Hindawi
International Journal of Intelligent Systems
Volume 2023, Article ID 2578286, 19 pages
https://doi.org/10.1155/2023/2578286

Research Article

Knowledge Graph-Enhanced Intelligent Tutoring System Based on Exercise Representativeness and Informativeness

Linqing Li[1] and Zhifeng Wang[1,2]

[1]*CCNU Wollongong Joint Institute, Central China Normal University, Luoyu Road, Wuhan 430079, China*
[2]*Faculty of Artificial Intelligence in Education, Central China Normal University, Luoyu Road, Wuhan 430079, China*

Correspondence should be addressed to Zhifeng Wang; zfwang@ccnu.edu.cn

Received 15 July 2023; Revised 23 September 2023; Accepted 3 October 2023; Published 16 October 2023

Academic Editor: Vasudevan Rajamohan

In the realm of online tutoring intelligent systems, e-learners are exposed to a substantial volume of learning content. The extraction and organization of exercises and skills hold significant importance in establishing clear learning objectives and providing appropriate exercise recommendations. Presently, knowledge graph-based recommendation algorithms have garnered considerable attention among researchers. However, these algorithms solely consider knowledge graphs with single relationships and do not effectively model exercise-rich features, such as exercise representativeness and informativeness. Consequently, this paper proposes a framework, namely, the Knowledge Graph Importance-Exercise Representativeness and Informativeness Framework, to address these two issues. The framework consists of four intricate components and a novel cognitive diagnosis model called the Neural Attentive Cognitive Diagnosis model to recommend the proper exercises. These components encompass the informativeness component, exercise representation component, knowledge importance component, and exercise representativeness component. The informativeness component evaluates the informational value of each exercise and identifies the candidate exercise set (E_C) that exhibits the highest exercise informativeness. Moreover, the exercise representation component utilizes a graph neural network to process student records. The output of the graph neural network serves as the input for exercise-level attention and skill-level attention, ultimately generating exercise embeddings and skill embeddings. Furthermore, the skill embeddings are employed as input for the knowledge importance component. This component transforms a one-dimensional knowledge graph into a multidimensional one through four class relations and calculates skill importance weights based on novelty and popularity. Subsequently, the exercise representativeness component incorporates exercise weight knowledge coverage to select exercises from the candidate exercise set for the tested exercise set. Lastly, the cognitive diagnosis model leverages exercise representation and skill importance weights to predict student performance on the test set and estimate their knowledge state. To evaluate the effectiveness of our selection strategy, extensive experiments were conducted on two types of publicly available educational datasets. The experimental results demonstrate that our framework can recommend appropriate exercises to students, leading to improved student performance.

1. Introduction

Online education has emerged as a significant supplementary learning strategy for students [1, 2]. Many students rely on online exercise recommendations to support their learning. With the vast amount of educational materials available, the challenge lies in recommending appropriate exercises for effective learning [3–6]. Exercises play a crucial role in personalized educational services by serving as a powerful tool to assess students' mastery of concepts [7]. However, given the abundance of exercise resources, it is nearly impossible for students to complete them all within a limited time. Therefore, assisting students in finding suitable exercises becomes a significant problem. An exercise recommendation system has been proposed to address this issue by leveraging students' historical answer sequences [8–10].

Knowledge graphs (KGs), also known as cognitive maps, provide graphical representations where concepts or words are organized into nodes and connected by vectors representing relationships. The application of knowledge graphs to learn the order of skills has shown promising results (e.g., [11–14]). The arrangement of concepts in a knowledge graph significantly impacts learning ability [15–18].

The cognitive diagnosis model functions as an effective method to discover the knowledge state of students by analyzing the past interactions of students. As illustrated in Figure 1, the student firstly picks up some exercises (e.g., e_1, e_2, and e_3). Then, the response of the student in each exercise can be obtained. Finally, the cognitive diagnosis model analyzes the response of students to estimate the student knowledge state on each concept.

Researchers have recognized that exercise and skill features in knowledge graphs greatly influence the quality of learning when recommending exercises [19, 20]. Various methods have been developed to learn the features of knowledge graphs and recommend appropriate exercises, resulting in improved student performance [21–24]. These methods effectively explore skill and exercise features to enhance learning efficiency. Additionally, high-quality exercises contribute to learners' comprehension of the learning material. Consequently, the research community strives to create a high-quality exercise set to enhance e-learners' performance. Previous research [11, 25] applied KGs to consider the dependencies of learning objects in exercise recommendations. However, these works only focused on basic relationships to establish links between KGs, without further investigating exercise features during the recommendation process. As a result, these methods fall short of meeting the requirements of modern e-learning.

This paper presents an innovative framework called Knowledge Graph Importance-Exercise Informativeness and Representativeness (KI-EIR) to address diverse learning needs based on KGs. To recommend exercises with high learning quality, the KI-EIR framework combines multidimensional KGs with exercise features to define the recommendation goal and enhance exercise quality. The KI-EIR framework consists of four innovative components and a novel cognitive diagnosis model called the Neural Attentive Cognitive Diagnosis (NACD) model, which facilitates exercise recommendation to achieve the recommendation goal. Specifically, the NACD model estimates the student knowledge state by analyzing the past interaction of students. Then, according to the cognitive diagnosis results, different types of students can be correctly distinguished. Finally, the intelligent tutoring systems recommend the proper exercises to students and improve their knowledge proficiency. The four components are the informativeness component, exercise representation component, knowledge importance component, and exercise representativeness component. The recommendation goal involves recommending exercises with high informativeness and representativeness.

Specifically, the informativeness component aims to select exercises with high informativeness from the untested exercise set (E_U) to the candidate exercise set (E_C). The exercise representation component incorporates a graph neural network with two types of attention mechanisms to generate exercise and skill embeddings. The knowledge importance component utilizes an innovative knowledge point extraction algorithm that incorporates skill embeddings to extract knowledge points based on the multidimensional KG. Five skill features of these knowledge points are discussed to generate skill importance weights. Subsequently, the exercise representativeness component selects exercises with high knowledge coverage from the candidate exercise set (E_C) to the tested exercise set (E_T) to achieve representativeness objectives. Finally, the NACD model predicts student performance on the tested exercise set and estimates their current knowledge state.

The main contributions of this paper can be summarized as follows:

(1) We propose a novel exercise recommendation method, KI-EIR, which selects exercises with high informativeness and representativeness. By incorporating the structural information of knowledge concepts, KI-EIR recommends exercises to students, thereby improving their overall cognitive level during the recommendation process.

(2) We design four innovative components and a novel cognitive diagnosis model, NACD, including the informativeness component, exercise representation component, knowledge importance component, and exercise representativeness component. The informativeness component estimates the informativeness of each exercise and generates the candidate exercise set. This exercise set serves as input to the exercise representativeness component, which selects exercises with high knowledge coverage based on the knowledge importance component. The knowledge importance component incorporates a multidimensional KG and a knowledge point extraction algorithm with five skill features to generate skill importance weights. Finally, the cognitive diagnosis model predicts student performance and estimates their current knowledge state based on exercise and skill relations.

(3) We evaluate the KI-EIR framework on two public educational datasets, including Assistment 2009-2010 and Eedi2020, using the AUC (informative metric) and knowledge coverage rate (KCR). The experimental results demonstrate that KI-EIR outperforms other methods such as RAND and EM.

The rest of the paper is structured as follows. Section 2 provides an overview of related works on cognitive diagnosis models, relation modeling, and exercise recommendation. Section 3 presents important terminologies and defines the goal and problem statement of this study. Section 4 describes the methods proposed in this paper. Section 5 presents the experimental evaluation of our framework using two different metrics. Section 6 concludes the paper and discusses future research directions.

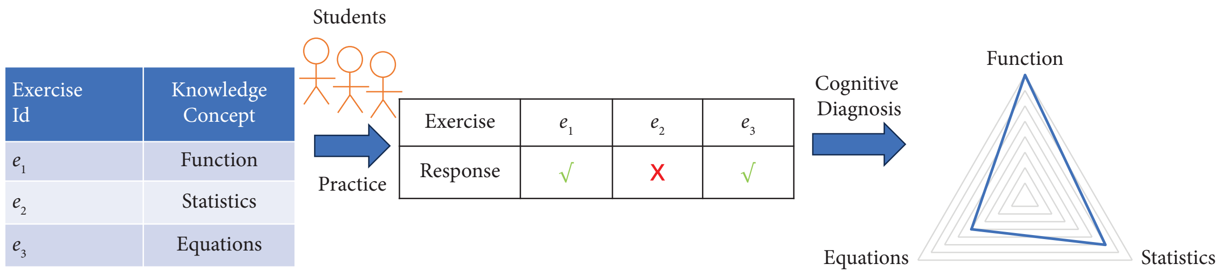

Figure 1: Toy example of cognitive diagnosis. The student firstly chose three exercises to test. Then, according to the test results, the cognitive diagnosis model estimates the student proficiency on each knowledge concept.

2. Related Work

2.1. Cognitive Diagnosis Model. Cognitive diagnosis plays a crucial role in various real-world scenarios, including games [26], medical diagnosis [27], and especially education [28]. The primary objective of cognitive diagnosis is to discover the latent trait characteristics of learners based on their testing records. These discovered characteristic features have been applied in tasks such as resource recommendation [29] and performance prediction [30]. Early approaches to cognitive diagnosis mainly depended on psychological evaluation [31]. The two most traditional cognitive diagnosis models, namely, the Item Response Theory (IRT) [32] and the Deterministic Input, Noisy And Gate (DINA) model [33], model the response generated by a learner answering an item as the interaction between the learner's trait features and the item. Ackerman [34] extended the characteristic features into a multidimensional space by proposing the Multidimensional Item Response Theory (MIRT). In recent years, deep learning has been incorporated into cognitive diagnostics by several researchers [35, 36]. Wang et al. [30] introduced NeuralCD, which utilizes neural networks to autonomously learn the interaction function. However, these cognitive diagnosis models overlook the deep relations between exercises, skills, and students when estimating students' knowledge state.

2.2. Relation Modeling. Based on psychological research, the relationship between exercises and skills has been extensively explored in numerous studies that measure students' knowledge levels (e.g., [37, 38]). Many researchers employ Q-matrices to model the relationship between exercises and skills, where exercises related to the same knowledge concept are considered connected in the Q-matrix. Additionally, researchers investigate the relationship between two exercises or skills based on exercise embeddings (e.g., [39, 40]). Semantic similarity scores of exercises are computed using prior interactions to model the significance of these interactions. However, these relation modeling methods do not consider the heterogeneous interactions between students, exercises, and skills. Therefore, this paper incorporates multiple dimension knowledge graphs (KGs) and Graph Convolutional Networks (GCNs) to establish exercise and skill relations and delve into exercise features such as informativeness and representativeness to recommend the proper exercises to students.

2.3. Exercise Recommendation. Traditional recommendation systems employ collaborative filtering, which can be categorized into nearest-neighbor collaborative filtering and model-based collaborative filtering. Nearest-neighbor collaborative filtering includes user-based collaborative filtering [41] and item-based collaborative filtering [42]. Model-based collaborative filtering involves mining hidden or explicit features to mitigate data sparsity and achieve good scalability [43]. When applying traditional recommendation methods to exercise recommendation in the educational field, students are treated as users and exercises as items. Thus, nearest-neighbor collaborative filtering can be further classified as exercise-based and student-based. Considering the impact of multiple dimensional knowledge graphs, recent research has proposed knowledge graph-based recommendation methods for exercise recommendations (e.g., [21, 25]).

Recent exercise recommendation methods that leverage knowledge graphs help mitigate misunderstandings in learning content descriptions. Inspired by this idea, Wan and Niu [44] introduced a learner-oriented exercise recommendation method based on knowledge concepts, represented as nodes, and the relationships between them as edges in knowledge graphs. Ouf et al. [45] developed exercise recommendation methods by incorporating knowledge graphs with semantic web to merge personalized concepts. To organize learning resources in a sequential manner, Shmelev et al. [46] proposed a method that integrates evolutionary methods and knowledge graph technology. Chu et al. [47] created an e-learning system based on a conceptual map that can generate learning paths using the connections in the concept map. Recognizing the need for diverse learning paths in different settings, Zhu et al. [25] presented a method for recommending learning paths using prebuilt learning scenarios. They developed an approach that requires the definition of starting and ending nodes to construct learning paths. However, all of them ignore to discuss the rich exercise features contained in the KG. Therefore, the KI-EIR model is proposed to incorporate the multiple dimension KG with the exercise features including the representativeness and informativeness to recommend the exercises to meet specific students' needs.

3. Preliminaries

This section is divided into three parts. The first part presents the problem addressed in this paper. The second part provides definitions for several terminologies used throughout the paper, including exercise informativeness, exercise representativeness, and heterogeneous interactions. The third part outlines the goals of the paper. Some important mathematical notations in the paper are summarized in Table 1.

Table 1: Important mathematical notations.

Notations	Descriptions
KC	The knowledge concept
KG	The knowledge graph
P	The learning paths in the knowledge graph
N_e	The total number of exercises
K	The total number of skills
E_U	The untested exercise set
E_C	The candidate exercise set
E_T	The tested exercise set
Level (KC)	The level of KC
P_n	Response matrix
$\widehat{E}$	Dissimilarity matrix

3.1. Problem Definition. Exercise recommendation aims to recommend exercises that enhance students' knowledge proficiency. The problem of this paper is how to recommend appropriate exercises that meet the specific requirements of each student. In this paper, two measurements are defined to evaluate exercise quality: exercise representativeness and exercise informativeness. Thus, the specific problem addressed in this paper is how to recommend exercises with high representativeness and informativeness from a large pool of exercises. To solve this problem, we propose the Knowledge Graph Importance-Exercise Informativeness and Representativeness (KI-EIR) framework, which comprises four components and a cognitive diagnosis model. Specifically, the informativeness component selects exercises with high informativeness from the untested exercise set (E_U) to the candidate exercise set (E_C). The exercise representation component and the knowledge importance component are designed to generate skill and exercise embeddings, as well as skill importance weights. The outputs of the exercise representation component and the knowledge importance component serve as input to the exercise representativeness component, which selects exercises with high representativeness from E_C to the tested exercise set (E_T). Then, the Neural Attentive Cognitive Diagnosis (NACD) model predicts students' performance on E_T and diagnoses their current knowledge state. Finally, according to the different student knowledge, different exercises are recommended to students to improve their knowledge proficiency.

3.2. Terminologies

Definition 1 (informativeness). In general, a valid exercise is expected to reduce the level of uncertainty in an examinee's knowledge state. Thus, the informativeness of an exercise can be defined as the amount of information that the underlying cognitive diagnosis model (M) can acquire from the exercise to update the estimate of knowledge states. Selecting the most informative exercises acts as an effective approach to achieve the informativeness goal. After the student completes the test, the performance of the student with M on the entire tested exercise set is predicted, and the performance is evaluated by using a metric such as the Area Under the Curve (AUC), denoted as Inf(S).

Definition 2 (representativeness). If a set of exercises achieves a certain knowledge coverage rate, it is considered as representative. The knowledge coverage rate functions as an evaluation metric to measure representativeness. Selecting a group of exercises with the highest coverage of knowledge concepts is essential to achieve the representativeness objective. The coverage, Cov(S), can be computed as the percentage of knowledge concepts covered by the tested exercise set, E_T, after the test.

Definition 3 (heterogeneous interaction). When answering exercises, there exist various interactions among students, exercises, and skills. Heterogeneous interaction, denoted as HI $= (V; E)$, consists of an object set, V, and a link set, E. The object types in V include students, exercises, and skills. E is a collection of relational types in the form $E = (r_A, r_C)$, where r_A represents the relation it answers and r_C represents the relation it contains. Figure 2 provides a toy example illustrating this definition.

3.3. Goals. The goal of this paper is to recommend exercises with high representativeness and informativeness to improve student performance in subsequent interactions. Informativeness is measured by using the Area Under the Curve (AUC), while representativeness is measured by the knowledge coverage rate when predicting the corresponding exercises.

4. Proposed Method

In this section, we present our proposed framework called Knowledge Graph Importance-Exercise Informativeness and Representativeness (KI-EIR). The framework aims to model exercise features and skill features to generate exercises based on their informativeness and representativeness. KI-EIR consists of four components: the informativeness component, the exercise representation component, the exercise representativeness component, and the knowledge importance component. The overall structure and components of KI-EIR are depicted in Figure 3. In order to efficiently run the KI-EIR framework, we train this framework based on two ideas. The first involves integrated training and unified optimization for all components. The second is exchanging space for time. We pretrain the exercise representation component to obtain the skill embedding and exercise embedding instead of running the exercise representation component in each epoch.

The KI-EIR framework operates as follows. Given a user $e_i \in E$, the framework models exercise features and skill features to generate exercises that are both informative and representative. At each step t, KI-EIR selects one exercise

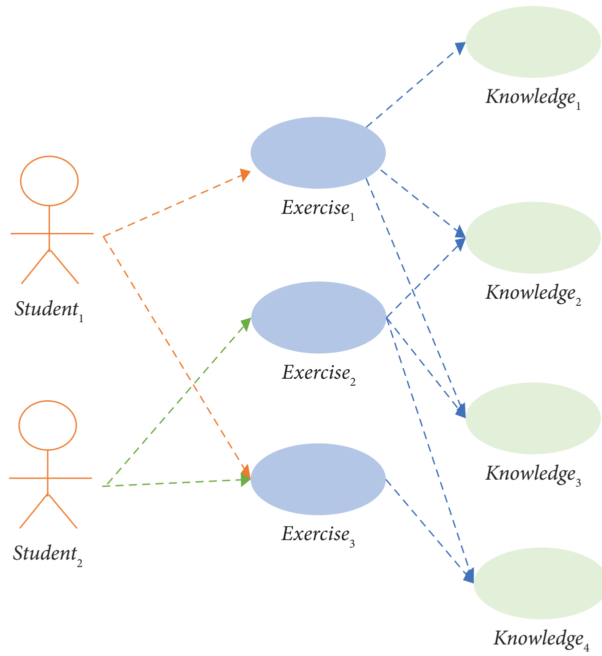

Figure 2: Toy example illustrating the heterogeneous interaction between students, exercises, and skills.

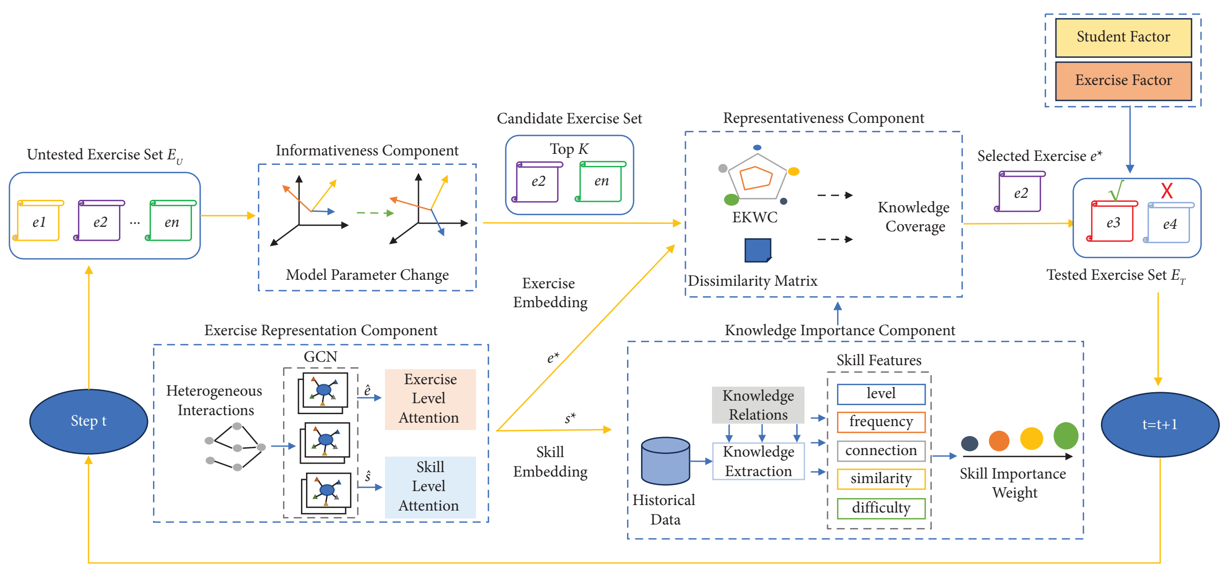

Figure 3: The overall framework of the KI-EIR selection strategy. The framework consists of four components: the informativeness component, the exercise representation component, the exercise representativeness component, and the knowledge importance component. The informativeness component selects high-informativeness exercises to form the candidate exercise set E_C. The exercise representation component extracts exercise and skill embeddings through a graph neural network. The exercise representativeness component selects an exercise with high representativeness to maximize the marginal gain. The knowledge importance component assesses the relevance of knowledge concepts. The NACD model predicts student performance based on the exercise embedding. The optimization tricks of this framework are integrated training and unified optimization and exchanging space for time in the exercise representation component.

from the untested exercise set E_U to be added to the tested exercise set E_T. The framework can be divided into four components:

Informativeness component: This component is responsible for selecting a candidate exercise set E_C from E_U based on informativeness. The selection is performed using a score function called Model Parameter Change (MPC). MPC estimates the user's knowledge states by observing their answers and quantifies the extent to which an exercise alters the diagnosis. The top-K highly informative exercises are selected to form E_C.

Exercise representation component: In this component, the framework extracts information on the heterogeneous interactions between users, exercises, and skills using a graph neural network. The exercise representation component generates the exercise embedding e^* and skill embedding s^*. These embeddings serve as inputs for the subsequent components.

Exercise representativeness component: The representativeness component selects an exercise with high representativeness from E_C to maximize the marginal gain. The selection process takes into account the exercise embedding e^* obtained from the exercise representation component.

Knowledge importance component: To enhance the selection process, the knowledge importance component assesses the relevance of knowledge concepts. It explores the relevance of knowledge concepts using the knowledge point extraction algorithm and incorporates five skill features to generate the skill importance weight.

Finally, the NACD model predicts student performance and estimates their state based on the exercise embedding e^*.

4.1. Informativeness Component. The informativeness component is the first step of the KI-EIR framework, where we select a candidate exercise set E_C consisting of top-K high-informativeness exercises. To measure the informativeness of an exercise, we propose a score function called Model Parameter Change (MPC).

MPC leverages the information contained in an exercise by estimating the user's knowledge states after answering the exercise. The parameter change in the abstract cognitive diagnosis model (CDM), denoted by θ in M, represents the knowledge states in KI-EIR. The amount of change in the CDM parameters reflects the amount of information gathered from the exercise. If the θ parameters change significantly, the exercise is considered more informative; otherwise, it provides little information. According to the different cognitive diagnosis results, the KI-EIR framework can recommend different types of exercises to meet the specific students' requirements.

The MPC function calculates the probability of correctly answering an exercise, which can be predicted by the cognitive diagnosis model M. Let $\Delta M = |\theta(R \cup r_{\mathrm{ij}}) - \theta(R)|$ represent the parameter change in our model when adding the record $r_{\mathrm{ij}} = \langle e_i, q_j, a_{\mathrm{ij}} \rangle$. Here, $\theta(R_i)$ represents the parameters obtained from the current interaction R_i of student e_i, and $\theta(R_i \cup r_{i,j})$ represents the parameters after adding the interaction. For each exercise e_j, the MPC function is defined as follows:

$$\begin{aligned} \mathrm{MPC}\left(q_j\right) &= E_{a_{\mathrm{ij}} \sim p} \Delta M\left(\langle e_i, q_j, a_{i,j} \rangle\right), \\ p &= M\left(e_i, q_j \middle| \theta(R_i)\right). \end{aligned} \tag{1}$$

$\Delta M(r_{\mathrm{ij}})$ is approximated by the gradient caused by r_{ij}. This approach is particularly effective for models trained using gradient-based methods, such as neural models.

Based on the MPC score function, we select exercises from the untested exercise set to form the candidate exercise set E_C. We calculate the MPC for each exercise and select the top-K exercises with the highest informativeness.

At the same time, this component also provides the KI-EIR framework with exercise informativeness interpretability, which means that the KI-EIR framework just needs to select the exercise with high informativeness and representativeness to recommend exercises.

4.2. Exercise Representation Component. The exercise representation component is the second step of the KI-EIR framework, where we extract exercise embedding (e^*) and skill embedding (s^*) by considering the heterogeneous interactions between students, exercises, and skills.

We employ the Graph Convolutional Network (GCN) model to generate embedding representations of exercises and skills, capturing their static relationships. Before applying the GCN model, we define the neighbors of exercises and skills based on three meta-relationships: exercise-student-exercise (eSe), exercise-skill-exercise (eKe), and skill-exercise-skill (kEk). In the eSe and eKe relationships, the exercise neighbors are exercises answered by the same student or covering the same skill. In the kEk relationship, the skill neighbors are skills contained in the same exercise. To propagate information in the GCN, we use two matrices: the exercise relation matrix (R^E) and the skill relation matrix (R^S), which capture the high-order information. Then, we apply the GCN model to generate the hidden embedding representations of exercises ($\hat{e}$) and skills ($\hat{s}$).

Each convolutional layer in the GCN model updates the nodes based on their own state and the state of their nearest neighbors. Let node_i denote the state of an exercise or skill and $\mathrm{Node}(i)$ denote a group of nodes representing the neighbors of node_i. The exercise at the l-th layer can be computed as follows:

$$\mathrm{node}_i^{\mathrm{l}} = \mathrm{RELU}\left(\sum_{j \in i \cup \mathrm{Node}(i)} w_i^{\mathrm{l}} \mathrm{node}_j^{\mathrm{l}-1} + b_i^{\mathrm{l}}\right), \tag{2}$$

where w^{l} and b^{l} represent the weight matrix and bias of the GCN layer, respectively, and RELU ( ) denotes the activation function used in the GCN model.

The hidden embedding representations obtained from the GCN model capture the static relationships between exercises and skills. However, they do not consider the similarity among exercises and skills when generating their embeddings. To incorporate the deep semantics of exercises and skills, we use exercise-level attention and skill-level attention mechanisms. These attention mechanisms learn the semantic relationships between students and exercises, generating the final embedding representations e^* and s^*. The formulation for e^* is as follows:

$$\begin{aligned} \alpha_E &= \mathrm{softmax}\left(\frac{\left(\hat{e}W^Q\right)\left(\hat{e}W^K\right)}{\sqrt{d_K}}\right), \\ \beta_E &= \delta_a \alpha_E + \left(1 - \delta_a\right) R^E, \\ e^* &= \beta_E \hat{e} W^v, \end{aligned} \tag{3}$$

where R^E is the exercise relation matrix, $\sqrt{d_K}$ is the scaling factor, and W^K, W^Q, and W^V are projection matrices.

The process for obtaining s^* is similar to that of e^*. The difference lies in using the hidden embedding representation of skills as input to the attention mechanism, and the skill relation matrix is used instead of the exercise relation matrix when calculating β_E.

At the same time, the connection between students, exercises, and skills is demonstrated by the heterogeneous graph. The information encompassed in the heterogeneous graph is subsequently extracted using a graph neural network, which produces the matching skill embedding and exercise embedding. As a result, the KI-EIR framework enables interpretation of the learning interactions.

4.3. Knowledge Importance Component. After the procedure of selecting exercises from the untested exercise set to the candidate exercise set, the knowledge importance component aims to compute the knowledge importance weight W_K as input for the next selection procedure: the representativeness component. Previous studies [48] have shown that organizing educational resources into different classes helps students understand learning profiles and enables them to logically organize and recall knowledge. Therefore, in our knowledge graph, we separate knowledge concepts into different classes to learn the weight of knowledge concepts (KCs). There also exists some incompleteness or outdated information in the knowledge graph, which results in the suboptimal recommendation results. Therefore, we also invite two domain experts to correct and validate the information in the knowledge graph to reduce the impact of the data problem.

We categorize learning objects into three classes:

(1) Subject knowledge: this class contains KCs at the subject level, such as "math," "physics," and "biology," supporting basic knowledge areas like "Ratio," "Geometry," and "Standard Form."
(2) Basic knowledge: It is the core of the framework. This class includes specific knowledge fields such as "Proportion" and "Negative Numbers" that are essential for solving specific tasks.
(3) Task: This class encompasses practical educational problems like "Factorising into a Single Bracket" and "Expanding Single Brackets." The task level is the bottom level of our knowledge graph framework.

Figure 4 presents a visual representation of the multidimensional KG framework we employ in our paper. Each class in this framework consists of a hierarchy and associated learning object instances. The learning objects represent meta-learning resources that are incorporated into the hierarchy and connected by semantic relationships, while the hierarchy reflects the knowledge structure of the current class. We establish various relationships, dividing them into intraclass relationships and interclass interactions, to illustrate the semantic connections between learning objects. Intraclass relationships link learning objects within a class, while interclass relationships provide links between educational resources from different classes (see Table 2). Our knowledge graph expands the accessibility of learning objects across classes and strengthens connections between cross-class learning objects. This graph of knowledge showcases how learned information can be practically applied, deepening e-learners' understanding of the studied information and helping them comprehend how theoretical knowledge can be used in practical scenarios.

4.3.1. Knowledge Point Path Extraction Algorithm. To determine the relevance of KCs, we explore all possible learning paths through the target learning object and the learning need of the e-learner. We have designed the knowledge point path extraction algorithm, which is based on the multidimensional knowledge graph, to accomplish this task. The algorithm consists of two phases.

The first phase involves calculating the relationship constraints φ based on the learning need. The getRelation() function is used to determine the relationship constraints $\varphi = (\alpha, \beta, \gamma \ldots)$ to meet the specific students' requirements.

In the second phase of the algorithm, a learning path is constructed using the relationship restrictions. Starting with the target learning item, the algorithm generates the learning path by searching for the next learning object associated with a relationship that satisfies the constraints. The associated learning object serves as a continuation of the search. The initial learning object of the current learning route is the chosen target learning object. If the current learning item has no related learning objects, the path will consist of only one KC. The algorithm performs a greedy search starting from the target learning object.

For a detailed description of the algorithm, refer to Algorithm 1.

4.3.2. Knowledge Importance Weight Extraction Algorithm. The knowledge importance weight extraction algorithm aims to extract the weights of KCs based on five skill features. In previous work on quantifying algorithms [49], the feature set F of KCs was proposed to select important KCs, including the level (f_1), frequency (f_2), connection (f_3), similarity (f_4), and difficulty (f_5) of the corresponding KCs.

(1) The level feature (f_1) is designed to extract the level of a KC. By applying the knowledge point path extraction algorithm (KPE), which transforms the one-dimensional knowledge graph into a multidimensional one, the levels of KCs in all related learning paths can be extracted. For example, if the output of the KPE is "A-B-C," where the levels of A, B, and C are 0, 1, and 2, respectively, the different knowledge levels of KCs can be extracted based on different learning paths. The level of the particular KC within all learning paths that encompass this KC can be defined as follows:

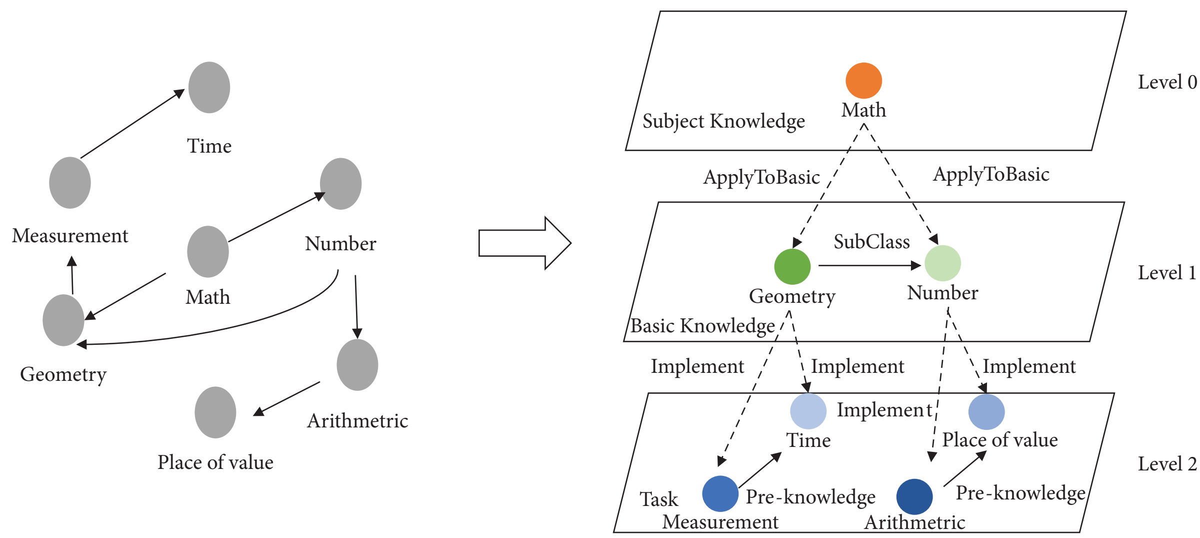

FIGURE 4: One-dimensional KG framework used in previous studies (left). Multidimensional KG framework employed in our paper (right). Dotted lines represent links between classes, solid lines denote interactions within classes, and nodes of various colors represent learning items in different classes. This transformation process is assisted by two domain experts in our university.

TABLE 2: Designed relationships in the knowledge graph.

Knowledge relationship	Type	Description
Subclass	Intraclass	Indicates that the current LO possesses a subclass
Implement	Interclass	Indicates that the current LO can implement subsequent LO
Preknowledge	Intraclass	Indicates that prior knowledge exists that should be learned before the current LO (basic knowledge)
ApplyToBasic	Interclass	Indicates that the current LO can be applied in the target LO

Input: Students' historical response dataset: $D = s_1, s_2, \ldots s_N$,
  $s_i = (e_i, s_i, t_i)$; The knowledge level graph G;
Output: The all possible learning paths: P.
While findAllPaths (KC_u) **do**
  R = getRelations (KC_u) (find the relations connected with KC_u)
  if $\{r \mid r \in R, r \in \oslash\}$ **then**
    P.addPath(p); Add the new path into the path set
  else while all $r \in R$ **do**
    list(KC_u) = getConnectObject (KC_u) (obtain the connected objects of KC_u with r)
    p.addElement (list(KC_u)) (put the element in the learning path)
    Recursively apply findAllPaths (KC_u);
  end
end

ALGORITHM 1: Knowledge point path extraction algorithm.

$$f_1(\mathrm{KC}) = \frac{\sum_{i=0}^{N_C} \mathrm{Level}(\mathrm{KC})_i}{N_C}, \tag{4}$$

where N_C is the number of learning paths that contain the KC.

(2) The frequency feature (f_2) focuses on extracting the frequency of a KC in all learning paths. A greedy algorithm is used to search for the frequency of the KC across all learning paths. The frequency of the KC can be defined as follows:

$$f_2(\mathrm{KC}) = \frac{N_C}{N}, \tag{5}$$

where N indicates the total number of learning paths.

(3) The connection feature (f_3) considers the connections between KCs. When KCs occur in the same learning path, they are considered connected. For example, in a learning path "A-B-C," skill A is connected to skills B and C. The connectivity can be calculated as follows:

$$f_3(\mathrm{KC}) = \frac{|\mathrm{ConnectSet}(\mathrm{KC})|}{K}, \tag{6}$$

where ConnectSet represents the list of connected KCs and K is the total number of KCs.

(4) The similarity feature (f_4) is used to explore the similarity between each skill. It utilizes the skill representations (s^*) and calculates the dot product to measure the similarity between skills:

$$f_4(\mathrm{KC}) = \frac{s_i^* \cdot s_j^*}{|s_i^*||s_j^*|}. \tag{7}$$

(5) The difficulty feature (f_5) leverages students' interactions to indicate the difficulty of skills. It models the cognitive difficulty of skills based on students' behavior when they attempt exercises containing the same skill at different timestamps. The cognitive difficulty of a skill set for each student (S_i) at timestamp t is represented by $\pi_{S_i,\mathrm{KC},t}$:

$$\pi_{S_i,\mathrm{KC},t} = \begin{cases} \left[\dfrac{|\{A_s == 0\}|_{0:\,t}}{|Q|_{0:\,t}} * 4\right], & \text{if } |N_v|_{0:\,t} \geq 5, \\ 5, & \text{otherwise,} \end{cases} \tag{8}$$

where $A_s = 0$ represents the set of exercises where the student answered incorrectly for the exercises containing the KC. The cognitive difficulty of the KC is divided into five levels if a learner has performed fewer than five attempts to answer the exercise. The average cognitive difficulty for different learners on the KC is defined as the difficulty feature of the KC:

$$f_5(\mathrm{KC}) = \frac{\sum_{i=0}^{N_S}\sum_{j=0}^{N_T}\pi_{S_i,\mathrm{KC},t_j}}{N_S \times N_T}, \tag{9}$$

where N_S is the number of students and N_T is the time consumption.

In this paper, we consider the novelty and popularity of KCs. To satisfy different learning preferences, we apply a weighted method (W) to combine the five skill features using equation (12). Each weight (w) in our learning preference options corresponds to a certain feature (f_i). The weight distribution details are as follows.

(6) Novelty: When considering the novelty of exercises, we set $w_1 = 0.5$, $w_2 = 0$, $w_3 = 0$, $w_4 = 0$, and $w_5 = 0.5$. We consider the level and difficulty of skills to represent the inherent novelty of exercises.

(7) Popularity: When considering the popularity of exercises, we set $w_1 = 0, w_2 = 0.6, w_3 = 0.1, w_4 = 0.3$, and $w_5 = 0$. We consider the frequency, connection, and similarity of skills to model the popularity of skills.

$$W = \sum_{i=0}^{5} W_i \times f_i(KC). \tag{10}$$

The weighted method calculates the weight of novelty (W_{nov}) and popularity (W_{pop}) for each KC.

Finally, by combining novelty and popularity, the skill importance weight (W_K) is obtained. The formulation is as follows:

$$\begin{aligned} W_{\mathrm{skill}} &= W_{\mathrm{nov}} + W_{\mathrm{pop}}, \\ W_K &= \mathrm{Tanh}(W_{\mathrm{skill}}), \end{aligned} \tag{11}$$

where $\tanh = (e^z - e^{-z})/(e^z + e^{-z})$.

The knowledge importance weight extraction algorithm allows us to determine the weights of KCs based on their skill features, incorporating novelty and popularity considerations. These weights play a crucial role in assessing the importance of KCs within the learning context.

As a result, we further discuss the relationship between skills, which contributes to develop the next exercise factor in the cognitive diagnosis model and representativeness evaluation metrics.

At the same time, the interpretability of the knowledge importance component can be discussed from two aspects. The first aspect is domain knowledge interpretability. As the mentioned above, two experts in this field from our university have labeled the knowledge graph to update any inaccurate or missing information. The second aspect is the skills importance interpretability. We delve more into the five qualities of skills: level, frequency, connection, similarity, and difficulty. These five metrics are interpretable in the knowledge importance component. Therefore, we believe that the knowledge importance component is also interpretable by incorporating the knowledge graph interpretability with skills interpretability.

4.4. Exercise Representativeness Component. After collecting a candidate set E_C of highly informative exercises and obtaining skill importance weights and exercise embeddings, this section focuses on designing the exercise representativeness component. The goal is to select exercises from E_C into the tested exercise set E_T that exhibit high representativeness. To assess the informativeness of exercises, a novel scoring function is proposed to evaluate the knowledge coverage of E_T. An approach is then devised to gradually add more exercises to E_T until it achieves the highest coverage score.

The knowledge coverage of the tested exercise set E_T can be estimated by checking whether the corresponding KCs exist in E_C. Therefore, a straightforward knowledge coverage function, denoted as SKC, is designed as follows:

$$\begin{aligned} \mathrm{Cov}(\mathrm{KC}, Q_c) &= 1 \text{ when } \exists \mathrm{KC} \in Q_C, \\ \mathrm{SKC}(Q_c) &= \frac{\sum_{i=0}^{K}\mathrm{Cov}(\mathrm{KC}, Q_c)}{|K|}, \end{aligned} \tag{12}$$

where $\mathrm{Cov}(\mathrm{KC}, Q_C) = 1$ indicates that the KC is involved in E_C. However, SKC has two obvious flaws. Firstly, it considers all KCs equally and fails to distinguish the importance of each KC. Secondly, the value of Cov is binary and does not reflect the number of exercises. For example, if the math quiz focuses on "Fractions" rather than "Real Numbers," it is more appropriate to select more fractions-related problems rather than simply covering both topics equally. Choosing nine exercises about "Real Numbers" and one exercise about "Fractions" should be equivalent to choosing five exercises from each.

To address these flaws, the exercise weight knowledge coverage function (EWKC) is proposed to calculate the knowledge coverage of the tested exercise set E_T. Specifically, the EWKC function combines the number of exercises to generate the knowledge coverage of E_C. Moreover, to account for the importance of exercises and KCs, skill importance weights obtained from the knowledge importance component are incorporated. The EWKC function is defined as follows:

$$\mathrm{cnt}(\mathrm{KC}, Q_T) = \sum_{q \in Q_C} 1\left[(q, \mathrm{KC}) \in Q_T\right],$$

$$E\mathrm{Cov}(\mathrm{KC}, Q_T) = \frac{\mathrm{cnt}(\mathrm{KC}, Q_T))}{1 + e^{(-\mathrm{cnt}(\mathrm{KC}, Q_T))}}, \tag{13}$$

$$\mathrm{EWKC}(Q_T) = \frac{\sum_{k \in K} W_k E\mathrm{Cov}(k, Q_T)}{\sum_{k \in K} W_k},$$

where W_K represents the exercise weight for the concept k, which is discussed in the knowledge importance component. The ECov function counts the occurrence of a KC in E_T and applies a sigmoid function to ensure the coverage value lies within the range of 0 and 1. Finally, the EWKC function calculates the weighted average knowledge coverage over all KCs, with weights determined by the importance of the corresponding skills.

However, the EWKC function only considers the impact of skills and ignores the influence of exercises. To better define the representativeness of exercises, the response matrix (P_n) and dissimilarity matrix ($\widehat{E}$) are introduced.

4.4.1. Response Matrix. The response matrix P_n of size $|S| \times N_e$ is designed, where each element is defined as follows:

$$P_n(i, j) = \begin{cases} a_j^i, & \text{if } j \leq N_e, \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

The matrix P_n stores the probability of students answering the next exercises correctly. $|S|$ represents the number of students, $|C|$ represents the number of exercises done by each student, and N_e represents the total number of exercises. If an exercise was not done by a student, the corresponding columns are filled with zeros. These columns correspond to the $N_e - |C|$ hypothetical exercises that students cannot answer correctly and will be replaced by other exercises in the future.

4.4.2. Dissimilarity Matrix. To consider exercise representativeness, the dissimilarity between exercises, denoted as $\widehat{E}$, is defined as follows:

$$\widehat{E}_{\mathrm{ij}} = 1 - \frac{\widehat{e_i^*} \cdot \widehat{e_j^*}}{\left|\widehat{e_i^*}\right|\left|\widehat{e_j^*}\right|}, \tag{15}$$

where $\widehat{e^*}$ represents the exercise representation based on the exercise representation component.

The final knowledge coverage combines skill features and exercise features to measure the representativeness of exercises. It is defined as follows:

$$R_{\mathrm{ij}} = \alpha_1 \sum_{KC \in Q_T} \mathrm{EWKC}(\mathrm{KC}) + \alpha_2 P_n(i, j) + \alpha_3 \widehat{E}_{\mathrm{ij}}, \tag{16}$$

where α_1, α_2, and α_3 are hyperparameters in the model. The representativeness of exercises is evaluated based on the weighted sum of the knowledge coverage of KCs in E_T, the probabilities stored in the response matrix P_n, and the dissimilarity matrix $\widehat{E}$.

The exercise representativeness component aims to select exercises from the candidate set E_C into the tested exercise set E_T with high representativeness. The knowledge coverage of E_T is evaluated using the EWKC function, which incorporates skill importance weights and exercise numbers. The response matrix P_n and dissimilarity matrix $\widehat{E}$ are introduced to consider exercise features and representativeness. The final knowledge coverage is calculated by combining skill and exercise features. The hyperparameters α_1, α_2, and α_3 control the relative importance of these features in measuring exercise representativeness.

This component also provides the KI-EIR framework with another aspect of exercises' representativeness interpretability. It means that the KI-EIR model can evaluate the quality of exercise based on informativeness metric and representativeness metric. Then, according to the quality of exercises, different exercises are recommended to different types of students.

4.5. Cognitive Diagnosis Model. This section introduces a novel cognitive diagnosis model, NACD, within the KI-EIR framework. The NACD model aims to estimate the knowledge state of students and make predictions on the tested exercise set. To achieve accurate diagnosis, the NACD model incorporates student factor modeling and exercise factor modeling. The student factor modeling focuses on capturing students' behavior during exercise training, specifically their slipping behavior and guessing behavior. Additionally, the exercise factor is modeled based on the output e^* generated by the exercise representation component. The exercise factor aims to explore the relationship between exercises and skills and utilizes the exercise-skill relation matrix as input for a relative-distance attention mechanism to generate the exercise factor representation.

4.5.1. Exercise Factor. To model the relationship between exercises and skills, an exercise-skill relation matrix Q is constructed to map exercises to skills. In order to consider the relationship between the skills, the skill importance weight: W_K, obtained from the knowledge importance component, is considered to generate the exercise-skill relation matrix. Specifically, when an exercise e_i contains a knowledge point k, the corresponding position will be replaced by W_K, which considers the popularity and novelty of skills. Based on the exercise-skill relation matrix, the knowledge point vector of exercise e can be obtained as follows:

$$K^V = x^e \times Q^T, \tag{17}$$

where $Q \in \mathbb{R}^{N_e \times K}$ and $x^e \in 0, 1^{N_e \times 1}$ represents the one-hot representation of exercises. The exercise embedding, composed of the corresponding K^V values, is then used as input for the relative distance mechanism. The relative distance between input sequences, represented by x_i and $x_j = (x_1, x_2, \cdots, x_{n-1})$, is captured using edge vectors $a^V_{i,j}$ and $a^K_{i,j}$. To prevent unbounded values, the edge vectors are clipped using the function $\text{clip}(x, k) = \max(-k, \min(k, x))$, where k represents the maximum absolute value. The associated relative position representations for W^K and W^V are defined as $W^K = (W^k_{-k}, \cdots, W^K_k)$ and $W^V = (W^V_{-k}, \cdots, W^V_k)$. Finally, the relative position attention mechanism outputs the exercise factor representation F^E. The following equations describe the process:

$$\begin{aligned} a^k_{i,j} &= W^k_{\text{clip}}(j - i, k), \\ a^V_{i,j} &= W^V_{\text{clip}}(j - i, k). \end{aligned} \tag{18}$$

The edge vectors are then utilized as input for the attention mechanism. The attention weights $a_{i,j}$ are calculated based on the relative distances $e_{i,j}$, which are computed as

$$\begin{aligned} a_{i,j} &= \frac{\exp\left(e_{i,j}\right)}{\sum_{i=1}^{n} \exp\left(e_{i,k}\right)}, \\ e_{i,j} &= \frac{x_i W^Q \left(x_j W^k\right)^T + x_i W^Q \left(a^K_{i,j}\right)^T}{\sqrt{d_{F^E}}}, \\ F^E_i &= \sum_{j=1}^{n} a_{ij}\left(x_j W^V\right). \end{aligned} \tag{19}$$

In the above equations, W^Q, W^K, and W^V represent the query, key, and value matrices, respectively, and d_{F^E} denotes the dimension of F^E.

4.5.2. Student Factor. The student factor, denoted as F^S, models the representation of students based on their knowledge proficiency vectors in diagnosing their states. The formulation for the student representation is as follows:

$$H^S = \text{sigmoid}(x^s \times A), \tag{20}$$

where $x^S \in 0, 1^{S \times 1}$ represents the one-hot encoding of students and A is a trainable matrix within the framework.

Next, we introduce two factors related to student behavior: the slipping factor and the guessing factor. The slipping factor captures situations where a student attempts to complete an exercise but provides an incorrect answer due to careless mistakes. The guessing factor represents instances when a student may guess an answer because they have not fully mastered the corresponding skills. The formulations for the slipping factor and the guessing factor are as follows:

$$\begin{aligned} H^{\text{Slipping}} &= \text{sigmoid}(x^e \times B), \\ H^{\text{Guessing}} &= \text{sigmoid}(x^e \times C), \end{aligned} \tag{21}$$

where B and C are trainable matrices.

To generate the student factor representation F^S, we incorporate H^S, H^{Slipping}, and H^{Guessing} as inputs to a two-layer linear network. The input for the linear network is defined as follows:

$$X = Q \times \left(H^S - H^{\text{Slipping}}\right) \times H^{\text{Guessing}}. \tag{22}$$

Subsequently, X serves as the input for the linear network:

$$\begin{aligned} f_1 &= \sigma(W_1 \times X + b_1), \\ F^S &= \sigma(W_2 \times f_1 + b_2), \end{aligned} \tag{23}$$

where $\sigma(\,)$ represents the sigmoid activation function.

4.5.3. Student Performance Prediction. To predict student performance on a tested exercise set, E_T, we combine the student factor F^S and the exercise factor F^E. The formulation for this prediction is as follows:

$$p = \text{sigmoid}\left(W_s \times F^S + W_e \times F^E + b_p\right), \tag{24}$$

where W_s and W_e are weighted matrices, b_p is the bias vector, and p represents the likelihood that the student will answer the subsequent interaction exercise, denoted as e_{N_e+1}, correctly.

5. Experiments

In this section, we conduct experiments using two public educational datasets: the Assistment dataset and the Eedi dataset to investigate the performance of our selection strategy: KI-EIR. The experiments are organized into five aspects. First, we compare the novel cognitive diagnosis model, NACD, with baseline models in terms of AUC and ACC matrices to validate the effectiveness of the NACD model. Second, we compare the performance of our selection strategies with the random strategy and EM strategy using the informativeness metric. The experimental results demonstrate that our strategy outperforms other selection strategies. Third, we discuss the performance of our strategies compared to other strategies using the representativeness metric. Next, we present the visualization of the recommendation process of the KI-EIR strategy, EM strategy, and random strategy using heatmaps to highlight the excellent performance of the KI-EIR strategy. Finally, we explore the key components of the KI-EIR method further on the Eedi dataset.

5.1. Dataset Descriptions. We use two datasets in this paper: the Assistment (ASSIST) dataset and the Eedi dataset. The Assistment dataset is generated by collecting information from the Assistment Online Tutoring Systems. It is an open-source dataset for researchers to perform cognitive diagnosis tasks. The experiments in this paper are conducted on the problem bodies of this dataset.

The Eedi2020 dataset is obtained from the NeuralIPS platform, which collected 1,380,000 records from 4918 students. Each student participated in an average of 280 workouts. For this study, we use problems 3 and 4 from the NeuralIPS dataset to compare the performance of our model.

The statistical information of the Assistment and Eedi2020 datasets is shown in Table 3.

5.2. Baselines and Selection Strategies. To evaluate the KI-EIR method, we test it based on four standard cognitive diagnosis models: IRT, MIRT, NCDM, and KaNCDM. The details of these models are as follows:

(1) IRT [50]: This is the most popular CDM in computerized adaptive learning. IRT and conventional approaches focus on developing and applying multiitem scales to assess "latent variables" (hypothetical constructs).

(2) MIRT [51]: MIRT is a multidimensional extension of IRT that demonstrates its potential for estimating several characteristics of ability. The IRT-based methods have also been expanded to accommodate MIRT.

(3) NCDM [30]: This cognitive diagnosis model is the most standard model in the educational data mining field. The NCDM model employs neural networks to learn the complex relationships of exercises in order to produce accurate and understandable diagnosis results.

(4) KaNCDM [52]: This framework is further developed based on NCDM to estimate the current knowledge state of students. The KaNCDM improves upon NCDM in terms of feasibility, generality, and extensibility to make predictions. Extensibility is further discussed from two aspects: content-based extension and knowledge-association-based extension.

We also apply two selection strategies to compare the performance of our selection strategy. The details of these selection strategies are as follows:

(1) Random strategy (RM): This strategy serves as the baseline for the selection strategies. It randomly selects exercises from the exercise set without considering the overall performance.

(2) Expectimax strategy [53] (EM): Expectimax is a tree-based, brute-force MDP search algorithm that determines the expected utility of each action. It assumes that the agent will always choose the option that maximizes utility and that the environment will generate a subsequent state using a stochastic process after an action has been taken. Specifically, we treat the exercises as the state and the knowledge concept importance as reward to recommend the exercises to students.

However, after observing the volume of datasets and the overall structure of the KI-EIR framework referring to Figure 3, we can find that the KI-EIR framework concerns about two types of networks including the graph neural network in the exercise representation component and fully connected network in the cognitive diagnosis model. The Eedi datasets possesses a large volume of data containing 138W student records. Therefore, in order to efficiently run the KI-EIR framework in the intelligent tutoring system, the following optimization tricks for this framework are introduced. The first trick is based on the idea of exchanging space for time. We pretrain exercise representation component to store the exercise embedding and skill embedding rather than running this component in each epoch when we train the KI-EIR framework. The second trick involves integrated training and unified optimization for all components.

The framework settings for the KI-EIR framework are described in this part, as illustrated in Table 4.

5.3. Results and Discussion. In this paper, we evaluate the prediction task of the cognitive diagnosis model based on whether an exercise was successfully answered in the next interaction. We use the Area Under the Curve (AUC) and Accuracy (ACC) metrics to measure students' performance in making predictions. A higher AUC or ACC value indicates better cognitive diagnosis performance, while a value of 0.5 suggests random selection. The cross-entropy loss function is used.

A binary value represents the effectiveness of exercise recommendation. We measure the performance of our selection strategy with the metric of informativeness metric and coverage metric. The informativeness metric (Inf(s)) is used for measuring the informativeness of the selection strategy in the exercise recommendation. The AUC metric is adopted to indicate the informativeness of the selection strategy referring to the following formula:

$$\mathrm{Inf}(s) = \mathrm{AUC}\Big(M\big(e_i, q_j \big| \Theta\big) \big| e_i \in E, q_j \in Q\Big). \tag{25}$$

The coverage metric (Cov(s)) is accepted to measure the representativeness of the selection strategy. Cov(s) is computed based on the percentage of knowledge concepts covered by the strategy-selected exercises.

$$\mathrm{Cov}(s) = \frac{1}{|K|} \sum_{k \in K} 1\left[k \in Q_c\right]. \tag{26}$$

5.3.1. Student Performance Prediction. Table 5 presents a comparison of the results of baseline models with the Neural Attentive Cognitive Diagnosis (NACD) model. The NACD model outperforms all baseline models in terms of AUC and ACC on both the ASSIST and Eedi datasets. The MIRT model demonstrates better performance than the IRT

Table 3: Statistics of the Assistment and Eedi datasets.

Statistic	ASSIST	Eedi2020
Number of records	267415	1382726
Number of students	2493	4918
Number of exercises	17671	948
Avg record/student	107.2	280

Table 4: The framework setting for the KI-EIR framework.

	ASSIST	Eedi2020
Attention embedding size	200	200
Top-K	5	5
Dropout rate	0.2	0.2
Learning rate	0.002	0.002
Number of epochs	100	100
α_1	0.7	0.7
α_2	0.15	0.15
α_3	0.15	0.15

Table 5: Comparison of results of baseline models with the Neural Attentive Cognitive Diagnosis (NACD) model.

	ASSIST		Eedi2020	
	AUC	ACC	AUC	ACC
IRT	0.652	0.618	0.672	0.618
MIRT	0.666	0.645	0.673	0.621
NCDM	0.751	0.728	0.720	0.660
KaNCDM	0.768	0.728	0.773	0.705
NACD-FE	0.769	0.733	0.751	0.703
NACD-FS	0.656	0.636	0.773	0.705
NACD	0.772	0.735	0.775	0.707

NACD outperforms all baseline models in terms of AUC and ACC.

model by extending it into the multidimensional space to predict students' performance. The NCDM model utilizes neural networks to further improve predictions and estimate the current state of students. The KaNCDM model enhances the NCDM model in terms of feasibility, generality, and extensibility, resulting in improved performance. The NACD-FE and NACD-FS models analyze the importance of exercise and student factors in predicting student performance. The results indicate that the student factor is more influential than the exercise factor, and the NACD model, which comprehensively considers both factors, outperforms all models.

In order to discuss the performance deeply, the train-test curve of the NACD model is introduced to measure the performance. The details are as follows. As illustrated in Figure 5, the vertical dimension means the AUC of the current NACD model in the corresponding epoch and the horizontal dimension presents the epoch of current training or testing process. We can observe that the performance of this model is relatively stable on both the training set and the test set. On the training set, the accuracy of the model remains around 84.5% and 80.3% on ASSIST and Eedi, respectively. On the test set, the model's accuracy fluctuated around 77.2% and 77.5%, respectively. According to the performance of train set and testing set, we can also draw a conclusion that the NACD model does not have overfitting problems.

5.3.2. Informativeness Comparison. In this part, we focus on comparing the different informativeness performance of strategies using the AUC metric (equation (25)). Figures 6 and 7 present the results at the middle (step $t = 10$) and final (step $t = 20$) stages of the tests. We compare the KI-EIR strategy with the EM strategy and random strategy by applying different cognitive diagnosis models: NACD, IRT, and MIRT.

The random strategy performs the worst among all strategies on the Eedi dataset and provides the baseline accuracy for the experiment. The EM strategy, which utilizes a Markov decision process, outperforms the random strategy on the Eedi dataset by considering the impact of interactions between exercises and students. However, the EM strategy performs worse than the IRT model on the ASSIST dataset due to the large number of exercises, which leads to inaccurate predictions when each exercise is treated as a state. The KI-EIR strategy, which incorporates exercise and skill features from the knowledge graph, outperforms all models on both datasets, indicating its effectiveness in achieving the informativeness goal. The specific reasons are as follows. The first is that two innovative exercise evaluation metrics are designed including the representativeness and informativeness. By analyzing two exercise metrics, the quality of exercises is correctly modeled to recommend the proper exercises to students. The second is that the KCs in the knowledge map are comprehensively explored to generate the skill importance weight. The last is that the KI-EIR strategy enhances the recommendation system by making the process more flexible without requiring modifications to the general methodology. Therefore, the KI-EIR framework outperforms other selection strategies on both datasets.

5.3.3. Representativeness Comparison. The representativeness comparison focuses on exploring the performance of different selection strategies in terms of the coverage metric. As illustrated in Figures 8–10, the EM strategy performs better than the random strategy because it considers the impact of behavior when selecting exercises on the Eedi dataset. However, the EM strategy performs worse than the random strategy on the ASSIST dataset due to the fact that there exist too many states to result in the inaccurate prediction of Markov decision process.

Compared with previous two selection strategies, the KI-EIR clearly measures the quality of the exercises by defining two exercise quality metrics such as the representative and informativeness metrics. This framework also considers the correlation between KCs in the multiple dimension knowledge graph to recommend the related exercises. Therefore, the KI-EIR framework combines the exercises' quality metrics with the KCs to recommend the exercises and shows significant improvements in coverage metric compared with other strategies, reaching close to 0.8 and 1 on Eedi and ASSIST, respectively.

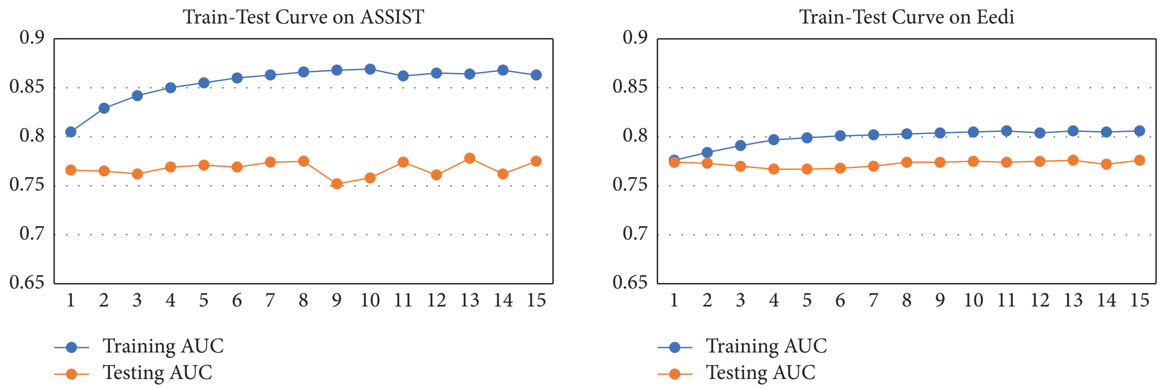

Figure 5: The train-test curve of NACD model on ASSIST and Eedi in terms of AUC metric.

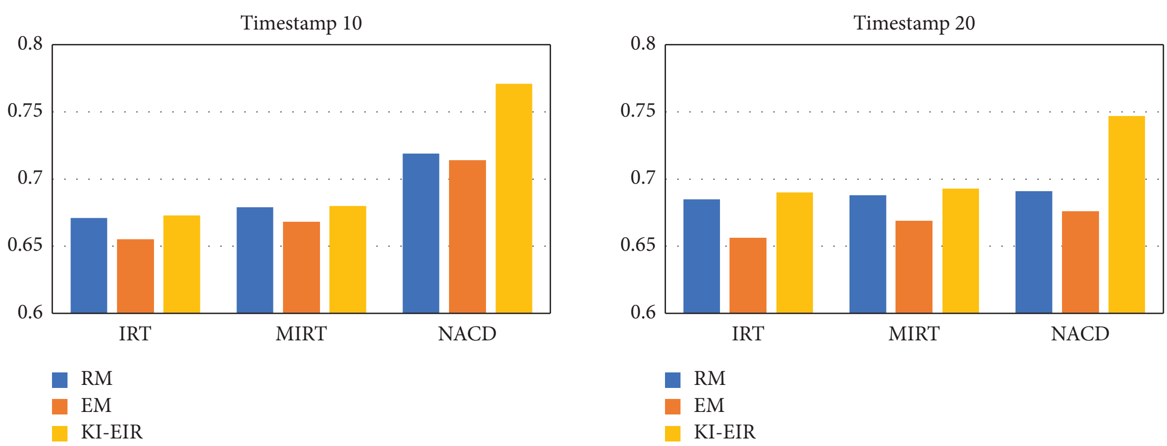

Figure 6: The selection strategies' performance on ASSIST datasets in timestamp 10 and timestamp 20.

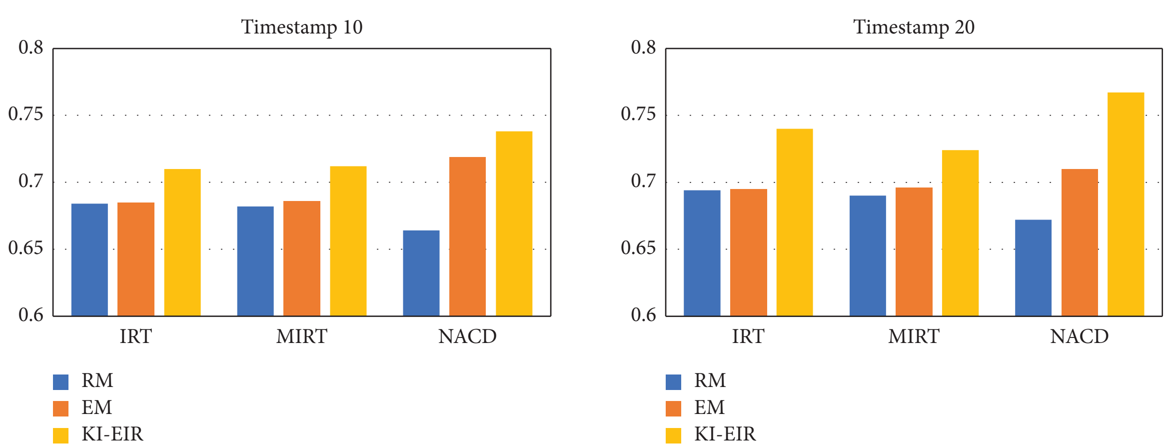

Figure 7: The selection strategies' performance on Eedi datasets in timestamp 10 and timestamp 20.

5.3.4. Visualization of Selection Strategies. In this section, we validate the performance of the selection strategies, including the KI-EIR strategy, EM strategy, and random strategy, in recommending exercises to improve student performance on the Eedi dataset. Heatmap is used to visualize the evolution of student performance, as measured by the AUC metric.

The heatmap in Figure 11 depicts the differences in performance based on different selection strategies (KI-EIR, EM, and random) by observing the color change. The vertical dimension represents the selection strategies (KI-EIR, EM, and random), while the horizontal dimension represents the different testing phases from 0 to 19. The color of the heatmap represents the performance of students when recommended with appropriate exercises, with stronger colors indicating a greater impact of the selection strategies.

According to Figure 11, the random strategy performs worse than the other selection strategies and is treated as the

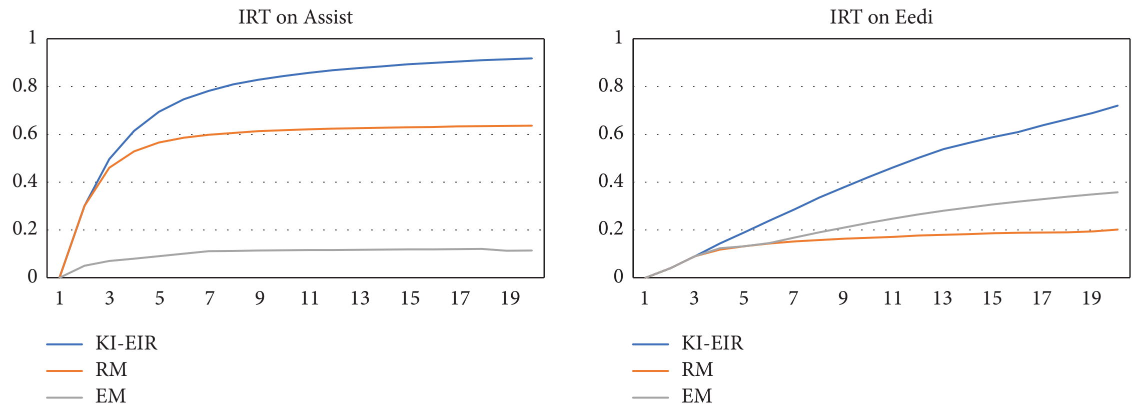

Figure 8: IRT on Eedi2020 and Assistment.

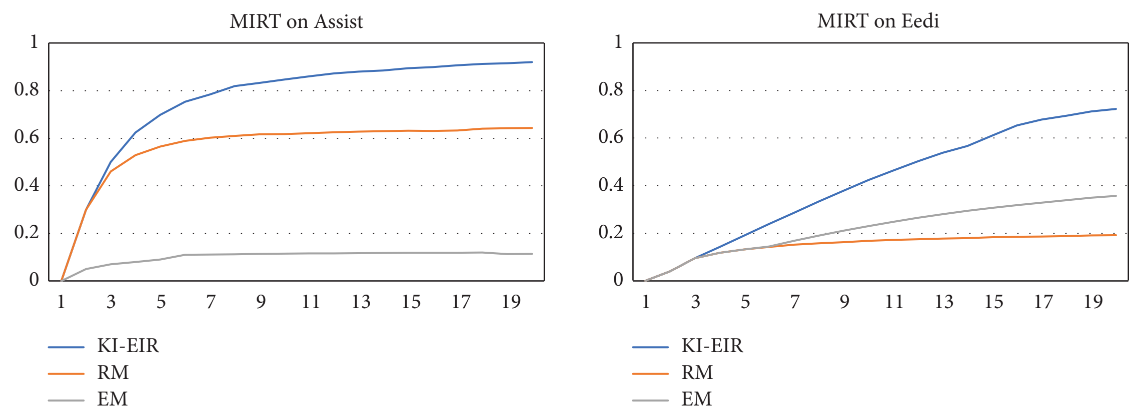

Figure 9: MIRT on Eedi2020 and Assistment.

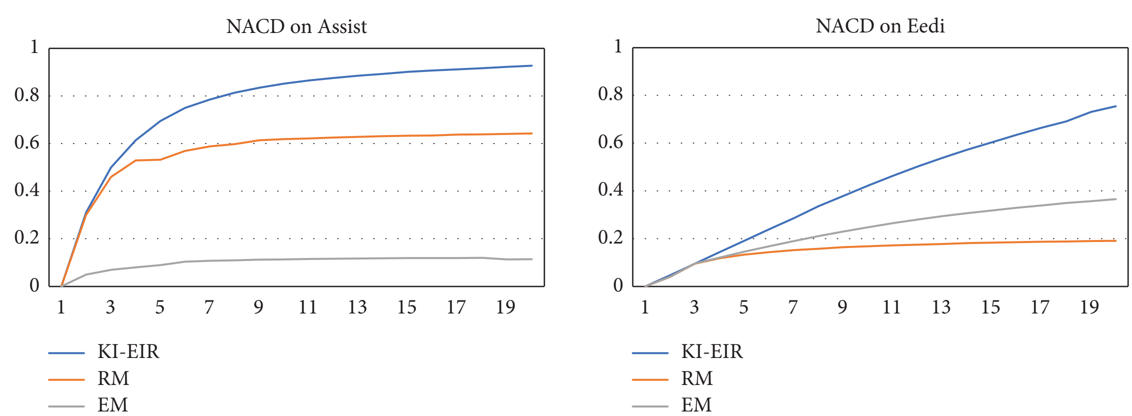

Figure 10: NACD on Eedi2020 and Assistment.

baseline of three selection strategies. The EM strategy, which considers the exercises' states and applies the Markov decision process to recommend exercises, outperforms the random strategy in all testing phases. However, the EM selection strategy ignores the exercises and skill features. Therefore, the KI-EIR framework further discusses the exercises' features by introducing two innovative quality metrics including representativeness and informativeness. This framework also deeply explores the relation between different KCs in KGs to assist in the recommendation of the intelligent tutoring system. As a result, the KI-EIR strategy achieves the best student performance among the selection strategies.

Overall, the results and discussions demonstrate the effectiveness and superiority of the proposed KI-EIR strategy in cognitive diagnosis and exercise recommendation. The KI-EIR strategy outperforms baseline models and other selection strategies in terms of informativeness and representativeness. It leverages exercise and skill features, along with the knowledge graph, to provide accurate and

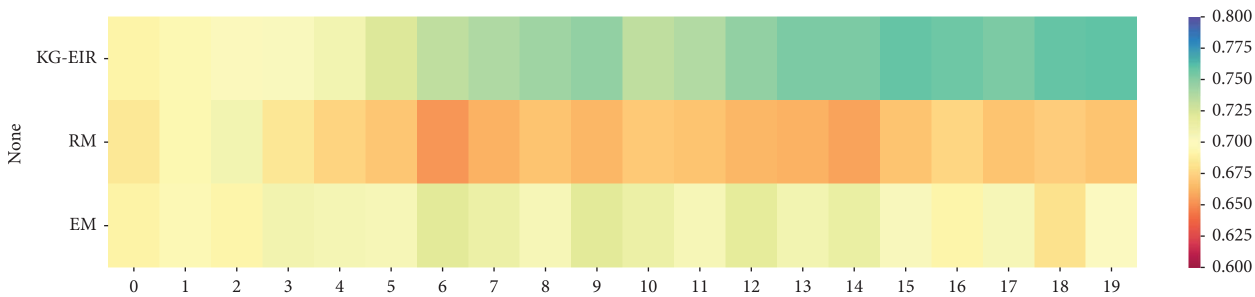

Figure 11: Heatmap illustrating the performance of selection strategies on the Eedi2020 dataset.

Table 6: Ablation study of the KI-EIR model based on the IRT model on two datasets.

Method	Eedi	ASSIST
Informativeness component (IF)	0.654	0.670
Exercise representativeness component (ER)	0.656	0.672
Informativeness component (IF) + knowledge importance component (ER)	0.661	0.676
Exercise representativeness component (ER) + knowledge importance component (KI)	0.662	0.675
KI-EIR	0.672	0.679

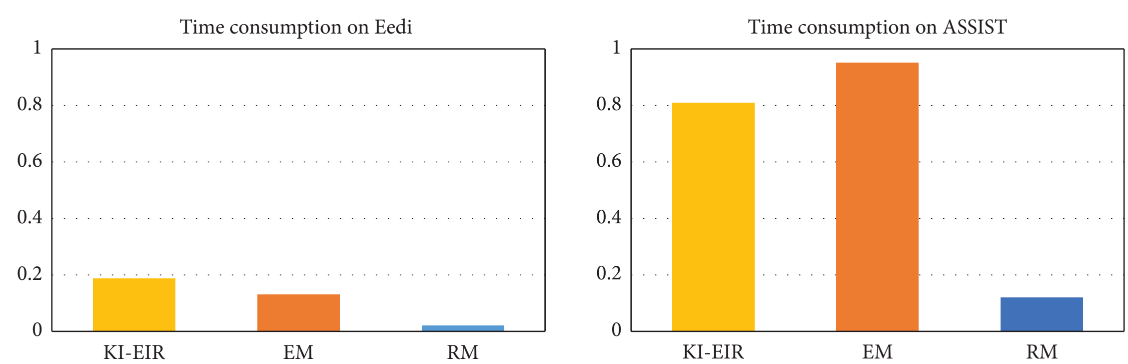

Figure 12: The time consumption comparison on Eedi and ASSIST.

informative exercise recommendations for students. The visualization of the selection strategies further supports the outstanding performance of the KI-EIR strategy. These findings contribute to the improvement of cognitive diagnosis models and the enhancement of recommendation systems in educational settings.

5.3.5. Ablation Experiments. This section aims to identify the key components of the KI-EIR model through a series of ablation experiments. Four variations of the KI-EIR model are considered, each incorporating one or more components. Specifically, “IF,” “ER,” and “KI” indicate that the KI-EIR framework only contains the informativeness component, exercise representativeness component, and knowledge importance component, respectively. “IF + KI” or “ER + KI” indicates that the KI-EIR framework only contains the informativeness component and knowledge importance component or exercise representativeness component and knowledge importance component, respectively.

The conclusions drawn from the experiments are as follows. Firstly, the individual components of informativeness, exercise representativeness, and knowledge importance do not yield satisfactory outcomes when used alone. The performance gradually improves as more components are incorporated into the KI-EIR method. Secondly, when the exercise representativeness component (ER) is involved in the KI-EIR framework, the performance becomes better than IF, increasing to 65.6% and 67.2%. Therefore, the exercise representation component (ER) is more important than the informativeness component (IF) according to the experimental results in this section. Thirdly, since the knowledge importance component (KI) provides the skill weight for ER, when the ER component is removed, the KI component is also removed. Consequently, the inclusion of the KI component leads to greater performance improvement compared to the IF component.

Table 6 presents the results of the ablation study of the KI-EIR model based on the IRT model on the Eedi and ASSIST datasets.

The results in Table 6 demonstrate that the KI-EIR model outperforms the ablated versions in terms of AUC on both the Eedi and ASSIST datasets. The inclusion of all components in the KI-EIR model leads to the best performance. The ablation study confirms the importance of the informativeness, exercise representativeness, and knowledge importance components in the KI-EIR model, with their combined effect resulting in improved performance.

In order validate the performance of these tricks for these key components in the KI-EIR framework, we test the time consumption of recommending exercises for a student. The details are shown in Figure 12.

The conclusions drawn from the experiments are as follows. Firstly, the response time of KI-EIR, EM, and RM strategies is all below 1 s compared with the maximum user-tolerable page loading time: 2 s. It means that the optimization tricks are proved as efficient to reduce the complexity of the KI-EIR framework. Secondly, three selection strategies spend more time on the ASSIST dataset than on the Eedi dataset due to the fact that the exercise number of the ASSIST dataset is much larger than the Eedi dataset.

6. Conclusion and Future Work

In this paper, we proposed a comprehensive framework, the KI-EIR (Knowledge Graph-Enhanced Exercise Item Recommendation) model, to address the challenge of providing informative and representative exercises in cognitive diagnosis tasks. The KI-EIR model consists of four key components: informativeness, exercise representation, knowledge importance, and exercise representativeness. We can also observe that the KI-EIR possesses two types of scalability. The first scalability is the model scalability. We can use IRT or MIRT to recommend the exercises though these cognitive diagnosis models do not provide excellent performance referring to the experiments: the informativeness comparison and representativeness comparison. The second scalability is the dataset scalability. We compare the large dataset: Eedi2020, with the small dataset: ASSIST, to validate our performance. The results indicate that our framework is extensible in both two types of datasets.

The informativeness component estimates the informativeness of each exercise and selects exercises with high informativeness from the untested exercise set to the candidate exercise set. The exercise representation component utilizes the Graph Convolutional Network (GCN) model and two types of relation attention mechanisms to generate skill embeddings and exercise embeddings. The knowledge importance component applies the knowledge point extraction path algorithm and knowledge importance weighted algorithm to calculate the skill importance weight. Finally, the exercise representativeness algorithm combines the skill importance weight, exercise weight, knowledge coverage, response matrix, and dissimilarity matrix to select exercises from the candidate exercise set into the tested exercise set with high representativeness. The NACD model is then employed to accurately estimate the state of students based on the selected exercises. The recommendation process is also interpretable. The four aspects of interpretability of the KI-EIR framework can be discussed. The first aspect is the domain knowledge interpretability. The knowledge graph is labeled by the two domain experts in our university to correct these outdated or incompleteness information in Section 4.3. The domain knowledge in the validated knowledge graph is inherently interpretable because we can easily observe the specific knowledge concepts and the relationship between them. The second aspect is exercise informativeness and representativeness interpretability. In order to measure the quality of exercises and make the recommendation process interpretable, two metrics are proposed to measure the exercises including the informativeness and representativeness in Section 4.1 and Section 4.4. Therefore, the exercise recommendation process of KI-EIR is interpretable because the KI-EIR framework just needs to select the exercises with high informativeness and representativeness. The third aspect is the learning interaction interpretability. The heterogeneous graph in Section 4.2 provides the insight of the relationship between students, exercises, and skills. Then, the graph neural network is applied to extract the information contained in the heterogeneous graph and generate the corresponding skill embedding and exercise embedding. Therefore, the modeling of the learning interactions is interpretable in the KI-EIR framework. The final aspect is the skill importance interpretability. We further explore the five properties of skills in Section 4.3.2 including the level, frequency, connection, similarity, and difficulty. These five properties of skills are interpretable. As a result, the KI-EIR, which can demonstrate interpretability from four aspects, can provide the interpretable and transparent recommendations to students.

The KI-EIR model demonstrates promising results in improving cognitive diagnosis and exercise recommendation in educational settings. By leveraging the power of knowledge graphs and incorporating multiple components, our framework provides accurate and informative exercise recommendations for students, thereby enhancing their learning experience and academic performance. The proposed framework opens up new avenues for research and development in the field of educational data mining and cognitive diagnosis. However, the KI-EIR framework still possesses some limitations to recommend the exercises to students. The first limitation is that some metrics are crucial for the recommendation system, but it is hard for the KI-EIR framework to consider these metrics in current stage such as the user satisfaction, user engagement, and long-term learning outcomes. The reason is as follows. The first is that these metrics require a large number of real users involved after the platform is put into practice. The second is that data of other recommendation systems cannot be directly referenced because different platforms possess different user satisfaction, user engagement, and long-term learning outcomes so that other recommendation system datasets cannot be referenced directly. The second limitation is that some nuanced factors are valuable for KI-EIR, but these factors are hard to be quantified based on the datasets such as nuanced pedagogical methods for different teachers and individual learning styles.

In future research, we suggest exploring the application of reinforcement learning techniques such as Deep Q-Network (DQN) to further improve the selection of exercises with high representativeness and informativeness. This approach can help reduce the time required for the selection phase and enhance the efficiency and effectiveness of the cognitive diagnosis process. At the same time, we also need to further discuss the optimization methods of our intelligent tutoring system to reduce the time consumption of the recommending process. Finally, we will put the KI-EIR framework into practice and obtain some valuable metrics such as user satisfaction, engagement, and long-term learning resources to further improve our framework.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Disclosure

Our paper has been submitted in preprint form to arXiv [7].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (nos. 62177022, 61901165, and 61501199), AI and Faculty Empowerment Pilot Project (no. CCNUAI&FE2022-03-01), Collaborative Innovation Center for Informatization and Balanced Development of K-12 Education by MOE and Hubei Province (no. xtzd2021-005), and Natural Science Foundation of Hubei Province (no. 2022CFA007). We would like to sincerely thank Prof. Shi Dong and Prof. Mingzhang Zuo from Central China Normal University for their assistance in annotating and correcting the subject knowledge graph used in our research.

References

[1] Y. Cui, Z. Ma, L. Wang et al., "A survey on big data-enabled innovative online education systems during the covid-19 pandemic," *Journal of Innovation & Knowledge*, vol. 8, no. 1, Article ID 100295, 2023.

[2] L. Ma, X. Zhang, Z. Wang, and H. Luo, "Designing effective instructional feedback using a diagnostic and visualization system: evidence from a high school biology class," *Systems*, vol. 11, no. 7, p. 364, 2023.

[3] H. Ma, Z. Huang, W. Tang, and X. Zhang, "Exercise recommendation based on cognitive diagnosis and neutrosophic set," in *Proceedings of the 2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pp. 1467–1472, IEEE, Hangzhou, China, May 2022.

[4] Z. Wang, Y. Hou, C. Zeng, S. Zhang, and R. Ye, "Multiple learning features–enhanced knowledge tracing based on learner–resource response channels," *Sustainability*, vol. 15, no. 12, p. 9427, 2023.

[5] W. Kang, L. Zhang, B. Li, J. Chen, X. Sun, and J. Feng, "Personalized exercise recommendation via implicit skills," in *Proceedings of the ACM Turing Celebration Conference-China*, pp. 1–6, Chengdu China, May 2019.

[6] L. Lyu, Z. Wang, H. Yun, Z. Yang, and Y. Li, "Deep knowledge tracing based on spatial and temporal representation learning for learning performance prediction," *Applied Sciences*, vol. 12, no. 14, pp. 7188–7221, 2022.

[7] L. Li and Z. Wang, "Knowledge graph enhanced intelligent tutoring system based on exercise representativeness and informativeness," 2023, https://arxiv.org/abs/2307.1507.

[8] Z. Wu, M. Li, Y. Tang, and Q. Liang, "Exercise recommendation based on knowledge concept prediction," *Knowledge-Based Systems*, vol. 210, Article ID 106481, 2020.

[9] L. Li and Z. Wang, "Calibrated q-matrix-enhanced deep knowledge tracing with relational attention mechanism," *Applied Sciences*, vol. 13, no. 4, pp. 2541–2624, 2023.

[10] N. L. Stout, D. Santa Mina, K. D. Lyons, K. Robb, and J. K. Silver, "A systematic review of rehabilitation and exercise recommendations in oncology guidelines," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 2, pp. 149–175, 2021.

[11] I. Kamsa, R. Elouahbi, F. El Khoukhi, T. Karite, and H. Zouiten, "Optimizing collaborative learning path by ant's optimization technique in e-learning system," in *Proceeedings of the 2016 15th International Conference on Information Technology Based Higher Education and Training (ITHET)*, pp. 1–5, IEEE, Istanbul, Turkey, September 2016.

[12] S. Jia, X. Liu, P. Zhao, C. Liu, L. Sun, and T. Peng, "Representation of job-skill in artificial intelligence with knowledge graph analysis," in *Proceeedings of the 2018 IEEE Symposium on Product Compliance Engineering-asia (ISPCE-CN)*, pp. 1–6, IEEE, Shenzhen, China, December 2018.

[13] M. A. Pellegrino, M. Santoro, V. Scarano, and C. Spagnuolo, "Automatic skill generation for knowledge graph question answering," *Semantic Web: ESWC 2021 Satellite Events: Virtual Event*, vol. 18, pp. 38–43, 2021.

[14] A. J. Martin and M. M. Dominic, "Personalization of learning objects according to the skill set of the learner using knowledge graph," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 6, pp. 3974–3987, 2021.

[15] P. E. Sanderson, "E-Learning: strategies for delivering knowledge in the digital age," *The Internet and Higher Education*, vol. 5, no. 2, pp. 185–188, 2002.

[16] Z. Wang, W. Yan, C. Zeng, Y. Tian, and S. Dong, "A unified interpretable intelligent learning diagnosis framework for learning performance prediction in intelligent tutoring systems," *International Journal of Intelligent Systems*, vol. 2023, Article ID 4468025, 20 pages, 2023.

[17] A. Saxena, A. Kochsiek, and R. Gemulla, "Sequence-to-sequence knowledge graph completion and question answering," 2022, https://arxiv.org/abs/2203.10321.

[18] H. Wang, F. Zhang, J. Wang et al., "Exploring high-order user preference on the knowledge graph for recommender systems," *ACM Transactions on Information Systems*, vol. 37, no. 3, pp. 1–26, 2019.

[19] P. Lv, X. Wang, J. Xu, and J. Wang, "Utilizing knowledge graph and student testing behavior data for personalized exercise recommendation," in *Proceedings of the ACM Turing Celebration Conference-China*, pp. 53–59, Shanghai China, May 2018.

[20] L. Zhu, Y. Liu, X. Hei et al., "A study on exercise recommendation method using knowledge graph for computer network course," in *Proceedings of the 2020 International*

Conference on Networking and Network Applications (NaNA), pp. 436–442, IEEE, Haikou City, China, December 2020.

[21] D. Shi, T. Wang, H. Xing, and H. Xu, "A learning path recommendation model based on a multidimensional knowledge graph framework for e-learning," *Knowledge-Based Systems*, vol. 195, Article ID 105618, 2020.

[22] H. Bi, H. Ma, Z. Huang et al., "Quality meets diversity: a model-agnostic framework for computerized adaptive testing," in *Proceedings of the 2020 IEEE International Conference on Data Mining (ICDM)*, pp. 42–51, IEEE, Sorrento, Italy, November 2020.

[23] P. Lv, X. Wang, J. Xu, and J. Wang, "Intelligent personalised exercise recommendation: a weighted knowledge graph-based approach," *Computer Applications in Engineering Education*, vol. 29, no. 5, pp. 1403–1419, 2021.

[24] T. Zhao, Y. Huang, S. Yang et al., "Mathgraph: a knowledge graph for automatically solving mathematical exercises," in *Proceedings of the Database Systems for Advanced Applications: 24th International Conference, DASFAA 2019*, pp. 760–776, Springer, Chiang Mai, Thailand, April 2019.

[25] H. Zhu, F. Tian, K. Wu et al., "A multi-constraint learning path recommendation algorithm based on knowledge map," *Knowledge-Based Systems*, vol. 143, pp. 102–114, 2018.

[26] S. Chen and T. Joachims, "Predicting matchups and preferences in context," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 775–784, San Francisco, CA, USA, August 2016.

[27] J. Xu, C. Deng, X. Gao, D. Shen, and H. Huang, "Predicting alzheimer's disease cognitive assessment via robust low-rank structured sparse model," in *Proceedings of the Conference IJCAI*, Melbourne, Australia, August 2017.

[28] Q. Liu, R. Wu, E. Chen et al., "Fuzzy cognitive diagnosis for modelling examinee performance," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 9, no. 4, pp. 1–26, 2018.

[29] Y. Chen, X. Li, J. Liu, and Z. Ying, "Recommendation system for adaptive learning," *Applied Psychological Measurement*, vol. 42, no. 1, pp. 24–41, 2018.

[30] F. Wang, Q. Liu, E. Chen et al., "Neural cognitive diagnosis for intelligent education systems," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 4, pp. 6153–6161, 2020.

[31] Q. Liu, "Towards a new generation of cognitive diagnosis," in *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, pp. 4961–4964, Montreal, August 2021.

[32] F. Lord, "A theory of test scores. Psychometric monographs," 1952, https://www.psychometricsociety.org/sites/main/files/file-attachments/mn07.pdf?1576607452.

[33] J. De La Torre, "Dina model and parameter estimation: a didactic," *Journal of Educational and Behavioral Statistics*, vol. 34, no. 1, pp. 115–130, 2009.

[34] T. A. Ackerman, "Multidimensional item response theory models," *Wiley StatsRef: Statistics Reference Online*, Wiley, NJ, USA, 2014.

[35] E. Tsutsumi, R. Kinoshita, and M. Ueno, "Deep-irt with independent student and item networks," *International Educational Data Mining Society*, vol. 14, 2021.

[36] M. Wu, R. L. Davis, B. W. Domingue, C. Piech, and N. Goodman, "Variational item response theory: fast, accurate, and expressive," 2020, https://arxiv.org/pdf/2307.09891.pdf.

[37] X. Song, J. Li, S. Sun, H. Yin, P. Dawson, and R. R. M. Doss, "Sepn: a sequential engagement based academic performance prediction model," *IEEE Intelligent Systems*, vol. 36, no. 1, pp. 46–53, 2021.

[38] Y. Yang, J. Shen, Y. Qu et al., "Gikt: a graph-based interaction model for knowledge tracing," in *Proceedings of the Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2020*, pp. 299–315, Springer, Ghent, Belgium, September 2020.

[39] K. Nagatani, Q. Zhang, M. Sato, Y.-Y. Chen, F. Chen, and T. Ohkuma, "Augmenting knowledge tracing by considering forgetting behavior," in *Proceedings of the World Wide Web Conference*, pp. 3101–3107, San Francisco CA, USA, May 2019.

[40] Y. Huo, D. F. Wong, L. M. Ni, L. S. Chao, and J. Zhang, "Knowledge modeling via contextualized representations for lstm-based personalized exercise recommendation," *Information Sciences*, vol. 523, pp. 266–278, 2020.

[41] Y. Chuan, X. Jieping, and D. Xiaoyong, "Recommendation algorithm combining the user-based classified regression and the item-based filtering," in *Proceedings of the 8th International Conference on Electronic Commerce: The New E-commerce: Innovations for Conquering Current Barriers, Obstacles and Limitations to Conducting Successful Business on the Internet*, pp. 574–578, Fredericton New Brunswick Canada, August 2006.

[42] B. M. Kim, Q. Li, C. S. Park, S. G. Kim, and J. Y. Kim, "A new approach for combining content-based and collaborative filters," *Journal of Intelligent Information Systems*, vol. 27, no. 1, pp. 79–91, 2006.

[43] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, 2009.

[44] S. Wan and Z. Niu, "A learner oriented learning recommendation approach based on mixed concept mapping and immune algorithm," *Knowledge-Based Systems*, vol. 103, pp. 28–40, 2016.

[45] S. Ouf, M. Abd Ellatif, S. E. Salama, and Y. Helmy, "A proposed paradigm for smart learning environment based on semantic web," *Computers in Human Behavior*, vol. 72, pp. 796–818, 2017.

[46] V. Shmelev, M. Karpova, and A. Dukhanov, "An approach of learning path sequencing based on revised bloom's taxonomy and domain ontologies with the use of genetic algorithms," *Procedia Computer Science*, vol. 66, pp. 711–719, 2015.

[47] K.-K. Chu, C.-I. Lee, and R.-S. Tsai, "Ontology technology to assist learners' navigation in the concept map learning system," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11293–11299, 2011.

[48] F. Giunchiglia, B. Dutta, and V. Maltese, "Faceted lightweight ontologies," *Conceptual Modeling: Foundations and Applications: Essays in Honor of John Mylopoulos*, vol. 36,51 pages, 2009.

[49] G. Mester, "Rankings scientists, journals and countries using h-index," *Interdisciplinary Description of Complex Systems*, vol. 14, no. 1, pp. 1–9, 2016.

[50] S. E. Embretson and S. P. Reise, *Item Response Theory*, Psychology Press, London, UK, 2013.

[51] T. A. Ackerman, M. J. Gierl, and C. M. Walker, "Using multidimensional item response theory to evaluate educational and psychological tests," *Educational Measurement: Issues and Practice*, vol. 22, no. 3, pp. 37–51, 2005.

[52] F. Wang, Q. Liu, E. Chen et al., "Neuralcd: a general framework for cognitive diagnosis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 8, pp. 8312–8327, 2023.

[53] C. Piech, J. Bassen, J. Huang et al., "Deep knowledge tracing," *Advances in Neural Information Processing Systems*, vol. 28, 2015.