"Online + Offline" Hybrid Teaching Model in the Post Epidemic Era Based on Deep Reinforcement Learning

Shaolin Liang[1(✉)] and Pei Su[2]

[1] College of Mathematics, Sichuan University of Arts and Science, Dazhou 635000, China
uhdsfd545@sina.com
[2] Wuhan Railway Vocational College of Technology, Wuhan 430000, China

Abstract. In order to achieve students' in-depth understanding of the teaching content, in the post-epidemic era, an "online + offline" hybrid teaching model based on deep reinforcement learning has been designed. First, the basic data is preprocessed to remove interfering data and convert it into a form that can be directly used by the model. In the domain knowledge unit of the model, on the basis of determining the composition of the domain knowledge elements and their associated relationships, a structure in which the superordinate relationship and the subordinate relationship, the predecessor relationship and the successor relationship coexist is constructed; in the learner unit of the model, the deep reinforcement determines Based on the learning source, a block-based data management mechanism is established to jointly promote the operation of the model. The experimental results show that the "Online + offline" hybrid teaching model in the post epidemic era based on deep reinforcement learning has good performance and can achieve good teaching results.

Keywords: Deep reinforcement learning · Post epidemic era · "Online + offline" · Mixed teaching model · Domain knowledge unit · Learner unit

1 Introduction

A sudden outbreak of New Coronavirus disrupted people's normal study, work and life. As an emergency measure during the epidemic prevention and control period, "suspend classes and not stop learning" became a milepost event in the development of online education in the world. Today, the epidemic situation is gradually stable, and schools have resumed classes one after another, marking that online teaching has entered the "post epidemic era". When we re embrace the real classroom again, what has this unprecedented and the world's largest online education changed? What's left for us? Will our class go back to the past? Does the classroom teaching model return to the origin or seize the opportunity to change and transform? What is the future form of education? What direction will education develop in the future? These will become the problems that every educator should think and study. "Epidemic period" Online teaching has become

S.-H. Wang and Y.-D. Zhang (Eds.): ICMTEL 2022, LNICST 446, pp. 112–126, 2022.
https://doi.org/10.1007/978-3-031-18123-8_9

the main way to carry out education. Online teaching is a teaching form in which teachers and students use educational resources to realize interaction on the technical platform. Only by realizing the interaction among teachers, students, resources and technology can we build efficient online teaching. In the large-scale online education practice launched due to the epidemic, the majority of teachers and students from everything in good order and well arranged life, teachers have too much understanding, and students have different experiences. From teachers' and students' perspectives, it is of great significance to summarize the experience and problems of online teaching in the epidemic period. It has far-reaching significance for teachers to study the transformation of classroom teaching mode and promote the integration of online and offline teaching in the era of epidemic [1–3]. With the Internet plus education, With the deepening development of, online teaching based on MOOC, superstar learning link, wisdom tree and other teaching platforms is more and more widely used in colleges and universities. In recent years, the hybrid teaching mode of combining online and offline has gradually become a trend in college teaching [4, 5]. These large-scale open online course platforms not only break the constraints of learning time and place, but also provide learners with rich and diverse learning resources. Students produce a large number of learning behavior data in the process of online learning. By analyzing these learning behavior data, we can explore the learning laws behind the data, and teachers can improve teaching according to these laws. It provides personalized management and teaching for students with different learning situations, which is of great significance to both teachers' teaching and students' learning [6–8]. Based on MOOC and SPOC, the online and offline mixed teaching mode is discussed. The specific implementation steps of SPOC curriculum are described in detail. The design idea and practice process of mixed teaching mode and flipped classroom under line and online are given. The effectiveness and feasibility of the combination of teacher student evaluation and curriculum goal comparison are verified, and the results of questionnaires before and after the course are given. The analysis and reflection provide valuable experience for promoting teaching reform [9]. At present, the analysis and research on the relationship between relevant data in the learning process and teaching effect are basically aimed at the time and energy paid by learners in a certain learning behavior, that is, quantity and quality. These rough data can not accurately and completely reflect the actual situation of teaching, nor can they accurately define and analyze the teaching effect [10]. At the same time, compared with learning participation and learning input, the behavior sequence actually generated by learners in the learning process can better reflect learners' behavior path and cognitive process. With the help of the influence of learners' behavior sequence on learning effect, teachers can determine the key behavior sequence for learning process analysis, so as to monitor learners' learning status and implement teaching intervention in time, the purpose of improving the learning effect is to combine online and offline blended teaching as a teaching mode of university courses. In recent years, it is gradually moving towards campus. Especially under the influence of the New Coronavirus epidemic in 2020, teachers who are not yet ready are learning how to use MOOC, super star learning, intelligent tree, online teaching platform, QQ group classroom, sharing screen and so on. Online and offline hybrid teaching was originally a teaching mode combining the advantages of online teaching and traditional teaching. Through the organic combination of the two teaching organization forms, learners' learning is led from shallow to deep to deep learning. The ultimate purpose of hybrid teaching is not to use the online platform, build digital teaching resources, or renovate teaching activities, but effectively improve the learning depth of most students.

Based on this, this paper designs a "Online + offline" hybrid teaching model in the post epidemic era based on deep reinforcement learning, and evaluates its teaching effect.

2 Data Preprocessing

In order to achieve the purpose of deep reinforcement learning, it is necessary to mine teaching resources and student data. Data preprocessing is an important step in the process of data mining, because the data in the real world is dirty, or incomplete (the attribute of interest has no value), noisy (there are errors or exceptions in the data, that is, data deviating from the expected value) and inconsistent (inconsistent data connotation) [11]. Data must be preprocessed before mining. The preprocessing of model data in this paper includes data preparation, data cleaning and data transformation.

2.1 Data Preparation

Whether data mining is successful or not, data preparation is very important, and it is the premise of realizing the application of data mining. Data preparation contents: first, determine the data source to mine data and collect the original data; second, merge and sort the data from different data sources into the same database. This model takes the original data in Teachers' basic information table, teachers' evaluation information table, student achievement table and so on as the research object. For example, in the data of teacher evaluation information table, a teacher may correspond to multiple teaching evaluation records. Therefore, it is necessary to average the attribute value of teaching evaluation score in the teacher's N records, so that the teacher has only one record in the database. According to the purpose of data mining, filter out valuable data and establish data source table.

2.2 Data Cleaning

Due to various data quality problems, the data may contain incorrect values. When integrating from multiple different data tables, we must pay attention to the consistency of data between different data tables. Data cleaning is to improve data quality by eliminating tuples such as errors, noise, defects and inconsistencies in the original data set. In this paper, the incorrect and inconsistent data are processed by manual correction. For example, for the record with empty score in the student's grade sheet, ① if the student has transferred or dropped out, the record will be deleted directly, and the student's record will be deleted in the student's basic information data sheet; ② If the teacher missed or mistakenly entered the score in the "positive teaching management model", the correct score will be entered. And the students' scores: first, the calculation of students' make-up examination scores. According to the relevant regulations of our school, students who fail in the examination shall take a make-up examination, and the make-up examination result shall be included in the student file according to 60 points. There is a certain difference between the make-up examination result and the original result. In order to ensure the accuracy and rationality of data mining results, the result of failed students shall be subject to the original result rather than calculated according to the make-up examination result. Second, the calculation of the scores of students who do not normally take the exam. According to the regulations of our university, students may not take the examination with the approval of relevant departments of the University for special reasons, but they must take the make-up examination, and the make-up examination results shall be included in the student file according to the normal examination results.

2.3 Data Conversion

Data transformation is to transform data into a description form suitable for data mining. The transformation of this paper mainly includes the following contents.

(1) Smoothing: if the current data point is null or noisy data, take out the weighted average of K (K can be customized) data points before (after) the current point and replace them.
(2) Aggregation: summarize and summarize data. It mainly constructs the data side for multi granularity data analysis.
(3) Data generalization: data generalization is to replace low-level or data level data objects with more abstract or higher-level concepts.

In this paper, the method of data generalization is mainly used for data conversion. First, convert the attribute value of birth date in the data into the corresponding age segment, convert the attribute value of workload into the corresponding workload segment, convert the attribute values of middle school students' admission average score and students' school average score in the data into the corresponding grade, and convert the attribute value of teaching evaluation score in the data into the corresponding teaching evaluation grade. According to the principle that the teaching evaluation score is not less than 60 points, There are three grades of design evaluation: excellent, medium and qualified. In the specific transformation process, the nodes experienced by the data form a set of model meta nodes. There is a data transmission link between each two nodes. Under this condition, the model can be expressed as:

$$B = (b_{i,j})n * n \in \{0, 1\}^{n*n} \tag{1}$$

In the formula, $b_{i,j}$ represents the node location of the model data center. If there is link connectivity between the two nodes, the corresponding element is 1, otherwise it is 0. It is assumed that there is a non ring path link L between two nodes. If the link passes through the node, there is $l \in b_{i,j}$. At this time, the implementation process of the three transformations is:

$$
\begin{aligned}
C_l &= \sum_{i=a}^{n}\sum_{j=a}^{n} c_{i,j} \cdot l_{i,j} \\
D_l &= \sum_{i=a}^{n}\sum_{j=a}^{n} d_{i,j} \cdot l_{i,j}
\end{aligned}
\tag{2}
$$

$$
H_l \leq \sum_{i=a}^{n}\sum_{j=a}^{n} l_{i,j} \tag{3}
$$

$$
s(l_{i,j}) = h(l_{i,j}) - H_l \leq 0 \tag{4}
$$

In the formula, H_l represents the link effectiveness factor corresponding to the link represented by the individual, and C_l, D_l, $h(l_{i,j})$ and $s(l_{i,j})$ are different attributes of the data respectively. Through this condition, it provides a reference for the construction of hybrid teaching model.

3 Design of "Online + Offline" Mixed Teaching Model Based on Deep Reinforcement Learning

3.1 Framework Design of "Online + Offline" Hybrid Teaching Model

Based on the project response theory, social comparison theory and metacognition theory, and referring to the hypermedia general model of deep reinforcement learning education, this study designs a "Online + offline" hybrid teaching model based on deep reinforcement learning for learning experience. According to certain standards, the domain knowledge unit and learner unit are designed and developed. The basic laws of learning and teaching are as follows: first, learning is a process in which learners actively participate; Second, learning is a gradual process of experience accumulation. Third, different types of learning have different processes and conditions. Fourth, for learning, teaching is the external condition of learning. Effective teaching must be an activity that gives timely and accurate external support to learners according to the law of learning. Starting from the emphasis on "learners' individual characteristics and learning needs", dynamically track learners' knowledge status, knowledge level and learning behavior, and dynamically update learners' units by using coverage modeling technology and data-driven technology under the coordination of deeply strengthened learning mechanism according to the domain knowledge unit and learner unit driven teaching model, Present open learner unit, open social learner unit and good adaptive learning content for learners, and finally realize personalized service and trigger learners' learning experience in metacognition and social comparison.

3.2 Domain Knowledge Unit

Domain knowledge unit is the foundation of deep reinforcement learning hybrid teaching model. It points out the application field and learning content of deep reinforcement learning hybrid teaching model, and provides domain structure and information that needs to be adapted. Domain knowledge unit describes the knowledge unit, knowledge point, learning object, association relationship between knowledge units, association relationship between knowledge units, association relationship between knowledge points, association relationship between knowledge points and association relationship between knowledge points and learning objects involved in the application field. Each knowledge point corresponds to multiple learning objects, and each learning object has text, video Test questions and other forms. By analyzing the learning object metadata celts-3.1 defined by the Educational Information Technology Standards Committee of the Ministry of education of China, this study outlines the structure diagram of domain knowledge units, determines the attributes of domain knowledge elements, gives the reference specifications of domain knowledge units, and constructs domain knowledge units.

3.2.1 Construction of Domain Knowledge Unit Structure

Domain knowledge unit is a collection of domain knowledge elements and their relationships. Domain knowledge elements have different names and relationships in different models. Therefore, in the domain knowledge unit, it is necessary to determine the composition and correlation of domain knowledge elements. Learning objects are the support of learning tasks and learning activities in the learning process. Learning objects present learners with learning content suitable for their personality characteristics with their rich and diverse types. The types of learning objects include text, video, audio, pictures, tests, test questions, examples, animation and demonstration, courseware, teaching cases, FAQs, etc. The learning objects designed in this research include text, video and test questions. Text, video and test questions are related to domain knowledge units through knowledge points. It can be seen that the domain knowledge elements contained in the domain knowledge unit include disciplines, primary knowledge points, secondary knowledge points, courses, chapters, sections, knowledge points, learning objects, texts, videos and test questions. In a broad sense, this study refers to disciplines, primary knowledge points, secondary knowledge points, courses, chapters and sections as knowledge units. The relationships between domain knowledge elements in a domain knowledge unit are defined as the following.

(1) *Superior relationship and inferior relationship*
The domain knowledge element a at the upper level is more integrated, contains more knowledge content and expresses more abstract content; The lower domain knowledge element B is more localized and the content expressed is more specific. Generally, it only reflects one aspect of the upper knowledge point, that is, a contains B and B is A part of A. For example, disciplines, primary knowledge points, secondary knowledge points and knowledge points form a superior subordinate relationship in turn; Courses, chapters, sections and knowledge points form a superior subordinate relationship in turn; The learning object is the upper knowledge point of text, video and test questions, and the text, video and test questions are the lower knowledge points of the learning object.

(2) *Precursor relationship and follow-up relationship*

The antecedent relationship and follow-up relationship indicate that there is a logical sequence of domain knowledge elements. If you must master the premise domain knowledge element B before learning A domain knowledge element A, that is, B precedes A, then the antecedent of a is B and the follow-up of B is A. For example, the relationships between primary knowledge points and primary knowledge points, between secondary knowledge points and secondary knowledge points, and between knowledge points form precursor and follow-up relationships; Precursor and follow-up relationships are formed between courses, between chapters, and between sections. A more specific example is that before learning the knowledge point "rational division", we must master the knowledge point "reciprocal", and we must first learn the knowledge point "rational addition" before learning the knowledge point "rational subtraction". The core element of domain knowledge unit is knowledge points. The domain knowledge elements contained in disciplines, courses and learning objects are centered on knowledge points and represent well structured domain knowledge through top-down hierarchical relations. Such a domain knowledge structure is conducive to the reuse of knowledge and learning resources.

3.2.2 Attribute Division of Domain Knowledge Elements

After determining the domain knowledge elements and their relationship, it is necessary to refine the attributes of domain knowledge elements according to certain metadata standards.

The first is learning object metadata. The Educational Information Technology Standards Committee of the Ministry of education of China defines learning object metadata celts-3.1, which is conducive to educators, learners or automated software to find, acquire, use and evaluate learning objects, and in the process of knowledge representation in the field of design and development, It will enable developers to fully consider the cultural and linguistic diversity of learning objects and their metadata in the use context, so as to promote the exchange and sharing of learning objects. The data elements of learning objects can be described in nine different categories:

(1) General class refers to the general information that describes the learning object as a whole;
(2) Lifetime class refers to the attribute information, learning history and personal and organizational information related to the learning object in the current state;
(3) Meta metadata class refers to some information of metadata instance itself;
(4) Technology refers to the information on the technical requirements and technical characteristics adopted by the learning object;
(5) Education refers to the information about the learning object and the characteristics of education and teaching;
(6) Rights refer to the information of the learning object in terms of intellectual property rights and use conditions;
(7) Relationship class refers to the relationship information between learning objects and other related learning objects;
(8) Commentary refers to the information that evaluates the learning object in terms of teaching use, including the author and creation time of the learning object;

(9) Classification class refers to the relationship information between learning objects and one or some specific classifications.

3.3 Learner Unit

Learner unit is an important basis for deep reinforcement learning model to realize deep reinforcement learning. It records learners' personal information, knowledge state, knowledge level, learning behavior, knowledge point planning, learning object review and other attribute information. Here, except that personal information is static information, other information is dynamic information. Based on the domain knowledge unit and the information that learners interact with the model, this study uses coverage modeling technology and data-driven technology to track the dynamic information in real time to ensure that the dynamic information is always in the latest state. Dynamic information provides a basis for adaptive presentation of learning objects and learning contents.

3.3.1 Learner Unit Construction

In the mixed teaching model designed in this paper, learner unit construction mainly includes the following contents:

To determine the source of deep reinforcement learning, that is, to model the learner unit according to the learner's personality characteristics, which are called the source of deep reinforcement learning. According to the relationship between personality characteristics and domain, personality characteristics are divided into domain related and domain independent personality characteristics, as shown in Fig. 1.

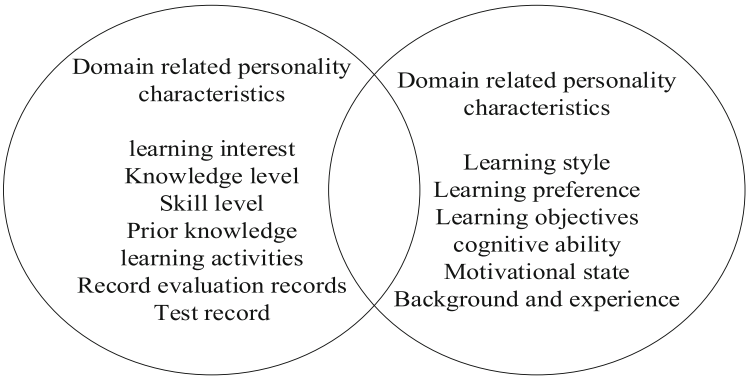

Fig. 1. Domain related and domain independent personality characteristics

It can be seen from Fig. 1 that the domain-related personality characteristics refer to the information recorded by the model is directly related to the learned domain knowledge, and is a reflection of the learner's knowledge level and skills in the learning situation, such as learning interest, knowledge level, skill level, prior knowledge. Knowledge, learning activity records, assessment records, and test records; domain-independent personality characteristics refer to those information that are not directly related to the knowledge in the field of study, but have indirect guiding significance for the learner's learning process, such as learning styles, learning preferences, and goals., cognitive abilities, motivational states, background and experience. According to whether the personalized features change with time, the personalized features are divided into static and dynamic personalized features, as shown in Fig. 2.

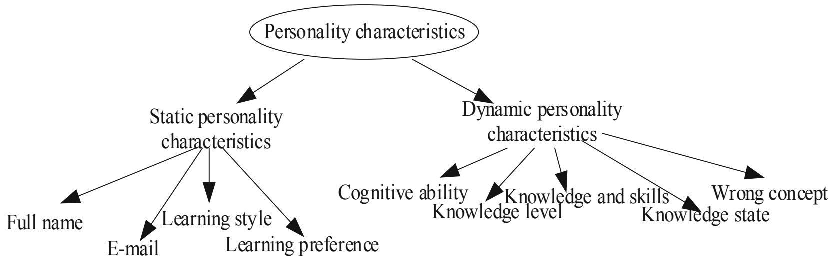

Fig. 2. Static and dynamic personality

According to Fig. 2, static personalization features are all set before the learning process occurs and usually remain unchanged throughout the learning process, such as name, email, age, native language, learning style, learning preference, etc., which can be determined by the learner Available directly through the options menu, and also through the use of a questionnaire. The dynamic personalization feature is the information collected in the learning process, such as the interaction between the learner and the model, the performance behavior and the learning history, which need to be determined by certain rules and algorithms, and constantly updated in the model, such as learning interest, cognitive ability, knowledge level, knowledge skills, knowledge mastery state, wrong and lost concepts, learning behaviors, emotional factors, metacognition and other personalized characteristics. The dynamic personalized characteristics of learners constitute the basis of each learner's needs in the model of in-depth reinforcement learning, which can be measured by questionnaire and the tests that learners must complete in the learning process. The deep reinforcement learning sources of this study include personal information, subject knowledge status, curriculum knowledge status, knowledge level, learning behavior, knowledge point planning, text and wrong question review.

3.3.2 Learners' Choice of Unit Modeling Technology

Learner unit modeling technologies include coverage modeling, lead plate modeling, perturbation modeling, machine learning technology, constraint based model, fuzzy learner modeling, Bayesian network, project response model, ontology based learner modeling and so on. Different modeling techniques are usually combined to infer and modify the dynamic personality characteristics in the learner unit, so as to continuously improve the learner unit and ensure the accuracy and diversity of information in the learner unit. This study uses coverage modeling technology, item response model and weight algorithm to construct learner units. Through the visual processing of the designed learner unit, the information of the learner unit in the model is transmitted to the learners in the form of visualization. The designed visual view of learner unit helps learners understand the gap between what they have learned and what experts expect by presenting learners' knowledge status and learning progress, so as to promote learners to learn more knowledge.

3.3.3 Learner Unit Data Management

The purpose of data management of learner unit is to realize the sharing and migration of learner information between different models, improve the portability of learner information data, and ensure the privacy, security and integrity of learner data.

The first is the management of core data. The packaging of learner information core data in the model is to ensure that learner information is transmitted and exchanged among different learner information models, such as learning management model, human resource model, student information model, enterprise e-learning model, knowledge management model, resume database and so on. The main purposes of describing learners' personality characteristics based on the data model include: recording and managing learning history, learning objectives and learning achievements related to learning; Involve learners in the learning process; Discover learning opportunities for learners. In this paper, the core data of the model is divided into 11 categories: identity recognition; Accessibility; Qualification, certification and licensing; Objectives; Activities; Ability; interest; Affiliation; Performance report; Security key; Association relationship. The specific description is shown in Table 1.

Table 1. Model learner information core data management mechanism

Core data name	Core data description
Identification	Learner's personal information related to learning, such as ID, grade, E-mail, QQ number, wechat account, etc.
Target	Learning goals or strong desire to learn
Activity	In formal learning, informal learning, training, work experience, lifelong education, and all learning related activities in military or civil service, self-report, etc.
Qualification, certification and licensing	Certification certificates, licenses, qualification certificates, etc. issued by authoritative certification bodies
Ability	Describe the knowledge, skills and abilities acquired by learners in emotional, cognitive and psychological fields
Report card	The summary report of learners' academic achievements can be in a variety of forms

(*continued*)

Table 1. (*continued*)

Core data name	Core data description
Interest	Describe learners' hobbies, recreational activities and other interest information
Subordination	Members of professional organizations
Accessibility	Accessibility to learner information such as language competence, qualifications and learning preferences, learning preferences include cognitive preferences (such as learning style preferences), physical preferences (such as liking large fonts), and technical preferences (such as liking specific computer platforms)
Security key	The password and security key set when the learner interacts with the learner information model and service
Association relationship	Represents the association between the core components of the model

The second is the management of public and private information, which aims to standardize the semantic grammar of learners' units and describe learners' information and knowledge. In this paper, public and private information are divided into learner personal information, learner performance, learner relationship, learner security and learner document; Learners prefer six types of learner information to realize the construction of "Online + offline" mixed teaching model.

4 Analysis of Teaching Effect

Taking college mathematics teaching as an example, the model is solved by SPSS software to obtain the multiple linear regression model of teaching effect. The data obtained from the questionnaire and sorted out are imported into SPSS software, and the multiple linear regression method is used to obtain the relevant results, that is, the college mathematics teaching effect under the online and offline mixed teaching model. Analyze the coefficient of each influencing factor in the model, fully discuss with teachers, and then readjust the proportion of each influencing factor online and offline, so as to analyze the teaching effect of the model in college mathematics courses.

4.1 Data Sources

This paper collects and arranges the relevant data of college mathematics courses taught by online and offline mixed model in three colleges and universities, and designs a questionnaire for the relevant influencing factors to obtain the relevant data. Through the preliminary investigation, the factors affecting the teaching effect are analyzed and summarized into the following 10 influencing factors: online teaching video (X1), online teaching PPT (X2), online homework (X3), online test (X4), online Q & A (X5), offline classroom teaching (X6), offline homework (X7), offline test (X8), offline Q & A (X9), and participating in college students' mathematical modeling activities (X10), the dependent variable Y represents the total score. The questionnaire is prepared by using the questionnaire star app, as shown in Table 2.

Based on the questionnaire in Table 2, a survey was conducted in three universities that adopted the online and offline mixed teaching model for mathematics courses. 328 questionnaires were distributed and 328 questionnaires were recovered, of which 328 were valid. The obtained data were processed. Single choice questions corresponded options 1, 2, 3 and 4 to 100, 80, 60 and 40 respectively, and the data were statistically

Table 2. Questionnaire

Influence factor	Problem	Option 1	Option 2	Option 3	Option 4
X1	Whether to watch the online teaching video carefully and completely	Very serious	Quite serious	Commonly	Unclear
X2	Whether to watch online teaching ppt carefully and completely	Very serious	Quite serious	Commonly	Unclear
X3	Do you complete online work carefully and independently	Very serious	Quite serious	Commonly	Unclear
X4	Do you complete the online test carefully and objectively	Very serious	Quite serious	Commonly	Unclear
X5	Do you seriously participate in online Q & A	Very serious	Quite serious	Commonly	Unclear
X6	Whether offline classroom teaching is carried out seriously	Very serious	Quite serious	Commonly	Unclear

(*continued*)

Table 2. (*continued*)

Influence factor	Problem	Option 1	Option 2	Option 3	Option 4
X7	Whether the offline operation is completed carefully and independently	Very serious	Quite serious	Commonly	Unclear
X8	Have you completed the offline test carefully and objectively	Very serious	Quite serious	Commonly	Unclear
X9	Do you seriously participate in offline Q & A	Very serious	Quite serious	Commonly	Unclear
X10	Whether to participate in college students' mathematical modeling activities	Participate many times	Attend once	Understand	Unclear

analyzed, That is, each line represents a questionnaire, a total of 328 valid data, as shown in Table 3.

Table 3. Questionnaire data

Total score		Influence factor				
Serial number	Y	X1	X2	X3	…	X10
1	82	80	60	70	…	60
2	85	100	60	70	…	60
3	93	90	80	90	…	80
4	79	80	90	60	…	80
…	…	…	…	…	…	…
328	69	90	100	80	…	70

4.2 Analysis of Teaching Effect

It can be seen from Table 2 that 10 influencing factors are counted in the questionnaire, the data in Table 3 are standardized, and the following multiple linear regression model is established:

$$Y = a_0 + a_1X1 + a_2X2 + a_3X3 + \ldots a_{10}X10 + \varepsilon \tag{5}$$

In the formula, $a_i\,(i = 1, 2, \ldots, 10)$ and ε represent adaptability error. The stepwise regression method with SPSS software can be used to select independent variables. These six variables can be removed from the x 2 online teaching ppt, x 5 online Q & A, x 7 offline homework, x 8 offline quiz, x 9 offline Q & A and x 10 participating in college students' mathematical modeling activities, so as to eliminate the remaining four variables X 1 online teaching video, x 3 online homework, x 4 online quiz X 6 offline classroom teaching is analyzed to obtain the multiple linear regression model of the effect of online and offline mixed teaching model:

$$Y = a_0 + a_1X1 + a_2X2 + a_3X3 + \ldots a_6X6 + \varepsilon \tag{6}$$

The regression model was tested by F-test, and the $P < 0.05$ was obtained, so it passed the significance test. The parameters of this model are estimated by the least square method to obtain the following regression equation:

$$\hat{Y} = 35.25 + 0.225X1 + 0.213X3 + 0.075X4 + 0.575X6 \tag{7}$$

To sum up, the mixed teaching model designed in this paper has an obvious positive effect on the four factors affecting the teaching effect: online teaching video, online homework, online test and offline classroom teaching, and the offline classroom teaching has the greatest impact on the teaching effect of these four factors, followed by online teaching video, with a test $P < 0.05$, The analysis results clearly show that the model can give full play to the irreplaceable offline classroom teaching, and maximize the teaching effect in the new era network environment with the help of online teaching video.

5 Conclusion

This study proposes a hybrid teaching model of deep reinforcement learning for learning experience. In the process of demonstration and evaluation, the metacognition and social comparative learning experience triggered by the teaching model have a good positive response, and give good feedback to the teaching institutions provided. With the deepening of research and continuous reflection in the research process, it is found that there are still two aspects of research to be carried out:

(1) Optimization of domain knowledge unit

In this study, domain knowledge unit is not the focus of the research, but it plays an important role in the research. The domain knowledge unit of this study only provides learning objects such as videos, texts and test questions of each knowledge point in the rational number chapter. In order to conduct more in-depth and lasting research, we need to further improve the learning resources of the remaining chapters, which is more conducive to demonstration and evaluation.

(2) Expansion of deep reinforcement learning source

The learner unit of this study focuses on knowledge status, knowledge level and learning behavior. Although it evaluates the metacognition and social comparison learning experience caused by two open learner units and teaching strategies, it does not take metacognition and social comparison as the two dimensions of the learner unit. On the basis of this study, It is necessary to expand the source of deep reinforcement learning to psychological personality characteristics such as metacognition and social comparison, so as to provide more comprehensive learning services for the deep reinforcement learning model.

Fund Project. Sichuan Provincial Key Research Base of Philosophy and Social Sciences-China West Normal University Sichuan Education Development Research Center Funded/Project "Research on the Implementation Path of'Internet + Education' Helping Higher Education Curriculum Reform in the Post-epidemic Context" (Project Number: CJF21005).

References

1. Zalat, M.M., Hamed, M.S., Bolbol, S.A.: The experiences, challenges, and acceptance of e-learning as a tool for teaching during the COVID-19 pandemic among university medical staff. PLoS ONE **16**(3), e0248758 (2021)
2. Dias, S.B., Hadjileontiadou, S.J., Diniz, J., et al.: DeepLMS: a deep learning predictive model for supporting online learning in the Covid-19 era. Sci. Rep. **10**(1), 19888 (2020)
3. Schlesselman, L.S.: Perspective from a teaching and learning center during emergency remote teaching. Am. J. Pharm. Educ. **84**(8), 8142 (2020)
4. Snook, A.G., Schram, A.B., Jones, B.D.: Faculty's attitudes and perceptions related to applying motivational principles to their teaching: a mixed methods study. BMC Med. Educ. **21**(1), 188 (2021)
5. Doukanari, E., Ktoridou, D., Efthymiou, L., et al.: The quest for sustainable teaching praxis: opportunities and challenges of multidisciplinary and multicultural teamwork. Sustainability **13**(13), 1–21 (2021)
6. Giddens, J., Curry-Lourenco, K., Miles, E., et al.: Enhancing learning in an online doctoral course through a virtual community platform. J. Prof. Nurs. Off. J. Am. Assoc. Coll. Nurs. **37**(1), 184–189 (2020)
7. Jia, Y.: Research on the practice of college English classroom teaching based on Internet and artificial intelligence. J. Intell. Fuzzy Syst. **2021**(1), 1–10 (2021)
8. Rodriguez-Rodriguez, E., Sanchez-Paniagua, M., Sanz-Landaluze, J., et al.: analytical chemistry teaching adaptation in the COVID-19 period: experiences and students' opinion. J. Chem. Educ. **97**(9), 2556–2564 (2020)
9. Baghal, T.A.: The effect of online and mixed-mode measurement of cognitive ability. Soc. Sci. Comput. Rev. **37**(1), 89–103 (2019)
10. Pp, A., Kr, B.: Implicit adversarial data augmentation and robustness with Noise-based learning. Neural Netw. **141**(9), 120–132 (2021)
11. Xu, N., Fan, W.H.: Research on interactive augmented reality teaching system for numerical optimization teaching. Comput. Simul. **37**(11), 203–206+298 (2020)