J. COMBIN. MATH. COMBIN. COMPUT. 127b (2025) 5455--5470

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Research and Practice on Optimizing the Innovation of Teaching Methods for Civic and Political Education in Colleges and Universities Based on Deep Neural Networks

Guoqi Lin[1, ✉]

[1] *Shaoxing University, Shaoxing, Zhejiang, 312000, China*

ABSTRACT

This study focuses on the innovation of teaching methods for Civic Education in colleges and universities, and provides a structured knowledge framework for teaching by constructing a Civic Knowledge Mapping and integrating course knowledge points. On this basis, a new classroom teaching mode is designed to integrate online and offline teaching resources to enhance student interaction and participation. A knowledge tracking model of key-value memory network (MKVMN) based on multi-feature fusion is proposed to accurately track students' mastery of Civics and Politics knowledge by capturing students' multi-dimensional learning behavior characteristics. To optimize the recommended path for students' personalized learning, an improved ant colony algorithm is introduced to generate personalized learning paths based on students' individual differences. The experimental results show that when the number of learning units is 0-10 (pre-study period), the improved ACO algorithm model does not have obvious advantages for students' learning, but when the number of learning units reaches 11-50, the difference between the experimental group students' learning performance and the control group becomes more and more obvious, so it can be seen that the improved ACO algorithm can obviously improve the students' Civic and Political Science learning performance. In addition, the IACS-PRA algorithm is especially effective in long path recommendation, which finds the optimal personalized recommendation path through a gradual approach to help students learn Civics and Political Science more efficiently, and provides a practical demonstration for the digital transformation of Civics and Political Science education in the new era.

Keywords: MKVMN, IACS-PRA, Knowledge Graph, Personalized Recommendation, Civic and Political Education

✉ Corresponding author.
E-mail address: hester_2025@sina.com (G. Lin).
Received 03 March 2024; Revised 10 May 2024; Accepted 30 November 2024; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-302

1. Introduction

With the continuous development and change of society, the teaching reform of Civic and Political Education in colleges and universities has become a hot topic in the field of education nowadays [1-2]. As an important course to cultivate students' comprehensive quality and social responsibility, Civic education plays an important role in college education [3-4]. However, in the current teaching reform of Civic education in colleges and universities, there are a series of problems such as single teaching content, obsolete teaching methods and insufficient construction of teachers. The necessity of the teaching reform of Civic Education in colleges and universities is obvious, only by constantly improving and innovating the teaching content and methods can we better adapt to the development needs of today's society, and contribute to the cultivation of socialist builders and successors who are all-rounded in morality, intelligence, physical fitness, aesthetics and labor [5-8].

Deep neural network is a multilayer perceptron consisting of an input layer, multiple hidden layers and an output layer [9]. The input layer of the network receives raw data, such as text, images and audio, etc., and then through the nonlinear transformation and feature extraction of multiple hidden layers, the final output is the target result [10-12]. The key of deep neural network in the optimization of teaching methods of Civic Education in colleges and universities is that it can automatically learn features from data.Traditional optimization methods may need to extract features manually, which is not only time-consuming and laborious,but also may miss some important information[13-14]. While deep neural network algorithm can directly deal with the original data, automatically mining the complex relationship hidden in the data through the multi-layer network structure, so as to achieve the optimization of teaching methods [15-16].

In this paper, we first constructed the Civics knowledge map to realize the association and integration of knowledge points. A new classroom teaching mode of online map guidance and offline contextual interaction is designed to realize the deep integration of theoretical lecturing and practical exploration. At the level of knowledge tracking, we creatively propose a multi-feature fusion key-value memory network knowledge tracking model (MKVMN), which integrates multi-dimensional features such as knowledge mastery and cognitive behaviors to update students' knowledge state representations. For the learning path optimization problem, a dynamic pheromone updating mechanism and an adaptive heuristic function are introduced into the ant colony algorithm, which is combined with the knowledge graph output to generate a personalized learning path recommendation scheme. The effectiveness of the personalized learning path recommendation (IACS-PRA) algorithm model under the improved ant colony algorithm is demonstrated through experiments on four real data sets. Finally, the effectiveness of the IACS-PRA model in recommending students' learning paths is verified in the teaching of Civic Education in colleges and universities.

2. Personalized learning path recommendation model based on improved ant colony algorithm

2.1. *Construction of Knowledge Mapping of Teaching Methods for Civic and Political Education in Colleges and Universities*

In the research system of "Deep Neural Network Optimization of Civic and Political Education in Colleges and Universities", curriculum teaching is one of the important means to implement the innovation of teaching methods, and it is necessary to deepen the reform of classroom teaching and build a teaching mode oriented to competence-based education.The Civic and Political Education course of BJ University is a first-class course at the national level, and the teaching team is a provincial teaching team, which is led by famous teachers. The teaching team is a provincial teaching team led by famous teachers, and the online open course has been built. In the teaching of the Civic and Political

Education Program, the initial classroom teaching reform based on knowledge mapping has been carried out, and the knowledge mapping has been carried out throughout the program.

2.1.1. Building a repository of quality teaching resources. Centering on students' development and based on ideological education and political requirements, we associate multimedia courseware, ideological elements, question banks, videos, teaching objects, teaching case banks, teaching materials, scientific research papers, MOOC, SPOC and innovation competitions with knowledge points. Accurately construct a high-quality teaching resource base based on the knowledge points, and provide students with materials to prepare for the construction of the knowledge framework.

2.1.2. A new classroom teaching model based on knowledge mapping. With the requirements of ideological and political education to determine the knowledge points, curriculum and classroom teaching objectives of Civics and Politics, combining knowledge mapping with classroom teaching, optimizing the blended teaching design based on BOPPPS through guide-teach-learn-evaluate, constructing a deep learning mode based on knowledge mapping, the classroom teaching mode based on knowledge mapping is shown in Figure 1.

1) Learning knowledge online. In the form of cases and independent learning tasks, through MOOC learning, testing and feedback, students initially master the knowledge points and realize the low-order knowledge learning in Bloom's cognitive law.

2) Raising competence offline. Teachers first show the knowledge map of each chapter, so that students can learn on the basis of online learning, through the knowledge map learning, grasp the teaching objectives of the course and the key points and difficulties from the overall situation, and then clarify the knowledge points that need to be learned, the a priori knowledge and its connection, and plan the learning path.

3) Group participation in innovation. Teachers analyze the learning records of students in the first two sessions, and based on constructivist theory, give lectures and answer questions. Students conduct group-based, project-based and inquiry-based learning based on Feynman Learning Method to further construct the knowledge system, and through group learning and flipped learning, conduct after-school expansion to map out the knowledge of each chapter.

4) Figure-number double-drive assessment. Through students' self-assessment, group mutual assessment, teacher evaluation and objective evaluation, students' learning is monitored while the knowledge points are measured and assessed in order to achieve personalized and precise teaching. Through the above links, a new form of Civics classroom teaching led by Civics and centered on knowledge mapping is formed to create a high-quality and efficient classroom.

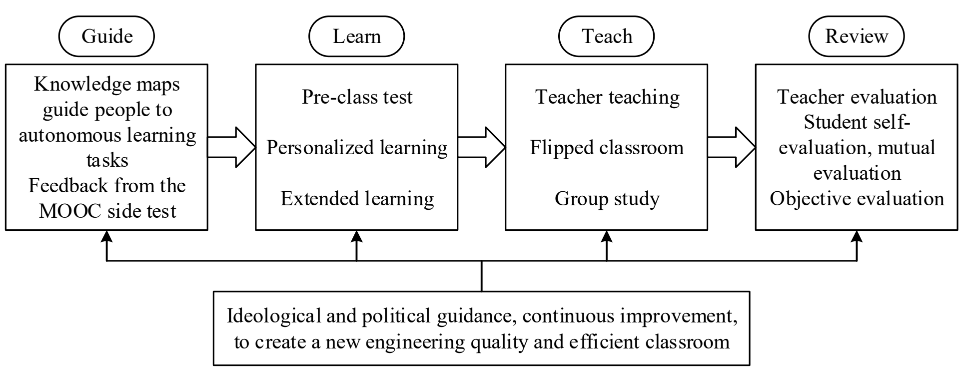


Fig. 1. Classroom teaching model based on knowledge map

2.2. Knowledge Tracking Model for Key-Value Memory Networks with Multi-Feature Fusion

In this paper, we propose a knowledge tracking model for key-value memory networks (MKVMN) with multi-feature fusion [17]. Different from the knowledge tracking method that only uses the features of knowledge concepts, the model in this paper obtains two features, namely, problem difficulty information and students' personal learning ability, by means of data preprocessing, and cross-fertilizes the problem difficulty information with knowledge concepts, and the students' personal learning ability with the problem responses, in order to improve the dynamic key-value memory network. The goal of knowledge tracking is to track students' knowledge proficiency status through their past learning sequences. Based on the student's previous historical interaction sequence $X=\{X_1,X_2,...,X_{t-1}\}$, the probability that the student will correctly answer the current question with the knowledge concept belonging to c_i, i.e., $P(r_t=1|c_i,d_i,a_i,X)$, at time step t is predicted.

2.2.1. Calculation of problem difficulty and learning ability. In this paper, the classical model IRT [18] in psychometrics is utilized to construct new ways of coding questions and question-response coding. In the IRT model, two scalars, question difficulty and student ability, are usually used to describe the probability of a student answering a question correctly.

1) Question difficulty calculation. In order to measure the difficulty coefficients of various different questions in the whole dataset d_t, in this paper, we usually use the correct rate of students correctly answering a certain question to characterize the difficulty level of the question, i.e.:

$$d_t=\frac{\Sigma_1^N r_t==1}{N} \tag{1}$$

where r_t refers to the question response information with index t, N refers to the total number of questions in the dataset, $\Sigma_1^N r_t=1$ refers to the number of times the student answered the question correctly, and d_t refers to the difficulty coefficient of the question resulting from the calculation.

2) Learning ability calculation. A student's historical interaction sequence is a time series from which changes in a student's individual learning ability can be inferred. A student's individual learning ability a_t can be defined as the difference between the student's rate of correctness and error according to the interaction sequence, with the following formula:

$$C_{0:t}=\sum_{r=0}^{t-1}\frac{r_r==1}{t} \tag{2}$$

$$E_{0:t}=\sum_{r=0}^{t-1}\frac{r_r==0}{t} \tag{3}$$

$$a_t=C_{0:t}-E_{0:t} \tag{4}$$

where r_r refers to the question response information at index r, $r_r==1$ and $r_r==0$ refer to students answering the question correctly and answering the question incorrectly, respectively, $c_{0:t}$ refers to the rate of students answering the question correctly in index interval $[0,t)$, $E_{0:t}$ refers to the rate of students answering the question incorrectly in index interval $[0,t)$, and a_t refers to the difference between the rate of correctness and the rate of incorrectness in index interval $[0,t)$, i.e., the student's individual learning ability at the current index t.

2.2.2. Coding layer:

1) Problem coding. This paper proposes a new way of problem coding, which works by cross-fertilizing problem difficulty information with knowledge concepts.The problem coding q_t constructed in this paper when indexed as t is as follows:

$$q_t = c_t \oplus d_t \square\ c_t \tag{5}$$

where c_t is the coding of the knowledge concept to which question q_t belongs and d_t is the difficulty factor for question q_t. This type of question coding is proposed so that questions covering the same knowledge concept are closely related, but with an important individual difference that should not be ignored.

2) Question-response coding. The question-response coding approach is similar to the question coding approach. In this paper, we define a student's individual learning ability characteristic and cross-fertilize it with the question response, and the question-response coding e_t constructed in this paper at index t is as follows:

$$e_t = s_t \oplus a_t \square\ s_t \tag{6}$$

$$s_i = embeding(c_i + r_i \cdot M) \tag{7}$$

where M is the total number of knowledge concepts, s_i is the way the knowledge concept-response is coded, s_t is the knowledge concept-response coding at index t, and a_t is the student's individual learning ability at index t.

2.2.3. Calculation of relevance weights. Calculating the relevance weights can indicate the magnitude of relevance between the current question to be answered by the student and all potential knowledge concepts. After obtaining the encoding vector q_t of the question, q_t is first multiplied with an embedding matrix A to obtain a continuous embedding vector k_t. k_t will do the inner product operation with each key slot $M^k(i)$ in the key memory matrix, and the relevance weight vector $w_t(i)$ between the question the student is currently answering and the potential knowledge concepts will be obtained by the softmax activation function, that is:

$$k_t = q_t^T \cdot A \tag{8}$$

$$w_t(i) = softmax(k_t^T M^k(i)) \tag{9}$$

where the softmax activation function is given by $softmax(x_i) = e^{x_i} / \Sigma_i e^{x_i}$. The MKVMN model will reuse the correlation weight matrix w_t.

2.2.4. Reading process. When a new problem q_t is input to the MKVMN model, the model obtains what is read from the value memory matrix r_t, i.e., the student's proficiency status for the current problem, by multiplying each value slot $M_t^v(i)$ in the value memory matrix with the relevance weight vector $w_t(i)$ for its weighted sum, i.e.:

$$r_t = \Sigma_i w_t(i) M_t^v(i) \tag{10}$$

Considering that each problem has its own level of difficulty, here the readings r_t are usually spliced with the continuous embedding vector k_t obtained above and fed into a fully connected layer via the tanh activation function to obtain an output vector f_t, i.e:

$$f_t = \tanh(W^T[r_t, k_t] + b) \tag{11}$$

where the tanh activation function operates with the formulas $\tanh(x_i) = (e^{x_i} - e^{-x_i})/(e^{x_i} + e^{-x_i})$, W and b are the parameters learned during the training process.

2.2.5. The writing process. In this paper, based on the obtained problem-response encoding e_t through a series of model components, it is written into the value memorization matrix of the MKVMN model. Based on the above weight matrix w_t, the forgetting vector o_t is computed as follows:

$$o_t = sigmoid(W_o^T g_t + b_o) \tag{12}$$

where W_o is a transformation matrix and o_t is a column vector whose elements are all in the range (0, 1).

After obtaining the forgetting vector o_t, the changed value memory matrix $\tilde{M}_{t+1}^v$ is as follows:

$$\tilde{M}_{t+1}^v(i) = M_t^v(i)[1 - w_t(i)o_t] \tag{13}$$

The elements corresponding to the value memorization matrix are set to 0 only when the weight vector $w_t(i)$ and the forgetting vector o_t are both 1.

Similar to the acquisition of the forgetting vector o_t, the knowledge growth vector m_t is computed as follows:

$$m_t = \tanh(W_m^T g_t + b_m) \tag{14}$$

where W_m is also a transformation matrix and m_t is also a column vector.

After obtaining the knowledge growth vector m_t, the final updated value memory matrix M_{t+1}^v is calculated as:

$$M_{t+1}^v = \tilde{M}_{t+1}^v + w_t(i)m_t \tag{15}$$

2.2.6. Prediction Layer and Model Training. In this paper, the output vector f_t obtained in the reading process step is used to obtain the response value p_t of the student answering the current question, i.e., the magnitude of the probability of answering question q_t correctly, by means of a sigmoid activation function, then:

$$p_t = sigmoid(W^T f_t + b) \tag{16}$$

where the sigmoid activation function operates as $sigmoid(x_i) = 1/(1+e^{-x_i})$, W and b are the parameters learned during the training process.

During the training period of the MKVMN model, parameters such as the embedding matrices A and B and W_o, W_m, b_o, b_m can be learned by minimizing the cross-entropy loss function of the predicted response value p_t and the true response value r_t, i.e:

$$L = -\sum_t (r_t \log(p_t) + (1-r_t)\log(1-p_t)) \tag{17}$$

2.3. Personalized Learning Path Recommendation Method under Improved Ant Colony Algorithm

2.3.1. Ant Colony Algorithm. Ant colony algorithm (ACS) [19] is one of the intelligent bionic algorithms derived from the inspiration of ants' foraging behavior in nature, and the main idea is to simulate the ants to leave pheromones between food and anthill to improve the efficiency of optimality seeking in a positive feedback way.

1) Path construction. From node i to node j, a pseudo-random state transfer probability is introduced, obeying the following rules:

$$S = \begin{cases} \arg\max_{j \in allowed} \left\{ \tau_{ij}^{\alpha} * \eta_{ij}^{\beta} \right\} & q \le q_0 \\ s & \text{else} \end{cases} \tag{18}$$

where: $allowed$ is the range of nodes that can be selected; τ_{ij} represents the pheromone concentration between two nodes; η_{ij} is the reciprocal of the distance between two nodes; α is the information heuristic factor; β is the expectation heuristic factor, which is used to regulate the degree of influence of η_{ij}; q, q_0 takes the range of values are $[0,1]$, q is the random number on the interval, q_0 is the adjustability parameter on the interval; s is the mechanism of using roulette wheel in choosing the path with high pheromone concentration and close distance when selecting pheromone concentration and proximity, can have a certain probability of selecting suboptimal edges, which allows the algorithm to explore the edges in a biased manner. Namely:

$$P_{ij}(t)=\begin{cases}\dfrac{\left|\tau_{ij}(t)\right|^{\alpha}*\left|\eta_{kj}^{B}\right|^{\beta}}{\sum\limits_{s\in \mathrm{A}allowed}\left|\tau_{is}(t)\right|^{\alpha}*\left|\eta_{rs}\right|^{\beta}} & j\in allowed \\ 0 & \text{else}\end{cases} \tag{19}$$

2) Pheromone updating. The path selection at the beginning is biased towards randomization, so the volatility factor ρ is introduced to erase the influence of some pheromones and prevent the algorithm from falling into a local optimum. τ_0 is the starting pheromone concentration on each edge, and the local pheromone update rule is:

$$\tau_{ij}(t+1)=(1-\rho)\tau_{ij}(t)+\rho\tau_0 \tag{20}$$

When all paths are constructed, a global pheromone update is performed, where $\Delta\tau_{ij}$ is the pheromone increment and L_{gb} is the length of the current globally optimal path, and the update rule is:

$$\Delta\tau_{ij}=\begin{cases}1/L_{gb0} & edge(i,j)\in L_{gb} \\ 0 & \text{else}\end{cases} \tag{21}$$

2.3.2. Personalized Learning Path Recommendation. The method of personalized learning path recommendation (IACS-PRA) optimized by ant colony algorithm, on the basis of subject-specific knowledge point map as the path, quantifies the difficulty of knowledge points through the sum of the number of knowledge sub-nodes and uses deep knowledge tracking model to get the knowledge level of different learners, and computationally classifies different ant colonies through similar learners, so that the recommendation results take into account the commonality and individuality of the learners, and to a certain extent, also enhances the interpretation of the recommendation. To a certain extent, the interpretability of recommendation is also enhanced.

1) Knowledge Graph Construction. Knowledge graph is a knowledge mover, which requires knowledge source, knowledge extraction and knowledge reasoning. The course team constructs the knowledge map of the course from the four levels of knowledge system, knowledge domain, knowledge unit and knowledge point, from top to bottom level by level, to form a structured knowledge system. The first step is to sort out the course knowledge points according to the syllabus. Based on the concept of OBE, taking graduation requirements as the main line, focusing on professional cultivation objectives and course objectives, paying attention to the relationship between prerequisite courses and subsequent courses, starting from the knowledge domain, associating knowledge units (chapters and sections), and deconstructing knowledge points at all levels (level 1 and level 2) to realize the “granularity” of refinement. The second step is to set the teaching objectives of the knowledge points. The goal of knowledge mapping is to explain and illustrate the knowledge. According to the characteristics of the course, the teaching objectives of the knowledge points are formulated from the three levels of

knowledge objectives, competence objectives and ideological and political objectives. The third step is to establish the relationship between knowledge points. Based on different logical structural relationships such as symbiosis, progression, dependence, etc., the knowledge points are linked together to facilitate the formation of the knowledge system.

2) Course knowledge point difficulty grading. For grading the difficulty of course knowledge points, the number of branch knowledge of a certain knowledge point within the course knowledge map is the breadth, and the number of all sub-knowledge points under the branch knowledge is the depth. Combined with the opinions of course experts r, α for the weighting factor, K for the knowledge point, according to the breadth and depth of the knowledge point to give the difficulty coefficient d. To d for normalization, d the closer to 1 represents the more difficult knowledge points. d at 0 to 0.3 is easy, 0.3 to 0.6 is medium, and 0.6 to 1.0 is difficult, for distances between nodes in the path are 1, 2, and 3, respectively, i.e:

$$d = \alpha \sum K_i + (1-\alpha) r \tag{22}$$

3) Training the learner model. The question answering of m student is collected $\{X_1, X_2, \cdots, X_m\}$, where $X_i = \{(q_{t1}, a_{t1}), (q_{t2}, a_{t2}), \cdots, (q_{tj}, a_{tj})\}$, q_t, a_t represent the questions and corresponding answers answered by the learner in t moments, the X input is used as the input to the model and the input sequence is one-hot coded, 0 means the question is answered correctly and 1 means the answer is incorrect, and there are j questions, the length of the coded sequence is $2j$. If the learner answers the question incorrectly on the ith question, the sequence corresponds to a value of 1 on the ith position and 0 on the rest. Value is 1 and the rest of the positions are 0. For the output Y obtained from the first t moments of training, the length is the total number of questions n, the knowledge level vector of each learner is calculated by the neural network after training to obtain the knowledge level vector matrix $U(m \times n)$ of all learners U. The cosine similarity between learners U_a, U_b is calculated thus dividing the subsequent ant colony species, i.e.:

$$\cos(U_a, U_b) = \frac{\sum_{i=1}^{n} u_{ai} u_{bi}}{\sqrt{\sum_{i=1}^{n} u_{ai}^2} \sqrt{\sum_{i=1}^{n} u_{bi}^2}} \tag{23}$$

The similarity threshold is set to 0.9, i.e., two learners are considered to be similar if the similarity is greater than 0.9 and vice versa.

When recommending personalized learning paths for learners, different categories of learners are abstracted as different ant colonies, and the pheromones left behind by the ants traveling on the path will affect the path selection of later ants. Based on the principle of mutual attraction of similar learners and mutual exclusion of dissimilar learners, the path selection is made and the recommendation result is given. In the solution process, the algorithm is divided into two parts, the initial path construction by the first generation of ants, while the second and subsequent generations of ants are divided into different species.

From the above algorithm flow, it can be seen that the maximum time complexity of this algorithm is $O(m * n^2)$, where m is the number of ants and n is the number of knowledge nodes. The maximum time complexity of traditional ant colony algorithm is $O(m * n^2)$, which shows that this algorithm does not increase the time complexity.

3. Analysis of the effect of the innovation of ideological education in colleges and universities

3.1. *Knowledge Trace Performance Testing*

3.1.1. Comparison of training hours. The MKVMN model is designed with multi-thermal information embedding network and temporal convolutional network instead of the step of self-encoding embedding and attention calculation in the transformer structure designed by the ACS model. In this paper, five comparative models "DKT, DKVMN, SAKT, SAINT, ACS" are selected for the experiments with this model (MKVMN). In this paper, the training time of each knowledge tracking model is recorded as the median of the three training times under the conditions of input sequence lengths of 100, 200, 400, 800, and 1000.

Figure 2 shows the model training lengths for different sequence lengths. In terms of training time, the MKVMN model drastically shortens the training time on each sequence compared to the SAINT and ACS models, under the condition that the type of input side information is basically the same. It can be seen that the model designed in this paper has excellent training time under the input condition of long sequences. However, because MKVMN sets up inputs with multiple types of side information, the model needs to process more information, and the training time of the model over the length of each sequence is slightly more than that of the DKT model, which requires fewer information inputs.

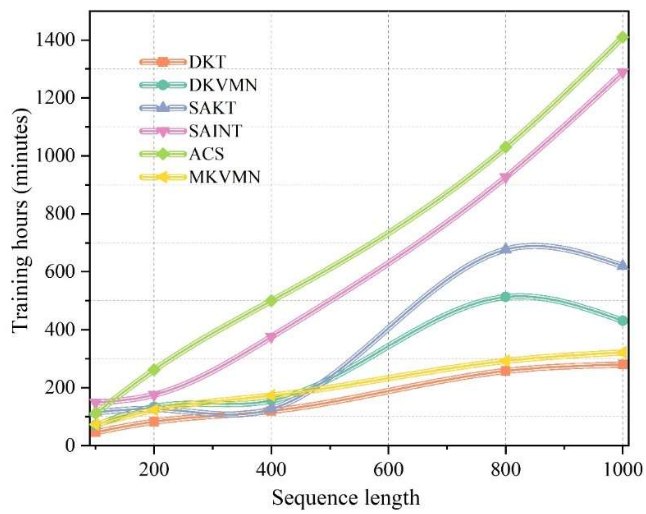


Fig. 2. The duration of the model training in different sequence lengths

3.1.2. Visualization of forecast results. In order to visualize the accuracy of the prediction results of the MKVMN model, this paper selected a student's historical question-answering interaction record data, and predicted his question-answering performance at the corresponding time step using the DKT, ACS, and MKVMN models, and the results of the different model prediction results were visualized as shown in Fig. 3. The prediction results were filtered and selected representative exercises 520 and 619. In the exercise information, 520 and 619 correspond to knowledge points 30, 135, 32, 22 and knowledge points 30, 135, 32, 87, respectively, which can be seen that these two exercises have a fairly high degree of similarity.

The model DKT gives a probability prediction of 0.534 for students being able to answer the first occurrence of Exercise 520 correctly, which rather predicts the result as a possible wrong or right answer in real doing, and subsequently gives probability predictions of 0.604 and 0.592 for the reoccurrence of Exercise 520 in the sequence. In real learning behavior, if students continuously answer a certain exercise correctly or incorrectly, it indicates to a certain extent that students have a poor grasp of the knowledge contained in this question. If the model can learn and grasp the state of students' knowledge mastery, it should change the predicted probability according to the performance of answering questions. In this sample, DKT performed poorly on the task of tracking students' knowledge mastery status with predictions of future performance.

Model ACS gave a probability prediction of 0.373 for a student being able to answer the first occurrence of Exercise 520 correctly, which comes from model ACS learning better about the correlation between information about the exercise and information about the answer interaction. For the first occurrence of Exercise 520, ACS's prediction favored that students would answer the question incorrectly. To some extent, the unique way of information embedding and the design of the attention mechanism of ACS help to improve the prediction performance of the model.

The model MKVMN designed in this paper has good prediction performance on this sample. When Exercise 619, which is highly similar to Exercise 520, appears for the first time, the MKVMN model tends to give a prediction of a possible correct answer with a probability of 0.615, based on previous interactions where students have answered Exercise 520 correctly. This is not the same as the prediction given by the model ACS. In terms of model design, MKVMN has multi-hot information embedded in the network for the information on the knowledge points of the exercises, and the model may have learned the correlation information between the exercises and the knowledge points to give more accurate predictions when encountering similar exercises.

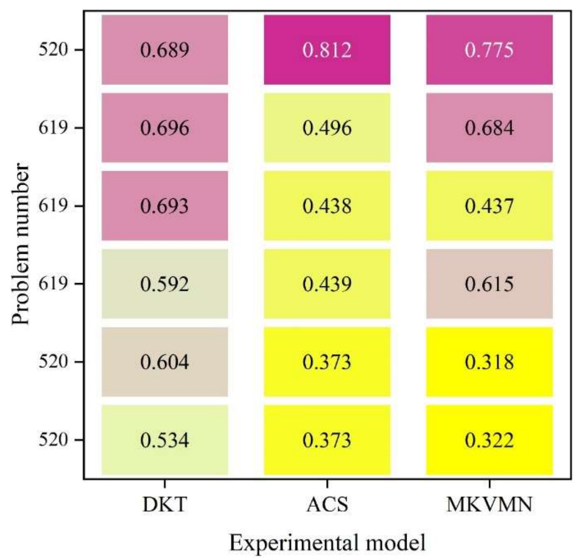


Fig. 3. Different models predict the results of the results

3.1.3. Comparison of parameter values. AUC refers to the area under the receiver operating characteristic (ROC) curve (AUC), which is usually used as a highly targeted knowledge tracking evaluation metric. For the experimental group, this paper tries to record the preservation of the optimal MKVMN model trained under the influence of different values of hyperparameters. The AUC values of the models under different hyperparameter values are shown in Table 1. Figure 4 shows the performance of the model under different values of hyperparameters. The experimental results prove that the adoption of the sigmoid function improves the model performance, and also verifies that the setting of the sigmoid function enables the model to enhance the ability to learn difficult samples during the training process, which helps to solve the problem of poor model fitting ability caused by the sample class imbalance problem in the knowledge tracking task.

Table 1. The AUC value of the model under different hyperparametric value

Hyperparametric value	0.9	1.1	1.3	1.5	1.7	1.9	2.1	2.3	2.5	2.7	2.9	3.1	3.3
AUC	0.756	0.779	0.819	0.847	0.846	0.853	0.856	0.853	0.848	0.846	0.840	0.835	0.774

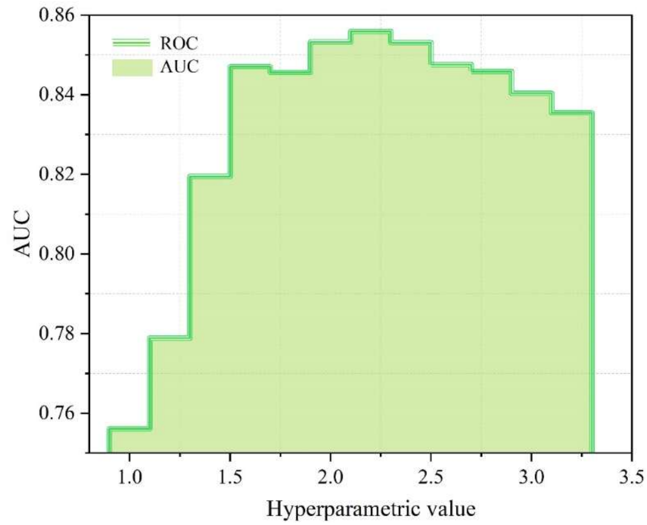


Fig. 4. Model performance of different hyperparametric value

3.2. *Effectiveness analysis of improved ant colony algorithm*

Resource preparation: This paper downloads the course resources from the remote guidance platform that meet the learning duration, defines the resource as a learning unit, and sets the number of learning knowledge of each unit to 5 to ensure that the learning is completed within 15 minutes. After screening and categorization, a total of 50 learning units and 250 learning knowledge quantities were obtained. In order to reflect the degree of improvement of teachers' and students' information technology teaching ability, the 50 learning units are divided into 3 learning areas: basic technology literacy, technical support learning and technology support learning.

Experimental subjects: 50 teachers and students with little difference in pre-existing knowledge level and related geographic areas are selected and divided into two groups for comparative validation experiments, the experimental group and the control group are both 25 people, and the 50 learning units are completed online. The control group chooses the initialized learning path to learn, and the experimental group chooses the personalized learning path with the improved ant colony algorithm to learn.

The test scores of teachers and students in the two groups in the three domains of basic technological literacy domain, technological support learning domain and technological support teaching domain are obtained as shown in Figs. 5 to 7, respectively. It can be seen that during the period of 1-10 learning units, the difference in the growth rate of the two groups' scores is small; during the period of 11-50 learning units, the experimental group's scores grow significantly larger than those of the control group. During the period of 31-50 units of study, the magnitude of change in the performance of the experimental group was relatively smaller compared to the period of 11-35 units. Teachers' and students' grades were generally improved and there was a positive correlation between teachers' and students' grades and the number of units in the learning domain.

From the analysis of the simulation results, it can be seen that: in the process of the pre-learning period, the experimental group showed no significant learning advantage, because the learning performance has a greater impact on the category of learning content. In the middle of the learning process, the experimental group has a more obvious growth compared with the control group. During the later stages of learning, the experimental group's learning performance became increasingly different from that of the control group. It is hypothesized that this is related to the fact that the Civics Knowledge Map provided under the algorithm of this paper helps students to form an efficient learning mode as well as a clear and explicit knowledge framework.

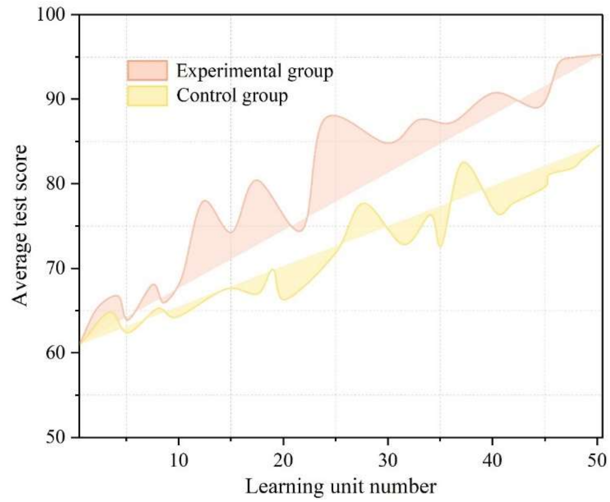


Fig. 5. Learning testing in basic technical literacy areas

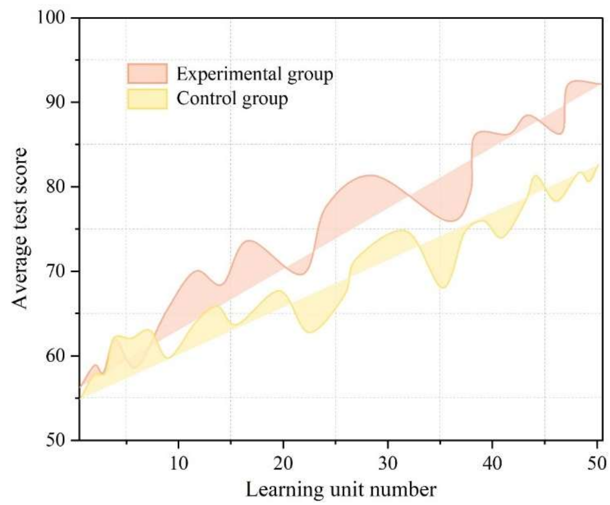


Fig. 6. Technical support learning test results

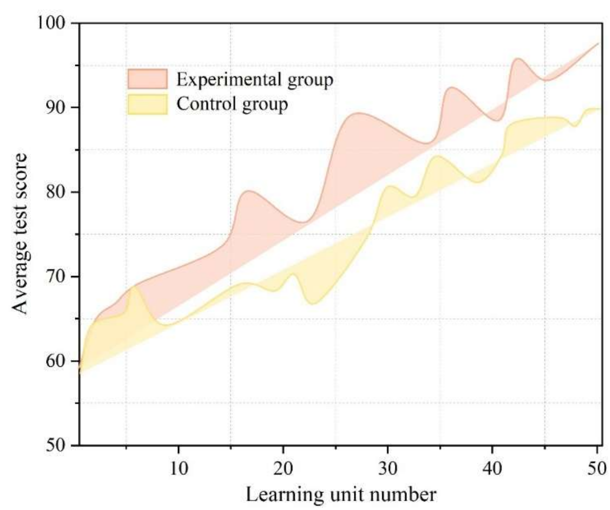


Fig. 7. Technical support teaching learning test situation

3.3. Recommendation of Learning Paths under the Innovation of Teaching Methods of Civic and Political Education in Colleges and Universities

In order to evaluate the learning path recommendation algorithm based on knowledge tracing proposed in this paper, four datasets, Mynereus, ASSISTments2020, ASSISTments2024 and Ednet, are chosen for experiments, which have been carefully selected to comprehensively cover many domains such as Mathematics, Computing, Programming, etc., and the specific information of the datasets is shown in Table 2.

Table 2. Specific information of the data set

Data set	Topic number	Learner number	Number of knowledge	Make a list of questions
Mynereus	171	230	64	87158
ASSISTments2020	13008	4137	94	326257
ASSISTments2024	9059	4248	50	160623
Ednet	11166	5009	202	1080766

In order to evaluate the learning path recommendation algorithm proposed in this paper, some traditional sequence prediction based learning path recommendation algorithms are selected as reference models for comparison:

KNN, as a classification algorithm, can be used to cluster learners, using the idea of collaborative filtering, calculating the target group closest to the learner in the learner community by cosine distance, and recommending the learning path of the current learner based on the doing records of the similar target group as the topic recommendation candidate set.

GRU4Rec is a sequence recommendation model modeled based on RNN, which utilizes all the information of the learner's historical learning records, and inputs the embedded representations of the topics in the historical learning records into the recurrent neural network sequentially, so as to make a prediction of the next jump of the learning path.

Learning path recommendation needs to consider the effectiveness of the learning path and the learner's interest, so in this experiment, the learning stage effectiveness is chosen as a measure.

3.3.1. Effectiveness of recommending learning paths of different lengths. In order to measure the difference in the effectiveness of each model for different lengths of learning path recommendation results, the length of 5, 15, 30, 40 learning paths were recommended, and the effectiveness results of the different lengths of learning paths recommended by each model are shown in Table 3. Through the experimental results, it can be seen that compared with the traditional learning path recommendation algorithm based on sequence prediction, the recommended learning paths significantly lag behind the learning path recommendation algorithm based on knowledge tracking proposed in this paper in terms of knowledge enhancement for learners. Moreover, the IACS-PRA learning path recommendation algorithm proposed in this paper especially improves significantly on long path recommendation, which may be due to the fact that the sequence prediction-based recommendation tends to recommend similar exercises continuously, thus causing the learning path to fall into some kind of localization.

Table 3. Recommend the effectiveness of different length learning paths

Model	Path length	Mynereus	ASSISTments2020	ASSISTments2024	Ednet
KNN	5	0.1084	0.1272	0.1306	0.1315
	15	0.1656	0.1848	0.1604	0.1777
	30	0.2155	0.2159	0.2154	0.2409
	40	0.2354	0.2533	0.2306	0.2854
GRU4Rec	5	0.1137	0.1314	0.1297	0.1636
	15	0.1908	0.1617	0.1535	0.1872
	30	0.2150	0.2038	0.1869	0.211
	40	0.2718	0.2239	0.1734	0.2261
IACS-PRA	5	0.1641	0.1823	0.1628	0.1928
	15	0.2124	0.2286	0.2262	0.2313
	30	0.2871	0.2987	0.2641	0.3538
	40	0.3694	0.3775	0.3513	0.4141

3.3.2. Analysis of Learning Path Recommendation Effectiveness under Different Methods. In this paper, the ASSISTMents2020 dataset is selected to test the effect of the improved ant colony algorithm on the visualization of students' learning path recommendation results. The traditional KNN and GRU4Rec recommendation algorithms are used in this paper to compare with the IACS-PRA algorithm in this paper to analyze the recommendation effect of learning paths under the three recommendation algorithms.

The prerequisite relationships between the knowledge points are shown in Figure 8. The visualization results of the learning path recommendation of three different methods are shown in Fig. 9, where (a) to (c) represent the KNN, GRU4Rec and IACS-PRA methods, respectively. The results show that the learner's latest learning record is situated on Knowledge Point 35, while the learner's learning goal is set on Knowledge Point 100, and the KNN, GRU4Rec and IACS-PRA methods are used to visualize the knowledge points to which the topics in the learner's recommended learning path belong. From the results, it can be found that the KNN and GRU4Rec methods have been recommending topics under one knowledge point, while the IACS-PRA method proposed in this paper recommends knowledge points in a gradual manner through the prerequisite relationship, and continuously consolidates and reviews the knowledge points just mastered during the learning process to obtain better learning results. All in all, the IACS-PRA recommendation method proposed in this paper has a significant improvement effect compared with the traditional sequence recommendation method in different lengths.

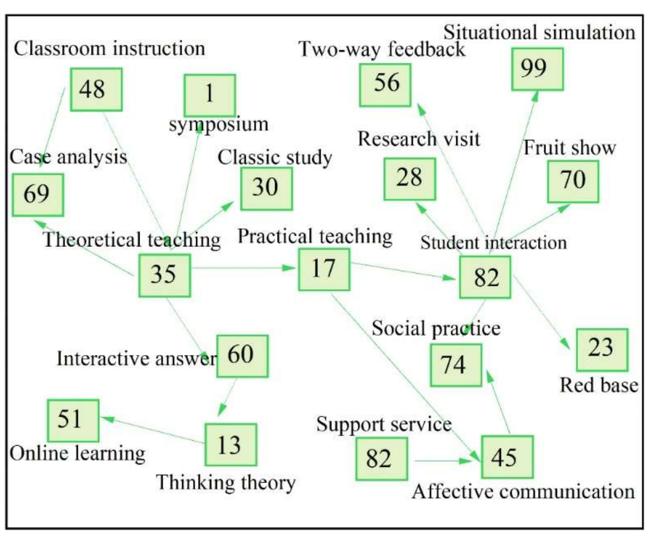
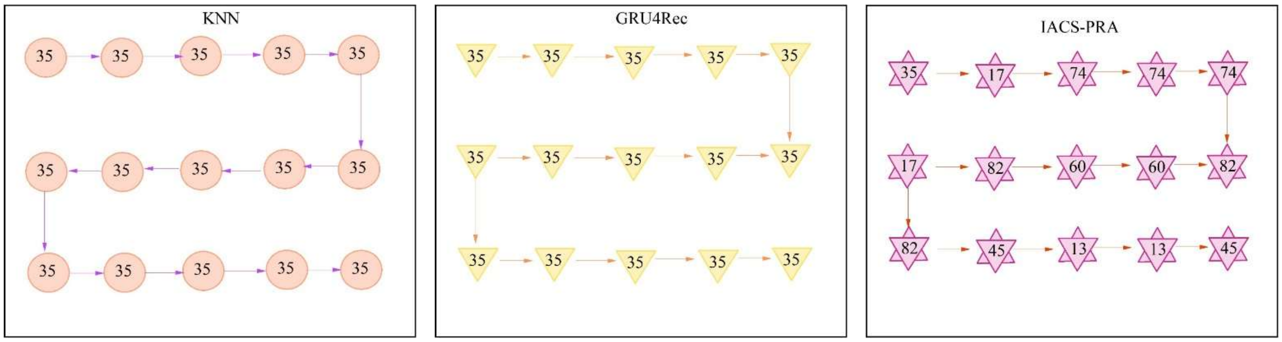


Fig. 8. Prerequisites between knowledge points



(a)KNN (b)GRU4Rec (c)IACS-PRA

Fig. 9. The visual results of the three different methods of learning paths

4. Conclusion

In this paper, under the deep neural network technology, it is proposed to combine knowledge mapping technology, knowledge tracking model and improved ant colony algorithm to construct a key-value memory network knowledge tracking model based on multi-feature fusion to research and personalized recommendation on the innovation of the teaching method of Civic and Political Education in colleges and universities. The results show that:

1) The result of training time comparison verifies the excellent performance of the model in knowledge tracking ultra-long sequence processing; the result of visualization of prediction results verifies that the design of the model's architecture and mechanism has an improvement effect on the model's fitting ability and prediction performance; the result of comparison of parameter values verifies that the design of the model's loss function is conducive to solving the problem of sample class imbalance that exists in the task of knowledge tracking, and it effectively improves the model's performance.

2) The application effect of the IACS-PRA model shows that during the period of 1-10 learning units, the difference between the two groups' performance growth is small; during the period of 11-50 learning units, the growth of students' performance using the IACS-PRA model is significantly larger than that of the control group, and it is obvious that the use of the algorithm in this paper can obviously improve students' Civic and Political Science learning performance.

3) The learning path generated by the learning path recommendation algorithm introduced in this paper is more helpful to improve the knowledge level of the learners; it recommends the Civics and Political Science knowledge points in a gradual and progressive way, thus obtaining better learning results.

References

[1] Wang, J. (2021). Analysis of challenges and countermeasures of ideological and political education in colleges and universities in the new era. J. High. Educ. Res, 2(6), 35.

[2] Wang, Z. (2019). Research on the Reform Method of Ideological and Political Courses in Colleges and Universities under the Background of Quality Education. In Institute of Management Science and Industrial Engineering. Proceedings of 2019 International Conference on Arts, Management, Education and Innovation (ICAMEI 2019) (Vol. 5). Institute of Management Science and Industrial Engineering:(Computer Science and Electronic Technology International Society).

[3] Zhu, G., & Tao, T. (2021). Strategies for Ideological and Political Education in Colleges and Universities from the Perspective of General Education. Academic Journal of Humanities & Social Sciences, 4(11), 18-23.

[4] Zhang, J. (2024). Innovative Pedagogical Approaches in University Ideological and Political Education. International Journal of Educational Research, 127, 63626

[5] Zhang, A., Sheng, L., & Wang, J. (2024). Empirical Research on Curriculum-based Ideological and Political Teaching of Universities. Int. J. Educ. Humanit, 13, 34-38.

[6] Sun, W., & Wang, H. (2023). Discussion on the Curriculum Organization and Teaching Method for Curriculum Ideological and Political Education. Frontiers in Educational Research, 6(23).

[7] Xinning, Y. (2022). Reflection and reconstruction of practical teaching mode of Ideological and political course in Colleges and Universities. International Journal of New Developments in Education, 2022, 4.0 (2.0).

[8] Cui, Y. (2023). Exploring the Competence Enhancement of College Ideological and Political Course Teachers. Frontiers in Educational Research, 6(15).

[9] Sze, V., Chen, Y. H., Yang, T. J., & Emer, J. S. (2017). Efficient processing of deep neural networks: A tutorial and survey. Proceedings of the IEEE, 105(12), 2295-2329.

[10] Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J., & Müller, K. R. (2021). Explaining deep neural networks and beyond: A review of methods and applications. Proceedings of the IEEE, 109(3), 247-278.

[11] Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. Neurocomputing, 234, 11-26.

[12] Ang, K. M., El-Kenawy, E. S. M., Abdelhamid, A. A., Ibrahim, A., Alharbi, A. H., Khafaga, D. S., ... & Lim, W. H. (2022). Optimal design of convolutional neural network architectures using teaching–learning-based optimization for image classification. Symmetry, 14(11), 2323.

[13] Gupta, V., & Pawar, S. (2022). An effective structure of multi-modal deep convolutional neural network with adaptive group teaching optimization. Soft Computing, 26(15), 7211-7232.

[14] Careem, R., Johar, G., & Khatibi, A. (2024). Deep neural networks optimization for resource-constrained environments: techniques and models. Indonesian Journal of Electrical Engineering and Computer Science, 33(3), 1843-1854.

[15] Reyad, M., Sarhan, A. M., & Arafa, M. (2023). A modified Adam algorithm for deep neural network optimization. Neural Computing and Applications, 35(23), 17095-17112.

[16] Abdolrasol, M. G., Hussain, S. S., Ustun, T. S., Sarker, M. R., Hannan, M. A., Mohamed, R., ... & Milad, A. (2021). Artificial neural networks based optimization techniques: A review. Electronics, 10(21), 2689.

[17] Zhengyang Wu,Yuqi Liu,Jianwei Cen,Zetao Zheng & Guandong Xu. (2024). A cross-domain knowledge tracing model based on graph optimal transport. World Wide Web,28(1),10-10.

[18] Leonie V D E Vogelsmeier,Irina Uglanova,Manuel T Rein & Esther Ulitzsch. (2024). Investigating dynamics in attentive and inattentive responding together with their contextual correlates using a novel mixture IRT model for intensive longitudinal data..The British journal of mathematical and statistical psychology,

[19] Simin Ghasemi,Vahid Rafe,Anvar Bahrampour & Reiko Heckel. (2025). Whole test suite generation from graph transformation specifications using ant colony optimization. Software Quality Journal,33(1), 10-10.