Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 18 No. 9 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijim.v18i09.49289

PAPER

An Intelligent Framework for English Teaching through Deep Learning and Reinforcement Learning with Interactive Mobile Technology

Jingan Hu(✉), Gaimin Jin

Department of Public Foreign Languages, Shijiazhuang College of Applied Technology, Shijiazhuang, China

RabbitAnn@126.com

ABSTRACT

As globalization deepens, the significance of English teaching in the educational landscape has become more prominent. Traditional teaching methods are increasingly inadequate for providing personalized and efficient learning experiences. This gap is being addressed by the rapid advancements in artificial intelligence, especially through deep and reinforcement learning. These technologies provide a framework for intelligent English teaching systems by mimicking human learning processes to customize personalized learning experiences, optimize learning paths, and enhance efficiency. However, challenges remain in fine-tuning teaching strategies to meet the varying needs of individual learners and dynamically adapting to their evolving interests in the short term. This study introduces a novel framework for an intelligent English teaching system that leverages the potential of interactive mobile technology alongside a deep Q-network (DQN) algorithm to dynamically adjust English teaching strategies. This approach enables real-time personalization of teaching strategies to create optimal learning paths for individual learners. Moreover, it incorporates a model based on neural collaborative filtering to capture and adapt to learners' short-term dynamic interests, thereby recommending relevant learning content in real-time. This framework enhances learning efficiency and personalizes content delivery, demonstrating considerable theoretical and practical value for the future of educational technology.

KEYWORDS

intelligent English teaching system, deep learning, reinforcement learning, deep Q-network (DQN), neural collaborative filtering, personalized learning pathways, dynamic interest modeling, interactive mobile technology

1 INTRODUCTION

Driven by the wave of globalization, English has emerged as an international language extensively utilized in academic exchange, business trade, and cross-cultural

Hu, J., Jin, G. (2024). An Intelligent Framework for English Teaching through Deep Learning and Reinforcement Learning with Interactive Mobile Technology. *International Journal of Interactive Mobile Technologies (iJIM)*, 18(9), pp. 74–87. https://doi.org/10.3991/ijim.v18i09.49289

Article submitted 2024-01-23. Revision uploaded 2024-03-17. Final acceptance 2024-03-18.

communication [1, 2]. The challenge of teaching English efficiently and effectively, thereby enhancing learners' linguistic capabilities, is significant in the field of educational technology [3, 4]. The rapid advancements in artificial intelligence, particularly the breakthroughs in deep learning and reinforcement learning, have gradually revealed their potential applications in intelligent teaching systems. By simulating the human learning process, intelligent English teaching systems are capable of providing personalized teaching plans and optimizing learning paths, thereby increasing learning efficiency [5, 6].

The current study on applying deep learning and reinforcement learning technologies to intelligent teaching systems aims to create a more intelligent learning environment for learners. It also focuses on continuously adjusting teaching strategies to meet learners' personalized needs [7]. Such research is of great importance for promoting the equitable distribution of educational resources and improving teaching quality. In the realm of English teaching, by precisely analyzing learners' behaviors and achievements, intelligent systems can offer more scientific and reasonable teaching suggestions, thereby advancing personalized and intelligent educational development [8–12].

However, existing research on the design and implementation of intelligent English teaching systems shows several deficiencies. For example, many systems lack the capability to dynamically adjust teaching strategies, thus failing to effectively address changes in learners' knowledge levels and interests [13–15]. Furthermore, the design of recommendation systems for managing learners' short-term dynamic interests is not adequately refined, resulting in a mismatch between the recommended content and the actual needs of the learners. These challenges limit the personalized effectiveness of intelligent English teaching systems and reduce teaching efficiency.

The main research content of this study consists of two parts. Firstly, a module for planning English teaching strategy change pathways has been designed based on the deep Q-network (DQN) algorithm. This module can dynamically adjust teaching strategies based on real-time feedback and learning outcomes from learners, planning the optimal learning paths for them. Secondly, a model for neural collaborative filtering of English learning recommendations based on short-term dynamic interests is proposed. This model can capture learners' immediate interests and make intelligent recommendations to enhance the relevance and attractiveness of learning content. These research components not only address the shortcomings of existing intelligent English teaching systems but also provide new theoretical support and technological solutions for personalized English learning, possessing high theoretical value and application potential.

2 ENGLISH TEACHING STRATEGY CHANGE PATHWAY PLANNING BASED ON DQN ALGORITHM

In the optimization requirements of the intelligent English teaching system framework, the most pressing issue is how teaching strategies can be dynamically adjusted based on learners' real-time performance and feedback to match their personalized needs while maintaining the attractiveness and effectiveness of the teaching content. Existing systems generally lack a profound understanding of and responsive approach to learner behavior. This results in rigid teaching strategies that find it challenging to adjust to changes in learners' knowledge levels and interests, thereby significantly restricting learning efficiency and experience. The study, based on the DQN algorithm, aims to address this core challenge by developing a module for planning teaching strategy change pathways. This module is capable of autonomous learning and optimization, enabling real-time and dynamic adjustment of teaching strategies. Figure 1 presents the framework of the teaching strategy change pathway planning module.

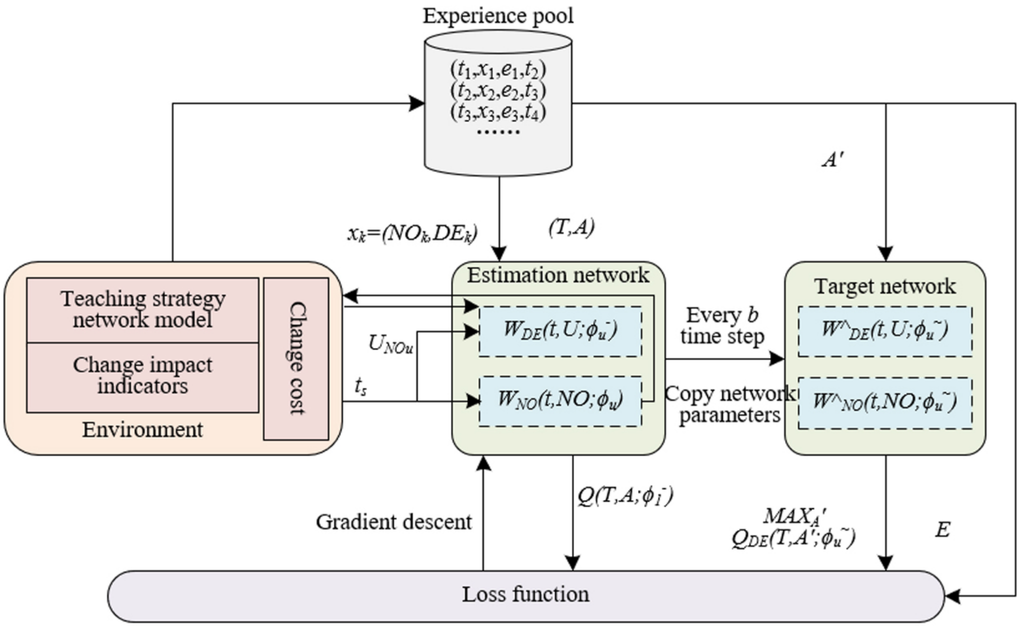

Fig. 1. Framework of the teaching strategy change pathway planning module

The method initially defines teaching impact indicators on the constructed teaching strategy network model to quantify the impact of a specific teaching activity change on the entire teaching system. Subsequently, a simulated teaching environment is constructed using the teaching strategy network model, teaching impact indicators, and teaching costs as a foundation for designing the environmental states and agent actions in the DQN algorithm. The DQN algorithm is utilized to explore the optimal teaching strategy path, with the goal of addressing the initial teaching strategy adjustment request by minimizing the total cost of teaching changes. The complexity of teaching strategy pathway planning lies in the necessity to consider multiple factors such as the coherence of teaching content, the appropriateness of learning difficulty, and the learners' capacity to assimilate information. Trained agents will be able to autonomously plan a path that thoroughly considers teaching effectiveness in response to initial requests for teaching strategy changes. This will enhance the efficiency of teaching strategy adjustments and reduce the negative impacts of teaching changes.

To precisely depict the impact of changes at a certain teaching strategy node NO_u on other nodes NO_k within the entire teaching system, it is necessary to calculate the change impact indicators for teaching strategy pathway nodes. Modifications in teaching strategy will first affect the node's direct "neighbor" nodes, i.e., the teaching activities or concepts associated with it. This impact will propagate outward as the teaching process progresses, affecting other teaching units. The intensity and range of propagation depend not only on the logical association between teaching units but also on factors such as teaching content difficulty, learners' cognitive abilities, and learning styles, which are similar to the connection strength between nodes. Therefore, this article proposes using the interdependency of teaching activities as a factor for decrement to simulate the propagation of teaching strategy changes. The propagation will cease when the impact strength diminishes to a point where it no longer significantly affects the learning process.

Assuming the neighbor node set is represented by NE_u, the connection strength between nodes by $E_u = (e_{uk}, e_{uj}, \ldots)$, the decrement factor of propagation by e_{xy}, the propagation pathway from node u to node k by PA_{uk}, and the connection strength from node x to node y by e_{xy}, the impact degree of change at node u on node j is denoted by IM_{uj}. The specific calculation process for an initial change at a certain node is illustrated in equations (1) and (2).

$$IM_{uk} = \prod_{x,y \in PA_{uk}} e_{xy} t \tag{1}$$

$$U_u = (IM_{u1}, IM_{u2}, \ldots, IM_{uj}, \ldots, IM_{uv}) \tag{2}$$

In the intelligent English teaching system, the DQN algorithm tailored for teaching strategy change pathway planning must accommodate the unique high-dimensional state spaces and continuous action spaces characteristic of the educational domain. Compared to traditional Q-learning algorithms, the advantage of the DQN algorithm lies in its utilization of deep learning networks to estimate the value function. This approach avoids the need for the direct storage and updating of large Q-tables, which is especially crucial for managing complex decision-making scenarios in teaching.

The implementation of the DQN algorithm for intelligent English teaching strategies involves the construction of two neural networks: the estimation network $W(t, x; \varphi)$ and the target network $\hat{W}(t, x; \varphi^-)$, which share the same architecture but differ in their parameter update strategies. The estimation network is responsible for generating the Q-values of the current policy, while the target network is used to calculate the target Q-values. The target network's parameters are updated at fixed intervals denoted by C to stabilize the learning process. Additionally, an experience replay pool (F) is maintained to store experiences generated from interactions between the agent and the environment. This helps to break the temporal correlations between experiences and enhance learning efficiency. The capacity of the experience pool is set to "f," limiting the required storage space and maintaining the diversity of experiences.

At the start of each training episode, the environment state t is initialized, and the agent evaluates the policy based on the current network state before taking action. The selection of each action is based on an ε-greedy strategy to balance exploration and exploitation. After executing an action, the agent observes a new environment state t_{s+1} and immediate reward e_s. This information is compiled into a new experience (t_s, x_s, e_s, t_{s+1}) and stored in the experience pool F. When the experience pool reaches full capacity, old experiences are replaced on a first-in, first-out basis. During the training process, a batch of experiences is randomly sampled from the experience pool. These experience samples, along with the target network, are used to compute the loss function. The parameters of the estimation network are updated through gradient descent. This training method, which relies on experience replay and the target network, helps improve the stability and efficiency of the intelligent English teaching system by planning strategy change pathways. Assuming the parameters of the estimation network $W(t, x; \varphi)$ at the u-th timestep are denoted by φ_u, and the parameters of the target network $\hat{W}(t, x; \varphi^-)$ at the u-th timestep are denoted by φ_u^-. The formula for calculating the loss function is as follows:

$$M_u(\varphi_u) = R_{(t,x,e,t')\sim I(T,X,E,T')}[(e + \varepsilon MAX_{x'} \hat{W}(t', x'; \varphi_u^-) - W(t, x; \varphi_u))^2] \tag{3}$$

Gradients with respect to the parameters φ_u are calculated, and the parameters of the estimation network $W(t, x; \varphi)$ are updated using direction of gradient descent. The formula for calculating the gradient is as follows:

$$\begin{aligned}&\frac{\partial M_u(\varphi_u)}{\partial \varphi_u}\\&= R_{(s,a,r,s')\sim U(S,A,R,S')}[(e+\varepsilon MAX_{x'}\hat{W}(t',x';\varphi_u^-)-W(t,x;\varphi_u))\nabla_{\varphi_u}W(t,x;\varphi_u)]\end{aligned} \tag{4}$$

In the teaching strategy change pathway planning based on the DQN algorithm, the actions of the agent represent potential modifications to the teaching strategy. These actions can include introducing new teaching materials, adjusting course difficulty, changing teaching methods, increasing or decreasing homework volume, providing additional learning resources, etc. Each action should aim to improve learning outcomes, increase student engagement, or optimize teaching efficiency. In designing the action space, it is necessary to ensure that it is rich enough to cover all possible teaching strategy adjustments while being moderate enough to avoid overcomplicating the problem.

The environmental state of the intelligent English teaching system reflects the students' learning situation and the current state of the environment. This can include students' learning progress, test scores, course interactions, study habits, feedback information, and more. The design of the environmental state needs to comprehensively reflect all important information in the teaching process, enabling the agent to make reasonable strategy adjustments based on this information. The state representation should be simplified to reduce the computational burden on the value function network and improve the policy's effectiveness while still retaining significant features for decision-making. Assuming the change node selected by the agent at time s is represented by NO_s, the change amount at time s for change node NO_s is denoted by DE_s, and the degree of change at node NO_j at time s compared to the initial state is denoted by ΔNO_{js}, the definitions of action x_s and environmental state t_s are given as follows:

$$x_s = (NO_s, SE_s) \tag{5}$$

$$t_s = (\Delta NO_{1s}, \Delta NO_{2s}, \Delta NO_{3s}, \ldots \Delta NO_{vs}) \tag{6}$$

The value function network, a core component of the DQN algorithm, is used to estimate the Q-values of state-action pairs. These values represent the expected return of choosing a particular action in a specific state. In the context of English teaching strategy change pathway planning, the value function network must be able to handle high-dimensional inputs and output Q-values for all possible actions. The design of the network should consider the characteristics of the teaching environment, such as the potential need to process sequential data (e.g., a sequence of student responses), which may require the adoption of structures such as recurrent neural networks (RNN) or long short-term memory networks (LSTM). Training the network must be conducted in a stable and effective manner to ensure accurate predictions in the dynamically changing educational environment. The node with the maximum Q-value is denoted by $NO_s = argmaxW(t_s, NO)$, with the corresponding node change impact indicator denoted by U_{NO}. For a timestep s, the environmental state t_s is input into $W_{NO}(t, NO; \varphi_u)$, with the network outputting Q-values for all nodes and selecting NO_s as the change node. U_{NO} is concatenated with t_s as the input for $WDE(t, U; {}^-\varphi_u)$, with the network outputs the change amount DE_s for NO_s.

The reward function defines the immediate feedback obtained by the agent after taking an action in the environment. In an intelligent English teaching system, the design of the reward function is crucial because it should accurately reflect the long-term effects of teaching strategy changes. Rewards can be based on several dimensions, such as improvement in students' learning progress, an increase in test scores, or enhanced classroom participation. When designing the reward function, these factors need to be considered collectively to ensure alignment with teaching objectives. This means that maximizing rewards should result in optimizing teaching outcomes. Furthermore, the reward function must avoid the interference of short-term rewards to promote the enhancement of long-term learning effects. The total change cost at time *t* is represented by $COSs_s$, with the initial cost of making an initial change request denoted by $COSs_{IN}$, the timestep interval from the initial change at time *s* denoted by $SEPo_s$, and the maximum number of steps for the product change pathway denoted by $SEPo_{MAX}$. The reward function e_s can be calculated as follows:

$$es_t = \frac{COSs_{s-1}COSs_s}{COSs_{s-1}} + \frac{SETo_s}{SETo_{MAX}} * \frac{COSs_{IN} - COSs_s}{COSs_{IN}} \tag{7}$$

3 NEURAL COLLABORATIVE FILTERING MODEL FOR ENGLISH LEARNING INTERCONNECTION RECOMMENDATIONS BASED ON SHORT-TERM DYNAMIC INTERESTS

In response to the dynamic nature of learners' interests and the attractiveness of English teaching content over time, as well as the issue of inter-course relevancy within intelligent English teaching systems, a neural collaborative filtering model for English learning interconnection recommendations based on short-term dynamic interests is proposed. This model utilizes interaction graphs between learners and teaching content, learners' social network graphs, and implicit association graphs of teaching content as inputs. During the learning phase, an attention mechanism combined with LSTM is utilized to capture learners' long-term preferences and short-term interest changes, as well as the long-term value and short-term trends of the teaching content. Consequently, these dynamic and static factors are integrated to generate latent representations of both learners and teaching content.

3.1 Model inputs

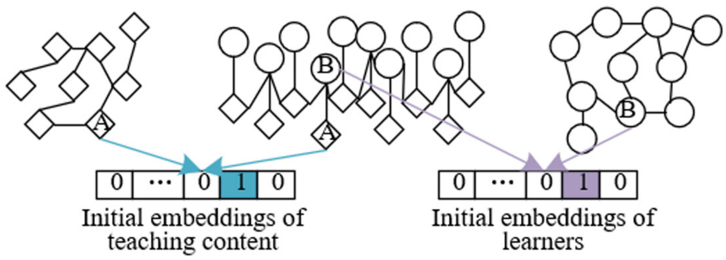

Fig. 2. Framework of the model input module

In the intelligent English teaching system, the neural collaborative filtering model for English learning offers personalized recommendations based on short-term dynamic interests. It utilizes the interaction matrix E between learners and English teaching content, the internal social network H_I of learners, and the implicit association network H_N among teaching content as inputs. Figure 2 displays the framework of the model input module. The implicit network H_N of teaching content is constructed by analyzing learners' evaluations and interactions with English teaching content, using cosine similarity to measure the association between teaching contents.

The similarity coefficient between any teaching content u and k is defined by t_{uk}. If $t_{uk} > \gamma$, with a fixed threshold represented by γ, then the teaching content u and k are considered similar. The formula for calculating similarity is as follows:

$$t_{uk} = \frac{E(\cdot,u)^S E(\cdot,j)}{\|E(\cdot,u)\|_2 \|E(\cdot,j)\|_2} \tag{8}$$

In this model, the embeddings of learners are represented not only by personal characteristics but also include their history of interactions with teaching content; similarly, the embeddings of teaching content also incorporate information about the learners who have evaluated it. Additionally, learners' ratings e of teaching content (such as completion, mastery level, etc.) are embedded into an f-dimensional vector and combined with the embeddings of learners and teaching content through a multi-layer perceptron (MLP) to produce interaction embeddings. This design aims to uncover learners' short-term dynamic learning interests and the immediate appeal of teaching content, thus offering more personalized and dynamically adaptive English learning recommendations for each learner.

The embedding of ratings is denoted by r_e, assuming the initial embeddings of learners and teaching content are represented by i_u and n_k respectively, and the interaction embeddings of learners and teaching content through MLP are represented by h_i and h_n, the relevant calculations are as follows:

$$w_{ke} = h_i(r_e \oplus n_k), o_{ue} = h_n(r_e \oplus i_u) \tag{9}$$

3.2 Representation of learners' short-term dynamic interests and long-term static interests

The model focuses on extracting learners' long-term static interests from the perspectives of social networks and teaching content, as well as the enduring value of teaching content. Initially, the model retrieves the original static embeddings of learners and teaching content, which capture the inherent characteristics of learners and the fundamental attributes of teaching content.

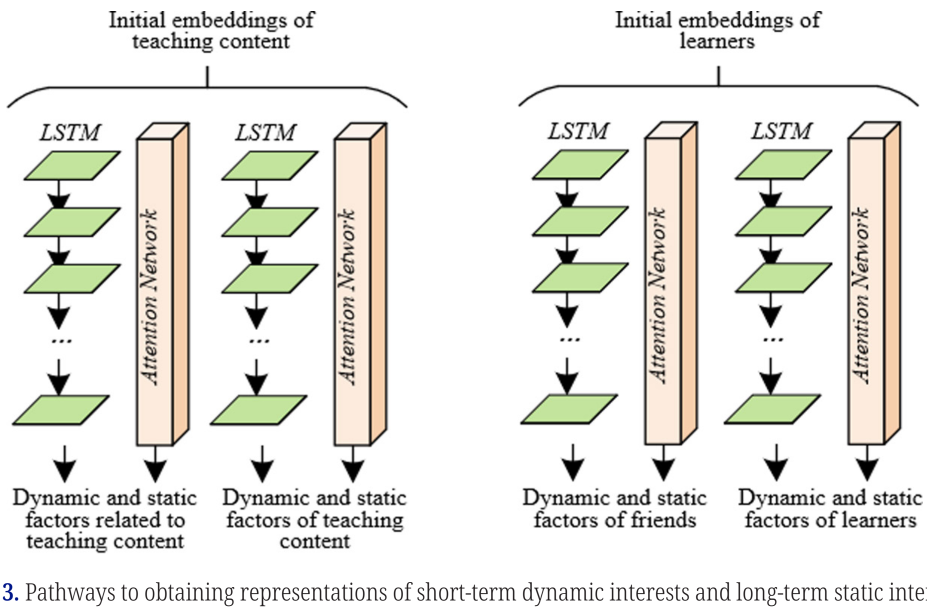

Fig. 3. Pathways to obtaining representations of short-term dynamic interests and long-term static interests

Through graph neural network (GNN) layers, the model aggregates information for both learners and teaching content, while utilizing an attention mechanism to distinguish the importance of different neighbors. This process generates refined long-term static latent factors G_u^M and G_k^M for each learner and teaching content, as detailed in Figure 3. Assuming the obtained static embeddings of learners and teaching content are represented by w_{ue} and o_{ue}, the nonlinear activation function by δ, and the network's weights and biases by Q and y, the static modeling calculation formula is as follows:

$$G_u^M = \delta\left(\sum\nolimits_{k\in Z(u)} \beta_{uk} Q_i w_{ke} + y_i\right) \tag{10}$$

$$G_k^M = \delta\left(\sum\nolimits_{k\in Z(k)} \beta_{ku} Q_n o_{ue} + y_{in}\right) \tag{11}$$

The weights and biases of the first and second layers of the attention network are denoted by (Q_1, y_1) and (Q_2, y_2), with the calculation formulas for β_{uk} and β_{ku} as follows:

$$\beta_{uk} = \frac{\exp(Q_2^S \cdot \delta(Q_1 \cdot (Q_i i_u \oplus Q_i w_{ke}) + b_1) + b_2)}{\sum_{j\in Z(u)} \exp(Q_2^S \cdot \delta(Q_1 \cdot (Q_i i_u \oplus Q_i w_{ke}) + b_1) + b_2)} \tag{12}$$

$$\beta_{ku} = \frac{\exp(Q_4^S \cdot \delta(Q_3 \cdot (Q_n i_k \oplus Q_n o_{ue}) + b_3) + b_4)}{\sum_{s\in Y(u)} \exp(Q_4^S \cdot \delta(Q_3 \cdot (Q_n i_k \oplus Q_n o_{ue}) + b_3) + b_4)} \tag{13}$$

For the social network of learners, the model calculates the static latent factor G_p^M for each learner i_u's friend i_p, and the static latent factor G_j^M for other teaching content n_j associated with the target teaching content n_k. Such long-term static representations help the model capture stable interests and preferences, providing more accurate and consistent English learning content recommendations aligned with learners' learning histories. Unlike traditional social recommendation models, the proposed model also takes into account the latent connections between teaching content and learners' progress, thereby improving the applicability and effectiveness of the recommendation system in educational settings. The specific formulas are provided as follows:

$$G_p^M = \delta\left(\sum\nolimits_{p\in V(u)} \beta_{up} Q_0 i_o + y_0\right) \tag{14}$$

$$G_j^M = \delta\left(\sum\nolimits_{j\in L(k)} \beta_{kj} Q_0 i_j + y_0\right) \tag{15}$$

In the model, the calculation of short-term dynamic representations focuses on capturing the immediate interaction state between learners and teaching cleaners' dynamic learning content. This reflects interests and the current attractiveness of the teaching content. For each learner i_u, the model considers their interactions with teaching content as a time series $T(u)$, comprised of interaction embeddings w_{ke} within a time window; similarly, for each teaching content vj, there exists a series $T(k)$ composed of interaction embeddings w_{ke} with various learners. To dynamically model these sequences for learners and teaching content, LSTM, a variant of RNN capable of processing sequence data, is employed. This approach efficiently captures and updates the dynamic latent factors G_u^t for learners and G_k^t for educational content. Compared to traditional neural collaborative filtering social recommendation models, the constructed model places greater emphasis on the dynamism of teaching content and the temporal aspect of learners' behaviors. This enables the recommendation system to respond in real-time to learners' current learning needs and changes in teaching content, providing more timely and personalized learning resource recommendations. The related calculations are provided as follows:

$$\begin{cases} a_v = \delta\left(Q_a\left[g_{v-1}, w_{ke(\pi=v)}\right] + y_a\right) \\ d_v = \beta\left(Q_d\left[g_{v-1}, w_{ke(\pi=v)}\right] + y_d\right) \\ p_v = \delta\left(Q_p\left[g_{v-1}, w_{ke(\pi=v)}\right] + y_p\right) \\ \tilde{z}_v = \tanh\left(Q_z\left[g_{v-1}, w_{ke(\pi=v)}\right] + y_z\right) \\ z_v = d_v \Phi z_{v-1} + a_v \Phi \tilde{z}_v \\ G_u^T = g_v = p_v \Phi TANg(z_v) \end{cases} \tag{16}$$

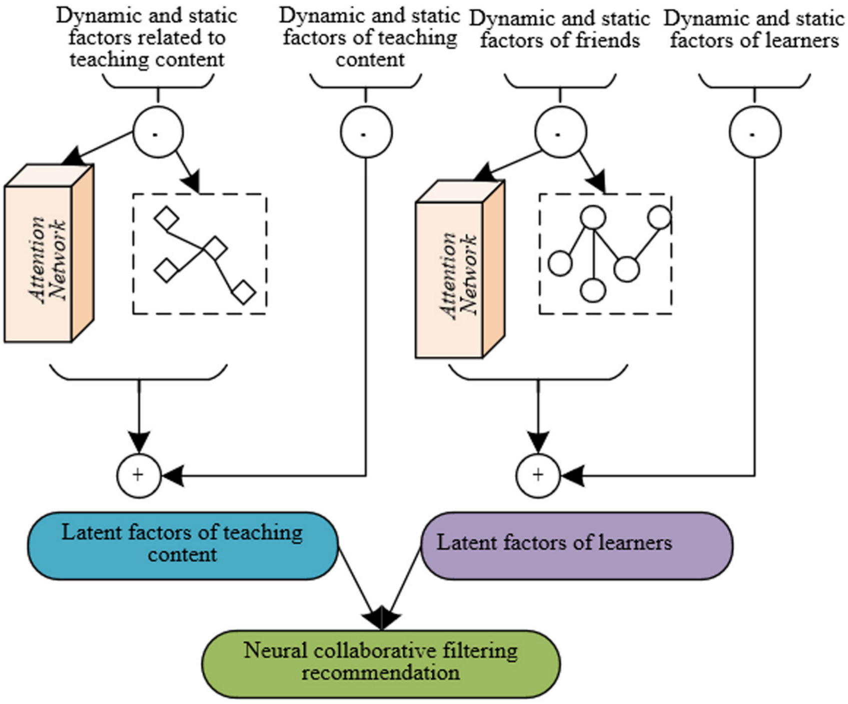

Fig. 4. Process of obtaining latent factors

In the model, the latent factors of learners and teaching content are calculated by integrating short-term dynamic representations with long-term static representations, forming a more comprehensive representation of learners' interests, g_u^U, and teaching content attractiveness, g_k^U. Figure 4 illustrates the process of obtaining latent factors. Specifically, the model initially captures the interaction behaviors of learners with teaching content through the interaction matrix E. It then examines the social relationships among learners and their influence on learners' interests through the social network H_I. Finally, it explores the interrelations among teaching content and their impact on the attractiveness of teaching content through the implicit network H_N. Within the social network, the model allocates weights based on the strength of friendships. Closer social connections lead to a greater influence on learners' interests, as reflected in the final calculation of the learner's latent factor, g_u. Similarly, in the implicit network of teaching content, the appeal of related teaching content is combined to create the latent factor g_k for teaching content. Compared to traditional neural collaborative filtering social recommendation models, the constructed model specifically targets the English learning scenario. It considers the dynamic interactions between learners and teaching content, as well as the learning-related behaviors of learners within the social network. This approach more accurately reflects learners' real-time learning needs and the educational value of teaching content, ensuring the effectiveness and personalization of the recommendation system within intelligent English teaching systems. The specific formulas are provided as follows:

$$g_u^U = G_u^M \Phi G_u^T, g_k^U = G_k^M \Phi G_k^T \tag{17}$$

$$g_p^U = G_p^M \Phi G_p^T, g_j^U = G_j^M \Phi G_j^T \tag{18}$$

Combining the influence of social network information on learners and teaching content network information on teaching content, g_u and g_k can be calculated through the following formulas:

$$g_u = h_i\left(g_u^U \oplus \left(\delta\left(\sum_{p\in V(i)} \beta_p W_t g_p^U + y_t\right)\right)\right) \tag{19}$$

$$g_k = h_n\left(g_k^U \oplus \left(\delta\left(\sum_{j\in V(k)} \beta_j W_U g_j^U + y_U\right)\right)\right) \tag{20}$$

Furthermore, g_u and g_k can be utilized as inputs for predicting ratings in the English learning interconnection recommendation.

4 EXPERIMENTAL RESULTS AND ANALYSIS

Figure 5 demonstrates that, during the process of changing English teaching strategies, the total cost of change shows a downward trend as the number of strategy adjustments increases. Initially, the total change cost is at 4500, while the initial change cost remains constant at 1800. When the running time reaches 600, the total change cost decreases to 1780. However, it subsequently rises to 2200 after 1000 runs, possibly due to the exploration involved in the strategy adjustment process. Thereafter, the optimal total change cost gradually declines, reaching its lowest point at 1200 after 1600 iterations and stabilizing thereafter. This indicates that with the continuous adjustment of strategies, the DQN algorithm is capable of effectively learning and planning lower-cost paths for teaching strategy changes. The analysis concludes that planning teaching strategy change pathways based on the DQN algorithm demonstrates significant effectiveness. Experimental data reveal that although total change costs may fluctuate during the early stages of learning due to the algorithm's exploration of new strategies, the DQN algorithm can gradually identify more optimized paths for teaching strategy changes. This process effectively reduces the total change cost, ultimately achieving a stable and low-cost level.

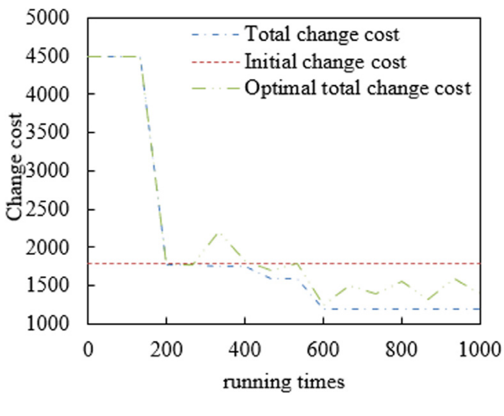

Fig. 5. Change curve of total change cost in English teaching strategy changes

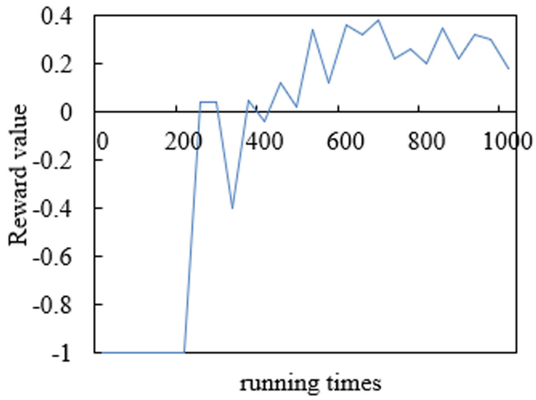

Fig. 6. Cumulative reward value change curve of agents in the DQN algorithm

The experimental results shown in adapting Figure 6 indicate that the performance of the DQN algorithm in English teaching strategies gradually improves with an increase in the number of iterations. During the initial 600 runs, the agents' cumulative reward value remains at –1, indicating that the algorithm has not yet identified effective paths for strategy changes during the exploration phase. However, from the 800th iteration onwards, positive reward values start to emerge, despite minor fluctuations during this process (e.g., the reward value is 0.04 at 800 iterations, then decreases to –0.4, and eventually rises again to 0.05). The overall trend shows an increase in reward values, reaching 0.36 by the 1000th running time and then maintaining within the range of 0.18 to 0.38. This change indicates that as the algorithm continues to learn and adjust, the agents begin to obtain higher cumulative reward values, signifying improvements in the effectiveness of teaching strategy change pathways. The analysis concludes that planning teaching strategy change pathways based on the DQN algorithm demonstrates significant applicability and effectiveness. Although the agents' exploration leads to negative cumulative reward values in the initial phase, as time progresses, the agents are able to learn from the environment and optimize their strategies, as evidenced by the continuously increasing positive reward values.

Table 1. Comparative experimental results of different nodes in English teaching strategy change pathway planning

Initial Change Request	Initial Change Cost	The Proposed Algorithm	DNM Algorithm
(*NO*4, 0.65)	1685.55	(*NO*4, 0.65) → (*NO*8, 0.244) → (*NO*15, 0.179) Total change cost: 1512.27	(*NO*4, 0.65) → (*NO*6, 0.234) → (*NO*6, 0.318) → (*NO*11, 0.27) → (*NO*19, –0.47) Total change cost: 1589.23
(*NO*9, 0.71)	923.56	(*NO*9, 0.71) → (*NO*12, 0.418) → (*NO*17, 0.12) Total change cost: 623.48	(*NO*9, 0.71) → (*NO*16, 0.741) Total change cost: 721.16
(*NO*14, 0.11)	1879.24	(*NO*14, 0.11) → (*NO*6, 0.132) → (*NO*17, 0.191) → (*NO*27, 0.174) Total change cost: 1212.27	(*NO*14, 0.11) → (*NO*19, –0.44) → (*NO*24, –0.547) → (*NO*12, –0.347) → (*NO*28, –0.284) Total change cost: 1379.23

The comparative experimental results presented in Table 1 illustrate the cost-effectiveness of the proposed English teaching strategy. The change pathway planning utilizes the DQN algorithm across multiple nodes. Taking the initial change request (*NO*4, 0.65) as an example, where the initial change cost is 1685.55, the pathway generated by the proposed algorithm is (*NO*4, 0.65) → (*NO*8, 0.244) → (*NO*15, 0.179), reducing the total change cost to 1512.27. In comparison, the pathway generated by the decision navigation model (DNM) algorithm results in a significantly higher total change cost of 1589.23. For other nodes, such as (*NO*9, 0.71) and (*NO*14, 0.11), the proposed algorithm also shows lower total change costs of 623.48 and 1212.27, respectively, compared to the DNM algorithm's costs of 721.16 and 1379.23. These data indicate that the DQN algorithm not only identifies lower-cost paths for teaching strategy changes but also performs better under various initial conditions than the DNM algorithm. The analysis concludes that the English teaching strategy change pathway planning method based on the DQN algorithm possesses distinct advantages. Given the DQN algorithm's ability to adjust teaching strategies based on real-time feedback and learning outcomes, it demonstrates higher efficiency and lower costs in designing optimal learning pathways compared to the traditional DNM algorithm.

Tables 2 and 3 compare the performance of different English learning recommendation models across three datasets using the mean absolute error (MAE) and root mean square error (RMSE) metrics. In Table 3, the MAE of the singular value decomposition++ (SVD++) and matrix factorization (MF) models on the interaction dataset are observed to be 0.7251 and 0.7269, respectively. On the social network dataset, they values are 0.7784 and 0.7895, while on the learning content information dataset, they are 0.5894 and 0.5741. In comparison, the model proposed in this paper achieves MAE percentages of 1.78%, 1.49%, and 1.38% across these three datasets, indicating superior performance in reducing errors. Similarly, in Table 3, the RMSE of the SVD++ and FM models on the interaction dataset is 0.9815 and 0.9568, on the social network dataset, they are 1.1256 and 1.1245, and on the learning content information dataset it is 0.7854 and 0.7546. The RMSE percentages for the model proposed in this study are 2.89%, 1.43%, and 1.73% on these datasets, respectively. These results further validate the effectiveness of the proposed model in reducing recommendation errors. The analysis concludes that the neural collaborative filtering model for English learning interconnection recommendation proposed in this study demonstrates lower MAE and RMSE across different datasets, indicating the model's accurate capture of users' immediate interests and ability to provide highly relevant recommendations. This performance advantage is attributed to the model's effective identification and application of short-term dynamic interest. This ensure that recommended content is closely related to users' current learning needs and interests.

Table 2. MAE comparison of recommendations for different English learning interconnection recommendation models

Model	Interaction Dataset	Social Network Dataset	Learning Content Information Dataset
SVD++	0.7251	0.7784	0.5894
MF	0.7269	0.7895	0.5741
The model proposed in this study	1.78%	1.49%	1.38%

Table 3. RMSE comparison of English learning interconnection recommendation models

Model	Interaction Dataset	Social Network Dataset	Learning Content Information Dataset
SVD++	0.9815	1.1256	0.7854
MF	0.9568	1.1245	0.7546
The model proposed in this study	2.89%	1.43%	1.73%

5 CONCLUSION

This study primarily focuses on two aspects: firstly, the design and implementation of an English teaching strategy change pathway planning module based on the DQN. Utilizing methods of reinforcement learning, this module dynamically adjusts teaching strategies based on learners' real-time feedback and learning outcomes, aiming to provide learners with personalized and optimal learning pathways. Experimental validation has shown the effectiveness of this module in terms of reducing teaching strategy change costs and increasing cumulative reward values. It demonstrates the capability to intelligently adjust teaching strategies according to varying learning scenarios. Secondly, a neural collaborative filtering model based on short-term dynamic interests, specifically for English learning interconnection recommendations, has been proposed. This model effectively captures learners' short-term dynamic interests, enhancing the relevance and appeal of learning content through an intelligent recommendation mechanism. Experimental results indicate that this model outperforms other variant models in terms of MAE and root mean square error.

However, this study also has certain limitations. The training process of the DQN algorithm may require a significant number of samples and time, and the algorithm's performance could be limited by particular state space designs and reward function settings. While the neural collaborative filtering model shows significant effectiveness in capturing short-term interests, further consideration of the stability of long-term interests and exploration of novel items is needed. Future research directions could expand in the following areas: first, attempting to combine the DQN with other types of reinforcement learning algorithms to further optimize the efficiency and effectiveness of teaching strategy adjustments. Secondly, explore models of recommendation systems that integrate short-term and long-term interests to achieve a better balance between recommendation accuracy and user satisfaction. Lastly, explore methods to decrease the time and resources required for model training, facilitating easier deployment of the model in practical applications.

6 REFERENCES

[1] Y. Wang, "Study of application of multimedia technology in English language teaching," in *7th International Conference on Information Technology in Medicine and Education (ITME)*, 2015, pp. 593–597. https://doi.org/10.1109/ITME.2015.136

[2] L. Liu, "The application and enlightenment of second language acquisition theory in English teaching," in *Proceedings of the 7th International Conference on Education, Management, Information and Mechanical Engineering (EMIM 2017)*, 2017, pp. 1239–1243. https://doi.org/10.2991/emim-17.2017.251

[3] M. Xiang, F. Mao, and L. Xiao, "A study on the integration of digital language teaching system into English teaching," *Journal of Computational Methods in Sciences and Engineering*, vol. 23, no. 2, pp. 913–920, 2023. https://doi.org/10.3233/JCM-226491

[4] H. Gao and Q. Ma, "Application of task-based language teaching in web-based college English teaching in China's context," *Advanced Materials Research*, vol. 271, pp. 1682–1687, 2011. https://doi.org/10.4028/www.scientific.net/AMR.271-273.1682

[5] M. Zulqarnain and M. Saqlain, "Text readability evaluation in higher education using CNNs," *Journal of Industrial Intelligence*, vol. 1, no. 3, pp. 184–193, 2023. https://doi.org/10.56578/jii010305

[6] D. Zhu, "Analysis of the application of artificial intelligence in college English teaching," in *Proceedings of the 2017 2nd International Conference on Control, Automation and Artificial Intelligence (CAAI 2017)*, 2017, pp. 235–237. https://doi.org/10.2991/caai-17.2017.52

[7] H. Guo, "Research on the current situation and countermeasures of English listening teaching based on multimedia intelligent-embedded processor," *Journal of Sensors*, vol. 2022, pp. 1–11, 2022. https://doi.org/10.1155/2022/9722209

[8] H. Huang and C. T. Hsin, "Environmental literacy education and sustainable development in schools based on teaching effectiveness," *International Journal of Sustainable Development and Planning*, vol. 18, no. 5, pp. 1639–1648, 2023. https://doi.org/10.18280/ijsdp.180535

[9] B. Asare, Y. D. Arthur, and F. O. Boateng, "Exploring the impact of ChatGPT on mathematics performance: The influential role of student interest," *Education Science and Management*, vol. 1, no. 3, pp. 158–168, 2023. https://doi.org/10.56578/esm010304

[10] L. Zhu, X. Yan, and J. Wang, "A recognition method based on speech feature parameters-English teaching practice," *Mathematical Problems in Engineering*, vol. 2022, pp. 1–11, 2022. https://doi.org/10.1155/2022/2287468

[11] O. Shkvyr, H. Dudchak, N. Kazakova, O. Polianovska, and N. Sivak, "Mathematical modeling of information technology integration in digital education: A regional perspective," *Ingénierie des Systèmes d'Information*, vol. 28, no. 3, pp. 603–610, 2023. https://doi.org/10.18280/isi.280308

[12] M. Saqlain, "Revolutionizing political education in Pakistan: An AI-Integrated approach," *Education Science and Management*, vol. 1, no. 3, pp. 122–131, 2023. https://doi.org/10.56578/esm010301

[13] J. Guo, "Practical problems and improvement strategies of English vocabulary teaching from the perspective of educational psychology," *Psychiatria Danubina*, vol. 33, no. S6, pp. 410–417, 2021.

[14] X. Wei, "Optimization design of teaching strategies for English teaching achievement improvement based on original algorithm," *Wireless Personal Communications*, vol. 102, pp. 1191–1201, 2018. https://doi.org/10.1007/s11277-017-5175-2

[15] M. Liu, "Analysis on implementation strategy of large-scale class teaching for college English under teaching paradigm reform mode," in *Proceedings of the 4th International Conference on Education Management and Information Technology (ICEMIT)*, 2018, pp. 1048–1052.

7 AUTHORS

Jingan Hu, who graduated from Northeast Normal University with an MA in English Language and Literature, is now working at Shijiazhuang College of Applied Technology. Her main research direction is English teaching and translation (E-mail: RabbitAnn@126.com; ORCID: https://orcid.org/0009-0009-7816-3145).

Gaimin Jin, graduated from Nanjing Normal University with an MA in English Language and Literature, is now working at Shijiazhuang College of Applied Technology. Her main research direction is English teaching and translation (E-mail: jingm_123@126.com; ORCID: https://orcid.org/0009-0003-4891-176X).