ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2 - 2026

Research on Adaptive Teaching Strategy Optimization of Smart Physical Education Course in Colleges and Universities driven by reinforcement Learning

Yan Yang[1] *and Zheng Li*[1,*]

[1] Department of Physical Education, Tianjin Chengjian University, Tanjin, 300384, Tian Jin, China

SUMMARY: *Aiming at the problems of insufficient identification of students 'differences, lagging exercise load adjustment and low teaching feedback accuracy in college physical education courses, an adaptive teaching strategy optimization method of smart physical education courses driven by reinforcement learning was constructed. The system fuses physical monitoring, sports performance, classroom participation and learning feedback data to form a dynamic state representation of students, and uses DDPG algorithm to optimize teaching content, exercise load and feedback methods. The experimental results show that the accuracy of model state recognition reaches 94.6%, the load safety control rate is 96.2%, the comprehensive achievement degree of the course is improved to 92.4%, and the comprehensive satisfaction degree is 93.2%. The results show that this method can improve the individual adaptability and intelligent regulation effect of college physical education curriculum.*

KEYWORDS: *Reinforcement learning; Smart physical education course; Adaptive teaching; Optimization of Teaching Strategies*

1 Introduction

In recent years, with the continuous advancement of digital construction of physical education courses in colleges and universities, smart physical teaching has obtained new development space in curriculum organization, exercise monitoring, process evaluation and learning support. Traditional college physical education courses rely on teachers 'classroom observation, unified teaching progress and stage performance evaluation. Although it can maintain the basic teaching order, it is difficult to accurately identify the differences in students 'physical ability foundation, movement mastery, exercise load tolerance and classroom participation. Especially in college physical education courses, there are obvious differences in students' physical fitness, sports interest, training experience and health status. If we continue to adopt a relatively fixed teaching program, it is easy to have some students 'exercise load insufficient, skills improvement slow, or some students 'training intensity too high, learning experience decline and other problems. It can be seen that how to dynamically adjust teaching content, exercise intensity and feedback method according to students 'real-time status has become an important direction of smart physical education curriculum optimization in colleges and universities.

With the development of wearable devices, motion sensors, campus sports management

*lizhenga2007045@163.com
https://doi.org/10.65102/is2026982 Published: 30 April, 2026

platform and classroom behavior collection technology, the multi-source data acquisition in the process of college physical education gradually has realistic conditions. For example, by collecting data such as heart rate changes, movement trajectories, movement completion quality, exercise times, classroom interaction records and learning feedback, students 'exercise performance and learning state can be more comprehensively described. These data can not only support teachers to judge students' current physical load and skill mastery level, but also provide a basis for computer algorithms to participate in the optimization of teaching strategies. As an intelligent decision-making method with state perception, action selection and reward feedback as the core, reinforcement learning can constantly revise the strategy in the continuous teaching process, so that the smart physical education curriculum turns from static arrangement to dynamic adaptation.

Based on the above background, this study focuses on the optimization problem of adaptive teaching strategies for smart physical education courses in colleges and universities, and constructs a reinforcement learning driven teaching decision-making framework. The framework takes students' physical ability level, sports performance and classroom participation status as input, and the action space is configured with teaching content selection, exercise load adjustment and feedback mode. The reward function is constructed by indicators such as learning effect, participation quality and safety load, so as to realize the continuous optimization of teaching strategies. The purpose of this study is to improve the personalization degree of college physical education courses, the accuracy of classroom control and the learning effect of students 'sports, and to provide computational, adjustable and verifiable method support for the construction of intelligent physical education system.

The main contributions of this research are reflected in the following aspects: (1) a multi-source data representation method for smart sports courses in colleges and universities is constructed, and student movement data, classroom behavior data and learning feedback data are integrated into a unified analysis framework to improve the integrity of student state recognition. (2) Design the state, action and reward mechanism in reinforcement learning environment, so that teaching content, training load and immediate feedback can be dynamically adjusted according to student performance; (3) Through ablation experiments, state recognition results analysis, comparison of different teaching strategies and satisfaction evaluation, the effectiveness of adaptive teaching strategies in improving sports performance, enhancing classroom participation and improving learning experience was verified. This research not only expands the application scenario of reinforcement learning in college physical education, but also provides a new practical path for smart physical education curriculum from experience regulation to data-driven optimization.

2 Related Research

In recent years, the application of reinforcement learning in the field of educational intelligence has gradually attracted attention, and related research mainly focuses on learning path recommendation, task sequence optimization, feedback generation and personalized intervention. Memarian and Doleck[1] gave a scope review of the application of reinforcement learning in education scenarios, and pointed out that reinforcement learning could provide dynamic decision support for personalized learning through learner state perception and strategy iteration. Vassoyan et al. [2] combined graph neural network with reinforcement learning to build a scalable adaptive learning model, indicating that the knowledge association and behavioral differences between learners can be transformed into graph structures that can be processed by computers. Azhar et al. [3] carried out research on problem ranking, and used reinforcement learning to optimize the order of learning task push, so that the system could adjust the subsequent learning content according to the students 'answer performance. Condor and Pardos[4] proposed the formative feedback method supported by deep reinforcement learning, emphasizing that feedback is not a single result output, but a strategy selection based on continuous changes in the learning process. Schutt et al. [5] focused on the dynamic adjustment of the difficulty of the intelligent tutoring system under the condition of small sample size, indicating that reinforcement learning still has certain adaptability in the teaching scene with limited sample size. The above research provides method inspiration for smart physical education courses in colleges and universities, but most of the results still focus on knowledge learning, online testing and intelligent tutoring systems, and lack of attention to physical data such as exercise load, action performance and classroom participation in physical education courses.

In the aspect of personalized learning and smart physical education, existing research has laid a foundation for the optimization of adaptive teaching strategies. Al-Badi and Khan[6] analyzed the perception of learners and teachers on artificial intelligence personalized learning, and pointed out that the effective application of intelligent systems needs to take into account learning experience, teacher acceptance and data interpretability. Fariani et al. [7] conducted a systematic review on personalized learning in higher education, and proposed that curriculum resources, learning rhythm, evaluation feedback and learning support can all be the objects of personalized adjustment. Molenaar[8] emphasized the importance of human-computer collaborative learning technology, and believed that teacher judgment and intelligent algorithm should form a complementary relationship, rather than the system completely replacing teacher decision-making. Jastrow et al. [9] sorted out the research on digital technology in physical education and found that wearable devices, video analysis and mobile platforms can enhance the data management ability of physical classroom. Wallace et al. [10] further analyzed the application of digital technology in physical education from the perspective of teachers and students, and pointed out that teachers 'digital ability directly affected the depth of integrating technology into the classroom. Knoke et al. [11] focus on the role of digital media in sports health promotion, and believe that data feedback can help students understand the connection between sports process and health status. Perez-Munoz et al. [12] and Putranto et al. [13] respectively discussed the application value of virtual reality, augmented reality and mixed reality technology in physical education and training, which provided a new implementation way for action learning and immersive teaching. Daoudi[14] 's research on learning analytics for serious games also shows that process data can be used to identify learning engagement and task completion quality. Although these achievements have expanded the field of digital research in physical education, the existing research still remains more at the level of data collection, resource display and result evaluation, and lacks an adaptive optimization framework that unifies multi-source data recognition, teaching action selection and reward feedback mechanism.

At the same time, the application research of artificial intelligence in higher education also suggests that smart physical education courses cannot only pursue algorithm performance, but also need to pay attention to teacher trust, explanation mechanism and teaching controllability. Celik et al. [15] pointed out that although the application of artificial intelligence in education can improve the efficiency of teaching analysis, it will also bring problems such as the reconstruction of teachers' roles, differences in technical understanding and ethical risks. Crompton and Burke[16] combed the development status of artificial intelligence in higher education, and found that learning prediction, resource recommendation and automatic evaluation were the most widely used directions. Nazaretsky et al. [17] proposed from the perspective of teacher trust that only when teachers understand the judgment basis of the system can they adopt intelligent suggestions in real teaching. Darvishi et al. [18] study on neurophysiological measurement of higher education shows that physiological data can be used to describe the learning state in a more detailed way. Fahd et al. [19] pointed out through meta-analysis that machine learning has a good application effect in college students' performance prediction, risk identification and learning support. Kizilcec[20] further stressed that the promotion of artificial intelligence education application needs to return to the actual needs of teachers. To further clarify the relationship between the existing research and the research in this paper, the key elements of the related research are now compared, as detailed in Table 1.

Table 1: Comparison of key elements of existing related studies

Research direction	Representative literature	Main content	Advantages	Limitations
Educational application of reinforcement learning	Memarian and Doleck [1]	Reviews the application scenarios of reinforcement learning in education	Clarifies the theoretical value of reinforcement learning in supporting personalized learning	Pays limited attention to physical data and classroom movement scenarios in physical education courses
Adaptive learning model	Vassoyan et al. [2]	Combines graph neural networks with reinforcement learning to optimize learning paths	Can process associations between learner relationships and knowledge structures	More suitable for knowledge learning scenarios, with insufficient modeling of sports performance
Learning task sequence optimization	Azhar et al. [3]	Uses reinforcement learning to optimize the sequence of question delivery	Can adjust subsequent tasks according to learning outcomes	The action space mainly focuses on question selection, making it difficult to cover physical education load regulation
Intelligent feedback generation	Condor and Pardos [4]	Generates formative feedback through deep reinforcement learning	Emphasizes the continuity and dynamic nature of feedback	Provides limited support for offline classroom behavior and movement process feedback
Digital technology in physical education	Jastrow et al. [9], Wallace et al. [10]	Studies the application of digital technology in physical education	Supports classroom data collection, movement monitoring, and teaching management	Most studies focus on technology use and lack strategy optimization algorithms
Smart physical education and health promotion	Knoke et al. [11]	Explores the role of digital media in promoting physical education and health education	Enhances students' understanding of exercise data	Insufficiently integrated with reinforcement learning decision-making mechanisms

It can be seen from Table 1 that the existing research has promoted the development of educational intelligence and physical teaching digitalization from different perspectives, but most of the results still stay at the level of single function improvement, such as learning path recommendation, sports data collection, virtual teaching environment construction or result evaluation analysis. For the course of smart physical education in colleges and universities, there is a continuous correlation between student state recognition, teaching action selection and teaching effect feedback, and it is difficult to complete the classroom process regulation by only relying on static evaluation or single recommendation. Based on this, this paper introduces reinforcement learning into the optimization of adaptive teaching strategies of smart physical education courses in colleges and universities. Through the unified design of state, action and reward mechanism, the teaching content, exercise load and feedback method can be dynamically adjusted with the change of students 'performance, so as to make up for the shortcomings of existing research in real-time decision-making and closed-loop optimization of teaching strategies.

3 Adaptive teaching framework of smart physical education course in colleges and universities driven by reinforcement learning

Aiming at the problems of insufficient identification of student differences, lag of teaching load adjustment, and lack of continuous optimization of classroom feedback in smart physical education courses in colleges and universities, this study constructs an adaptive teaching framework driven by reinforcement learning. Based on the multi-source data of physical education classroom, the framework transforms students' physical ability level, sports performance, classroom participation and learning feedback into computable state vectors, and selects the corresponding teaching content, exercise intensity and feedback method through the reinforcement learning model in the teaching process. Different from the way of relying solely on after-class performance evaluation, this framework emphasizes the continuous access of classroom process data, so that the system can update the teaching strategy in time when the students 'status changes, and then improve the individual adaptation ability and teaching control accuracy of the smart physical education course.

The overall framework is composed of data collection layer, state representation layer, strategy decision-making layer and teaching feedback layer. The data collection layer is responsible for accessing wearable devices, motion recognition terminals, classroom management platforms, student feedback questionnaires and other information. The state representation layer cleaned, synchronized, standardized and fused the data from different sources to form the student state description. The policy decision-making layer took the reinforcement learning algorithm as the core, and output the teaching action according to the current state. The teaching feedback layer calculates the reward signal according to the student's subsequent performance, and updates the policy model in the opposite direction. See Figure 1 for the detailed process.

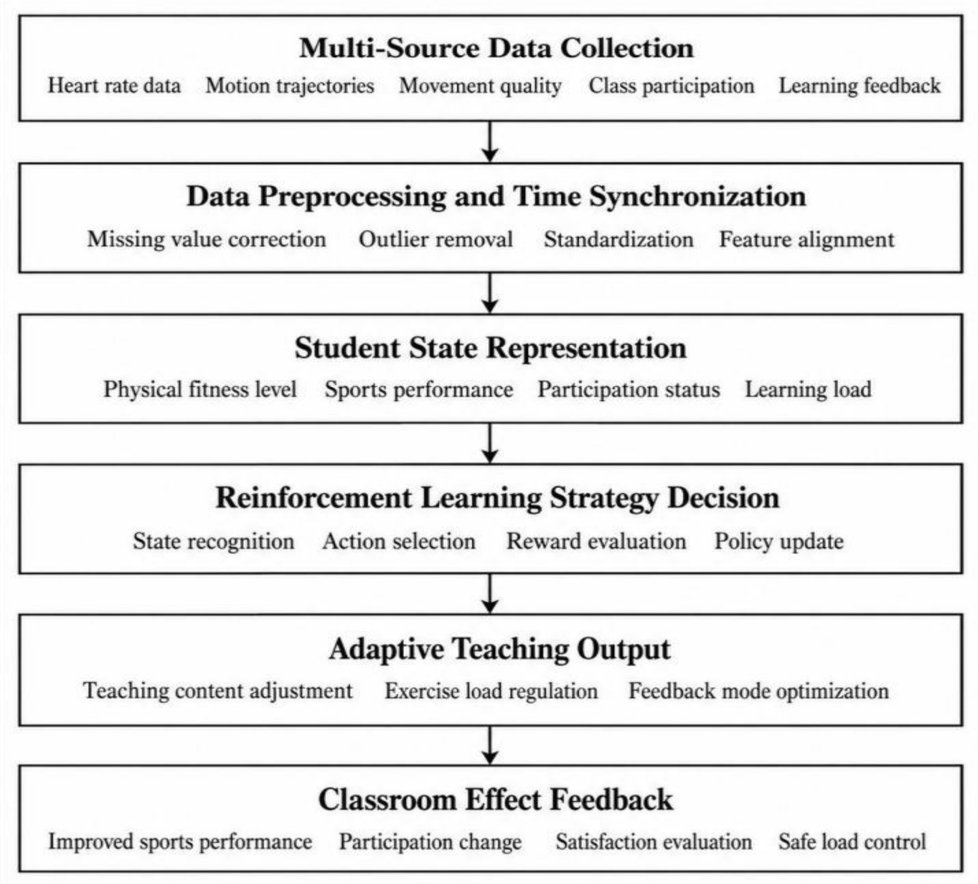

Figure 1: Adaptive teaching framework of smart physical education courses in colleges and universities driven by reinforcement learning

3.1 Multi-source data collection and student state representation of smart Physical Education Course

The adaptive optimization of smart physical education course relies on stable, complete and teaching meaningful data input. The college physical education classroom is different from the general online learning scene, and the students' status is not only reflected in the learning record, but also reflected in the exercise intensity, action completion quality, physical reaction and classroom participation behavior. To improve the accuracy of state recognition, data were collected from four dimensions: physical monitoring, sports performance, classroom participation, and subjective feedback. Physical monitoring data mainly came from smart bracelets, heart rate bands and sports watches, including heart rate, exercise duration, energy consumption and recovery heart rate and other indicators. Sports performance data were obtained through video recognition, inertial sensors and teaching platform records, including movement amplitude, speed change, number of exercises, completion rate and proportion of wrong movements. Classroom participation data were obtained from check-in records, interaction times, task completion and group collaboration performance. Subjective feedback data were collected through an after-class scale, including fatigue, level of interest, perceived learning difficulty, and course satisfaction.

In order to ensure that different data can serve the same teaching decision-making process, the system sets a unified student number, course number and time stamp in the collection stage. The data of each student in a physical education class is divided into several time Windows, and each window corresponds to a class segment, such as warm-up, skill explanation, group practice, physical training and relaxation. Through this division method, the system can place the body reaction, action performance and learning behavior in the specific teaching link to analyze, and avoid judging students 'learning status only by the total score. Table 2 shows the main data types and indicators of student state representation in the smart physical education course.

Table 2: Data types of student state representation in smart physical Education course

Data type	Collection source	Main indicators	Teaching implication
Physical fitness monitoring data	Smart wristbands, heart-rate belts, sports watches	Heart rate, recovery heart rate, energy expenditure, exercise duration	Determines students' exercise load tolerance and physical adaptation status
Sports performance data	Video recognition, inertial sensors, classroom recording platform	Action completion rate, movement amplitude, speed, practice frequency, error rate	Identifies students' skill mastery level and movement stability
Classroom participation data	Smart teaching platform, attendance system, teacher records	Attendance, interaction frequency, task completion, group collaboration performance	Reflects students' participation depth and classroom engagement
Learning feedback data	Post-class questionnaires, mobile feedback forms	Fatigue, interest level, perceived difficulty, satisfaction	Determines whether teaching strategies fit students' subjective learning experience
Environment and course data	Venue equipment, course management system	Venue conditions, course type, practice density, teaching stage	Helps explain differences in sports performance and changes in exercise load

According to Table 2, student status is not a result that can be completely described by a single indicator, but a dynamic feature formed by the joint action of multiple types of data. Increased heart rate may represent increased exercise load or may be related to nervousness, ambient temperature, or continuous practice density; Decreased movement completion rates may stem from inadequate skill understanding or from fatigue accumulation. If only one kind of data is analyzed, the system is easy to produce deviation judgment. Based on this, this study fuses multi-source data to map different indicators uniformly into the student state space.

In the data preprocessing stage, the system removes outliers, fills missing values, smoothen noise and standardises the original data. For continuous variables such as heart rate, speed and range of action, the standardization method is used to eliminate dimensional differences, and the calculation method is shown in Formula (1):

$$x_i' = \frac{x_i - \mu}{\sigma} \tag{1}$$

where, x_i represents the ith original observation, μ represents the mean of the index in the sample, σ represents the standard deviation, and x_i' represents the normalized feature value.

Through this processing, data of different dimensions can enter the same state vector, which facilitates the reinforcement learning model to calculate the state difference and policy payoff. After normalization, the system synchronized the multi-source data according to the classroom time window. Let the state vector of the student in the t-th time window be S_t, whose state representation can be defined as follows.

$$S_t = [P_t, M_t, E_t, F_t, C_t] \tag{2}$$

where, P_t represents physical ability monitoring features, M_t represents sports performance features, E_t represents classroom participation features, F_t represents learning feedback features, and C_t represents curriculum and environment features. The state vector contains not only the student's body response, but also the learning behavior and teaching situation, which can provide a more complete input basis for the subsequent reinforcement learning action selection.

3.2 Reinforcement Learning State, Action and Reward Design for Adaptive Teaching

In this study, the adaptive teaching process of smart physical education courses in colleges and universities was modeled as a Markov decision process. The system receives students' status information in each teaching time window, selects teaching strategies according to the current course stage and historical learning performance, and obtains feedback signals according to students' subsequent movement performance, changes in classroom participation and load safety. The teaching strategy is no longer regarded as a fixed arrangement, but is transformed into a decision problem that can be iteratively optimized by the computer, so that the teaching content, exercise intensity and feedback method can be adjusted with the change of students 'state. Let the decision-making process of smart physical education course in the t-th teaching segment be as follows.

$$M = (S, A, P, R, \gamma) \tag{3}$$

where, S represents the student state space, A represents the teaching action space, P represents the state transition probability, R represents the reward function, and γ represents the discount factor. The state space is composed of students' physical ability level, sports performance, classroom participation, learning feedback and curriculum environment. Combined with the state representation results in Section 3.1, the state of the student at the TTH time window can be further expressed as follows.

$$S_t = [p_t, m_t, e_t, f_t, c_t] \tag{4}$$

where, p_t represents physical characteristics such as heart rate, energy expenditure and recovery heart rate. m_t represents the motor performance characteristics such as movement completion rate, movement amplitude, number of exercises and error rate. e_t represents classroom participation characteristics such as attendance, interaction, task completion and collaboration performance. f_t represents the feedback characteristics such as fatigue, interest, difficulty perception and satisfaction. c_t represents contextual characteristics such as course type, teaching phase, site condition, and practice density. The state design can avoid evaluating students 'learning state simply by grades or attendance, and make the model input closer to the real process of physical education classroom.

Action space is used to describe the instructional control strategies that the system can output. The teaching action in the smart physical education course in colleges and universities is not a single instruction, but is composed of teaching content, exercise load and feedback method. Let the action vector for the TTH time window be as follows.

$$A_t = [a_t^{content}, a_t^{load}, a_t^{feedback}] \tag{5}$$

where, $a_t^{content}$ represents the adjustment of teaching content, including basic movement review, advanced skill practice, combined task training and group collaboration tasks, etc. a_t^{load} indicates exercise load adjustment, including the number of exercise groups, duration, interval duration, and intensity level. $a_t^{feedback}$ indicates the choice of feedback mode, including action error correction, rhythm prompt, load reminder, encouraging feedback and individual suggestions. For students with good physical status and stable action completion, the system can improve the task complexity. For students with slower heart rate recovery or elevated error rates, the system can reduce the practice density and increase the technical breakdown feedback.

Reward function is the key of reinforcement learning model for policy optimization. In this study, learning effect, participation status and safety load were incorporated into the unified reward calculation to avoid the model only pursuing sports performance improvement and neglecting classroom experience and physical safety. The reward function is defined as: The model input is closer to the real process of physical education class.

Action space is used to describe the instructional control strategies that the system can output. The teaching action in the smart physical education course in colleges and universities is not a single instruction, but is composed of teaching content, exercise load and feedback method. Let the action vector for the TTH time window be as follows.

$$R_t = \alpha G_t + \beta E_t - \lambda L_t - \eta Q_t \tag{6}$$

where, G_t represents the improvement degree of students 'exercise performance, E_t represents the improvement degree of classroom participation, L_t represents the degree of exercise load deviation from the safe interval, Q_t represents the degree of action error or fatigue feedback accumulation, α, β, λ, η are weight coefficients. The design can guide the system to improve the learning effect and control the movement risk at the same time, so that the adaptive teaching strategy has stronger educational adaptability. In the process of policy update, the system selects an action according to the current state, and updates the action value after getting the reward. The update can be expressed as follows.

$$Q(S_t, A_t) = R_t + \gamma \max Q\,(S_{t+1}, A_{t+1}) \tag{7}$$

where, $Q(S_t, A_t)$ represents the value of selecting a certain teaching action in the current state, and S_{t+1} represents the new state formed after the action is performed. Through continuous iteration, the model can gradually identify the matching relationship between different student states and teaching actions, so as to form the strategy optimization ability oriented to the classroom process.

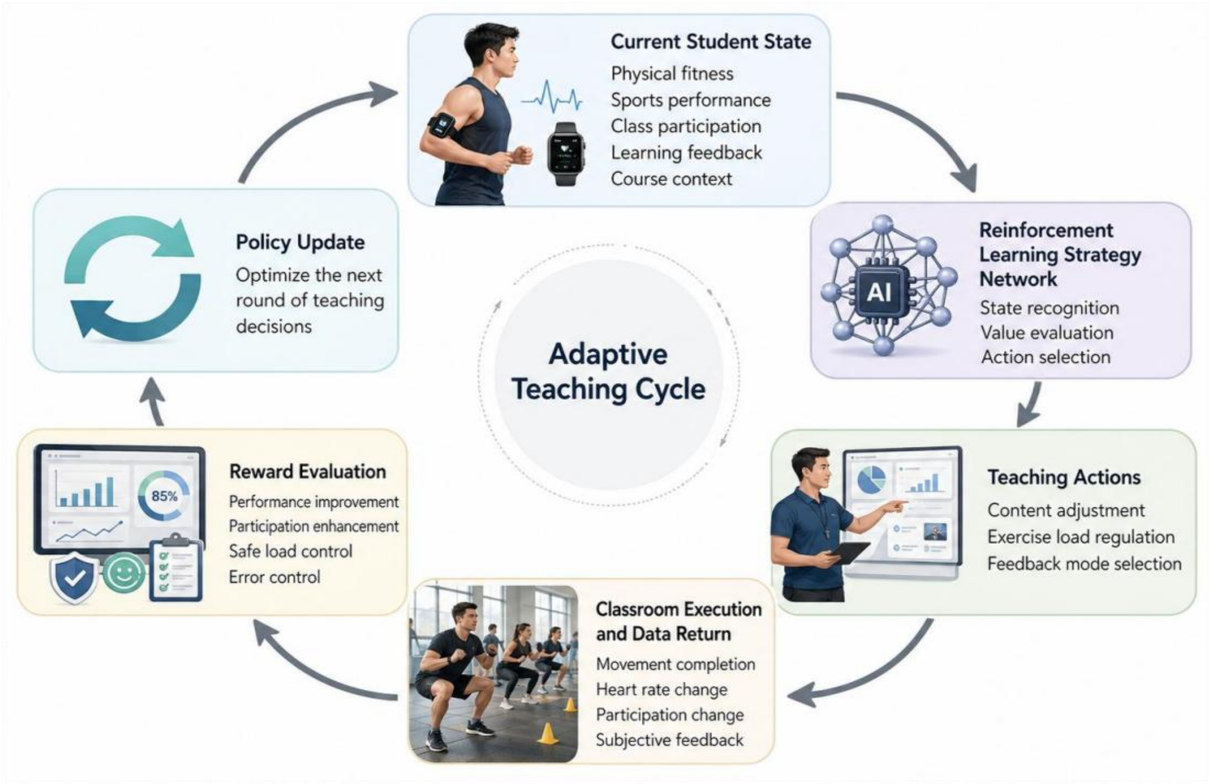

Figure 2: Reinforcement learning state, action, and reward design process

Figure 2 shows the operation logic of reinforcement learning in the smart physical education course in colleges and universities. The system does not directly replace teachers, but converts the changes of students 'state into computable basis and provides dynamic regulation suggestions for teachers. Teachers can confirm, modify or delay the system output according to the actual classroom situation, so that the algorithm decision and teaching experience form a synergy. In this way, the advantages of reinforcement learning in continuous decision making can be exploited, and the educational objectives, safety requirements, and classroom order of the physical education curriculum can not be weakened.

4 Adaptive teaching strategy optimization method of smart physical education course in colleges and universities

4.1 Identification of students' physical ability level, sports performance and classroom participation status

Student state recognition is the basis of adaptive teaching strategy optimization of smart physical education course in colleges and universities. The system takes wearable devices, sports recognition terminals and smart teaching platforms as data entry, takes heart rate, exercise duration, energy consumption, recovery heart rate and other indicators as physical ability level input, and takes action completion rate, action amplitude, speed change, practice times and proportion of wrong actions as exercise performance input. At the same time, attendance, number of interactions, task completion, group collaboration records, and after-class feedback were used as classroom participation inputs. In order to ensure that the recognition results can truly reflect the classroom process, the system firstly performs anomaly elimination, missing completion, time window alignment and standardization on the original data, and then integrates the data from different sources into a unified feature vector. The state recognition process is shown in Figure 3.

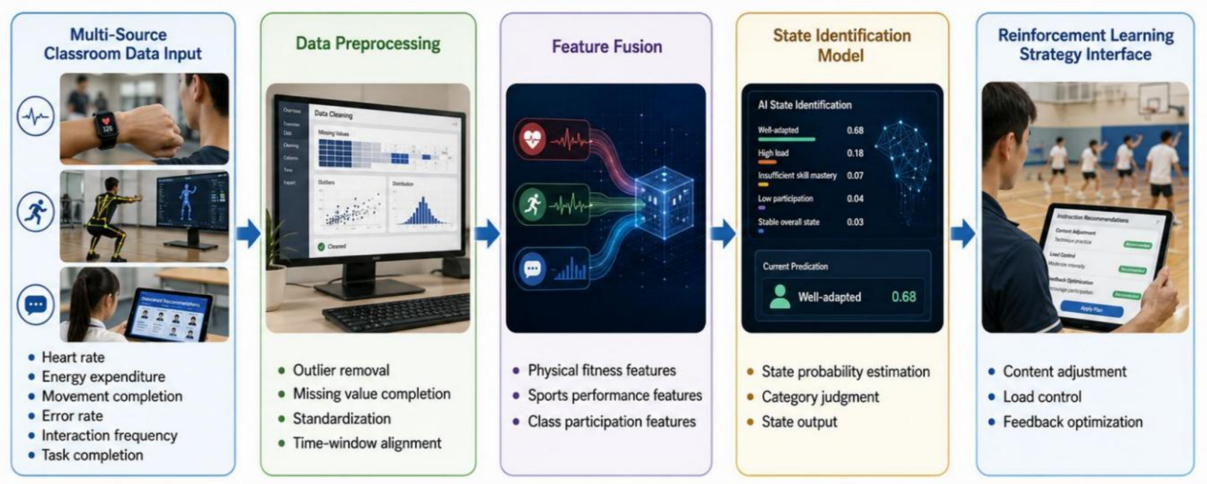

Figure 3: Identification process of students' physical ability level, sports performance and classroom participation status

In this process, the system does not judge the state of students according to a single sports score, but comprehensively analyzes the body reaction, action quality and classroom engagement. Let the fused feature vector of students in the TTH teaching time window be X_t, and the state recognition result can be expressed as follows.

$$\hat{Y}_t = \mathrm{softmax}(WX_t + b) \tag{8}$$

where, X_t represents the fusion feature consisting of physical ability, athletic performance, and classroom participation data, W represents the model weight matrix, b represents the bias term, and $\hat{Y}_t$ represents the predicted probability of the student status category. The status category mainly includes the types of good physical adaptation, high load, insufficient skill mastery, insufficient participation, and comprehensive stability. The system determines the current student state according to the maximum value of the predicted probability, and passes the recognition results to the reinforcement learning module, which provides a basis for the subsequent teaching content selection, exercise load adjustment and feedback mode optimization.

In order to evaluate the effect of state recognition, this paper uses accuracy and F1 value as the main indicators. Accuracy is used to measure the overall accuracy of the model, and is calculated as follows.

$$\mathrm{Acc} = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

The F1 score is used to measure precision and recall together and is calculated as follows.

$$F1 = \frac{2PR}{P + R} \tag{10}$$

where, TP represents the correctly identified positive class samples, TN represents the correctly identified negative class samples, FP represents the misjudged samples, FN represents the missed judged samples, P represents the precision rate and R represents the recall rate. The accuracy rate can reflect the overall recognition level of the model, and the F1 value can reduce the influence of the unbalanced distribution of categories on the evaluation results, which is more suitable for smart sports classroom scenarios with more types of student states.

With the above recognition method, the system can form a more refined judgment of student differences. When the student's heart rate was in a reasonable interval, the action completion rate was high and the task completion degree was stable, the model could be identified as a comprehensive stable state. When students have slow recovery of heart rate, increased error rate of movement and enhanced fatigue feedback, the model can be identified as a state of high load. When the interaction frequency of students is low, the task completion is insufficient and the physical indicators are normal, the model can judge that there is a problem of insufficient participation. Therefore, student state recognition no longer stays at the after-class statistical level, but is transformed into the pre-input of reinforcement learning strategy optimization, so that the smart physical education course in colleges and universities can complete more accurate teaching regulation according to students 'real-time performance.

4.2 Dynamic optimization of teaching content, exercise load and feedback strategy based on reinforcement learning

DDPG algorithm is suitable for policy optimization problems in continuous action space, and can establish a stable decision mapping between state input and teaching feedback. The teaching regulation of smart physical education courses in colleges and universities is not a simple choice of a fixed scheme, but a continuous combination of teaching content, exercise load and feedback methods. Based on the student state recognition results obtained in Section 4.1, the system takes the physical ability level, sports performance, classroom participation and subjective feedback as the input of the environment state, generates the teaching action by the Actor network, and evaluates the value of the action in the current classroom state by the Critic network, so as to complete the dynamic optimization of the adaptive teaching strategy. The specific structure is shown in Figure 4.

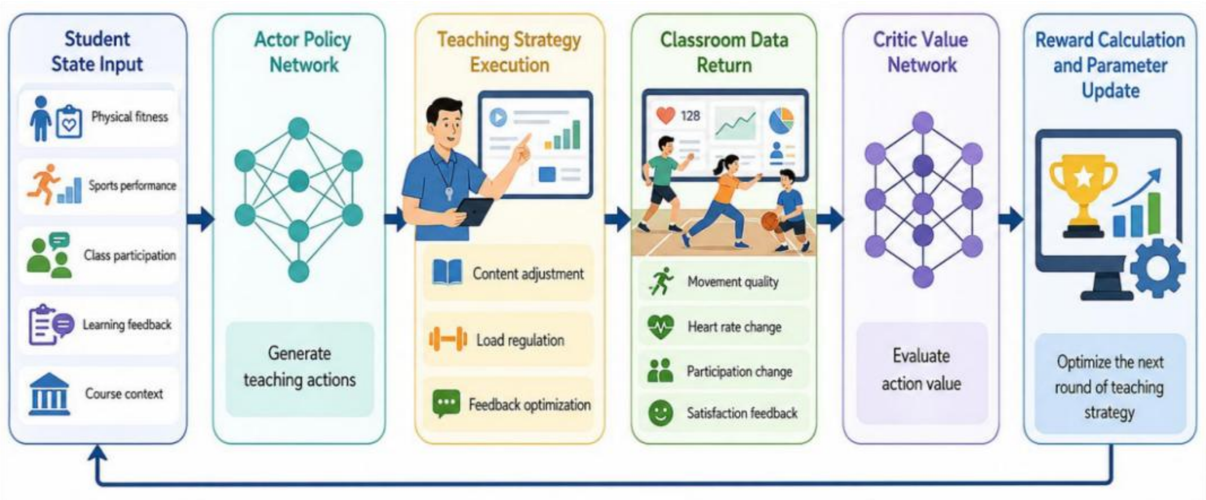

Figure 4: Optimization structure of teaching strategy of smart physical education course in colleges and universities based on DDPG

In this structure, the state space is composed of students 'multi-source features, and the action space corresponds to the teaching control content that can be implemented by teachers. The teaching content actions included basic action review, special skill strengthening, combination exercises and collaborative tasks. Exercise load action includes exercise duration, group number, interval time, intensity level and exercise density. Feedback actions include immediate error correction, action demonstration, rhythm prompt, load reminder and individual suggestions. In order to ensure the safety of teaching, the exercise intensity can be limited to 50%-85% of the maximum heart rate, the single round of exercise time can be set to 5-20 minutes, and the feedback frequency can be adjusted according to the error rate and participation of students. DDPG updates the policy through the objective value function, which is calculated as follows.

$$y_i = r_i + \gamma Q'(S_{i+1}, \mu'(S_{i+1})|\theta^{Q'}) \tag{11}$$

where, y_i represents the target value, r_i represents the reward obtained by the current teaching action, γ represents the discount factor, Q' represents the target Critic network, μ' represents the target Actor network, and S_{i+1} represents the new state formed after performing the teaching action. The Critic network updates its parameters by comparing the error between the predicted value and the target value, and its loss function is as follows.

$$L = \frac{1}{N}\sum_{i=1}^{N}(y_i - Q(S_i, A_i|\theta^Q))^2 \tag{12}$$

where N represents the number of samples and $Q(S_i, A_i|\theta^Q)$ represents the action value output by the current Critic network. The Actor network adjusts the direction of action generation according to the value gradient given by the Critic network, so that the system gradually selects the teaching strategy that is more conducive to the development of students. The reward function comprehensively considers the improvement of sports performance, improvement of classroom participation, load safety and feedback satisfaction, which is specifically expressed as follows.

$$R_t = w_1 G_t + w_2 E_t + w_3 F_t - w_4 L_t - w_5 M_t \tag{13}$$

where, G_t represents the improvement degree of students' action completion quality, E_t represents the change of classroom participation level, F_t represents students' satisfaction degree with the feedback method, L_t represents the penalty term when the exercise load exceeds the safe range, M_t represents the action error rate or the cumulative degree of fatigue, and w_1 to w_5 are the weight coefficients. The reward design can avoid the system simply pursuing sports performance, ignoring students' physical endurance and classroom learning experience. Through the above method, the system can output differentiated teaching plans according to different student states. For students with good physical adaptability and high movement completion rate, the model can improve the difficulty of the task and increase the proportion of combined exercises. For students with high load or high action error rate, the model can reduce the practice density, extend the interval time, and increase the action decomposition feedback. For students with insufficient participation but stable physical indicators, the model can be adjusted to a collaborative task or a goal challenge task to improve the level of classroom engagement.

5 Experimental results and analysis

5.1 Ablation experiments

In order to verify the effectiveness of each module in the adaptive teaching strategy optimization method of smart physical education course in colleges and universities, this study set up ablation experiments. The experimental data came from the teaching process records of smart sports course, and the samples included students 'physical ability monitoring data, sports performance data, classroom participation data and learning feedback data. All models are run on the same training and test sets, and the evaluation metrics include state recognition accuracy, F1 value, average reward value, and load safety control rate. Among them, the accuracy rate of state recognition is used to measure the judgment effect of the model on students 'state categories, the F1 value is used to evaluate the stability of multi-category state recognition, the average reward value reflects the comprehensive income after the optimization of teaching strategy, and the load safety control rate is used to judge the ability of the system to regulate students 'exercise intensity. The results of ablation experiments are shown in Table 3.

Table 3: Ablation experimental results of adaptive teaching strategy optimization model

Model configuration	State recognition accuracy/%	F1 score	Average reward value	Load safety control rate/%
Complete model (multi-source state representation + DDPG policy optimization)	94.6	0.941	0.873	96.2
Without physical fitness monitoring features	89.3	0.884	0.791	88.7
Without sports performance features	87.8	0.869	0.764	90.1
Without classroom participation features	90.5	0.897	0.806	92.4
Without learning feedback features	91.2	0.904	0.821	93.0
Without DDPG policy optimization module	84.9	0.836	0.692	86.5

It can be seen from Table 3 that the complete model has achieved better results in the four indicators, the state recognition accuracy reaches 94.6%, the F1 value reaches 0.941, the average reward value is 0.873, and the load safety control rate reaches 96.2%. This shows that after the combination of multi-source data representation and DDPG strategy optimization, the system can more accurately identify the student state, and output more appropriate teaching content, exercise load and feedback method according to the state change.

After removing the physical monitoring feature, the load safety control rate decreased to 88.7%, indicating that the indicators such as heart rate, energy expenditure and recovery heart rate had a direct impact on exercise intensity regulation. After removing the motion performance features, the accuracy of state recognition decreased to 87.8%, indicating that the action completion rate, error rate and practice quality were important basis for judging students 'skill mastery level. After removing the classroom participation feature, the average reward value dropped to 0.806, indicating that the participation state would affect the adaptation effect of teaching strategies on students 'learning engagement. After removing the DDPG strategy optimization module, the indicators dropped most obviously, and the average reward value was only 0.692, indicating that static rules were difficult to continuously adjust teaching strategies according to classroom feedback. The overall results show that each module has a supporting effect on the performance of the model, and the DDPG strategy update mechanism and sports performance characteristics have a more prominent influence on the adaptive teaching optimization.

5.2 Analysis of the identification results of students 'movement performance and classroom participation status

In order to evaluate the ability of the model to identify students 'sports performance and classroom participation status in the smart sports classroom in colleges and universities, this study selected four typical teaching scenarios of skill practice, physical training, group collaboration and comprehensive test for experiments. The same data preprocessing method and state recognition model are used in each scenario, and the evaluation metrics include recognition accuracy, recall and F1 value. The experimental results are shown in Figure 5.

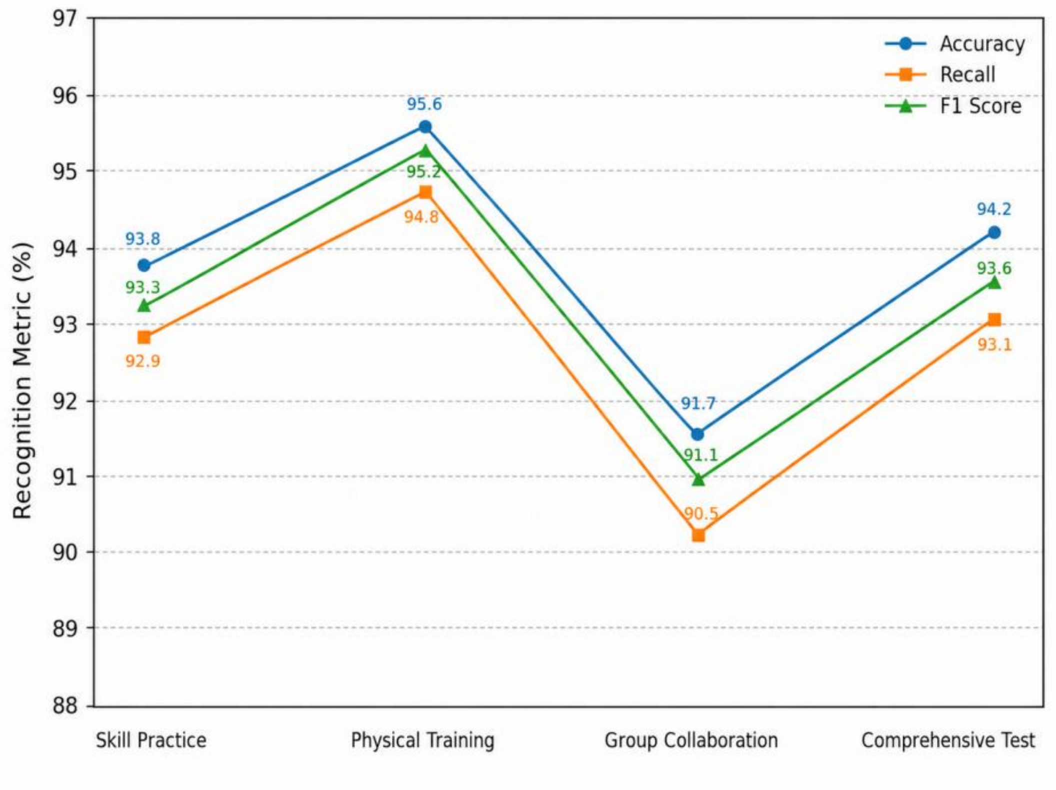

Figure 5: Student state recognition results under different teaching scenarios

Figure 5 shows that the model maintains a high recognition level in all four types of teaching scenarios. Among them, the recognition effect of physical training scene is the best, with an accuracy of 95.6% and an F1 value of 0.952. This is because the indicators such as heart rate, energy consumption, practice density and recovery speed of students change more obviously in the process of physical training, the boundary of data characteristics is clear, and the model can accurately distinguish between moderate load, high load and good physical adaptation. The accuracy of the comprehensive test scene is 94.2%, and the F1 value is 0.936, indicating that the model can still maintain a strong comprehensive judgment ability when the quality of action completion, the continuity of practice and the degree of classroom engagement change together.

The precision rate and recall rate of the skill practice scene are 93.8% and 92.9%, respectively, and the overall performance is relatively stable. In this scenario, students 'action range, error rate and completion rhythm have a great influence on state recognition, and the model can judge students 'skill mastery according to the change of action quality. The recognition result of group collaboration scene is relatively low, with an accuracy of 91.7% and an F1 value of 0.911. The reason is that there are differences in collaborative division of labor between students in grouping tasks. Although the movement data of some students do not change significantly, the interaction frequency and task participation are high, which is easy to cause inconsistency between movement performance and participation state.

5.3 Comparative analysis of course learning effects under different teaching strategies

In order to evaluate the influence of different teaching strategies on the learning effect of smart physical education courses in colleges and universities, this study set up four groups of comparative experiments, which were the traditional unified teaching strategy, the rule recommendation teaching strategy, the state recognition only teaching strategy and the reinforcement learning adaptive teaching strategy. The experimental period was 8 weeks, and the evaluation indexes included motor skill improvement rate, classroom participation improvement rate, exercise load adaptation rate and curriculum comprehensive achievement degree. The students in each group completed the same basic test before the experiment. After the experiment, the comprehensive evaluation was carried out according to the movement completion quality, classroom task records, heart rate load interval and learning feedback results. The comparison results of course learning effects under different teaching strategies are shown in Figure 6.

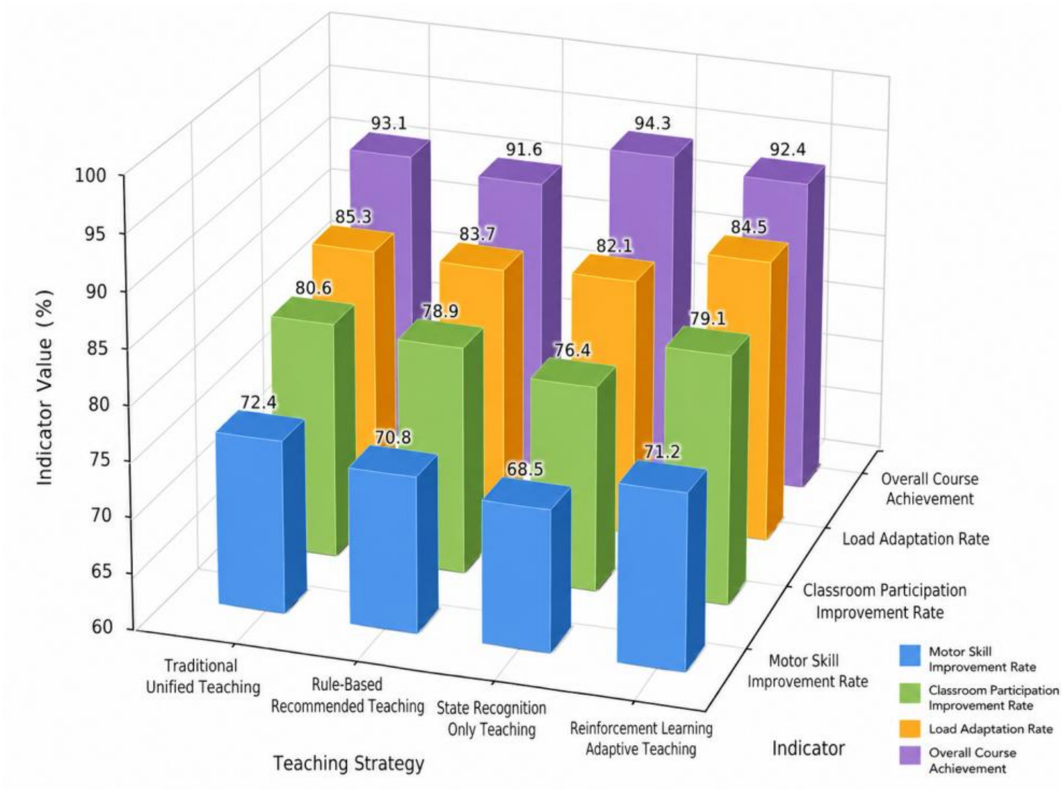

Figure 6: Comparison results of course learning effects under different teaching strategies

Figure 6 shows that the reinforcement learning adaptive teaching strategy achieves the highest results in all four indicators. Among them, the improvement rate of motor skills reached 93.1%, the improvement rate of classroom participation reached 91.6%, the load adaptation rate reached 94.3%, and the comprehensive achievement degree of courses reached

92.4%. This result showed that after combining the students 'state recognition results with the reinforcement learning strategy update mechanism, the system could timely adjust the teaching content and exercise load according to the changes of students 'physical ability, movement mastery level and classroom engagement, so as to make the course learning effect more stable.

The indicators of the traditional unified teaching strategy are relatively low, and the comprehensive achievement rate is 71.2%. The main reason is that the strategy uses the same teaching progress and practice intensity, and it is difficult to take into account the differences in physical ability and skill foundation between students. The teaching strategy of rule recommendation has improved compared with the traditional method, and the comprehensive achievement rate is 79.1%, indicating that the teaching adjustment based on fixed rules can improve some learning effects, but its adjustment basis is relatively limited, and it is difficult to deal with the continuous changes of students 'status in the classroom. The comprehensive achievement degree of teaching strategies only for state recognition reached 84.5%, indicating that multi-source data recognition could enhance teachers 'ability to judge students 'states, but without strategy iteration mechanism, teaching adjustment still relied on manual experience.

5.4 Analysis on the optimization effect of adaptive teaching Strategies and teaching satisfaction

In this study, teaching satisfaction was defined as a comprehensive student evaluation of course load scheduling, timeliness of teaching feedback, adaptability of task difficulty, and classroom participation experience. In order to test the actual effect of the adaptive teaching strategy of reinforcement learning, the experiment compared the traditional unified teaching strategy with the strategy of this paper, and calculated the performance of the two groups of students in terms of teaching strategy matching degree, exercise load adaptation rate, feedback timeliness and comprehensive satisfaction. All indicators are converted using the percentage system, and the results are shown in Figure 7.

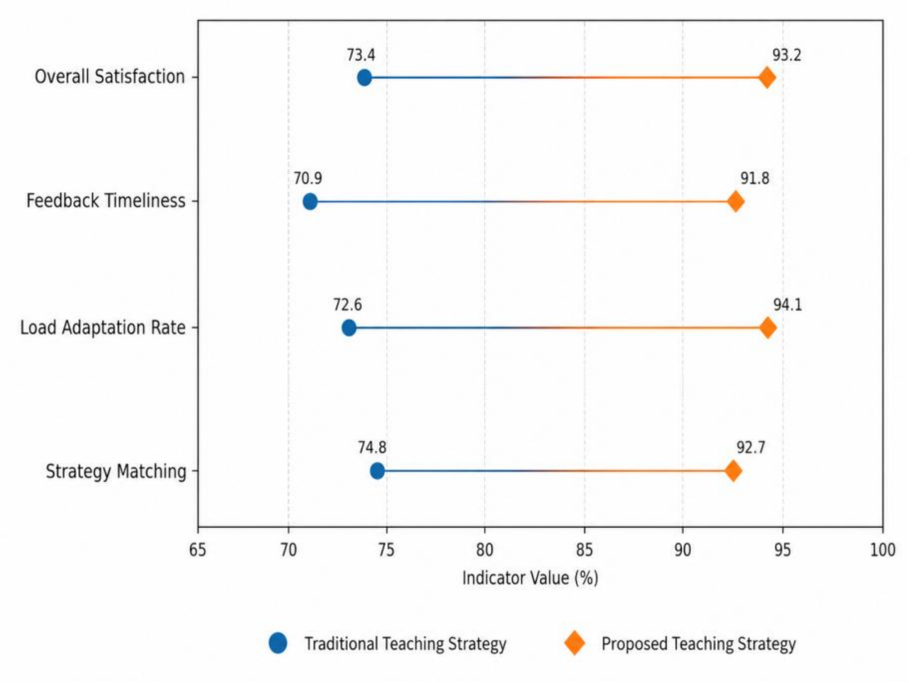

Figure 7: Comparison of optimization effect and satisfaction under different teaching strategies

Figure 7 shows that the proposed strategy is significantly higher than the traditional teaching strategy in all indicators. Among them, the matching degree of teaching strategies increased from 74.8% to 92.7%, indicating that the system could adjust teaching tasks according to students 'physical ability level, action completion quality and classroom participation status, and reduce the adaptation deviation caused by unified teaching progress. The adaptation rate of exercise load increased from 72.6% to 94.1%, indicating that the reinforcement learning model has a better adjustment effect in controlling the exercise intensity, interval time and task density, and can avoid students from repeatedly fluctuating between inefficient practice or excessive load.

The feedback timeliness increased from 70.9% to 91.8%, which reflected that the system could generate feedback suggestions in time when students had increased movement error rate, slow recovery of heart rate or decreased participation. The comprehensive satisfaction rate reached 93.2%, which was 19.8 percentage points higher than that of the traditional teaching strategy, indicating that students had higher recognition of the course rhythm, feedback methods and sports experience. After further paired sample t-test, the difference in comprehensive satisfaction reached a significant level ($t=10.46$, $p<0.001$), and Cohen's d was 2.87, indicating that the difference was not only statistically significant, but also had strong practical value. This shows that the adaptive teaching strategy of reinforcement learning can improve the accuracy of teaching regulation and improve students 'classroom experience at the same time, which provides effective support for the optimization of smart physical education curriculum in colleges and universities.

6 Discussion

The experimental results of this study show that the adaptive teaching strategy of smart physical education course in colleges and universities driven by reinforcement learning shows good application effects in student state recognition, course learning effect and teaching satisfaction. From the ablation experimental results, the state recognition accuracy of the complete model reaches 94.6%, the F1 value is 0.941, the average reward value is 0.873, and the load safety control rate reaches 96.2%. When the DDPG strategy optimization module is removed, the accuracy of state recognition is reduced to 84.9%, the average reward value is reduced to 0.692, and the load safety control rate is reduced to 86.5%. This result shows that the optimization of teaching strategies cannot be fully completed by relying on data recognition alone, and the strategy update mechanism in reinforcement learning plays a key role in the selection of teaching content, the adjustment of exercise load and the configuration of feedback methods.

From the recognition results of students 'sports performance and classroom participation status, the performance of the model in the physical training scene was the best, with an accuracy of 95.6% and an F1 value of 0.952. In the comprehensive test scenario, the accuracy is 94.2%, and the F1 value is 0.936, indicating that the model can better deal with the classroom situation where the quality of action completion, exercise load and participation state change together. In contrast, the accuracy of the group collaboration scenario is 91.7%, and the F1 score is 0.911, which is slightly less effective in recognition. The reason may lie in the different role division of students in the collaborative task, the movement data of some students does not change significantly, but the interactive participation is more, which leads to a certain dislocation between movement performance and classroom participation. It can be seen that the subsequent models still need to strengthen their ability to recognize collaborative behaviors, interaction quality, and task roles.

The comparison results of different teaching strategies further verify the advantages of the proposed method. Under the adaptive teaching strategy of reinforcement learning, the improvement rate of students 'motor skills reached 93.1%, the improvement rate of classroom participation reached 91.6%, the load adaptation rate reached 94.3%, and the comprehensive achievement degree of the course reached 92.4%, which were significantly higher than those of the traditional unified teaching strategy. The comprehensive achievement degree of traditional unified teaching was only 71.2%, and the load adaptation rate was 68.5%, indicating that the fixed progress and unified intensity were difficult to meet students 'differentiated learning needs. The reinforcement learning model can adjust the task difficulty, practice density and feedback form according to the change of students 'state, so as to make the teaching strategy more suitable for students 'actual performance.

In terms of teaching satisfaction, the comprehensive satisfaction of the proposed strategy reached 93.2%, which was higher than that of the traditional teaching strategy (73.4%), with an increase of 19.8 percentage points. Paired sample t-test results show that the difference in comprehensive satisfaction reaches a significant level, $t=10.46$, $p<0.001$, Cohen's d is 2.87, indicating that the difference is not only statistically significant, but also has strong practical value. This result indicates that the adaptive teaching strategy can improve students' acceptance of course pace, exercise load, and teaching feedback.

There are still some limitations in this study. The experimental samples are from specific college physical education courses, and the course types and student bases have a certain range. The applicability of the model in different schools, different sports and different teaching organization methods still needs to be expanded and verified. At the same time, the system depends on the data of wearable devices, video recognition and teaching platform. If there is signal delay, action occlusion or data noise, the state recognition results may be affected. Subsequent research can further expand the sample size, add ball games, physical energy classes and aerobics-dancing course scenes, and optimize the weight of the reward function to improve the interpretability and stability of the model in complex classroom environments.

7 Conclusions

Based on the teaching needs of smart physical education courses in colleges and universities, this paper constructs an adaptive teaching strategy optimization method driven by reinforcement learning. In this study, students' physical ability level, sports performance, classroom participation and learning feedback were incorporated into the unified state representation, and DDPG algorithm was used to dynamically adjust teaching content, exercise load and feedback method. Experimental results show that this method can improve the accuracy of students 'state recognition, improve the learning effect of the course, enhance the level of load adaptation, and improve students 'satisfaction with the smart sports course. The contributions of this research are mainly reflected in three aspects. Firstly, a multi-source data fusion framework for physical education classroom is established. The second is to realize the continuous optimization of teaching strategies; The third is to verify the application value of adaptive teaching in college physical education curriculum. The research is still affected by the sample scope, course type and data collection stability. In the future, the experimental scenarios can be expanded, the reward function can be optimized, and the interpretability of the model can be enhanced, so as to improve the promotion ability of the smart physical teaching system.

Author's Profile

MS YangYan, female, Department of Physical Education, Tianjin Chengjian University Master's candidate, Lecturer Research directions: School Physical Education, Sports Psychology.

Zheng, Li male, Department of Physical Education, Tianjin Chengjian University. Lecturer, bachelor’s academic background and master's degrees. Research directions: School Physical Education and Sports Psychology.

References

[1] Memarian B, Doleck T. A scoping review of reinforcement learning in education[J]. Computers and Education Open, 2024, 6: 100175.

[2] Vassoyan J, Vie J J, Lemberger P. Towards scalable adaptive learning with graph neural networks and reinforcement learning[J]. arXiv preprint arXiv:2305.06398, 2023.

[3] Azhar A Z, Segal A, Gal K. Optimizing Representations and Policies for Question Sequencing Using Reinforcement Learning[J]. International Educational Data Mining Society, 2022.

[4] Condor A, Pardos Z. A deep reinforcement learning approach to automatic formative feedback[C]//Proceedings of the 15th International Conference on Educational Data Mining. 2022: 662.

[5] Schütt A, Huber T, Aslan I, et al. Fast dynamic difficulty adjustment for intelligent tutoring systems with small datasets[J]. 2023.

[6] Al-Badi A, Khan A. Perceptions of learners and instructors towards artificial intelligence in personalized learning[J]. Procedia computer science, 2022, 201: 445-451.

[7] Fariani R I, Junus K, Santoso H B. A systematic literature review on personalised learning in the higher education context[J]. Technology, Knowledge and Learning, 2023, 28(2): 449-476.

[8] Molenaar I. Towards hybrid human-AI learning technologies[J]. European Journal of Education, 2022, 57(4): 632-645.

[9] Jastrow F, Greve S, Thumel M, et al. Digital technology in physical education: a systematic review of research from 2009 to 2020[J]. German Journal of Exercise and Sport Research, 2022, 52(4): 504-528.

[10] Wallace J, Scanlon D, Calderón A. Digital technology and teacher digital competency in physical education: a holistic view of teacher and student perspectives[J]. Curriculum Studies in Health and Physical Education, 2023, 14(3): 271-287.

[11] Knoke C, Woll A, Wagner I. Health promotion in physical education through digital media: A systematic literature review[J]. German Journal of Exercise and Sport Research, 2024, 54(2): 276-290.

[12] Pérez-Muñoz S, Castaño Calle R, Morales Campo P T, et al. A systematic review of the use and effect of virtual reality, augmented reality and mixed reality in physical education[J]. Information, 2024, 15(9): 582.

[13] Putranto J S, Heriyanto J, Achmad S, et al. Implementation of virtual reality technology for sports education and training: Systematic literature review[J]. Procedia Computer Science, 2023, 216: 293-300.

[14] Daoudi I. Learning analytics for enhancing the usability of serious games in formal education: A systematic literature review and research agenda[J]. Education and Information Technologies, 2022, 27(8): 11237-11266.

[15] Celik I, Dindar M, Muukkonen H, et al. The promises and challenges of artificial intelligence for teachers: A systematic review of research[J]. TechTrends, 2022, 66(4): 616-630.

[16] Crompton H, Burke D. Artificial intelligence in higher education: the state of the field[J]. International journal of educational technology in higher education, 2023, 20(1): 1-22.

[17] Nazaretsky T, Ariely M, Cukurova M, et al. Teachers' trust in AI-powered educational technology and a professional development program to improve it[J]. British journal of educational technology, 2022, 53(4): 914-931.

[18] Darvishi A, Khosravi H, Sadiq S, et al. Neurophysiological measurements in higher education: A systematic literature review[J]. International Journal of Artificial Intelligence in Education, 2022, 32(2): 413-453.

[19] Fahd K, Venkatraman S, Miah S J, et al. Application of machine learning in higher education to assess student academic performance, at-risk, and attrition: A meta-analysis of literature[J]. Education and Information Technologies, 2022, 27(3): 3743-3775.

[20] Kizilcec R F. To advance AI use in education, focus on understanding educators[J]. International Journal of Artificial Intelligence in Education, 2024, 34(1): 12-19.

[21] Viberg O, Cukurova M, Feldman-Maggor Y, et al. What explains teachers’ trust in AI in education across six countries?[J]. International Journal of Artificial Intelligence in Education, 2025, 35(3): 1288-1316.

[22] Del Gobbo E, Guarino A, Cafarelli B, et al. Automatic evaluation of open-ended questions for online learning. A systematic mapping[J]. Studies in Educational Evaluation, 2023, 77: 101258.

[23] Manhiça R, Santos A, Cravino J. The use of artificial intelligence in learning management systems in the context of higher education: Systematic literature review[C]//2022 17th Iberian conference on information systems and technologies (CISTI). IEEE, 2022: 1-6.