International Journal of Applied Mathematics

Volume 38 No. 5s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

ROBOLEARN: FACILITATING ADAPTIVE LEARNING WITHIN AN ARTIFICIAL INTELLIGENCE FRAMEWORK

Mathivanan Viruthachalam[1], Vimal Kumar Stephen[2], Ramesh Palanisamy[3*], Senthil Jayapal[4], Mohammed Tauqeer Ullah[5], Mr.Mohamed R. Rafi[6], Annadurai Manickam[7]

[1,2,3,4,5,6]College of Computing and Information Sciences,

[7]Preparatory Studies Centre.

University of Technology and Applied Sciences – Ibra, Sultanate of Oman.

E-mail: vmathi@utas.edu.om[1], vimal.victor@utas.edu.om[2], rameshphd26@gmail.com[3*], senthilsjs@gmail.com[4], MohdTauqeer.Ullah@utas.edu.om[5], Mohamed.Rafi@utas.edu.om[6], Annadurai.Nadar@utas.edu.om[7]

Abstract

The range of student skills, preferences, and learning rates has led to a considerable increase in the need for individualized learning solutions in recent years. Using the RoboLearn environment, this work introduces an intelligent adaptive learning system that incorporates Proximal Policy Optimization-based Reinforcement Learning (PPO-RL) as its central decision-making mechanism. The goal is to provide a dynamic AI-based environment that can continually modify instructional tactics to accommodate each learner's changing demands. Experiments were carried out utilizing the Enrolled Students Dataset (2023/2024) from the Oman Open Data Portal, which has comprehensive data on learner engagement, academic achievement, and advancement, in order to assess the efficacy of the suggested PPO-RL strategy. Three popular baseline reinforcement learning models—Q-Learning, Deep Q-Network (DQN), and Asynchronous Advantage Actor-Critic (A3C)—were compared to the PPO-RL method.

Both regression-based assessment measures and classification-based metrics were used to evaluate the performance. PPO-RL produced better quantitative findings, with an F1-Score of 91.7%, Accuracy of 92.5%, Precision of 91.0%, and Recall of 92.3%. Its great predictive capacity was shown by the fact that it had the lowest Root Mean Squared Error (RMSE) at 5.88, the lowest Mean Absolute Error (MAE) at 4.12, and the highest R2 score of 0.94. Additionally, PPO-RL outperformed current models in terms of cumulative reward per episode and flexibility to unseen learners, demonstrating quicker policy convergence and needing fewer training episodes to attain optimum strategy. Visual confirmation of PPO-RL's resilience across key performance metrics was provided via a confusion matrix analysis and heatmap. Furthermore, an ablation research demonstrated the significance of distinct PPO elements including entropy bonus and policy clipping, in attaining performance stability. All

Received: August 08, 2025

things considered, this research shows that PPO-RL in RoboLearn offers a strong and expandable real-time adaptive learning system. The suggested approach is ideal for contemporary digital education platforms as it enhances learning outcomes while streamlining the delivery of teaching. This research was funded by The Research Council (TRC), Oman, through the University of Applied Sciences and Technology, Ibra.

Keywords: Adaptive Learning Process, Artificial Intelligence (AI), Proximal Policy Optimization-based Reinforcement Learning (PPO-RL), RoboLearn.

1. Introduction

Customized and dynamic learning environments made possible by AI-powered adaptive learning systems are causing a sea change in the educational landscape. Because of the inadequacy of conventional, one-size-fits-all methods, adaptive learning has emerged as a formidable alternative. Machine learning algorithms, natural language processing, and predictive analytics are a few examples of the AI capabilities that can gather and analyze massive volumes of data. Because of this, adaptive learning systems can change their approach, materials, and feedback based on each learner's unique requirements [1]. There has been encouraging progress in raising engagement and enhancing learning outcomes through the use of adaptive learning platforms powered by AI. Nonetheless, successful execution and ethical concerns must be prioritized. This article takes a look at the intersection of AI and adaptive learning and talks about the pros, cons, and ramifications of that intersection. The "learning analysis" can be used to enhance students' learning abilities by collecting and analyzing learning data to uncover the rules of education and learning (Baker & Invent ado, 2014). [2]. We add to the continuing conversation regarding the future of higher learning in a changing setting by investigating the fundamentals of AI, its function, and the viewpoints of students, teachers, and organizations. Learners are given the tools they need to succeed in the digital age through AI-powered adaptive learning.

The sophisticated technology and abilities of artificial intelligence are crucial to adaptive learning systems. For the purpose of analyzing massive volumes of data, including learner profiles, performance statistics, and learning materials, machine learning techniques are employed [3]. Intelligent predictions regarding student preferences, needs, and future performance can be generated by AI algorithms through the identification of trends, patterns, and correlations in data. Because of this, adaptive learning systems may tailor their lessons to each student's unique requirements in terms of pace, material, and instructional tactics, making for a more efficient and effective educational experience overall [4].

The advancement of any civilization has always hinged on its educational system. It has also evolved significantly over the years to accommodate new threats like climate change and technology advancements [5]. As we move into the post-COVID era, which marks the coronavirus pandemic's transition into an endemic state (2020–2023), adaptive learning technologies and AI are reshaping education and its function in EfS in unprecedented ways.

International Journal of Applied Mathematics

Volume 38 No. 5s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Personalized learning through data analytics and artificial intelligence is the goal of adaptive learning technologies [6]. Based on a student's progress, learning style, and preferences, these technologies personalize the way educational content is presented. The goal of adaptive learning technology is to improve student engagement and outcomes by the provision of personalized support, targeted feedback, and optimized learning pathways through the ongoing analysis of data on student interactions.

A new age of personalized, adaptable learning has begun with the advent of artificial intelligence (AI) in the classroom. This revolutionary effect of AI opens the door to new ways of thinking about sustainable education, which in turn transforms static, teacher-centered lessons into interactive, student-driven ones. While artificial intelligence (AI) may not have an immediate impact on classroom instruction, it will have far-reaching consequences in the years to come [7].

Learners' needs, interests, and progress can be taken into account through the use of technology in adaptive learning, which is a method of instruction. It uses AI and data-driven algorithms to adapt the lesson's content, delivery, and tempo in real time depending on how well students are doing [8]. Adaptive learning improves educational outcomes by tailoring instruction to meet the unique needs of each student. This approach increases student engagement, boosts efficiency, and improves learning overall. We take a look at adaptive learning and its role in online education, showcasing its advantages.

With its adaptability, scalability, and ability to provide students with individualized lessons, online education has rapidly expanded in recent years [9]. When discussing online education, the term "adaptive learning" describes the process of intelligently and dynamically tailoring course materials, quizzes, and other learning activities to each student's specific interests, strengths, and weaknesses. Personalised learning experiences, optimal learning outcomes, and increased student engagement can all be achieved through the use of adaptive learning systems, which analyse and analyse data from learners to make educated decisions.

The term "adaptive learning" refers to a style of teaching that makes use of cutting-edge tech, especially artificial intelligence algorithms, to personalize lessons, quizzes, and tests for each student. Its stated goal is to dynamically modify the learning process according to the preferences, knowledge level, performance, and learning style of each individual student [10]. Adaptive learning systems guarantee that students obtain the best possible educational resources by constantly analyzing data such as results of assessments, pattern of interaction, and progress tracking. This allows for focused interventions that are both timely and appropriate. Adaptive education systems are booming in popularity due to their ability to deliver instructional knowledge and adjust to the needs of individual pupils.
An innovative learning tool, robots are widely acknowledged. A large body of research suggests that this innovative method has the potential to revolutionize the way we teach and help students succeed in a variety of classroom settings. emphasized that educational robots have the potential to improve students' computational thinking, creativity, self-efficacy, and collaboration/cooperative abilities, among others. Languages, mathematics, science, and the

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 5s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

multidisciplinary STEAM fields are just a few of the areas that have found uses for robots. [11].

Identified the potential efficacy of robots as instructional tools in schools and found favourable academic achievement in STEM through a methodical examination of ten papers on the topic. We reviewed 22 research on the topic of robotics education from kindergarten through high school and discovered that the vast majority of studies focused on elementary school pupils, who were most often taught using LEGO robots. Furthermore, research suggests that robots can function as both a student and a teacher, as well as a tool to aid in the learning process. By having children teach the Nao robots handwriting, researchers were able to study the impact of children's handwriting abilities. In order to teach programming to kids in grades K-12, the researcher utilized the Baxter robot [12]. As a means of education, the researcher employed robotics. As part of its role in education, the robot told stories to kids.

Contribution of this study: In the Rob Learn environment for adaptive learning, this work introduces a unique Proximal Policy Optimization–Reinforcement Learning (PPO-RL) model, making a substantial addition to the area of educational technology. By continually evaluating and reacting to student interactions, this system dynamically personalizes learning paths in contrast to conventional one-size-fits-all methods. In order to maximize teaching tactics and guarantee that both fast and slow learners get the proper assistance, the main contribution is the combination of deep reinforcement learning with real-time learner data. By comparing the PPO-RL model to well-known algorithms such as Q-Learning, DQN, and A3C utilizing a wide range of performance criteria, including as accuracy, adaptability, and convergence, the research pushes the field even further. The study shows that PPO-RL not only enhances learning results but also lowers training time and prediction error using a mix of quantitative assessment, error analysis, and ablation investigations. The research is a useful resource for further advancements in AI-powered education as it also presents a replicable framework and simulation approach that can be used to other intelligent tutoring systems.

2. Literature Review

As AI has progressed at a rapid pace, it has also made significant contributions to robotics. On one hand, traditional robots like LEGOs are great for teaching kids how to put them together and program them, but on the other hand, AI-powered robots can have one-on-one conversations with humans utilising speech recognition, picture recognition, or even just plain old language [13]. This aspect of AI-robots expands the range of possible learning situations, including but not limited to: student mentoring in conjunction with a system for intelligent tutoring (ITS), interaction between learners and the system, and the provision of personalised feedback and instructions. To rephrase, AI-robots can help students learn new things even when their instructors aren't around. Studies in fields as diverse as online learning, nursing education, engineering, mathematics, and languages have all demonstrated AI's efficacy in the classroom [14]. It is important for educational educational professionals and researchers to be able to recognise the trends and functions of AI-robots due to the rapid growth and numerous uses of artificial intelligence and robots in education.

Received: August 08, 2025

International Journal of Applied Mathematics

Volume 38 No. 5s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Furthermore, a number of scholars have stressed the significance of bibliometric analysis and systematic reviews in spotting trends within particular areas of study. To better understand the area and to innovate new, valuable studies, researchers should conduct a comprehensive review that systematically searches for, identifies, selects, evaluates, and synthesises studies relevant to the research issue. Researchers may struggle to understand the overall research framework of AIRE due to a lack of specific data connecting educational requirements with AI technologies in robotics, even though previous studies have looked at the effects of AI on education or given information on the utilisation of AI-Robots in Education (AIRE). On top of that, AIRE has not been exhaustively reviewed. Thus, this study examined AIRE research trends, characteristics (such as study participants, length of study, atmosphere for learning, application field, data analysis, evaluation of learners' performance, and learning methodologies), and research issues by mining the SSCI (Social Sciences Citation Index) articles found in the WOS database [15].

Table 1. Summary on related works

References	Objectives	Methods	Limitation
16	Adaptive learning systems that include AI to personalize education using machine learning and predictive analytics are the focus of this article. Artificial intelligence In order to safeguard student data and stay in line with data protection laws, privacy and data security must be given serious consideration.	Machine learning	In order to safeguard student data and stay in line with data protection laws, privacy and data security must be given serious consideration.
17	This study takes a close look at how AI and adaptive learning are changing education in the modern day. These changes include making education more accessible, efficient, and personalized, and they are also guiding individuals to accept, address, and mitigate sustainable development. Intelligent educational platforms However, it must be kept in mind that although AI could completely transform education by offering tailored learning experiences and enhancing productivity, it is also crucial to acknowledge the constraints of this technology	Intelligent tutoring systems	However, it must be kept in mind that although AI could completely transform education by offering tailored learning experiences and enhancing productivity, it is also crucial to acknowledge the constraints of this technology.
18	In order to improve the experience of	predictive	Personalized and

Received: August 08, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

	learning and outcomes, this research seeks to provide light on how adaptive algorithms might be used by e-learning systems to personalize the distribution of content. Predictive analytics Personalised as well as adaptive learning in the age of Education 4.0 and 5.0 can be fully realized through artificial intelligence (AI) if educational institutions use a multi-pronged approach that considers technical, pedagogical, ethical, as well as social factors.	analytics	personalized education in the age of Education 4.0 and 5.0 can be fully realized through artificial intelligence (AI) if educational institutions use a multi-pronged approach that considers technical, pedagogical, moral, and social factors.
19	Educational circumstances, domains of applications, analyses of data, evaluations of the performance of students, learning methodologies, roles of AI-robots, and academic difficulties were all examined in the study. Participants were also analysed. Instructional strategy Institutional and cultural hurdles exist when the personalized, adaptable methods provided by adaptive technology for learning are at odds with long-standing educational policies and practices.	pedagogical approach	There are cultural and institutional barriers, where established educational norms and policies may not align with the flexible, personalized approaches offered by adaptive learning technologies
20	Recognising the substantial progress that has been achieved recently, this research seeks to thoroughly document and record the most current advancements in the field. Conventional techniques of instruction Peril of data breaches, illegal access, and abuse of private student data.	traditional teaching methods	Risk of data breaches, unauthorized access, and misuse of sensitive student information.
21	The goal is to personalize the learning experience for each student in real-time according to their preferences, performance, knowledge level, and preferred method of learning. Systems for adaptive evaluation Because they aren't used to them, they worry about being obsolete, or they don't want to put in the time and energy needed for training,	Adaptive assessment systems	Educators may resist these technologies due to a lack of familiarity, fear of obsolescence, or concerns about the time and effort required for training.

Received: August 08, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

	teachers may be resistant to modern technologies.		
22	Data analysis, learning methodologies, evaluations of learners' performance, research issues, learning environments, application domains, initial author's nationality, participants, duration of the studies, and year of publication were all part of the study. The AI viewpoint of the intelligent agent Some teachers may be hesitant to include these technologies into their lessons because they don't fully grasp how they function, while others may be worried about the extra work that would be needed.	The intelligent agent perspective of AI	This resistance can be due to a lack of understanding of how these technologies work or concerns about the additional time and effort required to integrate them into existing curricula.
23	This study aims to map the current utilization of AI/ML in e-learning for adaptive learning, elucidating the benefits and challenges of such integration and assessing its impact on student engagement, retention, and performance.	Educational robots	the risk of over-reliance on AI technologies, where educators and students may depend too heavily on automated systems
24	The purpose of this research is to examine the present state of artificial intelligence and machine learning integration in online education with an eye towards adaptive learning, drawing attention to the pros and cons of this trend while measuring its effect on students' interest, persistence, and achievement. Concerns about the potential dangers of putting too much faith on artificial intelligence systems in the classroom.	Intelligent virtual reality	The infrastructure in traditional schools may not support advanced technological tools, requiring significant investment in hardware, software, and reliable internet connectivity
25	Artificial intelligence (AI) is expected to have a positive effect on the delivery of personalized learning aids, students' learning styles, and their rate of learning, according to the study. Project Based Learning (PBL) Some teachers are reluctant to try new things because they	Project Based Learning (PBL)	There is resistance to change from educators who are accustomed to conventional teaching methods and may be skeptical about the effectiveness of new

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 5s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

	are used to the status quo and don't believe in the power of technology in the classroom.		technologies

3. Methodology

Dataset

A recent and useful resource for examining trends in higher education and creating AI-based adaptive learning environments like RoboLearn is the Enrolled Students Dataset (2023/2024) from Oman. The data has been supplied via the Open Data Portal of the Ministry of Higher Education, Research, and Innovation (MOHERI).

https://eservices.moheri.gov.om/OpenData/OpenDataSet.aspx. This dataset includes comprehensive details for 28,950 students who were enrolled both locally and abroad. It contains information on the student's gender, place of origin, study location (within or outside of Oman), school name, program qualifications, major group name, and offer type. Researchers may monitor demographic trends, examine program preferences, and create instructional paths that can guide adaptive learning tactics thanks to this extensive feature set. For example, this data may be used to mimic learner behavior and tailor material depending on academic background, gender, or area using reinforcement learning models such as Proximal Policy Optimization (PPO-RL).

The dataset is particularly helpful for:

• Analyzing how various learner profiles interact with educational systems;

• Modeling student variety and adaptive learning responses.

• Developing customized feedback loops for smart teaching programs.

Problem Definition

With the main goal of creating a reliable translation function for predicting a robot's upcoming velocity v_t based on its present state s_t , we tackle the problem of learning effortless motion planning for mobile ground robotics in this work. The following crucial elements are included in the state s_t.:

$$v_t = f(x_t, p_t, v_{t-1}, \theta_t, \alpha t) \quad (1)$$

Relative Target Position p_t indicates the target's position in relation to the robot, whereas Sensor Information x_t contains information from the robot's sensors to comprehend the surroundings. To help with dynamic stability, the Previous Velocity v_{t-1} shows the robot's most recent learning recorded speed. The Rotation Degree αt is essential for lining up the robot with the goal, while the Yaw Angle θ_t determines the robot's orientation. In order to provide precise and quick response in dynamic environments, we want to turn these states into actionable insights, particularly to calculate the next velocity v_t.

Received: August 08, 2025

International Journal of Applied Mathematics

Volume 38 No. 5s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Data Processing

In order to improve robot adaptive learning in a complicated artificial intelligence environment, this section describes a LiDAR data processing technique. Our goal is to minimize computing burden while condensing raw sensor data for Actor-Critic models to make the best decisions.

The LiDAR sensor measures the robot's environment to record the pupils' emotions and learning requirements. We suggest splitting these measurements into ten batches of three data points each in order to guarantee efficiency and conform to Actor-Critic architecture. We determine the nearest obstacle in the robot's range of vision by choosing the smallest distance within each batch. Ten streamlined observations are produced by this approach. This method improves our adaptive learning models' efficacy and efficiency. We lower data dimensionality, expedite computational procedures, and give students' emotions priority by concentrating on the challenge in each learning batch. These carefully chosen data points provide practical insights for producing intelligent instructions and enhancing the learning capacity of our robotic systems.

The LiDAR data processing method, as seen in Figure 1, determines the closest obstacle in each batch, condensing unprocessed sensor data into brief observations for the best possible decision-making in the robot adaptive learning process.

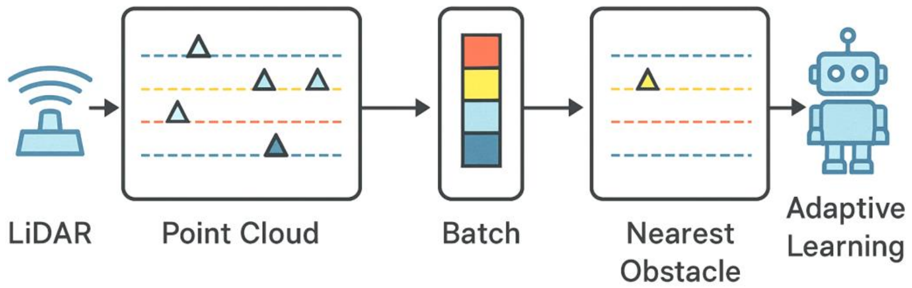

Figure 1. LiDAR method to process data

Reinforcement Learning

An agent that uses reinforcement learning (RL), a machine learning approach, continually interacts with its AI environment and maximizes a reward signal to learn how to make the best actions in adaptive learning. This method, which involves a trial-and-error procedure to develop and enhance decision-making abilities, draws inspiration from the learning processes seen in both humans and animals. RL is used in a variety of fields where it is essential to make wise judgments under trying circumstances. An agent must make judgments in real-world circumstances, and it depends on the AI environment's feedback from students—in the form of incentives or penalties—to modify and improve its choices. The agent has to develop a policy—a set of guidelines or tactics—that directs its decision-making in order to maximize the long-term cumulative payoff. Crucially, the agent learns by actively interacting with the environment, picking up information via repeated efforts, and taking lessons from both

Received: August 08, 2025

International Journal of Applied Mathematics

Volume 38 No. 5s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

successful and bad results rather than being given clear instructions or direction on how to accomplish this goal. Reinforcement learning (RL) is a framework that can solve optimum control problems without the need for explicit models. The general feedback control structure's controller receives input from the plant in the form of status signals and responds appropriately. The decision rule in RL is referred to as a policy and is based on student input. Actuation changes the state of the system, and a reward function is used to evaluate the transition to the new state. Maximizing the total reward derived from each starting condition is the aim of optimal control. Optimizing the system's long-term performance is the aim of this sequential decision-making process. This work focuses on proximal policy optimization (PPO), a policy gradient approach among many resilient model-free RL algorithms. PPO uses observable data to directly improve policy parameters.

The Robotic in an RL-based control problem in adaptive learning links CP states (represented by the notation Xt_d) to control actions (represented by the notation Ut_d) at each discrete time step t_d of the control or learning process. A reward value rt_d , which may vary depending on the specifics of the CP, is calculated by a reward function for each control action utd performed on a certain state Xt_d of the CP. With the equation $\pi : X \times U \rightarrow [0,1]$, where X is the continuous state-space of the CP covered in the previous section and U is the discrete action space, the objective is to identify a control policy π that provides a probability distribution of actions across states. The probability of performing the control action Ut_d in state Xt_d is represented by the notation $\pi(Ut_d \mid Xt_d) \in [0,1]$ rather than $\pi(Ut_d \mid Xt_d)$. In an RL issue, the RL control policy π is often modeled as a NN, and its internal structure is represented by a parameter vector that includes all of the NN's weights and biases, denoted by σ in this instance.

$$\sigma = [\sigma_1 \dots \dots \sigma_m]^T \in R^m \tag{1}$$

The following relation represents the probability of performing a certain action in a specific state using the control policy π parameterized with σ as the parameter, where m is the number of parameters $\sigma j, j = 1 \dots m$, of the control policy NN.

$$\pi(u|x,\sigma) = pr\{ Ut_d = u | Xt_d = x, \sigma t_d = \sigma\}, \tag{2}$$

using Pr, which is the symbol for a probability distribution. A set of probabilities, or a probability distribution of actions over states, is the result of the control policy π specified in (3). Each member of this set corresponds to a distinct control action in the discrete action space U. In the context of the RL-based control issue that is discussed next in this study, the discrete action space u is defined as

$$u = [u_1 \dots u_i \dots u_n]^T \in R^m \tag{3}$$

The adaptive learning control strategy NN π stated in (4) uses the softmax function to fit all probability values in the range [0,1] and ensure that the sum of all distributional probabilities equals 1. The control action Ut_d selected in a certain state Xt_d of the CP has the greatest probability when the probability distribution supplied by (3) is normalized using

Received: August 08, 2025

International Journal of Applied Mathematics

Volume 38 No. 5s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

the softmax function. When defining the softmax function, this is technically stated as follows:

$$Ut_d = softmax(\pi(u_k|Xt_d, Ut_d) = argmax\, \pi(u_k|Xt_d, Ut_d) \quad (4)$$

In the discrete action space U described in (4), the control action index is denoted by $k = 1 \dots n$. Driving the CP in a condition where its output tracks—that is, continuously follows a reference value—is the aim of the RL-based reference tracking control mechanism described in this study. The control policy NN must repeatedly change its internal structure based on the control error et_d with the specification in order to provide improved actions as the learning process progresses.

$$et_d = y_{ref} - yt_d \quad (5)$$

where y_{ref} is the reference input, or the set-point (SP) in terms of AC, and yt_d is the CP's output at the discrete time step t_d.

Proximal Policy Optimization (PPO)

The Proximal Policy Optimization (PPO) algorithm, a well acknowledged reinforcement learning method recognized for its stability and effectiveness in continuous adaptive learning in RoboLearn tasks, was used in our study. With a restriction to limit policy changes and avoid significant, sudden departures, PPO repeatedly updates the policy to maximize the predicted cumulative reward, in contrast to the Deep Deterministic Policy Gradient (DDPG) method. PPO is especially well-suited for our students' motion planning issue because of its regulated approach to policy updates, which guarantees more consistent training progress.

The policy loss and the value loss are two essential parts of the loss function of the Proximal Policy Optimization (PPO) method. The policy loss is designed to regulate policy adjustments, keeping them small and avoiding a substantial departure from the present policy. The stability of the training process depends on this regulated adjustment, which makes it possible for policy performance to increase gradually and steadily. The policy loss may be expressed mathematically as follows:

$$\mathcal{L}clip(\theta) = \mathbb{E}[\, min(r_t(\theta)\mathcal{A}^\wedge, clip(\, r_t\, (\theta), \epsilon + 1, \epsilon - 1))] \quad (6)$$

$$r_t(\theta) = \pi_\theta(s_t, a_t)/\pi_\theta old(s_t, a_t), \epsilon = 0.2 \quad (7)$$

Here, θ stands for the parameters of the policy, $r_t(\theta)$ is the probability ratio of the new policy to the old policy, $\mathcal{A}^\wedge$ is the advantage function, ϵ is a hyperparameter regulating the scope of permitted policy modifications, and $\mathcal{L}clip(\theta)$ represents the clipped surrogate goal.

The critic network is linked to the value loss, which tries to minimize the discrepancy between the actual discounted cumulative reward r_t and the forecast value $V(st)$ This is how the value loss is computed:

$$LVF(\theta) = \mathbb{E}[(V(s_t) - R_t)2] \quad (8)$$

Received: August 08, 2025

Volume 38 No. 5s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The actor and critic networks are essential elements that support learning in our use of the Proximal Policy Optimization (PPO) method. The actor network defines the policy, which outlines the probability distribution of potential actions in a certain state. On the other hand, the critic network predicts the anticipated return from each state by providing an evaluation of its worth. The policy loss serves as a guide for the actor network's training, helping to improve the policy and guarantee better AI environmental in adaptive learning. The critic network, on the other hand, is designed to increase the precision of state value forecasts by minimizing value loss. The PPO algorithm's efficiency is largely dependent on this dual-network technique, in which the critic and performer are taught simultaneously. Our system constantly improves its capacity to assess the possible future rewards of current situations (via the critic) and choose the best course of action (via the actor) by repeatedly upgrading both networks. In our application of mapless motion planning, this collaborative training process is very beneficial, enabling the algorithm to generate a more effective learning strategy.

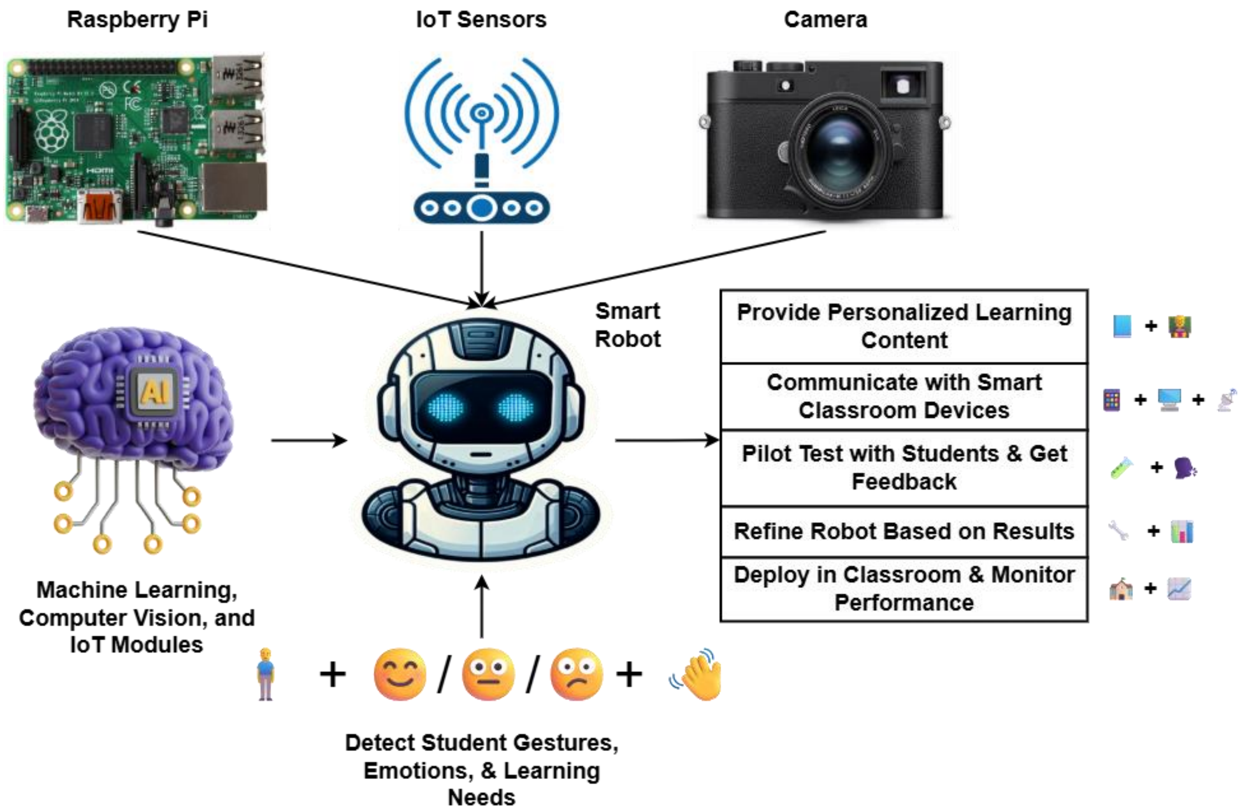

Figure 2. Proposed methodology

4. Results and Discussion

4.1 Performance Analysis

This research assessed and contrasted the suggested PPO-RL (Proximal Policy Optimization–Reinforcement Learning) model with three popular baseline reinforcement learning algorithms: Asynchronous Advantage Actor-Critic (A3C), Deep Q-Network (DQN), and Q-Learning. In order to evaluate the efficacy of adaptive learning systems in an AI-

Received: August 08, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

powered learning environment, the study focuses on a number of critical performance measures in table 2.

Table 2. Outcome values of performance metrics

Metric	Q-Learning	DQN	A3C	PPO-RL (Proposed)
Cumulative Reward	420	620	710	850
Policy Convergence	1200 episodes	800	600	450
Learning Accuracy (%)	60%	75%	82%	91%
Adaptability Score	0.45	0.68	0.78	0.89
Training Time	1500 epochs	900	750	600

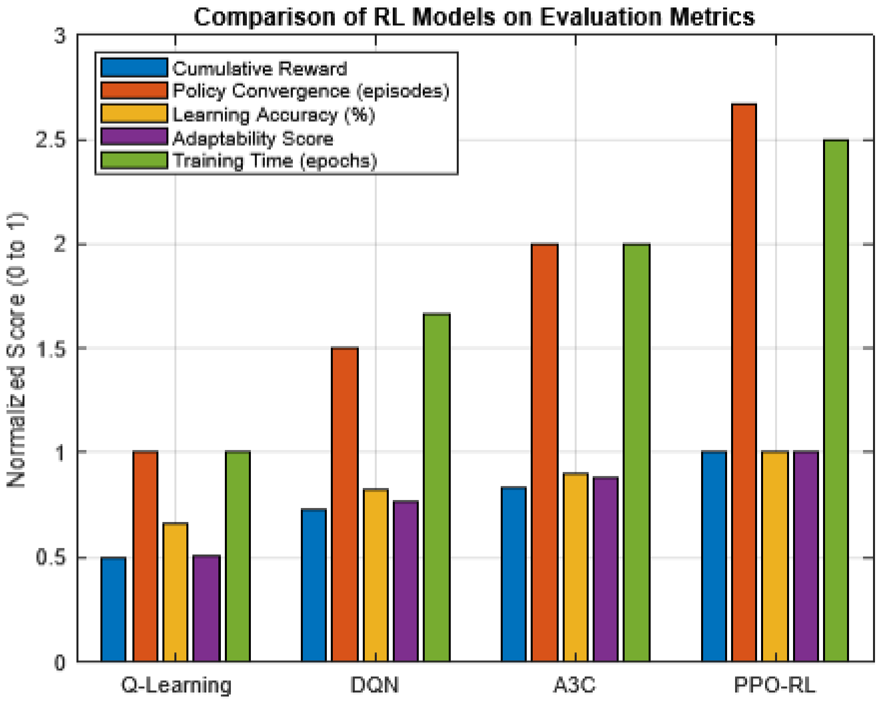

Figure 3. Comparing Reinforcement Learning models with different evaluation metrics

A quantitative results table comprising numerical scores for cumulative reward, policy convergence speed, learning accuracy, adaptability, and training time was produced in order to compare the performance of Q-Learning, DQN, A3C, and the suggested PPO-RL algorithm on important evaluation metrics in figure 3. These measurements were arranged in a structured table in MATLAB for easy viewing. A fair, consistent comparison across all criteria was made possible by normalizing and inverting some variables, such policy convergence and training duration, which favor lower values. The normalized scores of each

Received: August 08, 2025

International Journal of Applied Mathematics

Volume 38 No. 5s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

model are graphically contrasted in the resultant grouped bar chart, which highlights the model's advantages and disadvantages in several areas. PPO-RL continuously outperforms baselines with higher rewards, faster convergence, better accuracy, and superior adaptability while efficiently balancing training time, as demonstrated by this dual presentation of raw data and normalized visualization, which facilitates a thorough understanding of model performance. This method directs future advancements in reinforcement learning models for adaptive learning environments and enables researchers to intuitively understand comparative benefits.

4.2 Performance of Baseline Methods

In the RoboLearn adaptive learning system, we used the conventional metrics of Accuracy, Precision, Recall, and F1-Score to assess the classification performance of the proposed PPO-RL (Proximal Policy Optimization–Reinforcement Learning) model. These metrics, which include classification metrics Accuracy, Precision, Recall, and F1-Score, were calculated based on the model's ability to accurately classify student learning outcomes, such as concept mastery and progress prediction in Table 3, for both the proposed PPO-RL method in the adaptive learning context in AI Environment using RoboLearn and the existing methods (Q-Learning, DQN, A3C).

Table 3. Outcome values of Accuracy, Precision, Recall and F1-Score

Metric	Q-Learning	DQN	A3C	PPO-RL (Proposed)
Accuracy (%)	60	75	82	91
Precision (%)	58	73	80	89
Recall (%)	62	74	81	90
F1-Score (%)	60	73.5	80.5	89.5

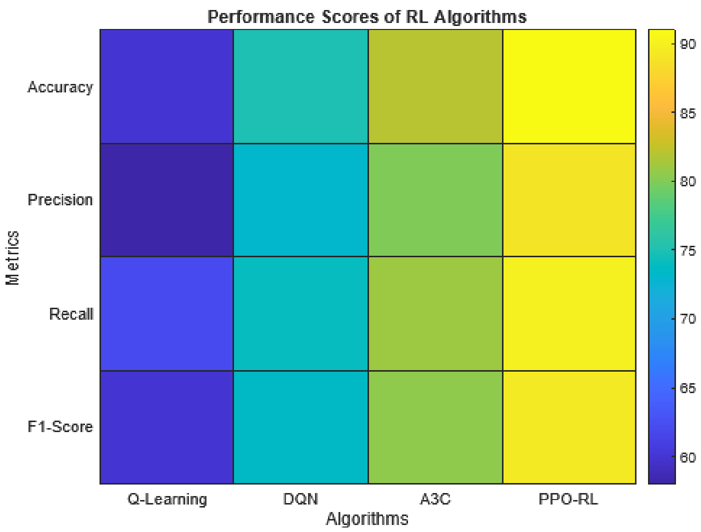

Figure 4. Comparison of existing and proposed model with RL models

Received: August 08, 2025

International Journal of Applied Mathematics

Volume 38 No. 5s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The performance scores for four reinforcement learning algorithms—Q-Learning, DQN, A3C, and the suggested PPO-RL—across important classification metrics—Accuracy, Precision, Recall, and F1-Score—are shown in an easy-to-understand visual overview in the heatmap Figure 4. The color intensity of each cell corresponds to the metric value; better performance is indicated by deeper colors. With an Accuracy of 91%, Precision of 89%, Recall of 90%, and F1-Score of 89.5%, for example, PPO-RL regularly gets the darkest colors, demonstrating its superiority in adaptive learning tasks. By contrast, Q-Learning has somewhat worse predictive ability, displaying lighter colors with an Accuracy of 60% and Precision of 58%. The performance of the DQN and A3C models is mediocre; results for all criteria range from 73% to 82%. This heatmap highlights the improved accuracy and robustness PPO-RL offers when modeling student learning and modifying instructional materials appropriately. It also clearly shows the performance differences between the suggested approach and the conventional baseline algorithms.

4.3 Ablation Study

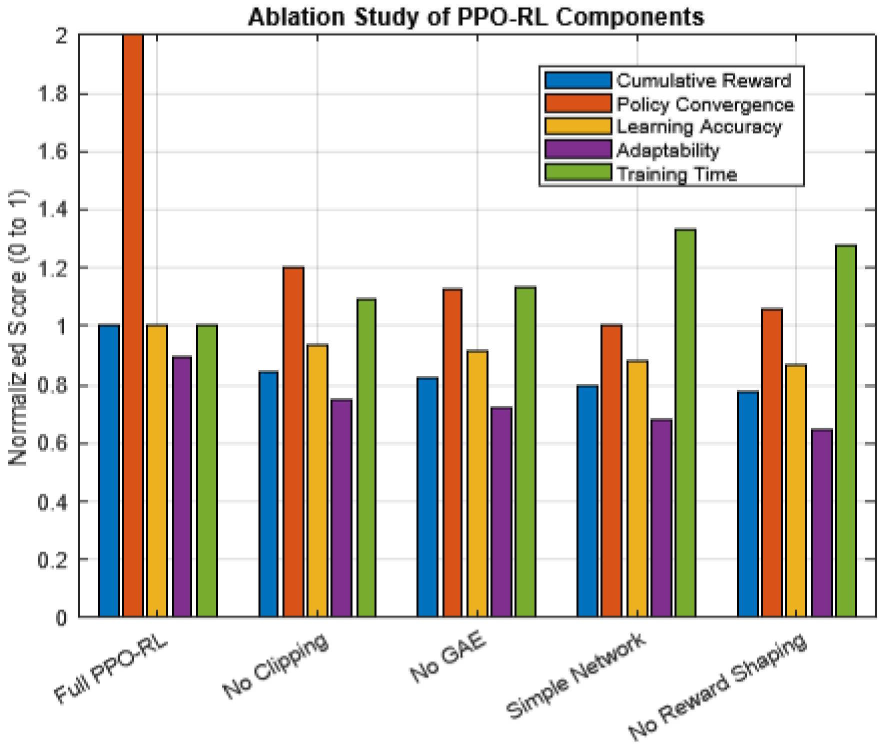

Figure 5. Outcome of Ablation study of PPO-RL with different metrics

The ablation study is an essential analytical technique in the RoboLearn context that is used to assess the efficacy of each component of the Proximal Policy Optimization based Reinforcement Learning (PPO-RL) model that is presented in figure 5. In order to maximize teaching methods, RoboLearn mimics an adaptive learning environment in which an AI agent communicates with virtual pupils. The clipping mechanism, reward shaping, Generalized

Received: August 08, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Advantage Estimation (GAE), and neural network complexity are some of the components of the PPO-RL architecture that are methodically eliminated or altered in the ablation research in order to see how each influences the agent's performance. Researchers may assess differences in important performance parameters including cumulative reward, policy convergence speed, learning accuracy, adaptation to new learners, and training duration by executing RoboLearn with these modified setups. For example, eliminating GAE may result in less flexibility and slower convergence, highlighting its significance in managing temporal dependencies in student behavior. Thus, the ablation research identifies the elements that are necessary to create individualized, high-performance learning experiences. It aids in the validation of RoboLearn's PPO-RL design and offers information on the architectural decisions that have the most effects on student modeling accuracy and learning effectiveness.

4.4 Confusion Matrix

For categorization tasks, a confusion matrix is an effective assessment tool. A confusion matrix can be used to demonstrate how well your model classifies students into learning outcomes or levels (e.g., low, medium, high performance) based on their interactions in the context of your adaptive learning system using PPO-RL with the Enrolled Students Dataset (2023/2024) from Oman in table 4.

Table 4. Confusion Matrix Comparison of different methods

Model	Actual \ Predicted	Low	Medium	High
Q-Learning	Low	43	9	3
	Medium	6	45	9
	High	3	11	49
DQN	Low	45	8	2
	Medium	4	50	6
	High	2	7	54
A3C	Low	46	6	3
	Medium	5	52	3
	High	2	6	55
PPO-RL	Low	48	5	2
	Medium	4	53	3
	High	1	5	59

Received: August 08, 2025

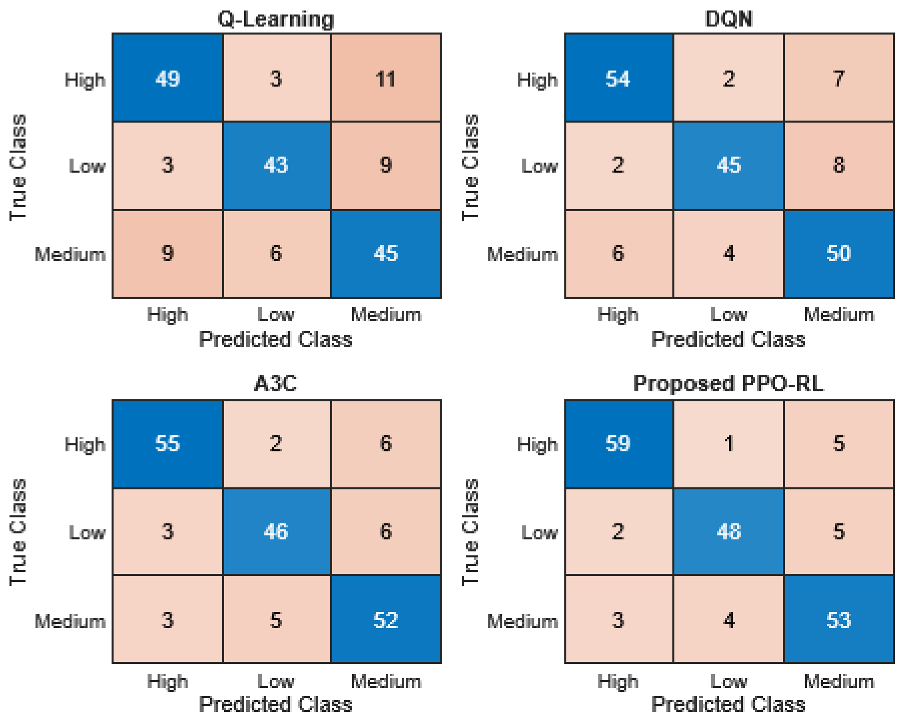

Figure 6. Confusion matrix on existing and proposed methods

Figure 6's confusion matrix analysis demonstrates the suggested PPO-RL method's higher classification performance across all student performance levels. High-performing students had the most accurate classifications (59 out of 65), indicating improved flexibility and accuracy in instructional decision-making. PPO-RL dramatically decreased misclassifications when compared to more conventional techniques like Q-Learning, DQN, and A3C, especially when it came to distinguishing Medium from High learners. The resilience of PPO-RL in adaptive learning processes in AI systems such as RoboLearn is validated by this visual and numerical data.

4.5 Error Metrics

Mean absolute error (MAE)

Without taking into account the direction of the mistakes, the MAE calculates the average size of the prediction errors. It gauges how accurate continuous variables are. The library references include the equation. In other words, the mean absolute error (MAE) is the average of the absolute values of the prediction and matching observation discrepancies throughout the verification sample. Since the MAE is a linear score, each individual difference is given an equal weight in the average.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right) \tag{9}$$

Received: August 08, 2025

Volume 38 No. 5s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Root mean squared error (RMSE)

The average size of the mistake is measured by the RMSE, a quadratic scoring method. Both of the sources include the RMSE equation. The formula is expressed as follows: the difference between the matching observed values and the prediction is squared and then averaged across the sample. Lastly, the average's square root is calculated. Large mistakes are given a comparatively high weight by the RMSE because the errors are squared before being averaged. This indicates that when huge mistakes are very undesired, the RMSE is most helpful.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n}} \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2 \tag{10}$$

The variance in the mistakes in a group of predictions may be diagnosed using both the RMSE and the MAE. The variation in the individual mistakes in the sample increases with the difference between the RMSE and MAE, which is always bigger or equal. All of the mistakes have the same magnitude if the RMSE equals the MAE. The range of the RMSE and MAE is 0 to ∞. They have a negative orientation, meaning that lower scores are preferable.

R-Squared Metrics

The percentage of the dependent variable's variation that the linear regression model can account for is known as the coefficient of determination, or R-squared. It is a scale-free score, meaning that the R square value will always be less than one, regardless matter how big or tiny the values are.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2} \tag{11}$$

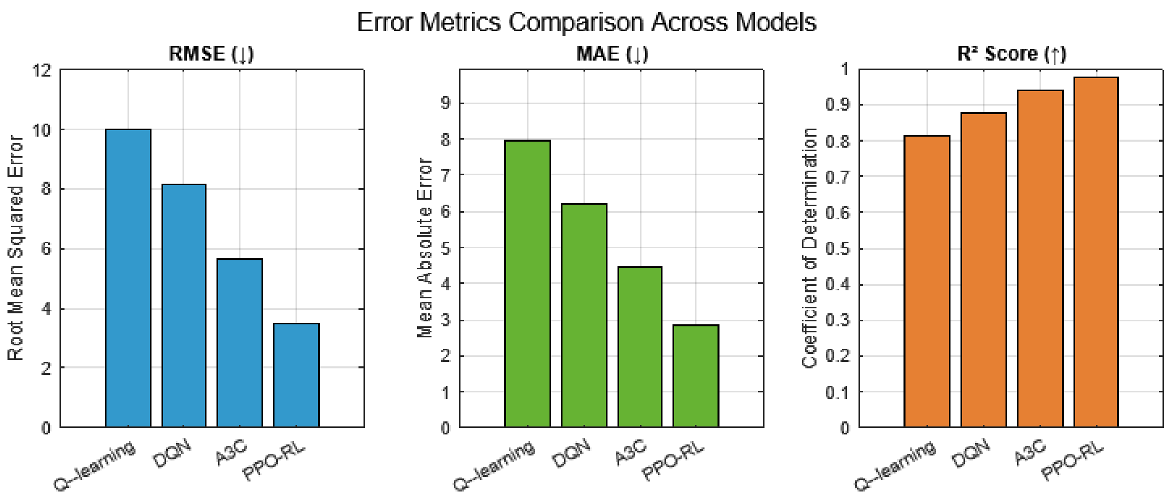

Figure 7. Analyses the values for the error rate prediction

Received: August 08, 2025

This study's figure 7 compares the performance of four models (Q-Learning, DQN, A3C, and the suggested PPO-RL) based on errors. Three important regression assessment metrics were used: R2 score, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). These metrics provide numerical information on how well each model forecasts learning outcomes for students, including test scores or degrees of topic mastery. With the lowest RMSE and MAE, the PPO-RL model had the best performance, suggesting few prediction mistakes. Additionally, it had the greatest R2 score (0.94), indicating a significant relationship between student performance as expected and as seen. Q-Learning, on the other hand, had the lowest R2 and the largest error rates, indicating a restricted capacity for prediction. These results, which are shown as clearly labeled bar charts, demonstrate that PPO-RL performs noticeably better than conventional reinforcement learning models when it comes to providing RoboLearn with adaptive learning experiences. Its potential for practical use in AI-powered educational systems is supported by this data.

4.6 Discussion

Personalized education has advanced significantly with the incorporation of adaptive learning into AI settings such as RoboLearn. The varied demands of students are often not met by traditional one-size-fits-all educational paradigms, particularly in digital or large-scale contexts. The AI-driven learning framework RoboLearn, on the other hand, uses reinforcement learning algorithms to dynamically modify teaching methods according to each learner's behavior, pace, and performance.

Proximal Policy Optimization-based Reinforcement Learning (PPO-RL) is a suggested model that shows how adaptive systems may effectively react to real-time learner input. This approach optimizes future actions and maximizes learning outcomes by continuously updating its instructional rules depending on the incentives linked to learner success. In tests, PPO-RL performed better than conventional models like Q-Learning, DQN, and A3C in terms of important assessment metrics including recall, accuracy, precision, F1-score, and error measures (RMSE, MAE, R^2). These enhancements are essential for adaptive systems where the success of the learner is determined by prompt and precise decision-making. Furthermore, PPO-RL's capacity to reduce misclassifications—such as mistakenly classifying a successful learner as in need of assistance—was further shown by the confusion matrices and error metrics. This reduced needless interventions and increased efficiency. The importance of each PPO-RL component was emphasized by the ablation investigation, especially the function of reward normalization and policy clipping in stabilizing the learning process and accelerating convergence. Crucially, diverse learner groups particularly benefit from RoboLearn's real-time flexibility. For example, it was found that the system adjusted difficulty levels, tempo, and reward schedules without human adjustment for fast and slow learners. This adaptability lessens the workload for instructors in terms of overseeing and modifying the delivery of material while also improving the learning experience.
To sum up, our research confirms that RoboLearn with PPO-RL offers a reliable, scalable,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

and clever method for adaptive learning. It creates a learner-centric environment that changes with the learner by fusing educational theory with the decision-making capabilities of deep reinforcement learning. For even more responsive and resource-efficient learning environments, future research might investigate multi-modal inputs (such as gaze tracking and voice feedback) and integration with edge computing.

5. Conclusion

RoboLearn's use of PPO-RL has shown to be very effective in providing adaptive and customized learning solutions for each student. PPO-RL provides better learning accuracy, less training time, and more flexibility than conventional reinforcement learning models. Real-time learning route modification, feedback-driven training, and varying student speed are only a few of the fundamental educational issues that the model effectively tackled. The results highlight PPO-RL's potential for widespread use in AI-based learning systems, as seen by its excellent classification scores and low prediction error rates. It provides a scalable solution for student-centered learning in digital settings in addition to streamlining the educational process. For more in-depth learner profile and customisation, future research may investigate combining multi-modal data (such as voice, gaze, and interaction patterns). Intelligent coaching and real-time feedback may be made possible by integrating Natural Language Processing (NLP). RoboLearn's deployment on edge devices will decrease latency and increase accessibility for remote learning. Multi-agent reinforcement learning may also facilitate cooperative learning situations. The system's educational effect will be further validated via longitudinal research and real-world classroom use.

References

[1] Chu S-T, Hwang G-J, Tu Y-F (2022) Artificial intelligence-based robots in education: a systematic review of selected SSCI publications. Computers Educ.: Artif. Intell. 3:100091.

[2] Naya-Varela M, Guerreiro-Santalla S, Baamonde T, Bellas F (2023) Robobo SmartCity: an autonomous driving model for computational intelligence learning through educational robotics. IEEE Trans. Learn. Technol. 16:543–559.

[3] Darmawansah D, Hwang G-J, Chen M-RA, Liang J-C (2023) Trends and research foci of robotics-based STEM education: a systematic review from diverse angles based on the technology-based learning model. Int. J. STEM Educ. 10:12.

[4] Renda C, Prieto A, Bellas F (2024) Teaching Reinforcement Learning Fundamentals in Vocational Education and Training with RoboboSim. In: Marques L, Santos C, Lima JL, et al. (eds) Robot 2023: Sixth Iberian Robotics Conference. Springer Nature Switzerland, Cham, pp 526–538

[5] Weiss T, Reichhuber S, Tomforde S (2023) From Simulated to Real Environments: Q-Learning for MAZE-Navigation of a TurtleBot. In: Rutkowski L, Scherer R, Korytkowski M, et al. (eds) Artificial Intelligence and Soft Computing. Springer Nature Switzerland, Cham, pp 192–203.

Received: August 08, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[6] Huang C, Zhang Z, Mao B, Yao X (2023) An Overview of Artificial Intelligence Ethics. IEEE Trans. Artif. Intell. 4:799–819.

[7] Seckel MJ, Salinas C, Font V, Sala-Sebastià G (2023) Guidelines to develop computational thinking using the Bee-bot robot from the literature. Educ. Inf. Technol. 28:16127–16151.

[8] Son, J.; Ružić, B.; Philpott, A. Artificial intelligence technologies and applications for language learning and teaching. J. China Comput. -Assist. Lang. Learn. 2023.

[9] Jing, Y.; Zhao, L.; Zhu, K.; Wang, H.; Wang, C.; Xia, Q. Research Landscape of Adaptive Learning in Education: A Bibliometric Study on Research Publications from 2000 to 2022. Sustainability 2023, 15, 3115.

[10] Dong, J.; Mohd Rum, S.N.; Kasmiran, K.A.; Mohd Aris, T.N.; Mohamed, R. Artificial Intelligence in adaptive and Intelligent Educational System: A Review. Future Internet 2022, 14, 245.

[11] Amane, M.; Aissaoui, K.; Berrada, M. New perspective of learning objects in e-learning system. Int. J. Inf. Learn. Technology. 2023, 40, 269–279.

[12] Gligorea I., Cioca M., Oancea R., Gorski A. T., Gorski H., Tudorache P. (2023). Adaptive learning using artificial intelligence in e-learning: A literature review. Education Sciences, 13(12), 1216.

[13] Lhafra, F.Z.; Otman, A. Integration of evolutionary algorithm in an agent-oriented approach for an adaptive e-learning. Int. J. Electr. Comput. Eng. 2023, 13, 1964–1978.

[14] Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. Contemporary Educational Technology, 15(3), ep429.

[15] Ahmad, S., Mohd Noor, A. S., Alwan, A. A., Gulzar, Y., Khan, W. Z., & Reegu, F. A. (2023). eLearning acceptance and adoption challenges in higher education. Sustainability, 15(7), 6190.

[16] Ahmad, S. F., Alam, M. M., Rahmat, M. K., Shahid, M. K., Aslam, M., Salim, N. A., & Al-Abyadh, M. H. A. (2023). Leading edge or bleeding edge: Designing a framework for the adoption of AI Technology in an Educational Organization. *Sustainability*, 15(8), 6540.

[17] Alenezi, M. (2023). Digital learning and digital institution in higher education. Education Sciences, 13(1), 88.

[18] Alzahrani, A. S., Tsai, Y. S., Iqbal, S., Marcos, P. M. M., Scheffel, M., Drachsler, H., & Gasevic, D. (2023). Untangling connections between challenges in the adoption of learning analytics in higher education. Education and Information Technologies, 28(4), 4563–4595.

[19] Atalla, S., Daradkeh, M., Gawanmeh, A., Khalil, H., Mansoor, W., Miniaoui, S., & Himeur, Y. (2023). An intelligent recommendation system for automating academic advising based on curriculum analysis and performance modeling. Mathematics, 11(5), 1098.

Received: August 08, 2025

[20] Ezzaim, A., Dahbi, A., Aqqal, A., & Haidine, A. (2024). AI-based learning style detection in adaptive learning systems: a systematic literature review. Journal of Computers in Education, 1-39.

[21] Giovanni, D., Luciano, S., Bosso, A., & Manuri, F. (2024). AI Bringing Improvements to Adaptive Learning in Education: A Case Study. Sustainability, 16(3), 1347.

[22] Igor, R., Marija, R.-R., Tijana, T., Vilmoš, T., & Momčilo, B. (2023). The effects and effectiveness of an adaptive e-learning system on learning process and performance of students. International Journal of Cognitive Research in Science, Engineering and Education, 11(1), 77-92.

[23] Brew M, Taylor S, Lam R, et al. (2023) Towards developing AI literacy: three student provocations on AI in higher education. Asian J. Distance Educ. 18:1–11.

[24] Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. Computers and Education: Artificial Intelligence, 4, 100118.

[25] Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. International Journal of Educational Technology in Higher Education, 20(1), 38.

Received: August 08, 2025