. (2019). Learning path recommendation based on knowledge tracing model and reinforcement learning. In 2019 *IEEE 5th International Conference on Computer and Communications (ICCC)* (pp. 1881–1885).

Chatzimparmpas, A., Kucher, K., Rossi, F., & Angelini, M. (2023). From sensable to sensible spaces: Stress-aware deep reinforcement learning in virtual environments. *ACM Transactions on Interactive Intelligent Systems, 13*(4), 1–23.

Dong, J., Mohd Rum, S. N., Kasmiran, K. A., Mohd Aris, T. N., & Mohamed, R. (2022). Artificial intelligence in adaptive and intelligent educational systems: A review. *Future Internet, 14*(8), 245.

El Gourari, A., Raoufi, M., Skouri, M., & Ouatik, F. (2021). The implementation of deep reinforcement learning in e-learning and distance learning: Remote practical work. *Mobile Information Systems*, 2021, 1–11.

Fahad Mon, B., Wasfi, A., Hayajneh, M., Slim, A., & Abu Ali, N. (2023). Reinforcement learning in education: A literature review. *Informatics, 10*(3), 74. https://doi.org/10.3390/informatics10030074

Jing, Y., Zhao, L., Zhu, K., Wang, H., Wang, C., & Xia, Q. (2023). Research landscape of adaptive learning in education: A bibliometric study on research publications from 2000 to 2022. *Sustainability, 15*(3), 3115.

Li, X., Xu, H., Zhang, J., & Chang, H. H. (2021). Optimal hierarchical learning path design with reinforcement learning. *Applied Psychological Measurement, 45*(1), 54–70.

Muniasamy, A., & Alasiry, A. (2020). Deep learning: The impact on future eLearning. International *Journal of Emerging Technologies in Learning, 15*(1).

Riad, M., Gouraguine, S., Qbadou, M., & Aoula Es-Sâadia. (2023). Towards an adaptive e-learning system based on deep learner profile, machine learning approach, and reinforcement learning. *International Journal of Advanced Computer Science and Applications, 14*(5). https://doi.org/10.14569/IJACSA.2023.0140528.

Riedmann, A., Schaper, P., & Lugrin, B. (2025). Reinforcement learning in education: A systematic literature review. *International Journal of Artificial Intelligence in Education*, 1–55.

Shi, K. (2025). Adaptive learning resource recommendation for ideological and political education using Q-learning. In 2025 *3rd International Conference on Data Science and Information System (ICDSIS)* (pp. 1–6). IEEE.

Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction (Vol. 1, No. 1, pp. 9-11). Cambridge: MIT press.

Sypsas, A., Paxinou, E., Zafeiropoulos, V., & Kalles, D. (2024). Virtual laboratories in STEM education: A focus on Onlabs, a 3D virtual reality biology laboratory. In D. May, M. E. Auer, & A. Kist (Eds.), *Online laboratories in engineering and technology education (Lecture Notes in Networks and Systems*, Vol. 1135). Springer, Cham. https://doi.org/10.1007/978-3-031-70771-1_16.

Tan, C., Han, R., Ye, R., & Chen, K. (2020). Adaptive learning recommendation strategy based on deep Q-learning. *Applied Psychological Measurement, 44*(4), 251–266.

Thomaz, A. L., & Breazeal, C. (2006). Reinforcement learning with human teachers: Evidence of feedback and guidance with implications for learning performance. In A. Cohn (Ed.), P*roceedings of the 21st National Conference on Artificial Intelligence (AAAI'06)* (pp. 1000–1005). AAAI Press.

Wiering, M. A., & van Otterlo, M. (2012). Reinforcement learning. *Adaptive Learning and Optimization, 12*(3), 729.

Zafeiropoulos, V. (2021). *Human-computer learning interaction in a laboratory environment* (Doctoral dissertation, Hellenic Open University, School of Science and Technology).