Proceedings of the 19th Conference on Computer Science and Intelligence Systems (FedCSIS) pp. 549–553 DOI: 10.15439/2024F302 ISSN 2300-5963 ACSIS, Vol. 39

Reinforcement Learning based Intelligent System for Personalized Exam Schedule

Marco Barone
Università Giustino Fortunato
Università degli Studi di Foggia
Email: marco.barone@unifg.it

Matteo Ciaschi
National Research Council (CNR)
email: matteo.ciaschi@cnr.it
ORCID: 0009-0009-5119-3563

Zaib Ullah
Università Giustino Fortunato
email: z.ullah@unifortunato.eu

Armando Piccardi
Università Giustino Fortunato
email: a.piccardi@unifortunato.eu

***Abstract*—Personalized learning has been proving to be useful concept in the learning of a student. Artificial Intelligence (AI) which has revolutionized many aspects of our lives has also been glowingly used in the education sector. One of the fascinating AI technique, the Reinforcement Learning (RL) is considered as the perfect tool to develop personalized solution in the education. RL algorithms have the ability to take into account personal characteristics of each student. This work presents the development of personalized exam scheduler using RL. The intelligent examination scheduler consider several parameters for training such as age, academic year, past education performance, discipline, number of courses, and gap between two exams. The trained RL agent then able to provide examination schedule to a student depending on a student personal record, interests and abilities. The preliminary results are encouraging and more research would bring useful contribution of AI in various aspects of learning process of a student.**

I. Introduction

Innovative educational approaches that involve the adapting of the educational process to the individual needs, interests and skills of each student can be defined as Personalized Learning [1]. The growing use of technologies like Artificial Intelligence (AI) and Machine Learning (ML), personalized learning allows to create customized learning experiences with the aim of increasing motivation, creating greater involvement, and improving the final result [2].

Widely recognized as the most disruptive innovation of recent years, AI is expanding its diffusion into an ever-increasing number of sectors: AI algorithms in healthcare are able to diagnose diseases through image analysis with high levels of accuracy and to create personalized treatment plans [3], [4]; in finance and smart offices, they allow to analyze large volumes of financial data to uncover trends and related investment opportunities [5]; self-driving cars and smart cities aspects are transforming mobility aiming to ensure better efficiency, sustainability and safety in the transport sector [6]; multi antenna communication including 5G and beyond networks [7], wireless sensor networks [8], dynamic treatment regimes [9], pervasive computing [10] and other significant progress many diverse areas including risk management in nuclear medicine department [11].

In the education sector, the concept of personalization goes beyond the simple adaptation of teaching materials to the needs of individual students, it includes an approach that recognizes and accommodates each student's unique learning style, interests and strengths. By personalizing the learning experience, teachers are able to create a more engaging and effective learning environment [12].

One aspect of personalization in education is customizing teaching to students preferred learning modes: some students may excel at visual learning, while others prefer auditory learning or other ways; by broadening the scope to additional teaching methods and resources, educators can ensure that all students have the opportunity to learn in the ways that best suit them. This personalized approach not only improves students' understanding, but also promotes a sense of autonomy in their educational journey [13].

Moreover, personalization in education can address students' individual academic needs and goals: for example, students can have different levels of proficiency in different subjects and thanks to personalized learning plans it is possible to identify specific areas where additional support is required. By thus aligning instruction with students' learning goals, educators make it easier for students to achieve academic goals at their own pace and avoid failure [14].

Another fundamental element, in addition to personalization, in maximizing the student learning process is represented by planning. Effective planning involves optimizing the allocation of time and resources to ensure that students have sufficient opportunities to learn and grow. This includes balancing teaching time with independent study, allocating time for collaborative activities and projects, and integrating breaks and reflection periods into the learning schedule.

In this context, the use of reinforcement learning (RL) algorithms in academic data favors innovative approaches to:

personalizing and optimizing the educational experience for students, adapting materials and content based on individual student progress and preferences;

optimization of allocation of educational resources, such as time, personnel, and materials. adaptive assessment systems

that can dynamically adjust the difficulty and format of assessments based on student performance;
intelligent tutoring systems, developing systems that provide personalized guidance and feedback to students and dynamically adjust strategies and activities.
In summary, the use of AI algorithms favors innovative approaches to personalize and optimize the students' educational experience; our contribution concerns the context of tutor systems to support students and in particular the implementation of a Reinforcement Learning based Intelligent System for personalized exam schedule with the aim of reducing exam fail.

II. Technical Background

The AI, Artificial Intelligence, can be defined as the ability to develop intelligence on programmable machines with the aim of imitating the human brain [15]. Machine Learning (ML) is a subfield of AI, which concerns the question of how to implement software agents that automatically improve with experience.
Machine Learning is divided into three categories such as supervised learning, unsupervised learning and reinforcement learning [16].
Supervised learning is a type of learning based on the training dataset, a set of labeled input data, data for which the correct output is known. Training dataset is provided by a domain expert, who takes on the role of external supervisor of the process.The main goal of supervised learning is to build models enable of generalizing the relationship between inputs and the corresponding outputs in order to make the most accurate predictions on data not yet seen, based on its previous training.
The second category, Unsupervised Learning, is based on an appropriate study of the dataset with the aim of extracting knowledge and hidden patterns, there is no supervisor.
In Reinforcement Learning (RL), the entire learning process is linked to a specific objective. An agent interacts with an unknown environment, according to the try and error scheme; just the same way children learn it makes actions and observes what it happens [17]. Following the actions taken, the agent will receive feedback from the environment respectively in terms of reward for positive actions and penalties for negative actions. Thanks to this feedback it trains and acquires experience and knowledge about the environment.
RL involves problems of finding the optimal action to take in various scenarios to maximize the cumulative reward. The RL agent must develop a strategy (a comprehensive correlation between scenarios and actions) by experimenting with actions independently, without guidance from domain experts, like many other machine learning approaches.
Another crucial aspect of RL problems is the constant trade-off between exploiting the agent's existing knowledge of the environment (repeating actions previously taken in a given scenario) and exploring new actions that haven't been attempted in that scenario before.
A first major distinction that can be made between the RL algorithms to use is that between single-agent and multi-agent, each with distinct applications and benefits. Single-Agent RL refers to scenarios in which a single agent interacts with the educational environment to optimize an individual student's learning path. In this context, the agent continuously analyzes the student's responses and performance to adapt the teaching content and improve the effectiveness of teaching, for example an intelligent tutor using Q-Learning or Deep Q-Networks (DQN) to personalize questions based on the student's answers. The goal is to maximize student progress by dynamically adjusting the difficulty of problems so that the student remains challenged but not overwhelmed. The benefit of this approach is the ability to create highly personalized learning paths, optimizing interaction with each student to improve their understanding and retention of the material.
Multi-Agent RL, on the other hand, involves multiple agents interacting with each other within the educational environment [18]. These agents can represent different students, tutors, or even different components of a complex educational system. Agents can collaborate or compete to achieve common or individual educational goals, for example an educational platform that uses multiple agents to simulate a collaborative learning environment, where students work together to solve complex problems.
Algorithms such as Multi-Agent Deep Deterministic Policy Gradient (MADDPG) can be used to coordinate the actions of agents so that each contributes effectively to the collective task. This approach facilitates collaborative and competitive learning, allowing students to benefit from interaction with their peers. Furthermore, it can improve virtual classroom dynamics by providing personalized support in a group learning context.
While single-agent RL focuses on optimizing the individual learning path, multi-agent RL supports both personalization by fostering group dynamics and promoting collaborative learning. However, multi-agent systems tend to be more complex to implement and manage than single-agent systems, as they require coordination and management of interactions between multiple agents. Finally, single-agent RL is ideal for personalized tutors and one-on-one learning assistants, while multi-agent RL is better suited for interactive and collaborative learning environments, such as classroom simulations and game-based learning platforms.
Q-Learning: This algorithm can be used to develop educational tutors who learn which actions (such as posing a question or reviewing a topic) lead to the best learning outcomes for students. Through a process of trial and error, the tutor updates a Q-table, which represents the value of each action in a given state, allowing the system to select the optimal action at any time.
Deep Q-Networks (DQN): DQN combines Q-Learning with deep neural networks to handle complex, continuous state space environments, such as those found in advanced e-learning platforms. This allows you to analyze a large amount of student data to make precise predictions on which educational content to propose next, optimizing learning in an

adaptive and personalized way.
Proximal Policy Optimization (PPO): This RL algorithm is particularly useful for managing stochastic policies in dynamic environments. In education, PPO can be used to create systems that not only decide which exercises or materials to propose, but also how to vary the difficulty and type of feedback based on the user's reaction in real time. PPO allows you to manage variations in teaching strategy in a more stable and efficient way, quickly adapting to the changing needs of students

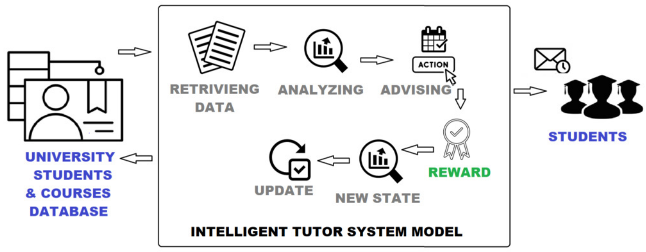

Fig. 1. System Model

III. Related Works

In this section, we will analyze the scientific literature regarding the use of AI and the development of adaptive systems in e-learning environments and the reinforcement learning approaches used to identify and model learning in education.
The authors of [1] [14] proposed the use of applications of the RL method in schools to improve learning activities and provide comfortable learning. The first system is personalized and adaptive e-learning platform based on Deep Q- Network RL and an online rule-based decision making implementation for elementary school. The second system is based on an intelligent educational environment for higher education in which the algorithm analyzes the change in student behavior, adapting teaching materials to improve the overall learning efficiency. In both, the intelligent learning created combines with the advantages of e-learning such as interactions, flexibility and experience.
In another work [12], authors presented a comparison between two models for customizing sequences of learning resources in a massive online open course (MOOC). The first model suggests sequences of learning resources that have been successful in the past, using case based reasoning and the Euclidean distance; while the second model recommends optimal learning resource sequences using Reinforcement's Q-Learning algorithm. A key role in the design of teaching resources is represented by the student's level of knowledge with respect to the level of complexity of the teaching resource with the aim of optimizing the MOOC learning process.
In the work of [13], the authors designed an intelligent adaptive e-learning system, based on RL that places particular importance on the adequacy of the student's real profile and its update compared to the one used in the learning path recommendation. The objectives of this system are the creation of a real profile of a given student, through the implementation of K-means and linear regression, and the recommendation of adaptive learning paths according to this profile, implementing the Q-learning algorithm.
Similarly, intelligent ambient assisted living systems are developed in [19] and [20]. The personalized systems are based on AI-powered tools like RL but are basically designed to provide personalized medication assistance and recommendation to the patients with cognitive impairments instead of students.
The next two quotes concern the use of gamification in the learning sector, in particular, the authors of the article [21] have demonstrated that gamification can be used to encourage student activities, increase engagement and evaluate their success; while the study [22] presents SEP-CyLE (Software Engineering and Programming Cyberlearning Environment), an online gamified tool designed to provide additional IT content to students.
In the article [23], to promote student involvement, satisfaction and performance, an integrated approach is developed that combines artificial intelligence with the result of the analysis of learning feedback through experimental research conducted to study the effects of learning.
The contribution of the authors of the article [24] is the development of an adaptive e-learning system capable of generating learning paths adapted to the profile of the individual learner. An approach is then proposed to dynamically compose adaptive online learning courses based on student activities, learning objectives and instructional design strategies using Q-learning. The algorithm gains knowledge by analyzing student behavior and provides the course content needed to achieve learning objectives based on positive and/or negative student feedback.
The work in [25] delves into the concept of an AI tutor that provides personalized learning paths and round-the-clock support to students. The AI tutor uses sophisticated algorithms and ML techniques to analyze a student's strengths, weaknesses, and learning style. By collecting data from various sources, such as assessments, quizzes and user interactions, the AI tutor ensures that students receive content and exercises proportionate to their individual progress promoting greater learning effectiveness.

IV. System Model

This section presents the proposed work by introducing each component of the system and their corresponding functionality as shown in Figure 1.
The first step is to retrieve data such that the system accesses the university database to retrieve the necessary information for each student, which includes: personal data like age, working status, qualifications, skills and preferences, and academic data like courses attended, exams taken, grades obtained, credits acquired.
We model this scenario as Markov Decision Process (MDP)

problem and use RL to solve this MDP. We describe the components of modelled MPD as:
the **Environment** in the porposed framework is the entire university ecosystem, students and provided courses;
the **Agent** is our personalized exam scheduler tutor such that the agent has to learn the optimal exam schedule for a student;
the **State** represents the current status of the student with all information retrieved such that all the possible combinations and conditions are considered as states and RL agent has to choose an action (exam schedule) for a a student;
the **Actions** are the possible actions that the system recommend to the students such that all the combinations that are available to a student to make selection for his/her examination;
the **Reward** is determined by the success or failure of the recommended actions, for example, passing the exam with a good grade is a reward, exam failure but also not showing up to the exam is a penalty for the system;
the **Policy** is the optimal strategy that a RL agent has to learn for a given environment. In our case, the RL agent (exam scheduler) has to learn the optimal exam schedule for a each students according to each student existing credentials.
Our implementation produces a continuous cycle of learning and updating, based on student feedback, starting from the observation of the current status, in which the system collects the most recent information on the student's academic progress, the system analyze it with the current policy and recommends the next action (exam to book).
After the student has followed the recommendation, the system get feedback on the outcome (positive or negative outcome) and then make updates: the student's status based on the feedback received, and the policy update; the reinforcement learning model updates the policy based on feedback to improve future recommendations.
Moreover, we also consider privacy and security aspects and ensure that all student personal and academic data is treated with the utmost confidentiality and compliance with privacy regulations; and personalization. We consider that each student has different needs, so the system must be flexible to adapt to the peculiarities of each student.

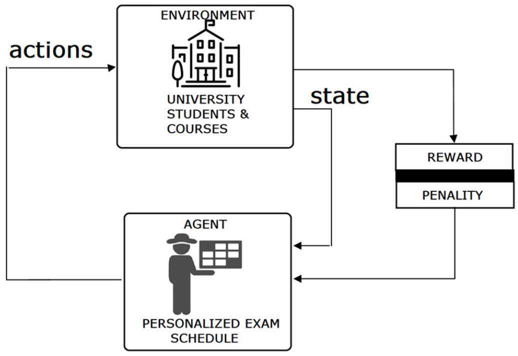

Fig. 2. Reinforcement Learning problem Model

V. Discussion

In this section, we discuss the effectiveness of our reinforcement learning framework. We started our experimental activity with a data subset comprising all the target university courses: two years Master's Degree in Applied Behavioral and Cognitive Psychology, Business Administration, Organization and Management of Services for Sport and Physical Activities; three years Bachelor's Degree in Transport Science and Technology, Psychological Science and Techniques, Educational Science, Business Law and Economics, Computer Engineering; five years Master's Degree in Law.
The objective of the experimental phase is to demonstrate the extent to which the intelligent tutor improves each student's learning path, facilitating final success and reducing the risk of failure.
By implementing a suitable RL algorithm, we aim to establish an optimal learning policy that minimizes the total number of learning steps. The reward function of the algorithm is structured to increase as the total number of learning steps decreases, thereby incentivizing more efficient learning performance.
Our evaluation will compare the outcomes of students who utilized the recommendations produced by our intelligent tutor with those who did not and we will employ a specifically defined metric that considers the relationship between the time taken and the results obtained. This comparison will help us understand the impact of our intelligent tutor on students' academic performance, specifically in terms of their ability to achieve learning objectives more efficiently and with a higher success rate.
To further substantiate our findings, we plan to conduct a series of statistical analyses, including hypothesis testing and confidence interval estimation, to determine the significance of the observed differences. We will also perform a detailed breakdown of the learning steps and rewards accrued by students in both groups, providing deeper insights into how the intelligent tutor influences learning behaviors and outcomes.
In addition to quantitative metrics, we intend to collect qualitative feedback from students and instructors to gauge their satisfaction and perceived effectiveness of the intelligent tutor. This approach will ensure a comprehensive evaluation of our reinforcement learning framework, highlighting both its strengths and areas for improvement.
Ultimately, the goal is to demonstrate that our intelligent tutor can serve as a valuable tool in educational settings, enhancing students' learning experiences and outcomes by providing personalized and effective learning recommendations.

As a case study in figure 3, for example, a student of age 25, taking a single course, and the difference between course enrollment and exam date is 164 days. The student followed a schedule where he attained 23 marks out of 30 and the

Age	AGE_NORM	Grade	GRADE_NORM	Ac_year	Univ_start_date	Exam_date	Date_diff	DIFF_NORM
40	0,9305	27	0,4669	2022/2023	24/10/2022	03/04/2023	161	0,072
22	0,2666	27	0,4669	2022/2023	10/10/2022	03/07/2023	266	0,1193
33	0,9454	28	0,607	2022/2023	29/08/2022	03/07/2023	308	0,1382
45	0,8377	30	1	2022/2023	14/11/2022	03/07/2023	231	0,1036
25	0,4901	23	0,184	2018/2019	22/10/2018	04/04/2019	164	0,0734
40	0,9305	27	0,4669	2021/2022	09/11/2021	04/04/2022	146	0,0653
53	0,4586	27	0,4669	2021/2022	01/03/2022	04/04/2022	34	0,0149
54	0,4139	24	0,3399	2021/2022	02/08/2021	04/04/2022	245	0,1099

Fig. 3. Data

proposed system should produced a schedule based on which student can achieve higher marks than 23 marks in the said experiment.

VI. Conclusion

Personalized learning has been demonstrated to be a useful idea in the education of a student. AI which has changed many facets of our lives has also been highly used in academia. RL is one of the intriguing AI approaches considered as an excellent tool for designing personalized solutions in education. RL algorithms can effectively take into account the student's characteristics. In this article, we explored the potential of utilizing RL for personalized exam scheduling. Our proposed framework considered many parameters including students' age, domain, number of courses, academic history, and the intermission between two consecutive exams. The trained RL agent will efficiently provide exam schedules based on their respective data. The initial results are inspiring and more research on this field would attain more valuable contributions in various aspects of personalized learning systems.

Acknowledgment

This research has been partially funded by the Regione Campania with the "PSR Campania 2014-2022 programme, Misura 16, Tipologia d'intervento 16.1.2." Project: "EVOO-LIO - L'Evoluzione dell'Olio EVO Sannita tracciato con la Blockchain.

References

[1] W. Sayed, A. Noeman, A. Abdellatif, M. Abdelrazek, M. Badawy, A. Hamed, and S. El-Tantawy, "Ai-based adaptive personalized content presentation and exercises navigation for an effective and engaging e-learning platform," *Multimedia Tools and Applications*, vol. 82, pp. 1–31, 06 2022.

[2] M. Naeem, S. T. H. Rizvi, and A. Coronato, "A gentle introduction to reinforcement learning and its application in different fields," *IEEE access*, vol. 8, pp. 209 320–209 344, 2020.

[3] M. Naeem and A. Coronato, "An ai-empowered home-infrastructure to minimize medication errors," *Journal of Sensor and Actuator Networks*, vol. 11, no. 1, p. 13, 2022.

[4] M. Naeem, A. Coronato, and G. Paragliola, "Adaptive treatment assisting system for patients using machine learning," in *2019 sixth international conference on social networks analysis, management and security (SNAMS)*. IEEE, 2019, pp. 460–465.

[5] A. Coronato, G. De Pietro, and M. Esposito, "A semantic context service for smart offices," in *2006 International Conference on Hybrid Information Technology*, vol. 2. IEEE, 2006, pp. 391–399.

[6] Z. Ullah, M. Naeem, A. Coronato, P. Ribino, and G. De Pietro, "Blockchain applications in sustainable smart cities," *Sustainable Cities and Society*, p. 104697, 2023.

[7] M. Naeem, S. Bashir, Z. Ullah, and A. A. Syed, "A near optimal scheduling algorithm for efficient radio resource management in multi-user mimo systems," *Wireless Personal Communications*, vol. 106, no. 3, pp. 1411–1427, 2019.

[8] A. Testa, A. Coronato, M. Cinque, and J. C. Augusto, "Static verification of wireless sensor networks with formal methods," in *2012 Eighth International Conference on Signal Image Technology and Internet Based Systems*. IEEE, 2012, pp. 587–594.

[9] S. I. H. Shah, A. Coronato, M. Naeem, and G. De Pietro, "Learning and assessing optimal dynamic treatment regimes through cooperative imitation learning," *IEEE Access*, vol. 10, pp. 78 148–78 158, 2022.

[10] M. Bakhouya, R. Campbell, A. Coronato, G. d. Pietro, and A. Ranganathan, "Introduction to special section on formal methods in pervasive computing," pp. 1–9, 2012.

[11] S. I. H. Shah, M. Naeem, G. Paragliola, A. Coronato, and M. Pechenizkiy, "An ai-empowered infrastructure for risk prevention during medical examination," *Expert Systems with Applications*, vol. 225, p. 120048, 2023.

[12] A. Flores, L. Alfaro, J. Herrera Quispe, and E. Cardenas, "Proposal models for personalization of e-learning based on flow theory and artificial intelligence," *International Journal of Advanced Computer Science and Applications*, vol. 10, 01 2019.

[13] R. Mustapha, G. Soukaina, Q. Mohammed, and A. Es-Sâadia, "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*, vol. 14, no. 5, 2023. [Online]. Available: http://dx.doi.org/10.14569/IJACSA.2023.0140528

[14] S. Fu, "A reinforcement learning-based smart educational environment for higher education," *International Journal of e-Collaboration*, vol. 19, pp. 1–17, 01 2023.

[15] U. b. Khalid, M. Naeem, F. Stasolla, M. H. Syed, M. Abbas, and A. Coronato, "Impact of ai-powered solutions in rehabilitation process: Recent improvements and future trends," *International Journal of General Medicine*, pp. 943–969, 2024.

[16] M. Jamal, Z. Ullah, M. Naeem, M. Abbas, and A. Coronato, "A hybrid multi-agent reinforcement learning approach for spectrum sharing in vehicular networks," *Future Internet*, vol. 16, no. 5, p. 152, 2024.

[17] M. Fiorino, M. Naeem, M. Ciampi, and A. Coronato, "Defining a metric-driven approach for learning hazardous situations," *Technologies*, vol. 12, no. 7, p. 103, 2024.

[18] M. Naeem, A. Coronato, Z. Ullah, S. Bashir, and G. Paragliola, "Optimal user scheduling in multi antenna system using multi agent reinforcement learning," *Sensors*, vol. 22, no. 21, p. 8278, 2022.

[19] A. Coronato and G. Paragliola, "A structured approach for the designing of safe aal applications," *Expert Systems with Applications*, vol. 85, pp. 1–13, 2017.

[20] M. Ciampi, A. Coronato, M. Naeem, and S. Silvestri, "An intelligent environment for preventing medication errors in home treatment," *Expert Systems with Applications*, vol. 193, p. 116434, 2022.

[21] F. Portela, "Towards an engaging and gamified online learning environment—a real casestudy," *Information*, vol. 13, no. 2, 2022.

[22] L. Zahedi, J. Batten, M. Ross, G. Potvin, S. Damas, P. Clarke, and D. Davis, "Gamification in education: a mixed-methods study of gender on computer science students' academic performance and identity development," *Journal of Computing in Higher Education*, vol. 33, pp. 1–34, 08 2021.

[23] F. Ouyang, M. Wu, L. Zheng, L. Zhang, and P. Jiao, "Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course: Revista de universidad y sociedad del conocimiento," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, p. 4, Dec 2023 2023.

[24] C. W. Fernandes, T. Miari, S. Rafatirad, and H. Sayadi, "Unleashing the potential of reinforcement learning for enhanced personalized education," in *2023 IEEE Frontiers in Education Conference (FIE)*, 2023, pp. 1–5.

[25] A. Robert, "Your ai tutor: Personalized learning paths and 24/7 support," 04 2024.