JRITM
JOURNAL OF RESEARCH AND INNOVATION IN TECHNOLOGY, COMMERCE AND MANAGEMENT

Journal of Research and Innovation in Technology, Commerce and Management
Vol. 3 Issue 9, September 2026, pp. 39108-39117
ISSN: 3049-3129 (Online)

AN INTELLIGENT ADAPTIVE LEARNING FRAMEWORK BASED ON REINFORCEMENT LEARNING FOR EDUCATIONAL TRANSFORMATION

Reena
School of Engineering and Technology
Sushant University Gurugram, India
Email: reenayadav8088@gmail.com

Sherry Verma
School of Engineering and Technology
Sushant University Gurugram, India
Email: sherryverma@sushantuniversity.edu.in

Cite as: Reena, & Sherry Verma. (2026). AN INTELLIGENT ADAPTIVE LEARNING FRAMEWORK BASED ON REINFORCEMENT LEARNING FOR EDUCATIONAL TRANSFORMATION. Journal of Research and Innovation in Technology, Commerce and Management, Vol. 3(Issue 9), 39108–39117.
https://doi.org/10.5281/zenodo.22806715

DOI: https://doi.org/10.5281/zenodo.22806715

Abstract — Application of reinforcement learning (RL) in the educational context can yield a major change in how students learn and interact with education and how instructors assess student achievement. Application of RL in education enables individualized and adaptive learning, through which level of difficulty can be tailored depending on the performance of a student. Consequently, it may lead to an increase in motivation and engagement by students. The paper will discuss the theoretical basis of the RL and its usage in the field of education with a particular focus on its potential to establish intelligent, adaptive, and personalized learning environments. The dynamic interplay of learners and educational systems that RL facilitates is reward-based and makes it possible to continuously optimize learning trajectories, content delivery, and pedagogical choices. The paper identifies important uses which include intelligent tutoring systems, customized learning processes, classroom control and dynamic educational systems like Reinforcement Learning in Adaptive and Intelligent Educational Systems (RLATES). Moreover, it discusses the advantages of RL in enhancing student motivation, benefiting a diverse group of learners, and maximizing the amount of resources.

Keywords—Reinforcement Learning, Education, Machine Learning, Student, Artificial Intelligence

I. INTRODUCTION

To humans, difficulties with reinforcement learning may appear to be routine. Maybe it's the first way our brains conceptualize education. This may have evolved in this way. We learn from our surroundings what we are and are not allowed to do as newborns. For the first time, a baby playing with a pet could look at it oddly and poke it since

it doesn't know what it is. As time goes on and the baby plays with the pet, they gradually figure out what they can and cannot do to the creature. A foundational, and possibly primitive, method of learning is the use of reinforcement and punishment to guide trial and error. Norms and standards for proper behaviour are often set by family and society. That is why RL is ingrained in our human upbringing[1].

Although most people have heard of RL, it could be difficult for academics to grasp the concept. Research on reinforcement learning (RL) focuses on a specific set of machine learning algorithms that use positive and negative reinforcement to teach agents (human or otherwise) how to interact with and learn from their environments. The use of machine learning and AI algorithms into educational systems is a recent development with the aim of bettering the educational process. According to several writers, RL has the potential to make students completely rethink their attitude to and involvement in the learning process.

An introduction to machine learning may help clarify RL. For machine learning, supervised and unsupervised learning are the two main categories. Learning from one's surroundings, mostly with the use of labelled data sets, is the main focus of supervised learning, as the name implies [2]. Two popular applications of supervised learning are classification and regression. In contrast, unsupervised learning does not make use of labels or instructions when processing data. Consequently, the unsupervised learning algorithm learns the proper way to categories data and can forecast its location. One popular technique in unsupervised learning is clustering. Many algorithms fall under the category of semi-supervised learning; RL is the most well-known example of this class. Real-time learning (RL) is a machine learning method that puts agents in an interactive setting where they may learn from their experiences and behaviors by responding to rewards and punishments.

Real-life learning (RL) outshines both supervised and unsupervised approaches. Achieving a goal and learning from one's surroundings are its stated objectives. Therefore, RL's actions are influenced by the environmental input it gets, and whether that feedback is positive or negative for the algorithm. The algorithm modifies its actions in response to the valence of the input in order to maximize positive reinforcement and minimize negative punishment. Instead of receiving a smaller reward immediately, the algorithm might be tweaked to wait for a larger one at time t + 1. Because of this, the likelihood of a larger cumulative reward at the end of the learning period is enhanced [3].

Looking at semi-supervised learning via a systems perspective can provide extra insights. From the perspective of a block diagram, we take into account the inputs and outputs of the system and strive to get optimal outputs by manipulating the inputs [4].

The system is closed-loop in supervised learning, with error feedback acting as both a supervisor and an educational tool. In contrast, reinforcement learning (RL) relies on signals, often numerical and representing reward or punishment, to guide the system's behaviour. The signals could be noisy. It is the responsibility of the RL algorithm to make the necessary revisions and adjustments in light of the ever-changing environment in order to achieve optimum outputs, which are defined as maximizing rewards and minimizing penalties.

Based on the RL algorithms' complexity and rigor, technical evaluations and classifications are typically carried out. But in its most basic and early form, reinforcement learning is just a collection of characteristics that can handle any

kind of problem—the kind that is easily understood, the kind that isn't, and even the most difficult ones [5].

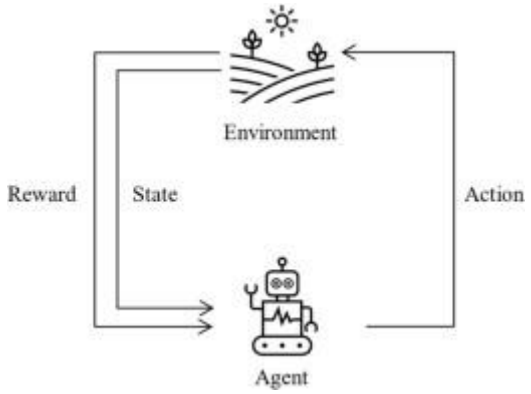


Figure 1 is a schematic depicting RL at a high level. An interactive setting is given to an agent. The environment is able to detect the agent's behavior and give feedback in the form of positive or negative incentives based on the agent's performance. As an added feature, the agent may record its surroundings after every activity. Therefore, the following are recognized as RL's core components:

Figure 1: High-level block diagram of RL

- **States:** The observations on the environment the agent can make upon acting in the environment
- **Action:** What an agent performs on the environment based on the state observations.
- **Reward:** The feedback the environment provides to the agent to shape their actions. A reward can be either positive (rewards) or negative (punishment).

II. REVIEW OF LITERATURE

Strielkowski, Wadim et al., (2024) Adaptive learning technologies and artificial intelligence (AI) are revolutionizing education in the modern era. They are making

education more personalized, accessible, and efficient while also guiding individuals to embrace, address, and mitigate sustainable development. This article examines these trends. The number of scholarly papers on this issue increased from 1 in 1990 to 636 in 2023, indicating that education has recently experienced a tremendous technological explosion spurred by numerous improvements in technology. The novel approach to valuing each student's distinct learning style and abilities, along with the ongoing digitalization and technological revolution in education, has opened the door for adaptive learning technologies. These innovative tools personalize educational experiences to meet the needs of individual learners. We can't hope to build a sustainable future without educating and informing our populace, which in turn fuels innovation and helps sustain economic development. From 1990 to 2024, a total of 3518 publications indexed in the Web of Science (WoS) database containing articles, proceedings papers, and book chapters utilizing the keywords "adaptive learning" and "AI" were analysed bibliometrically in our bibliographic study using VOS viewer. Our findings show that the COVID-19 pandemic's "digital surge" in education further strengthened the importance of recent technology innovations in revolutionizing adaptive learning. Adaptive education is an area that might benefit from our results if they are utilized by academics, researchers, relevant stakeholders, and policymakers [6].

Anne, Dan & Maina, Elizaphan. (2021) We provide an innovative three-stage model of adaptive online learning that takes into account various learner attributes. We construct an attribute model of a learner that includes the student's needs, a summary of their information, and the study area. We include learning style theories into the model to classify and identify individuals, with the goal of enhancing their online learning platform experience. The learner's emotional states may be retrieved from text inputs made during

platform interactions using the affective state extraction approach. After all that, we send the data the system retrieved to the adaptivity domain, where the learning route is organized into tutorials, lectures, and workshops based on established learning constraints using the off-policy Q-learning model free algorithm. The findings of the simulations demonstrate that when numerous qualities are present, the adaptability is greater than when only one trait is present. More than three traits should be considered in future studies [7].

Sfenrianto, Sfenrianto et al., (2018) Traditional learning approaches in teaching English assume that learners have same background, ability, and requirements. In this study, we propose an adaptive learning systems-knowledge level (ALS-KL) is a learning system that can personalize materials due to proficiency level of English language learners. A pre-test is given to measure the proficiency level of learners. The level is divided into elementary, intermediate and advanced. The ALS-KL provides the most suitable materials according to the needs of the learners. We applied the proposed system on 90 learners with different English proficiency level. The effectiveness of the system was evaluated using post-test. The results of the post-test showed that the proposed system was able to improve learners with intermediate-level and advanced level 1.5 times and 4 times respectively while reducing the number of learners with elementary-level by more than 50%. These indicate the effectiveness of the proposed system [8].

Iglesias, Ana et al., (2009) In the field of Adaptive and Intelligent Educational Systems (AIES), defining appropriate

pedagogical rules for student-centered tutoring ranks high among the most pressing concerns. In order for educational systems to automatically learn the most effective pedagogical policies for educating students, this study suggests incorporating Reinforcement Learning (RL) into the pedagogical module. Improving educational policies only via experience with other students who have comparable learning characteristics is one of the key features of this strategy. This report examines the educational system's learning performance from three key angles. To begin, there is a movement in the field of education toward more precise pedagogical policies. The second aspect to consider when applying RL to AIES is the function of exploration and exploitation tactics. A way to shorten the AIES training period is now available [3].

III. APPLICATIONS OF REINFORCEMENT LEARNING IN EDUCATION

Personalized, interactive, and highly effective learning are all possible because to reinforcement learning's (RL) useful applications in the classroom. Key applications of RL that are enhancing student experiences and educational achievements are highlighted below:

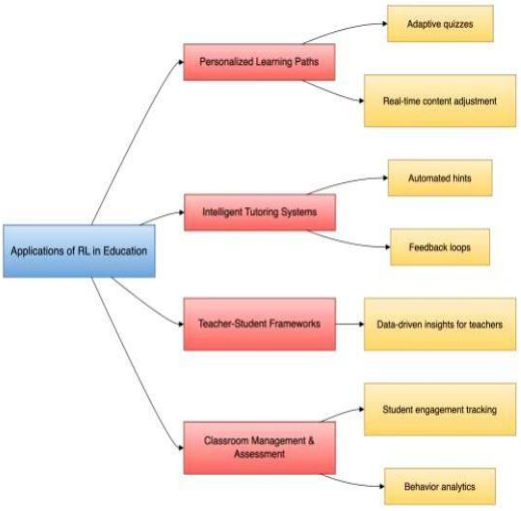


Figure 2: Applications of Reinforcement Learning in Education

Personalized Learning Paths

Continuous analysis of each student's progress, replies, and learning needs allows RL-based systems to generate personalized learning routes. Based on the student's performance, these systems modify the learning content and speed. For instance, the system can offer extra practice or divide the information into smaller pieces if a

learner is having trouble with a certain topic. Students are able to work at their own speed, which makes learning more interesting and less frustrating. Over time, individualized learning plans can facilitate consistent academic progress by progressively enhancing self-assurance and comprehension.

Intelligent Tutoring Systems

Students can get real-time feedback and assistance from RL-powered intelligent tutoring systems as they go through difficult courses. By analyzing a student's progress during the learning process, these systems can determine when the learner may benefit from additional support. For instance, the ITS may provide a helpful tip, propose an alternative method, or divide the issue into easier components if a student is struggling with a math's problem. Students can stay engaged and not feel overwhelmed by difficult ideas when RL-based teaching systems offer assistance at the perfect moment.

Teacher-Student Frameworks

Teachers and students in RL-based frameworks engage with course material as "agents," with the system providing feedback depending on the students' actions. The system gains a better picture of the student's success and potential areas of need with each encounter. Using this paradigm, educators may gain significant insights into how each student learns, allowing for more personalized lessons. For instance, if a student has recurring issues with a certain sort of difficulty, the instructor can tailor their help to address those specific needs. Teachers can enhance learning outcomes and provide more personalized education by integrating human assistance with findings from RL systems.

Classroom Management and Assessment

When it comes to managing classrooms and analyzing the quality of teaching, RL may also be useful. Robotic learning systems are able to monitor classroom dynamics in real-time by employing techniques such as Markov Chains. This allows them to detect trends in student behavior, engagement levels, and learning progress. For example, RL can reveal to educators the most productive methods of instruction and the moments when their pupils are most likely to be distracted. Teachers may use the assessment data that RL provides on their students and their teaching methods to make educated modifications that will help them better engage their students and assist their learning journey. RL also helps with classroom management.

IV. REINFORCEMENT LEARNING IN ADAPTIVE AND INTELLIGENT EDUCATIONAL SYSTEM (RLATES)

In recent years, there has been a surge in interest in distance learning. Learning by distance becomes crucial when both students and instructors are unable to physically be present in the same classroom. To meet the most fundamental needs of distant learning, students can at times rely on online materials (text, video, photos, voice, etc.) or teacher-provided online lessons. Nevertheless, students struggle to discover precise answers to queries through online resources, and the one-to-many teacher-student meetings fail to accomplish their

intended purpose—adaptive education. The high financial expense of one-on-one lessons makes it hard to apply this technique to all groups of

students, despite the fact that one-on-one classes are more productive and fulfilling than small group courses due to the fact that they allow for personalized teaching tactics. [9] In response to this need, the Adaptive Intelligence Educational System was created to empower students—those with access to a computer—to create their own unique approaches to learning. This way, they can experience the benefits of a low-cost, one-to-one teaching model and have their very own virtual instructor.

In order to learn student characteristics, AIESs utilize a range of machine learning approaches. The concept is to resequencing the course content modules depending on these qualities. Implementing reinforcement learning into the system enables the student to engage with the system and the system model to continually enhance its learning, leading to improved performance, as per the principles of AIESs. Reinforcement learning in AIESs is implemented using a system known as RLATES [10].

The pedagogical approach model and the knowledge model make up RLATES. Teaching material, such as which textbook chapters will be addressed and in what medium (video, audio, text, or graphics), is determined in the knowledge model. The pedagogical strategy model establishes the method of instruction by which the subject matter will be imparted. But you can't start teaching using RLATES right away. To begin, training the model entails providing it with training data; this will allow the system to understand how to best address students with varying characteristics in the classroom.

V. PROPOSED APPROACH

Main Components

As seen in Figure 3, the framework is composed of six primary parts. A user interface and the students' database are linked to the six components. Initialization of the system includes loading the state-action matrix history and the student's static information. This contains both static and dynamic data, such as gender, major, courses, and state-action-reward history, interaction level, log activities, and interaction logs. Step 2 involves loading the student's state. At the very least, this is her status as it was when she last used the system. A state that corresponds to her static data is assigned to her if this is her first time. In the third stage, a course of action is proposed, which often include suggested reading or entertaining content, such as a quotation, a bit of advice from her teacher in written or audio form, a video, or even a joke. Step four involves calculating the benefit of the activity. The learner is requested to assign a value based on her level of pleasure with the recommended content as a direct incentive. In addition, her pleasure and the amount of interaction with the system are used to determine the updated

incentive. To keep the proposed action's rewards up-to-date for the next time the learner uses the system, we give each of these activities a value between 1 and 5.

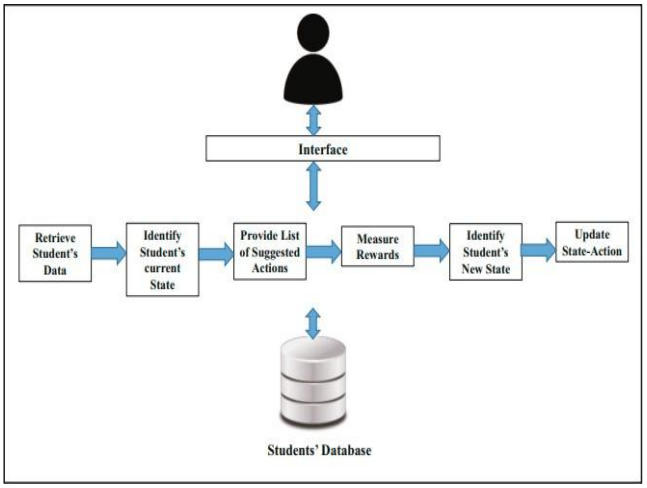


Figure 3: The framework of the proposed system

The results of the tests and homework she was given also provide an indirect metric. If the final grade (in points) is dropped, a negative reinforcement is added to the recommended course of action during the semester. As an illustration, the total reward that the set of suggested activities would earn will be -0.5 if the student's grade goes from 2.5 to
2. If there are five possible steps to take, the average value will be 0.1, which is negative. Accordingly, the system's primary objective is to ascertain, for each student's condition, the sequence of events that will maximize her engagement with the system during the semester while simultaneously improving her learning results. Identifying the student's new state is the fifth stage. To do this, just provide the student with a selection of states from which to select. If she feels like none of the existing states adequately account for her present situation, she is also asked about adding a new state to the system. In order to improve the system's performance, this stage involves adding a list of proposed actions for the newly-added states once an expert has analyzed them. This is done periodically, offline (for example, at the conclusion of each semester). Finding the right Trifecta of State, Action, and Reward is the biggest obstacle in RL-based personalized learning systems.

State-Action-Reward

One of the first and most important steps in building a personalized learning system is deciding which elements are important to include. By keeping track of these variables, the customization is carried out efficiently. A thorough and meticulous review of the research highlighting the elements that influence learning in various contexts was carried out to ascertain the components that should be incorporated into an efficient personalized learning system. The following groups
best describe the variables that need to be assessed according to this inquiry.

- Personal Factors
- Social Factors
- Cognitive Factors
- Structural Factors
- Environmental Factors

Simulation Setup

The simulation experiments make use of a system that has twenty states and twenty actions. An example of a state-action matrix might be one with dimensions 20 × 20. Additionally, Q-values (rewards) that follow the Normal distribution (mean = 0, standard deviation = 1) are initially populating the matrix. Furthermore, the action to be selected is determined using an e-greedy strategy, where e is set to 0.1. An e-greedy policy uses a probability of e to choose actions with maximum rewards. By avoiding the need to maximize rewards, this paves the way for exploration of the surroundings. Each stage has twenty different acts, and the rewards for each one are decided at random. Ten of the activities, however, had negative incentives attributed to them, while the other ten had good ones. The learning rate is also set to 0.1 along with the other parameters of the model. In addition, ten pupils' actions were replicated. A maximum of one hundred iterations were employed.

VI. RESULTS AND DISCUSSION

The number of iterations versus the number of actions that got positive rewards for each simulated student behavior (denoted by S1 to S10 in the legend) is shown in Table 1. An increasing number of iterations results in an increasing number of recommended activities receiving favorable rewards, as seen in the table. After enough iterations, the simulated system can determine the optimal actions to take for each student-state pair.

The figure 4 and Table 1 show the pattern of positive reward accumulation across ten simulated student instances (S1 to S10) over increasing runs from 10 to 100 iterations. The results indicate that in the initial stages, the number of positive rewards remains low and inconsistent, showing that the system is still in the learning phase. As the number of iterations increases, a gradual improvement is observed in most student instances, although some fluctuations appear in the mid-runs. A significant rise in positive rewards is clearly seen at 90 and 100 runs for all instances, indicating that the system reaches a learning convergence point and becomes more efficient in identifying optimal actions.

Table 1: Positive Reward Accumulation across Student Instances

Runs	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
10	5	3	2	3	2	4	3	2	3	2
20	6	4	3	4	3	5	4	3	4	3
30	12	10	4	5	3	14	11	4	5	4
40	27	5	3	4	2	6	5	3	4	3
50	43	4	2	3	2	5	4	2	3	2
60	44	5	2	3	3	6	5	2	3	2
70	49	4	3	4	2	5	4	3	4	3
80	55	5	3	4	3	6	5	3	4	3
90	85	80	78	82	75	88	84	79	81	77
100	90	85	80	88	78	92	87	82	84	79

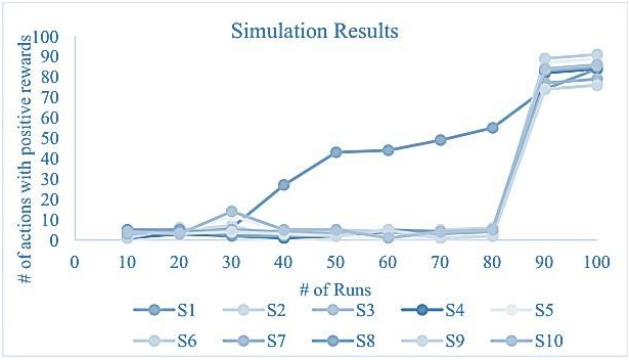


Figure 4: Positive Reward Accumulation across Student Instances

VII. CONCLUSION

Reinforcement learning is a strong paradigm to transform the current educational systems by having adaptive, personalized and intelligent learning environment. With the help of reward-based learning devices, RL-based systems are able to constantly optimize teaching methods, streamline content delivery, and dynamically adjust to the needs of individual learners. Nevertheless, to implement RL successfully in the educational field, it is necessary to discuss such issues as the scarcity of data, the absence of standardization, computational complexity, and

the necessity to have models that are robust and scaled-up. Being fair, inclusive, and accessible is also essential to avoid bias and ensure equitable learning opportunities. Further studies are essential to create lightweight, interpretable, and transferable RL models, better data collection structures, and multi-disciplinary collaboration. Reinforcement learning is an incredibly promising field that can transform the future of education by introducing the future of education between the traditional teaching processes and the smart and technology-oriented learning processes, which eventually can lead to more effective, inclusive, and sustainable learning processes.

REFERENCES

[1] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. 1999.

[2] Bennane, "Adaptive educational software by applying reinforcement learning," *Informatics in Education*, vol. 12, no. 1, p. 13, 2013.

[3] Iglesias, P. Martínez, R. Aler, and F. Fernández, "Learning teaching strategies in an adaptive and intelligent educational system through reinforcement learning," *Applied Intelligence*, vol. 31, no. 1, pp. 1–18, 2009.

[4] S. Narvekar, B. Peng, M. Leonetti, J. Sinapov, M. E. Taylor, and P. Stone, "Curriculum learning for reinforcement learning domains: A framework and survey," *Journal of Machine Learning Research*, vol. 21, no. 1, p. 82, 2020.

[5] S. Thrun and M. L. Littman, "Reinforcement learning: An introduction," *AI Magazine*, vol. 21, no. 1, pp. 103–110, 2000.

[6] W. Strielkowski, V. Grebennikova, A. Lisovskiy, G. Rakhimova, and T. Vasileva, "AI-driven adaptive learning for sustainable educational transformation," *Sustainable Development*, vol. 33, no. 2, pp. 1921–1947, 2024, doi: 10.1002/sd.3221.

[7] S. Sfenrianto, Y. Hartarto, and H. Akbar, "An adaptive learning system based on knowledge level for English learning," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 13, no. 2, p. 191, 2018, doi: 10.3991/ijet.v13i12.8004.

[8] D. Anne and E. Maina, "Reinforcement learning approach for adaptive e-learning based on multiple learner characteristics," *Open Journal for Information Technology*, vol. 4, no. 2, pp. 55–76, 2021, doi: 10.32591/coas.ojit.0402.03055o.

[9] E. Mousavinasab, N. Zarifsanaiey, S. R. N. Kalhori, M. Rakhshan, L. Keikha, and M. G. Saeedi, "Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods," *Interactive Learning Environments*, vol. 29, pp. 142–163, 2021.

[10] S. Suebnukarn, "Intelligent tutoring system for clinical reasoning skill acquisition in dental students," *Journal of Dental Education*, vol. 73, pp. 1178–1186, 2009.

[11] Sakthivelu, Karthigaiselvi, "Adaptive

Intelligence Learning Framework (AILF): A Functional Optimization Model FAI: (E×A) → (T,L) for Dynamic Educational Systems," *International Journal of Applied Mathematics*, vol. 38, no. 5, pp. 462–477, 2025, doi: 10.12732/ijam.v38i5.732.

[12] J. Hu and G. Jin, "An intelligent framework for English teaching through deep learning and reinforcement learning with interactive mobile technology," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 18, no. 9, pp. 74–87, 2024, doi: 10.3991/ijim.v18i09.49289.

[13] F. A. Dorça, L. V. Lima, M. A. Fernandes, and C. R. Lopes, "Comparing strategies for modeling students learning styles through reinforcement learning in adaptive and intelligent educational systems: An experimental analysis," *Expert Systems with Applications*, vol. 40, pp. 2092–2101, 2013.

[14] M. Liu, X. Shen, and W. Pan, "Deep reinforcement learning for personalized treatment recommendation," *Stat. Med.*, vol. 41, no. 23, pp. 1–23, 2022.

[15] Z. Fu, L. Yu, and X. Niu, "TRACE: Travel reinforcement recommendation based on location-aware context extraction," *ACM Trans. Knowl. Discov. Data*, vol. 16, no. 4, pp. 1–22, 2022.

[16] P. Wei, S. Xia, R. Chen, J. Qian, C. Li, and X. Jiang, "A deep-reinforcement-learning-based recommender system for occupant-driven energy optimization in commercial buildings," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6402–6413, 2020.

[17] S. Ji, Z. Wang, T. Li, and Y. Zheng, "Spatio-temporal feature fusion for dynamic taxi route recommendation via deep reinforcement learning," *Knowl.-Based Syst.*, vol. 205, p. 106302, 2020.

[18] R. Han, K. Chen, and C. Tan, "Curiosity-driven recommendation strategy for adaptive learning via deep reinforcement learning," *Br. J. Math. Stat. Psychol.*, vol. 73, no. 3, pp. 522–540, 2020.

[19] X. Wang, Y. Wang, D. Hsu, and Y. Wang, "Exploration in interactive personalized music recommendation: A reinforcement learning approach," *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 11, no. 1, pp. 1–22, 2014.

[20] J. Zhang, B. Hao, B. Chen, C. Li, H. Chen, and J. Sun, "Hierarchical reinforcement learning for course recommendation in MOOCs," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 1, pp. 435–442, 2019.