J. COMBIN. MATH. COMBIN. COMPUT. 127b (2025) 359--378

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

Resource Allocation and Optimal Path Selection of Civic and Political Education Curriculum Based on Multi-Objective Optimization

Wei Zheng[1], Qinghua Lu[2,✉]

[1] *Student Affairs Office, Hunan Railway Profession College, Zhuzhou, Hunan, 412000, China*
[2] *School of Marxism, Hunan Railway Profession College, Zhuzhou, Hunan, 412000, China*

ABSTRACT
This paper combines the project response theory to dynamically adjust and update the resources according to the learning effect and learning feedback in the process of Civic Education, so as to achieve the goal of matching the learners with the learning resources and realize efficient learning. The differential artificial raindrop algorithm based on perturbation mechanism is designed to realize the solution of multi-objective combinatorial optimization of learning resource allocation. Performance experiments show that the convergence curve of the resource allocation algorithm in this paper is gradually flattened, and the algorithm still has the evolutionary ability, the convergence curve is still decreasing, and the final characteristic difference value is also better than other BPSOR and GAR algorithms. In the case of the number of learning resources of 10, 20, 30, 50, 100, the time consumed is 207ms, 1602ms, 20506ms, 68430ms, 354687, all of which are the lowest, and the success rate is also the highest in the model. The optimal learning path is applied to an experimental class in a university for a 6-week teaching experiment, and the experimental class scores 87.2 points in the Civics test, which is much higher than the control class. This paper realizes the accurate capture of students' Civics learning problems and the recommendation of targeted teaching resources, which can improve the quality and effect of Civics teaching.

Keywords: cognitive diagnostic model, multi-objective optimization, differential artificial raindrop, Civics teaching

1. Introduction

The concept of education in the new period has changed the traditional education concept that education in the past was based on general education, and the Civic and Political Education in colleges

✉ Corresponding author.
E-mail address: hntd_zw@sina.com (Q. Lu).
Received 28 December 2024; Revised 10 February 2025; Accepted 25 March 2025; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-022

and universities, as an important part of the cause of educating people in the new period, is facing another transformation [1-2]. In order to better cultivate talents in the new era, it is necessary to strengthen the resource allotment of the Civic and Political Education in colleges and universities, so that the educators and the educated subconsciously strengthen the cognition of the importance of the Civic and Political Education [3-4]. At the same time, in order to better serve teachers and students, ideological education in colleges and universities must take a scientific, holistic and sustainable path, and must have sustainable utilization of resources and a correct scientific view of resources. Therefore, the deepening of resource consciousness has become the primary task of the reform of ideological and political education in colleges and universities [5-7].

Ideological and political education resource allocation is a support path to provide educational resources and educational tools for ideological and political education work, in order to guarantee that ideological and political classrooms can be carried out smoothly and realize the goals of Chinese ideological and political education work [8-9]. In China's educational environment, ideological and political courses in colleges and universities have always been the focus of work emphasized by the state, and the government and education authorities have set up perfect institutions, policies and work systems for ideological and political education, and given financial support to actively implement the work of ideological and political education in colleges and universities [10-12].

First of all, in the current curriculum planning of ideological and political education in colleges and universities, there is a situation of resource fragmentation. The so-called resource decentralization refers to the fact that the resource allocator configures each resource with a non-holistic idea, so that they can play their respective functions, but it is difficult to fail to produce the overall utility [13-14]. Secondly, another problem in the allocation of resources for ideological education in colleges and universities is the idleness of resources. Because of the lack of a correct scientific view of resource utilization, educators in colleges and universities often do not rationally plan, allocate and use the teaching resources that are allocated in the society and in the school itself [15-16]. Finally, the shortage of resources is a unique problem in the Civics education in colleges and universities. Compared with other courses, teaching administrators obviously do not pay enough attention to this aspect of Civic and Political Education, and Civic and Political Education also lacks the corresponding high-quality talents, educational funds and good educational equipment [17-18].

In this paper, we first algorithmically model the three types of features, namely learning objectives, learning styles and cognitive levels, and on this basis, we constructed a recommendation model for Civic and Political Education Resources and Learning Paths by means of mapping relationships. Aiming at the multi-objective resource recommendation and allocation problem that needs to be solved to design the differential artificial raindrop algorithm based on the perturbation mechanism, on the basis of the original artificial raindrop algorithm, the collision rate is used to regulate the parameters of the collision process, and the differential strategy with perturbation is introduced into the flow operator. The designed model is compared with two baseline models, BPSOR and GAR, and three classes are selected in a university for experiments to verify the effectiveness of optimal path generation.

2. Method

2.1. Cognitive Diagnostic Model

This section focuses on two commonly used models in the field of cognitive diagnosis, the item response theory model and the DINA model.

2.1.1. Theoretical model of project response. Item Response Theory (IRT) is a measurement model based on probabilistic statistics, which is mainly used to assess the level of competence of the examinees. In the IRT model, each exam question has a difficulty parameter and a differentiation parameter. The difficulty parameter reflects the degree of difficulty of the exam question, and the differentiation parameter reflects the degree to which the exam question is able to differentiate between subjects of different ability levels. By calculating the subjects' answers, the IRT model can more accurately infer their ability levels and assess the difficulty and differentiation of the questions.The main idea of IRT is that the difficulty and differentiation of the test questions are fixed, while the ability levels of the subjects are randomized. IRT can be used to estimate the ability level of subjects by establishing a link between their ability level and the difficulty of the test questions, and it can be used to assess parameters such as the difficulty and differentiation of the test questions based on the characteristics of the test questions. Therefore, IRT model has been widely used in the fields of educational measurement, psychometrics, and medical diagnosis.

Currently, there are three main IRT models, which are the one-parameter model (1PL), the two-parameter model (2PL), and the three-parameter model (3PL). Among them, the one-parameter model considers the differentiation parameter of the test questions to be a constant value, the difficulty parameter of the test questions to be the only variable, and the scores of the test questions to be related only to the ability level of the subjects and the difficulty of the test questions. The two-parameter model model, compared to the 1PL model, adds the differentiation parameter of the test questions, the differentiation parameter of the test questions represents the ability to differentiate between the abilities of the test questions, and the score of the test questions is not only related to the ability level of the subjects and the difficulty of the test questions, but also related to the differentiation parameter of the test questions. The three-parameter model adds a guessing parameter compared to the 2PL model, and the test-question guessing parameter represents the probability that subjects may guess by chance even when they have not mastered the knowledge point. The score of the test question is not only related to the ability level of the subject, the difficulty of the test question and the test question differentiation parameter, but also related to the test question guessing parameter.IRT model is one of the most widely used measurement models, of which the more representative ones are the Rasch model and the Logistic model. Among them, Rasch model is a one-parameter model in IRT model, where the difficulty parameter of the test questions is the only variable and the differentiation parameter of the test questions is a constant, so people call it as a one-parameter model, and the formula is shown in equation (1).

$$p_i(\theta)=\frac{1}{1+e^{-(\theta-b_i)}} \tag{1}$$

where θ represents the subject's potential trait, i.e., ability value; b is the difficulty of the question; and $p_i(\theta)$ is the probability that the student will answer the test question correctly [19].

Logistic model is a three-parameter model in IRT model, also known as three-parameter logistic model or 3PL model. Unlike the 1PL model and 2PL model, the Logistic model adds a guessing parameter, and the guessing parameter of a test question represents the probability that the subject may guess by chance even when he or she has not mastered the knowledge point. The score of the test question is not only related to the ability level of the subject, the difficulty of the test question and the test question differentiation parameter, but also related to the test question guessing parameter, whose formula is shown in (2).

$$p_i(\theta)=c_i+\frac{1-c_i}{1+e^{-Da_i(\theta-b_i)}} \tag{2}$$

where a,b,c denotes the topic parameters, respectively, differentiation, difficulty, and guessing; D is a constant.

The advantage of the IRT model is that it can take into account the differences in users' abilities and attributes, introduce more a priori knowledge, better adapt to the changes in the test question pool and the characteristics of the test subjects, and to a certain extent, it can more accurately estimate the ability level of the test subjects, so as to realize personalized recommendation. However, the IRT model has high requirements for test data and test questions, and in the case of small samples, a large error in ability estimation may occur.

2.1.2. DINA model. The DINA model is a statistical model applied to measurement, assessment, and testing, and it is a model constructed based on the findings of cognitive psychology theory.The DINA model takes into account the skill level of the subject and the characteristics of the test questions, and it attempts to infer whether or not a subject has mastered a particular skill through the performance of the test questions. Unlike the traditional IRT model, the DINA model assumes that the subject's skill level is binary, i.e., whether or not the subject has mastered a particular skill is determined, while the test questions can include multiple skill elements. In the DINA model, each test question has a skill matrix that identifies the skill elements included in the test question. The skill level of a subject can be estimated from the skill matrix and the performance of the test questions, while the skill matrix can be determined by methods such as maximum likelihood estimation.

A probabilistic model of students' cognitive mastery is constructed using the matrix of students' correct and incorrect answers to test questions and the Q matrix of knowledge units associated with the test questions. In this model, the student i's knowledge unit mastery vector a_i uses 0 and 1 to indicate whether the knowledge unit k is mastered or not, and the test question correlation vector q_a indicates whether the test question j examines the knowledge unit k. Then the student i's response in the test question j is n_{ij}, and its formula is shown in (3):

$$n_{ij} = \prod_{k=1}^{K} a_{ik}^{q_u} \tag{3}$$

where $n_{ij} = 1$ indicates that student i answered test question j correctly, which represents the knowledge unit contained in test question j that student i meta-mastered; 0 indicates that the answer is wrong, and at least one knowledge unit is not mastered for test question j. The formula of the DINA model is shown in (4), which represents the probability that student i answered test question j correctly.

$$P\left(x_{ij} = 1 \mid a_i, q_i, n_{ij}\right) = g_j^{\left(1-n_{ij}\right)} \left(1 - s_j\right)^{n_i} \tag{4}$$

where g, denotes the guessing rate, i.e., the probability of a student guessing correctly without fully grasping all the knowledge points of test question j; s_j denotes the probability of a student answering incorrectly while grasping the knowledge points examined in test question j; and x_{ij} denotes the scoring of a student i on test question j as 1 answered correctly and as 0 answered incorrectly.

The advantages of the DINA model are that it can accurately assess the subject's mastery on each skill element, and it can analyze the impact of the skill elements in the test questions on the subject's performance on the test.The DINA model has been widely used in the fields of educational measurement, human resource management, and clinical assessment.

2.2. Learning Resource Recommendation Model

2.2.1. Parametric representation of data. According to the learning resource recommendation process, the main attribute information involved in this paper includes learner attributes and learning resource attributes. The learner parameters are defined as follows.

1) Learners $L = \{l_1, l_2, \ldots, l_U\}$, U is the overall number of learners, l_u denotes the u th learner, $1 \le u \le U$.

2) Learners expect to learn the target knowledge points $K=\{k_1,k_2,\ldots k_K\}$, K is the number of target knowledge points, k_k indicates the learner's k th knowledge point to be learned, $1\le k\le K$.

3) Learning style is a personal characteristic of the learner's performance, which reflects the learner's preferred learning behavior. The learning style of a particular learner l_u satisfies the constraint formula $\sum_{q=1}^{8} le_{uq}=1$.

4) Generally speaking, in cognitive psychology, the process of modeling learning ability is called "cognitive diagnostic model", and by adopting the cognitive diagnostic model, we can obtain the learner's cognitive status of the knowledge points more clearly and accurately. Definition of cognitive level $lc=\{lc_1,lc_2,\ldots,lc_U\}$ represents the cognitive level of U learners, and $lc_u=\{lc_{u1},lc_{u2},\ldots,lc_{UK}\}$ represents the cognitive status of learner l_u on the k th knowledge point.

Learning resource attributes mainly describe the elements of resource characteristics, including resource difficulty, media type and relevance and other information, learning resource parameters are defined as follows.

1) Define learning resources $O=\{o_1,o_2,\ldots,o_N\}$, N as the number of resources, o_n as the n th learning resource, $1\le n\le N$, based on the knowledge point to be learned by the learner $\{o_{k1},o_{k2},\ldots,o_{KN}\}$, o_{kn} as the n th resource corresponding to the k th knowledge point.

2) Define the difficulty of the learning resources $D=\{d_{k1},d_{k2},\ldots,d_{KN}\}$, d_{kn} indicates that the k rd knowledge point corresponds to the difficulty information of the n th resource.

3) Define the type of learning resources $E=\{e_1,e_2,\ldots,e_P\}=\{\text{Text, Hypertext, Image, Audio, Video}\}$, each resource contains not a single type of performance, $Oe=\{oe_{km1},oe_{kn2},\ldots,oe_{kNP}\}$, oe_{knp} indicates the k th knowledge point corresponds to the type of resource contained in the n th resource, $\sum_{p=1}^{5} oe_{k\varphi p}=1$.

4) Define the relevance of the learning resource to the knowledge point $V=\{v_{k1},v_{k2},\ldots,v_{KN}\}$, v_{kn} indicates the relevance of the n th resource to the k th knowledge point.

2.2.2. Learner model. The learner model is one of the core components to achieve personalized learning resource recommendation.

1) Learning Objectives. According to one's own subjective will and learning planning, the learning objective mainly describes the target knowledge point that the learner expects to learn as well as the expectation value for the knowledge point.

2) Learning Style. Learning style is a personal preference for learning methods, which describes the personalized characteristics of learners and the main reasons for the formation of different personal differences.

Learning Behavior Patterns are defined as follows, with b_i referring to the learner's i nd behavior pattern.

$$b_i^u=\begin{cases}1 & b_i^u\in H\\ 0 & b_i^u\in M\\ -1 & b_i^u\in L\end{cases} \tag{5}$$

The type of learning style of a learner is not determined by one learning behavior; it is influenced by a combination of multiple learning behavior patterns, and based on the definition of behavior patterns and the quantification of learning behaviors, the learning style of Learner u is calculated as follows.

$$le_{uq} = \frac{\sum_{i=1}^{b} b_i^u}{b} \tag{6}$$

3) Cognitive level. Two matrices were used as inputs to DINA, i.e., R matrix and Q matrix.

q_{jk} represents the examination of the knowledge points in the test questions and $q_{jk} = 0$ indicates that the test questions were not examined. $a_{uk} = \{a_{u1}, a_{u2}, \ldots, a_{UK}\}$ describes the learner's mastery vector of the knowledge point, and $a_{uk} = 0$ indicates that the learner u has not mastered it. When the learner's knowledge point mastery vector is a known condition, Eq. (7) is utilized to obtain the learner's potential test question responses.

$$\eta_{uj} = \prod_{k=1}^{K} a_{uk}^{q_{jk}} \tag{7}$$

The DINA model equation is as follows.

$$P_j(a_i) = g_j^{1-\eta_{xj}} (1 - s_j)^{\eta_{xj}} \tag{8}$$

In this paper, we use the posterior probability of a under the condition of full permutation 2^k to obtain students' continuous cognition, which is between 0 and 1. The closer to 1 indicates that students' mastery of knowledge is higher, and students' u mastery of knowledge point k is calculated as follows.

$$\begin{aligned} lc_{uk} = Pa_{uk} = 11X_i &= \frac{\sum_{a_{si=1}} P(a_x \mid X_i)}{\sum_{x=1}^{2^i} P(a_x \mid X_i)} \\ &= \frac{\sum_{a_{ut}=1} \prod_{j=1}^{J} L(X_{ij} \mid a_x, \hat{s}_j, \hat{g}_j) Pa_x}{\sum_{x=1}^{2^i} P(a_x \mid X_i)} \end{aligned} \tag{9}$$

2.2.3. Learning Resource Recommendation Model Construction. In order to recommend personalized learning resources to learners, on the one hand, it is necessary to target learners' feature data, and on the other hand, it is also necessary to consider the feature data of resources. In this paper, we study the personalized learning resources recommendation optimization problem based on the mapping relationship.

1) Feature data matching degree calculation. Combined with the quantized learning style, the matching degree between it and the resource type information features is calculated as follows.

$$f_{es} = \frac{\sum_{q-1}^{Q} \sum_{p-1}^{P} \left| os_{np} - le_{uq} \right|}{P} \tag{10}$$

The formula for matching cognitive level to resource difficulty is as follows.

$$f_{dc}=\frac{\sum_{n=1}^{N}\left|d_{kn}-lc_{uk}\right|}{Q} \tag{11}$$

2) Learning cost. The learning behavior is divided into four categories, which are the learner's browsing length, number of views, downloads and favorites of the resource, and the learner's learning cost is calculated by the following formula.

$$f_c=\sum_{n=1}^{N}\left[\frac{1}{4}\left(\frac{\left|T-\sum_{i=1}^{n}t_i\right|}{T}+\frac{\left|R-\sum_{i=1}^{n}r_i\right|}{R}+\frac{\left|L-\sum_{i=1}^{n}l_i\right|}{L}+\frac{\left|H-\sum_{i=1}^{n}h_i\right|}{H}\right)\right] \tag{12}$$

3) Degree of relevance of learning resources to the knowledge points of learning objectives. The formula for calculating the degree of correlation between the resources and the knowledge points of the learning objectives is as follows.

$$f_r=\sum_{n=1}^{N}\left(x_{nu}\cdot\sum_{k=1}^{K}v_{kn}\right) \tag{13}$$

According to the above description, the mapping function is obtained from the sub-mapping function by weighting the coefficients with the following expression.

$$\min F(x)=\omega_{dc}\cdot f_{dc}+\omega_{es}\cdot f_{ex}+\omega_c\cdot f_c+\omega_r\cdot f_r \tag{14}$$

ω_{dc} , ω_{es} , ω_c , and ω_r denote the weighting coefficients, and the constraints of the objective function $\sum_{q=1}^{8}le_{uq}=1$, denotes the sum of the percentage of learning styles of the learner l_u is 1. $\sum_{p=1}^{5}oe_{knp}=1$, denotes the sum of the percentage of resource types of the learning resources corresponding to the knowledge point k is 1.

2.2.4. Overview of multi-objective optimization problems. In addition to multiple objective functions as the main components, the multi-objective optimization problem includes some equations and inequality constraints. Mathematically, it can be described as follows.

$$\begin{gathered}\min y=\left[y_1(x),y_2(x),\ldots,y_n(x)\right]\\ s_i(x)\le 0 i=1,2,3,\ldots,m\\ t_j(x)\le 0 j=1,2,3,\ldots,k\\ x=\left[x_1,x_2,x_d,\ldots,x_D\right]\\ \min x_d\le x_d\le \max x_d\end{gathered} \tag{15}$$

y represents the objective vector, x represents the decision vector, and n represents the total number of objectives to be optimized. Among the constraints, $s_i(x)\le 0$ represents the i th constraint and $t_j(x)\le 0$ represents the j th constraint. In the elements constituting the objective vector, $y_n(x)$ is the n th objective function and x is the space formed by the decision vector. $\min x_d$ and $\max x_d$ are the upper and lower bounds of the search space for each dimension of the vector, and $s_i(x)\le 0$ and $t_j(x)\le 0$ define the feasible domain.

2.2.5. Differential artificial raindrop algorithm based on perturbation mechanism. The differential artificial raindrop algorithm (ADARA) based on the perturbation mechanism is also a heuristic algorithm that simulates the variation of individual raindrops and transfers the effective information processing mechanism of the rainfall process to the optimization algorithm [20].

1) Coding design. The purpose of coding is to represent the feasible solution of the problem in a matrix of definite dimension, and in this paper, integer coding is chosen to initialize the raindrop population:

$$O^0 = \begin{pmatrix} o_{11}^0 & \cdots & o_{1D}^o \\ \vdots & \ddots & \vdots \\ o_{N1}^0 & \cdots & o_{ND}^0 \end{pmatrix} \tag{16}$$

The initial raindrop population is V^0, the number of individuals in the population is N, V^i is the i th generation population, and a certain individual vector is defined as follows:

$$o_i^G = \left[o_{i1}^0, o_{i2}^0, o_{i3}^0, \ldots, o_{iD}^0 \right] \tag{17}$$

o_i^G denotes the i rd individual of the Gnd generation population, which contains D structures, and the value at each structure is an integer, denoting the resource number corresponding to the resource, and all the raindrop individuals form the set of candidate solutions to the optimized problem [21].

2) Parameter tuning collision rate. Combined with the Civics resource recommendation problem, the potential energy of the raindrop reference after falling to the ground is the value of the objective function of the resource recommendation model, $F(x)$ is the resource recommendation objective function, and PE is the potential energy of the individual, so the raindrop operator corresponding to the smallest value of PE is the high-quality solution of this model.

$$PE = F(x) \tag{18}$$

Based on the above description, the collision rate is calculated as follows, where A is the optimal value of the underlying activation function.

$$P_c = \begin{cases} \dfrac{1}{1 + \exp\left(A \left(\dfrac{2\left(ColRain_r^G - ColRain_{avg}^G \right)}{ColRain_{bat}^G - ColRain_{avg}^G} \right) \right)} & f \geq f_{avg} \\ rand(0.5, 1) & f < f_{avg} \end{cases} \tag{19}$$

3) The flow operator introduces a perturbation strategy. Raindrop flow operator. The original flow operator is utilized for flow, and the flow operator is specified as follows.

$$Raindrop_i^G = \begin{cases} ColRain_{best}^G + rand_a \cdot \left(ColRain_{r1}^G - ColRain_{r2}^G \right) rand > 0.5 \\ ColRain_{r0}^G + rand_b \cdot \left(ColRain_{r1}^G - ColRain_{r2}^G \right) rand \leq 0.5 \end{cases} \tag{20}$$

A perturbation strategy needs to be added to the differential flow operator, and the computational expression is given below.

$$Raindrop_i^G = ColRain_{\min}^G + rand^2 \cdot \left(ColRain_{\max}^G - ColRain_{\min}^G \right) \tag{21}$$

Combining the above algorithms, Figure 1 shows the ADARA whole question flow.

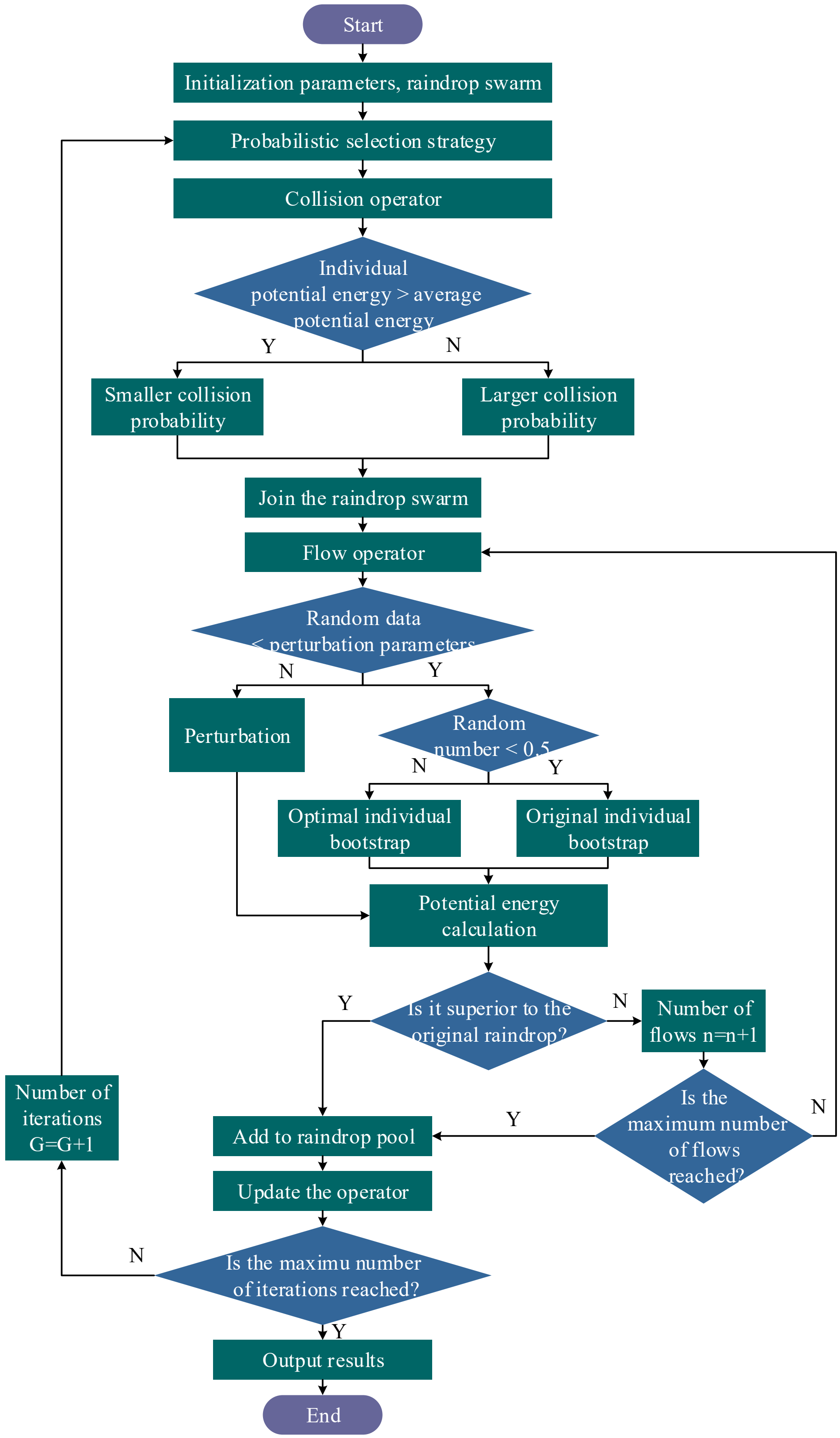


Fig. 1. The flow chart of ADARA

2.3. Adaptive Learning Path Generation Method

Firstly, we construct a knowledge graph model combining fine-grained knowledge elements and students' cognitive ability, then we generate a new goal subgraph through students' current knowledge mastery, then we obtain all paths from the current node to the goal node through topological sorting to compose a path state set, and finally, we obtain the optimal paths through the reinforcement learning method.

2.3.1. Domain Knowledge Mapping Models. Different from the traditional knowledge map with ternary as the basic constituent unit, the goal of the knowledge map constructed in this paper is to combine the fine-grained knowledge element with the cognitive ability of students. And at present, the simplest and most effective way to test the students' cognitive ability for this knowledge element is to understand it through the mastery of the exercises containing this knowledge element, so the knowledge base model can be defined as a hexadecimal group, denoted as:

$$G=\left(K1, relation, K2, QS, CA, chapter\right) \tag{22}$$

where: $K1$ and $K2$ denote fine-grained knowledge elements in a course; $relation$ denotes the relationship between $K1$ and $K2$; QS denotes the set of questions for knowledge element $K1$; CA denotes the cognitive ability to master $K1$; and $chapter$ denotes the chapter unit to which $K1$ belongs. Next, the concepts that appear in the model are defined one by one.

Definition 1 Fine-grained knowledge element is the smallest independent knowledge unit in a subject area that cannot be subdivided and can completely express knowledge, such as the “single linked table” in the data structure course is a fine-grained knowledge element.

Definition 2 $relation$ denotes the relationship between knowledge points, for which the relationship is defined as: dependent $\left(rely\right)$ denotes that a knowledge point depends on another knowledge point, and the two knowledge points have a necessary back-and-forth sequential relationship, or that the existence of a knowledge point must depend on the existence of another knowledge point; dependent $\left(b-rely\right)$ denotes that a knowledge point depends on another knowledge point, and the two knowledge points have a necessary back-and-forth sequential relationship; Synonymy $\left(syno\right)$ means that the two points have different names but refer to the same content; antonymy $\left(auto\right)$ means that the two points have opposite meanings; proximity $\left(si-mi\right)$ means that the two points have similar content.

Definition 3 QS the set of exercises for the knowledge element K1, which is dominated by objective questions such as multiple choice, fill-in-the-blanks, and judgment, and sets the constraint that only exercises containing knowledge point $K1$ and knowledge points prior to $K1$ will be populated into QS when the exercise expansion is performed.

Definition 4 CA evaluates the learner's ability to master knowledge point $K1$, which is an integer type. A learner is considered to have mastered knowledge point $>CA$ if the learner gets a score of 2 through the exercises in QS, otherwise it means no.

Definition 5 $chapter$ The chapter to which knowledge point $K1$ belongs is represented as a floating point type.

2.3.2. Adaptive learning path generation model based on deep reinforcement learning. After the domain knowledge graph is constructed, the optimal learning path will be dynamically generated for the learner taking into account the learner's cognitive ability. The problem is formulated as follows: there is a domain knowledge graph G, a subgraph of students' cognitive abilities G', and at this point the subgraphs T', $T'=G-G'$ can be derived for the learning goals under the current cognitive abilities of the learners. in addition, when generating T', only the $rely$-relationships, $syno$-relationships and $simi$-relationships in it are taken into account, and other nodes that do not have these relationships are recognized as same-parent knowledge points by their $chapter$-attributes. This ensures that the generated subgraph T' has either no relationship between knowledge points or the directionality of the relationships between knowledge points is ordered.

Next, topological sorting of T' can result in a set of paths $R=\{r_1,r_2,r_3,\cdots\}$ from which the optimal path r_i is found.

In view of the performance of deep reinforcement learning on the problem of finding the optimal path and its problems, this paper proposes a dynamic reinforcement learning optimization method to guide the optimal path selection.

First, the paths are vectorized to complete the initialization, a process that does not require adding additional information. After that, the path vectors are input to the evaluation network $evel\_net$. The initialized path vectors act on the environment by interacting with the environment, choosing connected actions according to the exploration strategy, and obtaining a new state. This process is repeated to get the optimal path. Path vectorization for the transformation of the path into a vectorized representation that can be recognized by deep reinforcement learning, in this paper, in the design of the intelligent body to select the action, the tree structure of the representation method, the leaf nodes are defined as the knowledge points in the path, the path will be represented in the form of a kind of binary tree. The specific description is as follows: take each knowledge point to be connected as a candidate, when the number of knowledge points in a path is n, n-dimensional vector action selection is used to express the action space, and in the process of connection, the vector value connects the corresponding subtree. When a knowledge point is added to the subtree, the action value of the corresponding position in the subtree is changed to $\frac{1}{h(n,i)^2}*CA$; if there is no corresponding knowledge point added to the subtree then the value is 0; h is the height value of the knowledge point Ki of the path in the subtree; CA is the cognitive ability to be possessed for mastering knowledge point Ki.

In this paper, we still use the DDQN network structure, which consists of two parts, the evaluation network and the target network, without adding additional networks. Among them, the purpose of the evaluation network is to mine the action corresponding to the maximum value, while the target network is responsible for the evaluation of the maximum action value. The target value is computed using a linear relationship consisting of $Q\left(s',a^*,\theta\right)$ and $Q\left(s',a^*,\theta^-\right)$ with the formula:

$$y^{DDQN}=r+\gamma\left[\beta Q\left(s',a^*,\theta\right)+(1-\beta)Q\left(s',a^*,\theta^-\right)\right] \tag{23}$$

where: β is defined as the weight, and the action selection calculates the action value by weighting both the evaluation network and the target network. $\beta\in[0,1]$, when $\beta=0$, the network is equivalent to the DDQN network, when $\beta=1$, the algorithm completely ignores the DDQN evaluation and uses only the DQN network to select the action, the weights β are calculated as:

$$\beta=\frac{\left|Q\left(s',a^*,\theta^-\right)-Q\left(s',a_L,\theta^-\right)\right|}{c+\left|Q\left(s',a^*,\theta^-\right)-Q\left(s',a_L,\theta^-\right)\right|} \tag{24}$$

where: c is a hyperparameter, which is used to compute the weights β, when performing the computation, the optimal value of c is different in problems with different characteristics, hyperparameter c affects the network to assign the weights, action a^* is denoted as the action that evaluates the network to have the maximum value of the action; and a_L is denoted to evaluate the action that evaluates the network to have the minimum value of the action, which is computed as:

$$a^*=\arg\max Q\left(s',a,\theta\right) \tag{25}$$

$$a_L=\arg\min Q\left(s',a,\theta\right) \tag{26}$$

Thus, $Q\left(s',a^*,\theta^-\right)$ denotes the maximum action value of the target network in state s' and $Q\left(s',a_L,\theta^-\right)$ denotes the minimum action value of the target network in state s'. When obtaining the value of Q, then the computation is done based on the cost model and the execution delay incurred for the action, and when updating the value of Q the difference of the loss function is computed through the mean square error, and the reward corresponding to the execution of the query action is obtained based on the cost model:

$$Q(s,a) \leftarrow Q(s,a) + \alpha\left[r + \gamma \max_a Q\left(s',a\right) - Q(s,a)\right], a \in E_{dp}(a) \tag{27}$$

where: r is the reward; α is the learning rate, which is used to determine the magnitude of error learning; γ denotes the discount factor, which represents the intended Q value; and $Q(s,a)$ denotes the estimated Q value. The Q value update process is:

$$L(\theta) = E\left[r + \gamma \max_a \cdot Q\left(s',a',\theta\right) - Q(s,a,\theta)\right] \tag{28}$$

3. Results and Discussion

3.1. Analysis of the Civics Resource Allocation Model

In order to verify the algorithm in detail, firstly, this paper formulates the evaluation criteria of the algorithm from the five main dimensions of convergence, efficiency, accuracy, stability and robustness of the algorithm, which refer to the main evaluation indexes of the evolutionary algorithm and recommendation algorithm.

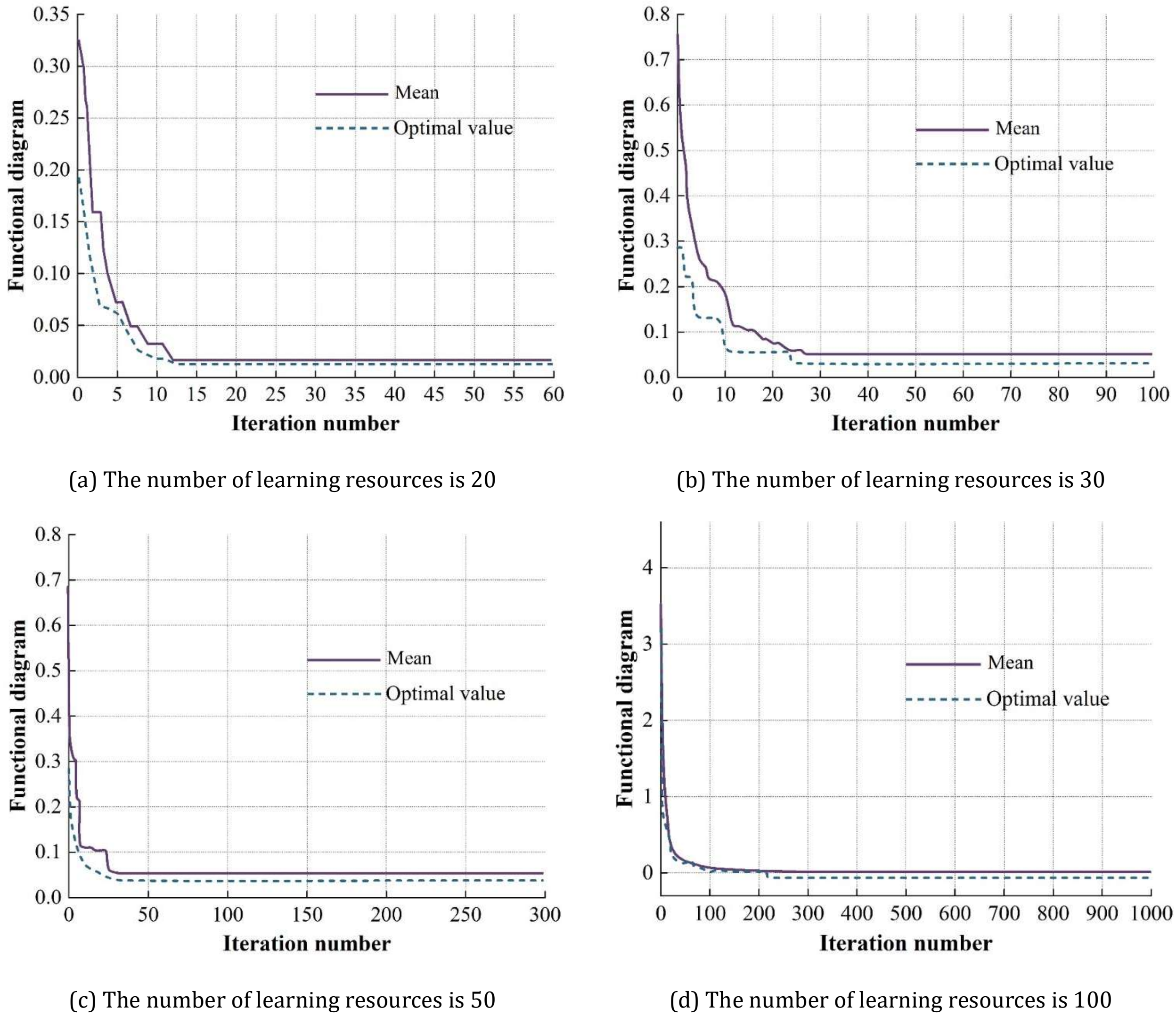
The same fusion of user-based collaborative filtering algorithm with binary particle swarm algorithm BPSO and genetic algorithm GA, called BPSOR and GAR algorithms, respectively, is applied as a baseline model in the field of learning resource recommendation, and the experimental environment and experimental parameters are set and described in a unified way. Finally, through the experimental data, the optimal value, mean, variance, success rate and execution time of the feature differences of the algorithms are compared with the evaluation criteria, and the accuracy and execution efficiency of the algorithms in this paper are verified.

3.1.1. Experimental environment. The basic environment of this experiment is divided into hardware environment and software environment, the hardware environment is as follows, CPU: Intel(R) Core (TM) i7-9600U, main frequency: dual-core main frequency of 2.80GHz, 2.90GHz, memory: 16GB, hard disk: 512G high-speed solid state drive. The software environment is as follows, operating system: Windows 10 64-bit Chinese Professional Edition, programming language: using python as the programming language, python base environment: Python2.7, 64-bit, Python programming environment IDE: PyCharm Community Edition, Java base environment: JRE 1.7 64, JDK 1.7 64-bit.

3.1.2. Convergence tests. For evolutionary algorithms, such as differential evolutionary algorithms, it is necessary to continuously converge and descend during the execution process in order to obtain the global optimal solution. For the learning resource recommendation system using differential, the feature difference values converge to the global optimal solution in order to obtain the personalized learning resource sequence that best meets the learner's current situation.

Therefore, the following experimental setup is carried out, firstly, different numbers of learning resources are defined by simulating the actual scenarios, and then the algorithm of this paper is subjected to 100 consecutive randomized trials, and the optimal value convergence curve and the mean value convergence curve of the feature differences of the algorithm are obtained finally, and the optimal value convergence curve and the mean value convergence curve of the feature differences of the algorithm are plotted based on the obtained values, so as to judge the basic convergence of the algorithm by the feature difference curves. The results of the difference value convergence curves are shown in Fig. 2. Where (a)~(d) are the difference value convergence curves when the number of learning resources is 20, 30, 50 and 100, respectively.

It can be seen that when the number of learning resources is 20, 30, 50, 100, the optimal value curve converges at 13th, 28th, 35th, 238th generation, all converging to the optimal value, and the characteristic difference mean value curve converges at 14th, 32nd, 72nd, 255th, respectively, with the mean value of convergence slightly higher than the optimal value. With different numbers of learning resources, both the optimal value and the mean value convergence curves converge faster, and gradually converge to a stable difference value at the later stage.



(a) The number of learning resources is 20

(b) The number of learning resources is 30

(c) The number of learning resources is 50

(d) The number of learning resources is 100

Fig. 2. Difference value convergence curve

3.1.3. Stability and accuracy analysis. First, the convergence accuracy data accuracy of the three algorithms is analyzed, and the higher the convergence accuracy algorithm, the more accurate its convergence function value. For personalized learning resource recommendation, its recommended learning resource sequence is more in line with the learner's personalized requirements. Therefore, in the experiment, the convergence mean value and the success rate of convergence to the global optimal solution obtained by different algorithms are compared and analyzed after 100 executions. The results are shown in Table 1.

When the number of learning resources is 10, 20, 30, 50 and 100, the mean values of feature differences of this paper's algorithm are 0.0798, 0.0991, 0.1405, 0.0596, 0.0687, and the success rates are 96%, 93%, 92%, 88% and 77%, respectively.The success rates of the BPSOR algorithm are 95%, 92%, 89% and 83%, 66%.GAR algorithm difference mean value is 0.1023, 0.1211, 0.1805, 0.0591, 0.0807, success rate is 91%, 86%, 85%, 82%, 66%. It can be seen that the algorithm of this paper has the smallest mean value of feature differences and the highest success rate, while the variance of feature differences is 0.0038, 0.0011, 0.0024, 0.0102, 0.0291, respectively, which are the lowest, indicating that the algorithm of this paper is more accurate and stable than the convergence of the fusion of feature differences computed by the BPSOR and the GAR, and it has good robustness and reliability, and it can accurately recommend the personalized learning resource sequences.

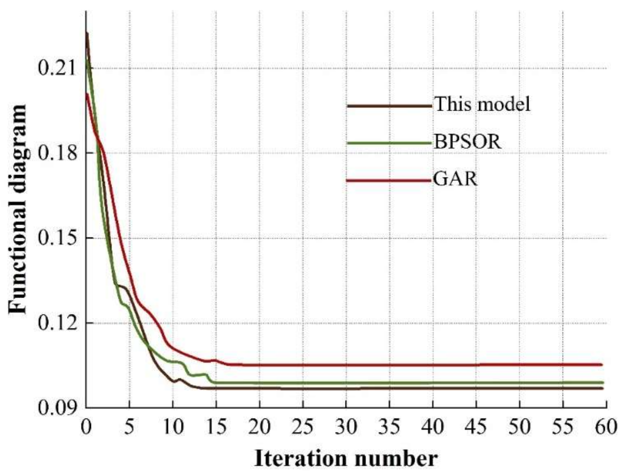
Table 1. Experimental results

Learning resources	Population size	Iteration	Algorithm	M	Variance	Success rate	Time consuming
10	30	30	This model	0.0798	0.0038	0.96	207
			BPSOR	0.0868	0.0065	0.95	232
			GAR	0.1023	0.0098	0.91	267
20	100	50	This model	0.099	0.0011	0.93	1602
			BPSOR	0.1239	0.0012	0.92	1960
			GAR	0.1211	0.0015	0.86	2009
30	150	100	This model	0.1405	0.0024	0.92	20506
			BPSOR	0.1475	0.004	0.89	20750
			GAR	0.1805	0.0035	0.85	21485
50	200	200	This model	0.0596	0.0102	0.88	68430
			BPSOR	0.0616	0.0119	0.83	68667
			GAR	0.0591	0.0137	0.82	69902
100	500	1000	This model	0.0687	0.0291	0.77	354687
			BPSOR	0.0807	0.034	0.66	383777
			GAR	0.081	0.0323	0.66	402933

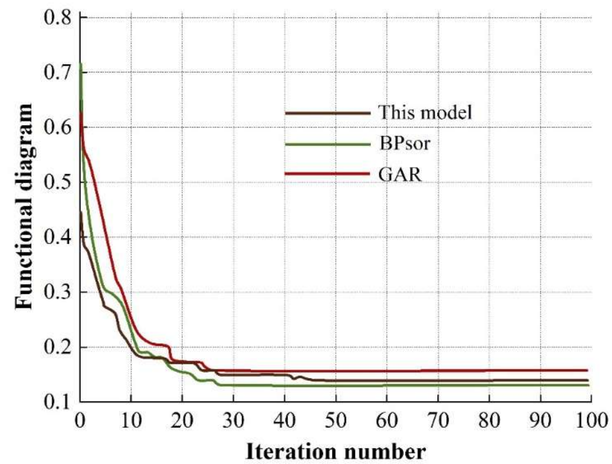
3.1.4. Analysis of the implementation process. This chapter analyzes the average convergence of the algorithms in detail by comparing the changes in the average convergence curves of the three algorithms under different conditions of the number of learning resources, and further analyzes the causes of the convergence curves through the changes in the convergence curves, especially makes a detailed analysis of the impact of the adoption of improvement strategies of different algorithms on the convergence of the algorithms. Figure 3 shows the difference mean convergence curve. Where (a)~(d) are the difference value convergence curves when the number of learning resources is 20, 30, 50 and 100, respectively.

First of all, by comparing the execution of the algorithm in the preliminary stage through Fig. 3, since this paper's algorithm and the BPSOR algorithm are both based on the difference between the population and the optimal solution and the individual optimal solution, directly searching in the vector space of the solution, and this paper's algorithm has a scaling factor parameter to adjust the speed of the search, and BPSOR has an inertia weight and an acceleration factor parameter to adjust the speed of the search, the two carry out a larger execution of the preliminary stage of the range of global exploration, which is conducive to the rapid convergence of the algorithm to a smaller value of the fitness function. The GAR algorithm, on the other hand, searches through the cross-variation of the population and is unable to search in the vector space of the entire population, so the convergence speed is slower than that of the above two algorithms in the early stage.

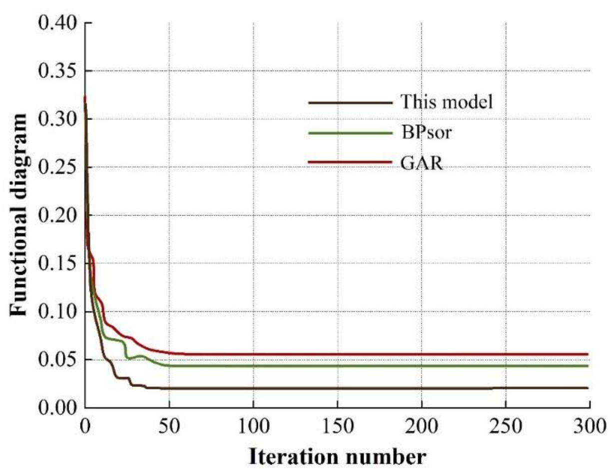
Secondly, by comparing the convergence of the algorithm execution to the middle and late stages in Fig. 3, the convergence curve of the proposed algorithm in this paper is gradually flattening out, but the algorithm still has the ability to evolve, and the convergence curve is still decreasing, and the final value of the feature difference is also better than other BPSOR and GAR algorithms. First of all, compared with the BPSOR algorithm, this paper's algorithm has the characteristics of the basic BPSOA algorithm searching in the population vector space, and has strong global search characteristics, which is conducive to the rapid search for the global optimal solution, and at the same time, the crossover operation of this paper's algorithm, which can be dynamically adjusted to the diversity of the population, preventing from falling into the local optimum, and searching for a global optimal solution, and the ability to dynamically adjust the diversity of the population is more important compared to the basic differential algorithm, can also speed up the convergence speed of the algorithm without affecting the adjustment of population diversity. The figure shows that the BPSOR algorithm has a higher convergence characteristic curve than this paper's algorithm in the late stage of execution and loses the ability to evolve, while this paper's algorithm is still able to evolve continuously and finely through diversity adjustment. Secondly, compared with the GAR genetic algorithm, this paper's algorithm has both the global search ability similar to the BPSOA algorithm, so that it can search for the optimal value globally, and at the same time, it has the genetic algorithm's variation ability to adjust the diversity of the population to prevent it from falling into the local optimum. Therefore, the GAR algorithm gradually has a larger gap between the feature difference convergence curve of the GAR algorithm and the algorithm of this paper in the late convergence period, while the feature convergence curve is also lower than that of the BPSOA algorithm.



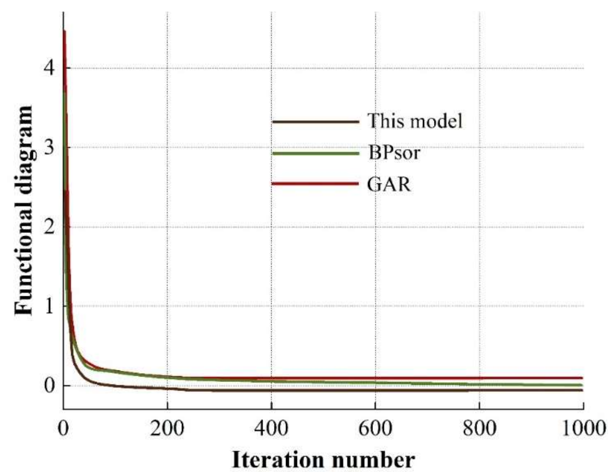
(a) The number of learning resources is 20



(b) The number of learning resources is 30



(c) The number of learning resources is 50



(d) The number of learning resources is 100

Fig. 3. Mean difference value convergence curve

3.2. *Optimal Learning Path Application*

In order to verify the application effect of the designed model, three classes of freshmen (50 students in Experimental Class I, 50 students in Experimental Class II, and 50 students in Control Class) majoring in software engineering in a university in Taiyuan City were used as experimental subjects. The first three chapters of the Civics teaching were used as the teaching content, which lasted for 6 weeks in total, and the teaching mode of the same teacher teaching in class and learners practicing with the adaptive learning system in class was adopted. Learners in the first class of the experiment used the system recommended path based on the model of this paper, learners in the second class of the experiment used the teacher's system recommended exercises, and learners in the control class chose their own exercises based on the concepts of the knowledge points to practice. After 6 weeks of study, learners in all 3 classes take a 1-hour mini-test.

3.2.1. Analysis of the results of teaching experiments. Because none of the learners in the 3 classes had any experience with Civics education, this study concluded that there was no significant difference in the academic performance of the learners prior to the experiment. Figure 4 is a box line plot of the overall distribution of learners' academic performance in the 3 classes, showing the overall data distribution of learners' test scores in the class. From the median, the experimental class 1 scored 87.2, the experimental class 2 scored 80.5, and the control class scored 76.7, which shows that the overall performance of the learners in the experimental class 1 was better than that of the other two classes; in terms of the IQR (the length of the box), the distribution of the performance of the experimental class 1 had a larger range, which implies that there was a greater variance in the performance of the learners in the class. The relatively centralized distribution of scores in Experimental Class 2 and the control class means that the learners in these two classes are more concentrated in their scores. In terms of outliers, no outliers were observed in the experimental class I and more low outliers existed in the experimental class II and the control class. From the median, IQR and outlier indicators of the box plot, it is clear that the experimental class I has better overall academic performance as compared to the other two classes.

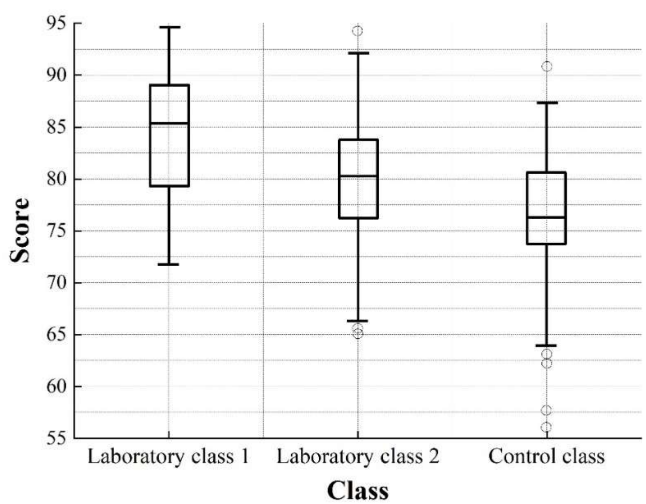


Fig. 4. 3 class learners' overall distribution box diagram

3.2.2. Case studies. This study also conducted a case study on the adaptive learning path of a randomly selected learner in the experimental class of a university whose learning trend is shown in Figure 5. In Fig. 5, the horizontal axis (x-axis) represents the path recommended by the exercise, the vertical axis (y-axis) represents the accuracy rate, the hollow point represents the average accuracy rate of all the learners who have answered this exercise in the system, and the solid point represents the student's answer in each exercise, if the answer is correct, y=1; if the answer is incorrect, y=0. As can be seen from the figure:

1) From the trend of the hollow point, the accuracy rate of the exercises recommended by the system for the learner is roughly wavy, when the student answered incorrectly on an exercise, the system recommends the exercises with higher accuracy rate for the learner, and when the learner answered correctly on an exercise, the accuracy rate of the exercises recommended by the system for the learner is getting lower and lower, and the results of this visualization show that the system can, based on the learner's answers, recommend the exercises with appropriate accuracy rate for the learner. This

visualization shows that the system can recommend exercises with appropriate accuracy rate, i.e., appropriate difficulty, for learners based on their answers.

2) When the average accuracy of an exercise is high, but the learner answers it incorrectly, the system will recommend an exercise with comparable accuracy again. If the second answer is still wrong, it means that the exercise is the student's weak knowledge concept; if both answers are correct, the previous error may be caused by a mistake or other reasons.

3) When the average accuracy of an exercise is low and the learner answers it incorrectly, the question may be a difficult one for the learner. In conclusion, from the case study of learners' adaptive learning paths, the systematic recommended learning paths modeled in this paper are able to locate weak knowledge point concepts, diagnose learning difficulties, and are able to recommend exercises of appropriate difficulty for learners.

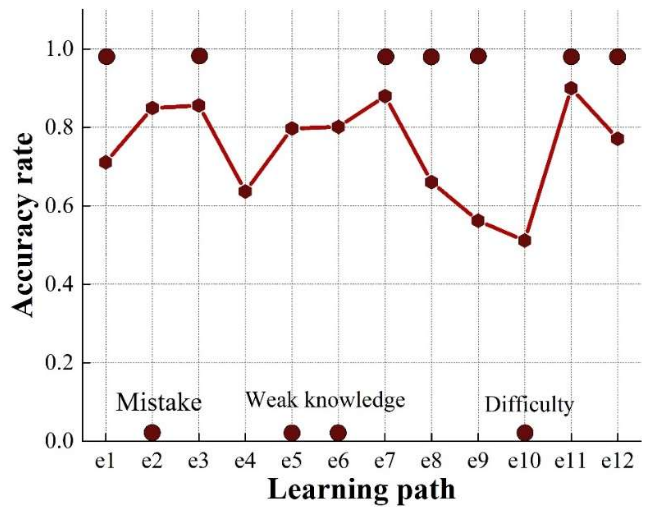


Fig. 5. Adaptive learning path case analysis diagram

4. Conclusion

Based on the cognitive diagnostic model, this paper constructs the Civic Education Course Resource Allocation and Optimal Path Recommendation Model, which realizes the personalized resource recommendation and learning path planning of Civic Education for college students. Experiments are designed to verify the performance of the model and the effect of practical application, and it is found that when the number of learning resources is 10, 20, 30, 50, 100 respectively, the mean value of the feature differences of this paper's algorithm is 0.0798, 0.0991, 0.1405, 0.0596, 0.0687, and the success rate is 96%, 93%, 92%, 88%, 77%, which is higher than that of BPSOR, GAR two baseline models. The experimental class using this paper's model was 6.7 and 10.5 points higher than the remaining two classes, respectively.

References

[1] Chen, X., & Tang, J. (2022, September). Remote Tutoring System of Ideological and Political Course Based on Mobile Client. *In International Conference on Advanced Hybrid Information Processing* (pp. 339-352). Cham: Springer Nature Switzerland.

[2] Zhang, H. (2023). Talk about the integration of curriculum ideology and politics in the. *Adult and Higher Education*, 5(1), 81-85.

[3] Lv, C., & Chen, B. (2019, August). Research on Ideological and Political Education from the Perspective of Educational Information Reception. *In 5th International Conference on Arts, Design and Contemporary Education* (ICADCE 2019) (pp. 743-747). Atlantis Press.

[4] He, F., Hu, Y., & Zhang, J. (2024). The Implementation Methods of the Ideological and Political Theories Teaching in Foreign Majors: A Case Study of Comprehensive English Course. *Journal of Education and Educational Research*, 7(3), 214-219.

[5] Sun, M., & Wang, J. (2024). Online allocation of teaching resources for ideological and political courses in colleges and universities based on differential search algorithm. *International Journal of Information Technology and Management*, 23(3-4), 278-290.

[6] Becher, T., & Maclure, S. (2024). The politics of curriculum change. *Taylor & Francis*.

[7] Duan, L. (2020). Discussing the Problems and Countermeasures of the Internet Age for Ideological and Political Work in Colleges and Universities. *TWP Series in Education, Sport Sciences and Physiology*, 1, 105-107.

[8] Dong, F., & Dong, S. (2023). Research On the Optimization of Ideological and Political Education in Universities Integrating Artificial Intelligence Technology Under the Guidance of Curriculum Ideological and Political Thinking. *ACM Transactions on Asian and Low-Resource Language Information Processing*.

[9] Jia, W. (2022). Research on the Development and Utilization of VR Technology in the Innovation Construction of Ideological and Political Theory Courses in Colleges and Universities. *Open Access Library Journal*, 9(5), 1-10.

[10] Guomin, H. (2023). Research on Teaching Strategies for Human Resource Management Courses in Higher Education Institutions under the Course Ideology of Integrating Politics and Values. *Advances in Vocational and Technical Education*, 5(13), 38-44.

[11] Ying, L. (2022). A study on the Curriculum-based Ideological and Political Education in Colleges and Universities from the Perspective of Social System. *Frontiers in Educational Research*, 5(2).

[12] Zhou, C. (2024). Coordinate and Optimize Political Ideology, Curriculum Reform and Local Governance to Promote Sustainable Development. *Lex localis-Journal of Local Self-Government*, 22(4), 459-476.

[13] Zhou, Y. (2022). The application of curriculum ideology and politics in the training of judicial vocational education talents. *Journal of Higher Education Research*, 3(2), 155-159.

[14] Yan, C. (2021). Practice and Exploration of “Full-Course Ideological and Political” in Colleges--Take Wenzhou City University as an Example. *Advances in Vocational and Technical Education*, 3(3), 38-42.

[15] Khoza, S. B. (2018). Can teachers’ reflections on digital and curriculum resources generate lessons?. *Africa Education Review*, 15(4), 20-35.

[16] Li, Y., Guo, Y., & Zhai, X. (2022). Research on the implementation path of red cultural resources into the ideological and political construction of college curriculum. *Advances in Education, Humanities and Social Science Research*, 1(2), 117-117.

[17] Ye, W., Qin, J., Mu, W., & Zhang, X. (2021). Analysis on the cohesion of red cultural resources into ideological and political education in primary and secondary schools. *International Journal of Frontiers in Sociology*, 3(3).

[18] Zhang, Y. (2020). On the integration of ideological and political elements into college English curriculum in the information age. *International Journal of Social Sciences in Universities*, 3(3).

[19] Nicolai D Ayasse & Cheryl D Coon. (2024). Investigating item response theory model performance in the context of evaluating clinical outcome assessments in clinical trials.. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation(prepublish),1-12.

[20] Qiaoyong Jiang,Lei Wang,Xinhong Hei,Guolin Yu & Yanyan Lin.(2016).The performance comparison of a new version of artificial raindrop algorithm on global numerical optimization.Neurocomputing1-25.

[21] Xin Bin,Wang Fan & Zhai Zhibo. (2021). Balwin-Teaching-Learning-Based Artificial Raindrop Algorithm for UAV Route Planning. Mathematical Problems in Engineering