Region - Educational Research and Reviews, 2025, Volume7, Issue5
DOI: 10.32629/rerr.v7i5.4082
ISSN Online: 2661-4634
ISSN Print: 2661-4626

Adaptive learning path planning based on reinforcement learning

Hui XU*

Wuhan Institute of Design and Sciences, Wuhan 430205, China
*Corresponding author
E-mail address: uniqueariel@163.com

Abstract: This article proposed the concept of "fit", focusing on the matching between learners and the learning environment, and digitally expressing the matching relationship from three levels: education, group, and technology. This study explores a reinforcement learning-based method for generating learning paths, aiming to realize the adaptive generation of learning paths. Experimental results showed this method feasible and effective, with the proposed adaptive BP neural network algorithm having the shortest path selection time (78 seconds) and higher optimization rate (38.75%) among the compared algorithms.

Key words: English digital education; adaptive learning path; reinforcement learning; fit relationship

1 Introduction

The rapid development of information technologies has positively impacted education, promoting changes in traditional teaching methods. This article explores adaptive learning path planning based on reinforcement learning and verifies the technology's feasibility through experiments.

2 Related works

Experts have long been committed to digital education for adaptive learning path planning. Benmesbah O proposed an adaptive genetic algorithm, reducing search space and improving efficiency [1]. Raj N S found that most studies focus on learning styles, preferences and knowledge levels rather than cognitive factors like social markers [2]. El-Sabagh H A showed that adaptive e-learning can attract student participation [3]. Shemshack A noted that learners' characteristics will determine personalized paths [4]. Wang S found individual adaptive learning improves math scores [5]. However, existing solutions lack personalized analysis and the full application of artificial intelligence, resulting in suboptimal outcomes.

3 Methods

3.1 Adaptive learning path

The learning path refers to a series of behaviors or resources set up to enable learners to acquire knowledge and skills in a specific field. Adaptive learning path refers to an ordered, personalized learning path that is based on individual characteristics such as knowledge level, learning methods, and resource preferences, aiming to meet individual learning needs and objectives.

Copyright © 2025 by author(s) and Frontier Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/

There are traditional path planning algorithms and methods that attempt to apply deep learning and reinforcement learning to solve path planning problems. Path planning algorithms can be classified based on the concepts of global, local, or a combination of global static and local dynamic. In addition to solving path planning problems, adaptive learning path planning usually includes the subproblem of environment mapping. The mapping methods used for this problem include visual mapping, Voronoi mapping, raster mapping, SLAM (simultaneous localization and mapping), etc.

Evolutionary educational psychology believes that human development is not only influenced by the social and cultural environment in which it operates, but also by the biological and genetic foundations that emphasize individual differences among learners. These individual differences are reflected in both intellectual and non intellectual factors, such as emotions, attitudes, motivation, attention, cognitive ability, and learning ability. These individual differences among students also lead to different teaching and learning methods, even if they are of the same age and have different levels of ability. Constructivist theory holds that students' knowledge is gradually formed through interaction and contact with the external environment. This constructivist process is based on the knowledge, experience, and psychological characteristics that students already possess. Due to the different experiences and psychological characteristics of different students, constructivism advocates individualized teaching and student-centered approach.

3.2 Implementation of adaptive learning

Adaptive learning implementation has three steps.

1) Personalized resource recommendation: System builds user profiles, labels resources, selects algorithms (e.g., content-based) to recommend matched resources, updating via feedback.

2) Personalized path guidance: It helps set goals, formulates paths, monitors progress for adjustments, provides guidance and evaluates outcomes.

3) Data collection and processing: It combines objectives, content and interactions, processes via modeling engine (e.g., recommending content for wrong answers).

3.3 English adaptive learning path based on reinforcement learning algorithm

(1) Setting variables related to learning paths

The fit between learners and the learning environment can be quantified through measurement scales and survey questionnaires. There are three quantifiable dimensions of fit between the learning environment and learners: mission difficulty (MD), learning ability (LA), and learning methodology (LM) [6][7]. The calculation method for the fit between learning tasks and learners is:

$$SLEF=\gamma_1 MD+\gamma_2 LA+\gamma_3 LM \quad (1)$$

Among them, γ_1 represents the weight of mission difficulty; γ_2 represents the weight of student learning ability; γ_3 represents the weight of the learning method; SLEF represents fit [8].

The learning effect (LE) of learners after selecting and learning resources is given through expert knowledge or testing. The learning return value R obtained by learners after selecting and studying learning resources is calculated using the following formula:

$$R = FW \times SLEF + LE \quad (2)$$

1) Learning resources are sorted and a directed graph of learning resources is established. Among them, nodes in the directed graph are used as learning resources, and linear segments are used as units to represent the learner's selection of the current learning resources.

2) Based on learning resources and learning performance charts, learning is initialized and corresponding performance tables are selected. The learning path selection function is the nearest neighbor matrix of n*n. Among them, n is the

number of learning resources, and the value of the matrix is represented by R_{ij}. When it is equal to -1, it indicates that the difficulty of the i-th learning task is independent of the j-th learner. If R_{ij} is greater than -1, it indicates that the current learner has selected the j-th learning resource. When the i-th learning resource is learned, the j-th learning resource can also be learned. When the value of R_{ij} is greater than -1, it indicates that students would improve their learning outcomes after completing the i-th resource.

The parameter space of the English learning path planner is too large, leading to poor direct teaching effectiveness and meaningless path planning in the early stages of learning. Therefore, this article divides the overall learning objectives into three small objectives and proposes a reward function for goal oriented learning. The reward and punishment design of neural network algorithms based on reinforcement learning is as follows:

1) Space search: The behavior of exploring unknown spaces and the behavior of successfully completing learning tasks are rewarded, and behaviors that exceed boundaries and linger in place are punished.

$$f = 4n - 0.2e^{\frac{1}{\Delta d}} - 0.8B \qquad (3)$$

2) Avoidance training: Successful avoidance of behaviors that are too difficult to achieve positive learning returns (static, dynamic) is rewarded, and behaviors that exceed boundaries and cause collisions are punished.

$$f=2n+2n_1-0.8B \qquad (4)$$

In the formula, n_1 represents the number of times during which student 1 successfully completed the learning task and achieved a positive learning return rate. Student 1's strategy of successfully avoiding obstacles can receive a large reward value.

3) Optimal path: Based on the second stage of training, the behavior of selecting the path with the highest learning return rate is rewarded.

$$f = 2n + 0.2\frac{SLEF}{v} - 0.2t - 0.8B \qquad (5)$$

According to Formula (5), individuals who take less time to complete the training objectives receive higher evaluations. At the same time, in order to avoid the high difficulty of learning tasks leading to longer time consumption, compensation scores are introduced, and the higher the individual student's fit with the learning task, the higher the score.

The calculated semantic similarity values are sorted to obtain the best information item, which is the ontology term with the highest matching degree with the current semantic context information. When constructing a subject domain ontology library, in addition to extracting ontology terms, the relationships between ontology terms are also annotated, including hierarchical relationships, attribute relationships, and related relationships. Based on these relationships, reasoning can form several learning paths centered around the best information item, represented by $L = (l_i, l_j, \cdots, l_m, l_n)$. Finally, based on the learning objective space and learning history process, the learning paths are screened to generate learning paths that meet the personalized characteristics of learners.

4 Results and discussion

Taking a certain knowledge point as an example, this study designed n learning tasks (with indicators such as duration and validity, where higher validity indicates better learning effects) and selected three students as research objects to verify the algorithm.

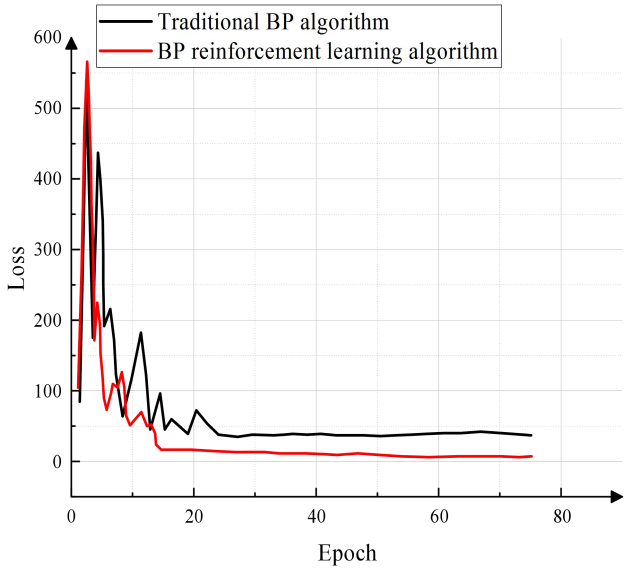

Figure 1. Traditional BP algorithm and reinforcement learning based on BP neural network

Compared with general reinforcement learning path planning methods, the learning method proposed in this paper can obtain better learning paths in a shorter time. The convergence speed and effectiveness of the BP neural network reinforcement learning in Figure 1 have been significantly improved compared to traditional BP algorithms.

Table 1. Learning path selection for different algorithms

Algorithm	Learning path options
Traditional BP Neural Network Algorithm	S1:A3→D2→A1→C2→C3→A2→B1→D1→B2→C1 S2:C2→D1→A1→C1→B1→A3→D2→C3→A2→B2 S3:A2→C2→B2→C1→A3→D1→C3→A1→D2→B1
Genetic Algorithm	S1:A1→C2→B1→B2→C3→A2→C1→D1→D2→A3 S2:B2→A3→A2→D2→C1→A1→D1→B1→C3→C2 S3:A1→D1→B1→C1→A2→B2→C2→A3→D2→C3
Adaptive BP Neural Network Algorithm	S1:D2→A2→C2→B2→C3→D1→B1→A3→C1→A1 S2:C2→A2→A3→D2→B2→C1→C3→B1→D1→A1 S3:C1→B2→D1→C3→A3→B1→D2→C2→A1→A2
Ant Colony Algorithm	S1:B2→A3→C2→D2→C3→A1→C1→A2→D1→B1 S2:C2→D1→C1→B1→C3→B2→A2→A3→A1→D2 S3:C3→A1→D1→B1→C1→A2→B2→C2→A3→D2

From Table 1, it can be seen that different algorithms have different choices for learning paths. The experiment also calculates the path selection time and optimization degree of the results of different algorithms, and the results are as follows:

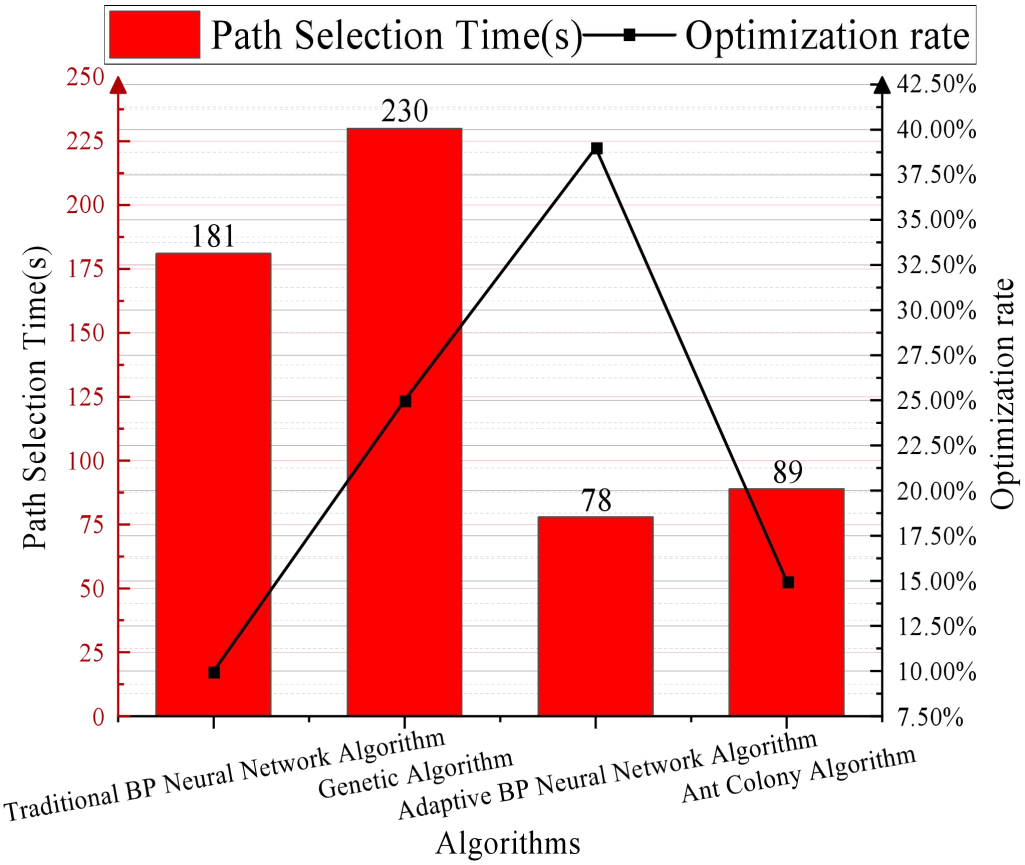

Figure 2. Results of this algorithm and comparison with other algorithms

From Figure 2, it can be seen that the adaptive BP neural network algorithm, which has the shortest path selection time and higher optimization rate among traditional BP neural network algorithms, genetic algorithms, adaptive BP neural network algorithms, and ant colony algorithms, is the algorithm proposed in this paper. Its path selection time and optimization rate are 78 seconds and 38.75%, respectively.

5 Conclusion

With AI development, adaptive learning algorithms emerge. Traditional teacher-centered education fails to meet diverse student needs. AI enables adaptive learning path planning, shifting to student-centered personalized education. Current research focuses on systems/models and platform comparisons, with limited core algorithm research.

Acknowledgement

This paper was part of the project of Research on Classroom Management Strategies for Foreign Language Teachers in Universities in the Age of Artificial Intelligence: An Exploratory Discussion Based on the Integration of Interactive Teaching and AIGC supported by 2024 Higher Education Research Project of China Association of Higher Education (Grant No. 24WY0415).

Conflicts of interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Benmesbah O, Lamia M, Hafidi M. 2023. An improved constrained learning path adaptation problem based on genetic algorithm. *Interactive Learning Environments*, 31(6): 3595-3612.

[2] Raj N S, Renumol V G. 2022. A systematic literature review on adaptive content recommenders in personalized learning environments from 2015 to 2020. *Journal of Computers in Education*, 9(1): 113-148.

[3] El-Sabagh H A. 2021. Adaptive e-learning environment based on learning styles and its impact on development students' engagement. *International Journal of Educational Technology in Higher Education*, 18(1): 1-24.

[4] Shemshack A, Kinshuk, Spector J M. 2021. A comprehensive analysis of personalized learning components. *Journal of Computers in Education*, 8(4): 485-503.

[5] Wang S, Christensen C, Cui W. 2023. When adaptive learning is effective learning: comparison of an adaptive learning system to teacher-led instruction. *Interactive Learning Environments*, 31(2): 793-803.

[6] Padakandla S. 2021. A survey of reinforcement learning algorithms for dynamically varying environments. *ACM Computing Surveys (CSUR)*, 54(6): 1-25.

[7] Rolf B, Jackson I, Müller M. 2023. A review on reinforcement learning algorithms and applications in supply chain management. *International Journal of Production Research*, 61(20): 7151-7179.

[8] Huang S, Dossa RFJ, Ye C. 2022. Cleanrl: High-quality single-file implementations of deep reinforcement learning algorithms. *Journal of Machine Learning Researc*h, 23(274): 1-18.

About the author

Hui Xu (1980-), female, from Anyang City, Henan Province; Master's degree from Huazhong University of Science and Technology; currently a teacher at the Department of Public Basic Courses, Wuhan Institute of Design and Sciences; Research interests: foreign linguistics and foreign language teaching research.