Revista de Psicodidáctica 29 (2024) 185–203

psicodidáctica

Revista de Psicodidáctica

www.elsevier.es/psicod

Original

Exploring the synergy of improved convolutional neural networks and attention mechanisms for potential STEM knowledge concept recommendation in MOOCs

Xiaona Xia [a,b,*] and Wanxue Qi [a,b]

[a] Faculty of Education, Qufu Normal University, Qufu, Shandong, 273165, China
[b] Chinese Academy of Education Big Data, Qufu Normal University, Qufu, Shandong, 273165, China

ARTICLE INFO

Article history:
Available online 14 May 2024

Keywords:
STEM
Knowledge concept
Potential interest recommendation
Deep learning
Graph convolution neural network
Positive learning psychology

ABSTRACT

The multi course association of STEM poses an important challenge to the learning background of learners. Once learners do not have sufficient understanding of knowledge association or do not implement the topological order of knowledge advancement, they are prone to burnout in the learning process, forming serious negative emotions, which is not conducive to learning effectiveness, and even premature dropout. This is clearly a psychological teaching problem, that is our research objectives. This study focuses on the STEM learning behaviors in MOOCs, and explores the deep learning routing. We design one novel method to process the context features and content features for knowledge concept recommendation. Multiple entities, features, and courses enable the construction and optimization of knowledge concept relationships. Then, an attention mechanism is used to achieve the knowledge concept propagation between different entities. The extensive experiments have proved this method might accurately capture potential interests of knowledge concepts, achieve the effective deep learning routing, and explore and guide the positive learning state, reduce or avoid the negative psychological outcomes, such as dropout or low pass rate. The entire study aims to enhance learning outcomes, improve learning motivation, optimize learning behaviors, and provide more effective suggestions for STEM education, that is very important for the interdisciplinary learning in higher education. The whole research might provide key support for tracking possible psychological changes in learners, improving learning behavior trends, and enhance learning quality during STEM learning, fully improve and optimize the learning state, construct effective decisions for positive learning interests.

Explorar la sinergia entre las redes neuronales de convolución mejoradas y los mecanismos de atención recomendados por posibles conceptos de conocimiento STEM en MOOC

Palabras clave:
STEM
Tallos
Concepto de conocimiento
Recomendación de interés potencial
Aprendizaje profundo
Red neuronal de involución
Aprendizaje activo de la psicología

RESUMEN

La asociación curricular múltiple de STEM plantea desafíos importantes para los antecedentes de aprendizaje de los estudiantes. Una vez que los estudiantes no tienen suficiente comprensión de la Asociación del conocimiento o no implementan el orden topológico de la promoción del conocimiento, son propensos a sufrir agotamiento durante el proceso de aprendizaje, formando emociones negativas graves, lo que no favorece el efecto del aprendizaje e incluso abandonan la escuela prematuramente. Este es obviamente un problema de enseñanza psicológica y nuestro objetivo de Investigación. Este estudio se centra en el comportamiento de aprendizaje STEM en MOOC y explora caminos de aprendizaje profundo. Hemos diseñado un nuevo método para procesar las características contextuales y de contenido de las recomendaciones conceptuales de conocimiento. Múltiples entidades, múltiples características y múltiples cursos pueden

DOI of original article: https://doi.org/10.1016/j.psicod.2024.04.003
* Corresponding author.
E-mail addresses: xiaxn@sina.com (X. Xia), wanxueqi@126.com (W. Qi).

https://doi.org/10.1016/j.psicoe.2024.05.004

construir y optimizar relaciones conceptuales de conocimiento. Luego, se utiliza el mecanismo de atención para lograr la difusión de conceptos de conocimiento entre diferentes entidades. Un gran número de experimentos han demostrado que este método puede capturar con precisión el interés potencial de los conceptos de conocimiento, lograr vías efectivas de aprendizaje profundo, explorar y guiar Estados de aprendizaje positivos, y reducir o evitar consecuencias psicológicas negativas como el abandono escolar o la baja tasa de aprobación. Todo el estudio tiene como objetivo mejorar los resultados del aprendizaje, mejorar la motivación para el aprendizaje, optimizar el comportamiento del aprendizaje y proporcionar asesoramiento más eficaz para la educación STEM, que es muy importante para el aprendizaje interdisciplinario en la educación superior. Todo el estudio puede proporcionar un apoyo clave para hacer un seguimiento de los posibles cambios psicológicos de los estudiantes, mejorar las tendencias del comportamiento de aprendizaje, mejorar la calidad del aprendizaje en el proceso de aprendizaje STEM, mejorar y optimizar de manera integral el Estado de aprendizaje y construir decisiones efectivas de interés activo en el aprendizaje.

Introduction

STEM education, which stands for Science, Technology, Engineering, and Mathematics, is an interdisciplinary educational approach that integrates the natural sciences with innovation in the context of information technology (Buckley et al., 2023). STEM education goes beyond simply integrating different subjects; it focuses on organically merging the distinct practical processes and intellectual aspects of each subject to enhance learners' multidisciplinary application and practice (Costello et al., 2023). At the same time, deep learning is a learning model that accompanies STEM education. It involves learners independently and completely completing the learning process. Based on their own learning background and professional needs, learners themselves develop suitable learning methods and task objectives (Gijsen et al., 2024). Deep learning is emphasizes active and self-directed learning, aiming to explore effective learning behaviors, cultivate problem-solving skills. It is a meaningful form of learning (Xia & Qi, 2023a). Learners actively and critically learn new knowledge through adaptive integration of related learning contents (Xia & Qi, 2022; Weston et al., 2023). STEM education can be linked with deep learning, as they are mutually compatible, promoting each other (Norris et al., 2023). In the process of STEM education, the integration of learning contents facilitates the formation of interdisciplinary topics (Anttila et al., 2023), helps learners actively acquire knowledge, improves learners' abilities in problem-solving, innovative thinking, and collaborative learning, thereby achieving deep learning (Bañeres et al., 2023). So the effective deep learning aligns well with STEM education. Therefore, STEM education and associated deep learning have high requirements for learners to construct positive learning behaviors and positive learning states, timely and effective learner psychological tracking and guidance should be integrated into the STEM education (Xia & Qi, 2024).

However, it is crucial to quantify the entire learning process, and essential for better and more comprehensive deep learning routing in STEM education (Chen et al., 2022; Pattison et al., 2020). In recent years, Massive Open Online Courses (MOOCs) have emerged as a public educational platform that achieves complete learning processes and a vast amount of learning resources. Platforms such as Coursera, edX, and Udacity have gained worldwide attention from both learners and teachers. MOOCs have encouraged significant participation from learners globally, leading to an accumulation of massive courses, resources, and learning behavior instances (Kleinschmit et al., 2023). The increasing richness of STEM-related learning contents and learners engagement in STEM have provided effective tools and technologies for tracking the learning process and predicting deep learning routing, enabling the learners' learning interests and behavioral trends, evaluating real-time learning states (Xia, 2021a). Consequently, MOOCs offer more flexible and autonomous means of personalized learning.

STEM education in MOOCs also presents several key challenges, also bring some opo. One significant challenge is how to attract learners to engage in the entire learning process online continuously and efficiently. To effectively address the difficulties and problems in learning contents, it is needed to explore solutions and learning materials that can better understand and capture learners' interests (Xia, 2022). The integration of deep learning and STEM education enables learners to identify appropriate learning behaviors and related learning resources. This self-directed learning behavior routing has become a hot issue of STEM, deep learning, and adaptive recommendation (Parviainen et al., 2020). Meanwhile, STEM education in MOOCs also brings some opportunities to higher or secondary education, especially in higher education, where there is a clearer understanding and implementation of STEM teaching objectives and learning needs. Learners have a strong sense of initiative in multi domain related learning and collaborative learning of STEM, forming many research topics. MOOCs fully share online resources and implement the entire learning process, enabling learners to learn about multi domains of STEM online learning processes, promoting effective propagation and limitation of knowledge concepts between multi courses. This has important implications for the interdisciplinary nature of STEM learning contents, which is the foundation for achieving deep learning for learners.

Related work

In order to track and predict the potential interests and deep learning routing related to STEM knowledge concepts, this study might effectively organize and analyze various relevant issues such as STEM course recommendation (Maric et al., 2023), learning behavior prediction (Xia & Qi, 2023b), and learners' learning tendencies (Xia & Wang, 2022). In the corresponding achievements, the recommendation mechanism for learning contents has been partially applied in MOOCs, where MOOCs can recommend corresponding courses based on learners' needs. However, learners' needs may also involve multi courses, and a course may associate many videos and learners' comments with specific explanations of knowledge concepts (Mubarak et al., 2022). Direct course recommendation may not fully satisfy learners' potential interests, because the vast amount of course-related resources might affect the effective learning behavior routing and achieve the deep learning (McCarthy et al., 2021).

For instance, different teachers teaching Linear Algebra may have different emphases, resulting in variations in knowledge concepts. A teacher engaged in big data research and service recommendation may focus on key analysis models, while a teacher involved in basic mathematics education may focus on the systematic and principled mathematical theorems and rules. These teaching approaches may not necessarily meet all learners who want to study Linear Algebra. Therefore, the tracking of potential interests in STEM knowledge concepts, as well as the efficient learning behavior routing, need to be based on knowledge concepts as direct recommendation items. Research on recommendation mechanisms for online learning processes mainly focuses on the following aspects:

(1) *Collaborative filtering recommendation mechanism based on learners' historical interaction and collaboration*: This is the most traditional recommendation method, which calculates the relevant items by identifying learners' potential interests in learning contents that are similar to their peers. This approach has been widely applied. However, the entire learning process in MOOCs is accompanied by massive sparse learning behavior instances, and the discrete relationships between learners and knowledge concepts directly affect the recommended effectiveness (Evenhouse et al., 2023). So it is necessary to expand the descriptive items of learning behavior instances during data analysis and feasibility prediction. When learners' social information, the relationships between learners and recommended items, learner profiles, knowledge graphs, and contextual information of learning process are all fully considered together (Mourdi et al., 2023), that might be used to reduce the negative impact of incomplete learning behaviors.

(2) *Recommendation mechanism based on potential features*: This approach has shown promising results in addressing the limitations of collaborative filtering recommendation. By leveraging the nonlinear fitting, it predicts and associates potential features to learn implicit relationships from learning behavior instances (Hsu, 2023). It describes learners' potential interests as feature vectors, ultimately generating a list of top-K recommended items (Aldowah et al., 2020). Various models such as Multilayer Perceptron, Neural Network Model, and Graph Neural Network can learn the representation of learners and recommended items in the embedding space. Implicit feedback mechanism is utilized to infer the preference relationships between learners and recommended items (Cetron et al., 2020), thereby capturing crucial potential features of learning behaviors. Some applications have demonstrated that incorporating attention mechanism into the process of potential feature mining can effectively differentiate learning interests, further enhancing the performances of recommendation mechanism.

(3) *Course recommendation mechanism based on learners' needs*: The course recommendation mechanism based on learners' needs is a direct application that combines the two aspects. It is the most widely used pattern that has received significant attention from researchers. By integrating recommendation mechanisms such as Recurrent Neural Network, Deep Belief Network, content-aware algorithm, and reinforcement learning algorithm, it can effectively and comprehensively utilize auxiliary information about learners and courses (Calvera-Isabal et al., 2023). Based on learners' access and usage behaviors related to learning resources in MOOCs during the learning process, it mines the potential information to represent learners' interest tendencies and behavioral routes, thus improving the recommendation process of learning resources (Dash et al., 2022). However, in cases where learners have specific needs for a relatively concentrated portion of some courses or learning needs, researchers might consult relevant knowledge concepts, principles, and rules to solve specific learning problems (Gomes et al., 2023), rather than learning whole learning contents, recommending a series of related courses can easily overwhelm learners with a massive amount of knowledge concepts, making it difficult to effectively filter out learning contents directly relevant to learners' current needs (Xia, 2021b). In such scenarios, one prediction and recommendation approach may not provide learners with critical guidance and item recommendation.

The present study

In fact, there are many types of entities and relationships for STEM learners and knowledge concepts in MOOCs (Gupta et al., 2022), it can further derive other associated entities and behavioral relationships. This representation of online learning process in MOOCs can provide rich learning behavior instances, enabling interaction and cooperation between learners and knowledge concepts (Xia & Qi, 2023c). Different knowledge concepts contain different content contexts (Lee et al., 2023), and relying solely on a single type of interaction and cooperation may overlook important entities and relationships (Flegr et al., 2023). However, many types of entities and relationships can increase the complexity of entire recommendation processes and the heterogeneity of information. It is necessary to support the heterogeneous information of learning behavior instances of STEM education and the end-to-end recommendation mechanism for knowledge concepts with multi learning behavior paths. Additionally, it should be able to track the entire learning periods to effectively analyze and demonstrate feasible and efficient deep learning routing strategies, and explore the psychological changes and behavioral tendencies of learners during the learning process, This is of great significance for online teaching in higher and secondary education (Xia, 2020a).

Therefore, the knowledge concept recommendation methods need to integrate the advantages of above three aforementioned aspects. It should achieve the hybrid recommendation mechanism that could support the explicit and implicit information feedback, and combine the learning needs for interdisciplinary and multi-course integration of STEM education in MOOCs, in order to design accurate knowledge concept recommendation solutions, effectively provide feedback on learning status and the psychological representation of learners.

Based on the massive learning behavior instances generated by MOOCs, this study focuses on the learning tasks related to multiple STEM courses and goals of deep learning, in order to build positive learning states and achieve sustainable effective learning psychology. Considering potential interests of knowledge concepts and explicit and implicit feedback, we design a hybrid recommendation mechanism. The learning behaviors are represented by heterogeneous information network, then are calculated and analyzed, the feasible and efficient deep learning routing strategies are mined from learning behavior instances. Helped by the fusion of multi-layer graph convolutional neural network and attention mechanism, the corresponding analysis model is designed to track the entire learning process, adaptively mine and predict STEM knowledge concepts in MOOCs and form efficient deep learning routing. The whole research might provide the critical decision and justification for STEM learners, achieve the sustainable and progressive learning process with a complete knowledge framework through MOOCs.

Method

Research design and participants

To argue the hybrid recommendation strategies that combine explicit and implicit feedback of potential interest in STEM knowl-

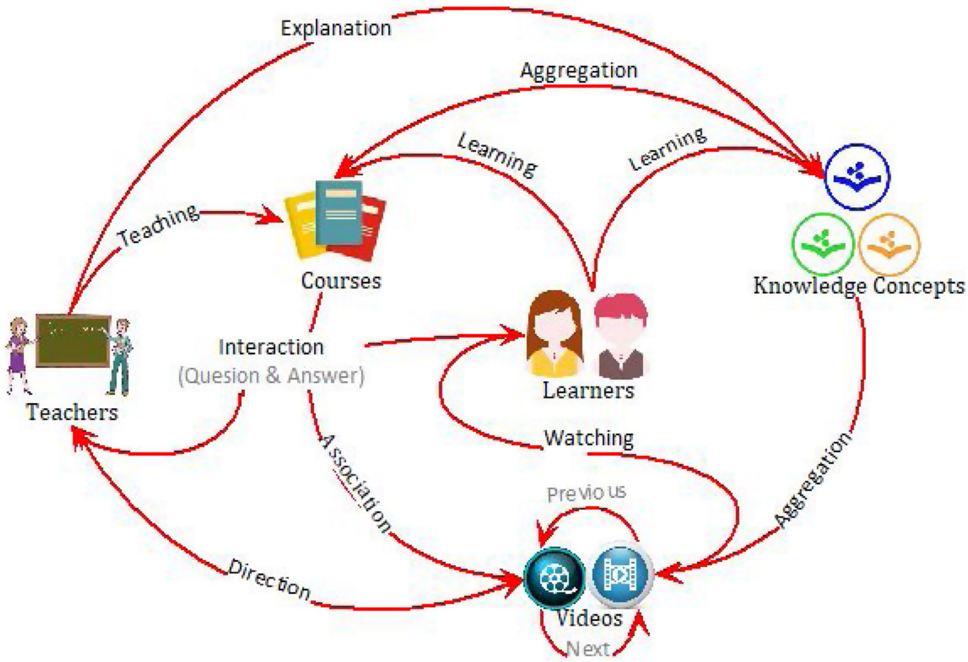

Figure 1. Entities and relationships of MOOCCube.

edge concepts, and to explore deep learning routing for STEM courses, this study focuses on one MOOC platform "XuetangX online". The factors and features associated with the entire learning process are referred to as MOOCCube, which forms one dataset that not only includes multi entities like learner, course, video, and teacher, but also generates rich and complex relationships between these entities. MOOCCube constitutes an open-source large-scale data repository composed by entities and relationships. Compared to the existing similar educational resources, MOOCCube has a massive scale, diverse features, abundant data, and sufficient attributes, enabling the more comprehensive description of learning behaviors. Relevant descriptive items include learning duration, frequency, video interval, and more. Since this MOOC platform has attracted active participation from over 200,000 learners, resulting in nearly five million learning behavior instances, which can be used to describe learner behavior tendencies and preferences. Additionally, MOOCCube registers a large number of courses along with corresponding videos, concepts, and reference materials. In MOOCCube, 706 courses and nearly 40,000 videos are selected. MOOCCube establishes the connections between learning behaviors and courses, associating relevant entities. At the same time, the learning behavior routing implies learning psychological trends, and the participation in interactive activities is to some extent a potential feedback of learning enthusiasm (Wu & Uttal, 2024).

This study takes knowledge concepts as the conditions to explore associated entities and their descriptive attributes. Knowledge concepts are defined as nodes, they form the relationships with corresponding descriptive items, and one item is the representation of entity, condition, rule, etc. Since items belong to different types, such as courses, learners, videos, teachers, etc., they are all associated with knowledge concepts, meanwhile, they might have different relationships. Thus, the relationships between knowledge concepts and different descriptive items, as well as between different descriptive items, form heterogeneous information network, which can enrich the characteristics of learners and descriptive items. By analyzing the heterogeneous information network, the similarities between learners or items can be discovered, further exploring potential feature vectors for efficient learning behaviors. Figure 1 illustrates the structure of heterogeneous information network in MOOCCube, that is used to describe the relationships among different entities, there are five entities: teachers, learners, courses, videos, and knowledge concepts, forming different types of directed relationships.

Table 1
Meta-paths centered around knowledge concepts

Entities	Meta paths
Learner	Learner→Knowledge Concept→Learner
Course	Course→Knowledge Concept→Course
Video	Video→Knowledge Concept→Video
Teacher	Teacher→Knowledge Concept→Teacher
Learner	Learner→Course→Knowledge Concept→Course→Learner
Learner	Learner→Video→Knowledge Concept→Video→Learner
Learner	Learner→Knowledge Concept→Teacher→Knowledge Concept→Learner
Teacher	Teacher→Course→Knowledge Concept→Course→Teacher
Teacher	Teacher→Video→Knowledge Concept→Video→Teacher
Course	Course→Video→Knowledge Concept→Video→Course

The entities and relationships in Figure 1 form meta-paths, that is constructed when entities serve as the starting node and end node, meta-paths are the fundamental relationships that constitute the topological graph. By centering around the knowledge concepts, different meta-paths can be formed, as shown in Table 1, it can be observed that learners, courses, videos, and teachers form four direct meta-paths around knowledge concepts, establishing the relationships through knowledge concepts. Learners, teachers, and courses, as intermediate nodes, are indirectly connected to other entities, forming the composite meta-paths centered around knowledge concepts. Although videos are served as direct descriptive media for knowledge concepts, that are viewed as the direct targets for learners, teachers, and courses, they do not form association tendencies starting from videos, thus not forming composite meta-paths. Therefore, the meta-paths centered around knowledge concepts might realize the complex connections among five entities in Figure 1. The knowledge concept propagation topology could directly influence the strategies of learning behavior routing.

Based on Figure 1 and Table 1, entities and learning behavior instances related to STEM education are obtained from MOOCCube,

Table 2
Statistics of entities and relationships

Entities	Statistics	Relationships	Statistical values
Course	273	Course-Knowledge Concept	921,307
Video	24,305	Video-Knowledge Concept	231,440
Knowledge Concept	78,116	Learner-Course	415,227
Learner	112,009	Course-video	25,512
Teacher	1,023	Teacher-Course	1,607

Table 3
Statistic values of knowledge concept association items

Items	Statistical values
Paper	547,435
Prerequisite-dependency	17,686
Taxonomy	3,152
Learning behavior Instances	2,900,470

which describe three years of complete online learning data, including 2016, 2017, and 2018. Focusing on the knowledge concepts, there are direct associations between two entities, and the corresponding statistical values are presented in Table 2. As knowledge concepts and relationships are subject to certain constraints and classification norms, they are also linked to relevant descriptive items, so we get corresponding data from Paper, Prerequisite-dependency, and Taxonomy of MOOCCube. The learning behavior instances mainly revolve around the complete record items of learners watching videos, and the related statistical values are shown in Table 3.

From Tables 2 and 3, it can be seen that MOOCCube contains vast amounts of data generated from STEM learning process of three years. The knowledge concepts form complex interrelationships and constraints. To effectively analyze the abundant data and predict potential interests, as well as provide adaptive interventions to deep learning routing, it is necessary to calculate the comprehensive correlations between entities and descriptive items. Given the interdisciplinary nature of STEM and the propagation of knowledge concepts, it is crucial to leverage the intrinsic relationships among knowledge concepts and design appropriate methods, then deduce corresponding key problems, that is used to enable effective data analysis and adaptive interventions.

By effectively analyzing the structure, elements, and features of STEM-related entities, items, and learning behavior instances, and based on the achievements of previous research (Khor & Dave, 2022; Khoushehgir & Sulaimany, 2023), this study aims to track the complete STEM learning process. Considering the demands for practical application and disciplinary correlation in STEM education, the hybrid recommendation strategy and deep learning routing are proposed, that are supported by both explicit and implicit features of potential interests in STEM knowledge concepts. As depicted in Figure 2, the analysis process involves classifying knowledge concepts across the domains of Science, Technology, Engineering, and Mathematics. By analyzing massive learning behavior instances and deriving learning patterns, we establish a set of problems to validate the findings and suggestions.

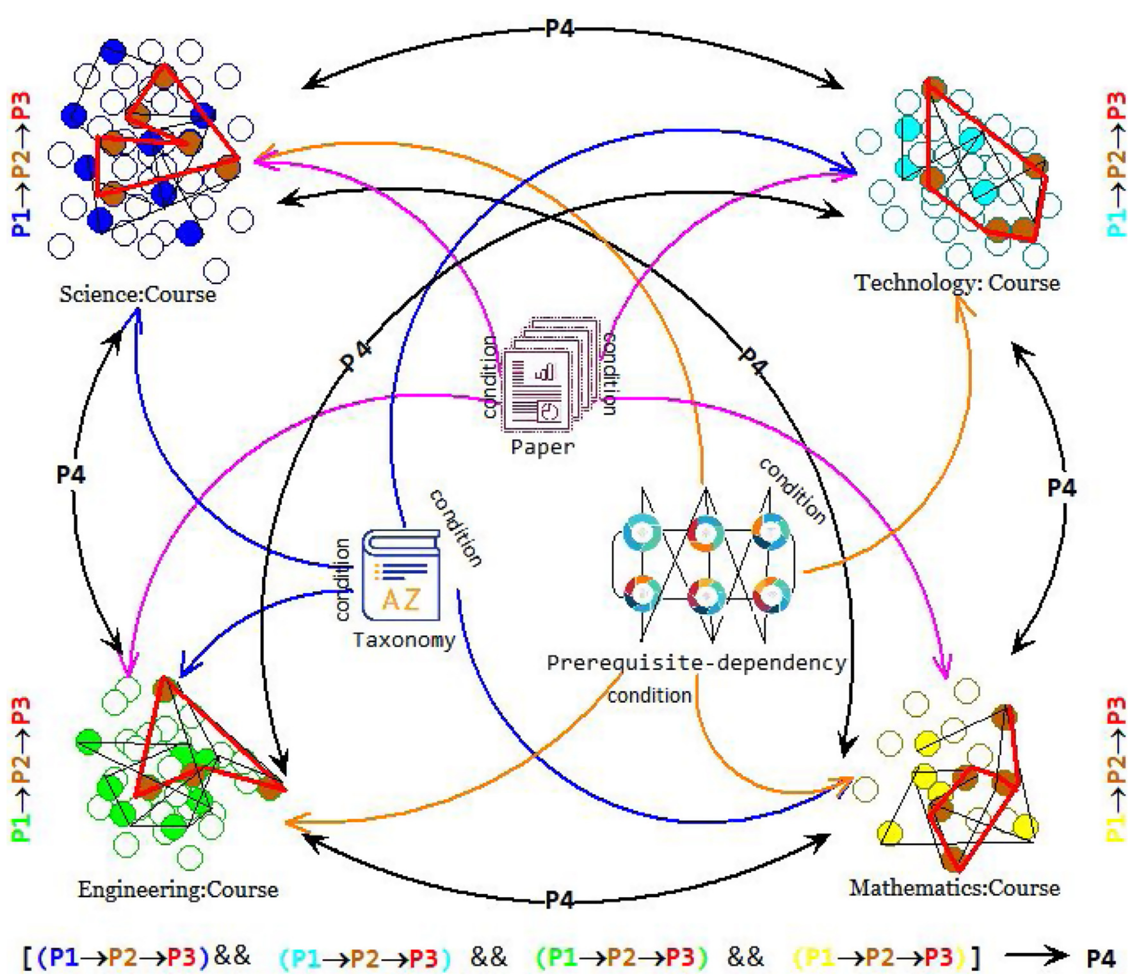

Figure 2. The potential relationships and related test problems of STEM knowledge concepts.

In Figure 2, four knowledge concept distribution for related courses are described. Under the constraints and conditions of Paper, Prerequisite-dependency, and Taxonomy, the knowledge concepts and their relationships are defined among different domains such as Science, Technology, Engineering, and Mathematics. By exploring the relationships between these domains, effective strategies for learning behavior routing can be developed and deduced, and corresponding learning state and psychological trends might be hidden in the changes and association of learning behavior routing in different domains. To validate the feasibility and reliability of hybrid recommendation strategy supported by both explicit and implicit features of knowledge concepts, this study identifies key courses related to STEM and proposes four test problems as follows:

P1. We select the key knowledge concepts for each domain and establish the corresponding knowledge concept topology, then verify its significant influence on the learning behavior routing of learners in each domain.

P2. Based on the validation results of P1, we define the related knowledge concept clusters for each domain to serve other three associated domains, then verify their significant influence on the potential interest tendency towards other domains.

P3. Based on the validation results of P2, we build the learning behavior routing of specific knowledge concepts for associating with other three domains, then verify their significant influence on the learning behavior routing of learners in terms of associating with other domains.

P4. Based on the validation results of P3, we verify the efficient deep learning behavior routing formed by learners among knowledge systems of different domains.

By arguing and analyzing these four problems, the feasibility and effectiveness of hybrid recommendation strategy and deep learning routing in promoting learner interests and behavior modes can be demonstrated.

Instruments

To achieve the hybrid recommendation mechanism that can support comprehensive explicit and implicit features for potential interests in STEM knowledge concepts, and to predict and explore efficient deep learning routing strategies, in this section, we will design a hybrid recommendation mechanism that combines explicit and implicit features. By effectively recommending a continuous sequence of knowledge concepts, the reliable deep learning routing strategies are tracked and predicted from massive learning behavior instances.

The corresponding approach is named as ReRo model (Recommendation for Potential Interests and Routing of Learning Behavior, ReRo). The entire analysis process is divided into the following three steps:

Step 1. The STEM learning behavior features and relationships are extracted. To extract the features and relationships of STEM learning behaviors, it is necessary to identify the features of knowledge concepts. After fully analysis and classification, we find that there are mainly two categories, content features and context features respectively: (1) *Content features*: The naming of STEM knowledge concepts provides an intuitive description of their contents, which contains rich semantic information. The keywords of knowledge concepts are generated from the subtitles of course videos and used as content features of knowledge concepts. Furthermore, the representation vectors of content features can be generated; and (2) *Context features*: In addition to using content features to represent the contexts of knowledge concepts, there are also rich contextual information, such as the relationships between different entities in the continuous learning process. In order to integrate the complex relationships between different entities, the context information is modeled as features when constructing relationships of knowledge concepts. Specifically, we consider the following four types of relationships in learning behavior instances.

Type 1: R_1^l, based on learning behaviors, R_1^l represented as a hit rate matrix of Learner→Knowledge Concept$A_1^l = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,m} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n,1} & c_{n,2} & \cdots & c_{n,m} \end{bmatrix}$, about $\mathrm{c}_{i,j} = 1$ means that learner i has clicked on knowledge concept j during the learning process.

Type 2: R_2^l, R_2^l describes the relationships between learners and courses, represented as a relationship matrix of Learner→Course $A_2^l = \begin{bmatrix} l_{1,1} & l_{1,2} & \cdots & l_{1,\mathrm{p}} \\ l_{2,1} & l_{2,2} & \cdots & l_{2,p} \\ \cdots & \cdots & \cdots & \cdots \\ l_{q,1} & l_{q,2} & \cdots & l_{q,p} \end{bmatrix}$, about $\mathrm{l}_{i,j}\,(1 \le i \le p, 1 \le j \le \mathrm{q}) \in \left\{0, 1\right\}$, $l_{i,j} = 1$ means that learner i has selected course j during the learning process.

Type 3: R_3^l, R_3^l describes the relationships between learners and course videos, represented as a relationship matrix of Learner→Video $A_3^l = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,\mathrm{r}} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,r} \\ \cdots & \cdots & \cdots & \cdots \\ w_{s,1} & w_{s,2} & \cdots & w_{s,r} \end{bmatrix}$, about $\mathrm{w}_{i,j}\,(1 \le i \le r, 1 \le j \le \mathrm{s}) \in \left\{0, 1\right\}$, $\mathrm{w}_{i,j} = 1$ means that learner i has watched video j during the learning process.

Type 4: R_4^l, R_4^l describes a learner who has studied a course taught by a specific teacher, represented as a interactive matrix of (Learner→Course) & (Teacher→Course) $A_4^l = \begin{bmatrix} t_{1,1} & t_{1,2} & \cdots & t_{1,\mathrm{x}} \\ t_{2,1} & t_{2,2} & \cdots & t_{2,x} \\ \cdots & \cdots & \cdots & \cdots \\ t_{y,1} & t_{y,2} & \cdots & t_{y,x} \end{bmatrix}$, about $\mathrm{t}_{i,j}\,(1 \le i \le x, 1 \le j \le \mathrm{y}) \in \left\{0, 1\right\}$, $t_{i,j} = 1$ means that learner i has learned some course taught by teacher j during the learning process.

Step 2. The meta-paths in heterogeneous information networks are selected. The four types of relationships obtained in Step 1 describe the interactive behaviors in the learning process, that enable the complex relationships of five different entities: teachers, learners, courses, videos, and knowledge concepts. These relationships are further influenced by the content features and context features of knowledge concepts, forming a heterogeneous information network. The network is represented by node set V and relationship set E, denoted as $G = \left\{V, E\right\}$ (Definition 1). A heterogeneous information network corresponds to the node type mapping function $\phi : V \to N$ and a relationship type mapping function $\varphi : E \to R$, where N and R are the sets of predefined node types and relationship types.

Based on the definition of G, the learning behaviors related to STEM in MOOCs can be transformed into a heterogeneous information network. The nodes in this network include five types of entities: teachers, learners, courses, videos, and knowledge concepts. They form four types of relationships R_1^l, R_2^l, R_3^l and R_4^l. Based on this heterogeneous information network, we can construct a multi framework for STEM learning behaviors, denoted as $S = (N, R)$ (Definition 2). This framework describes G and two mapping functions, ϕ and φ, it is a directed graph that incorporates the semantic information and constraints of STEM education.

Based on the definition of S, the semantic information centered around knowledge concepts between entities forms the meta-paths MP, that are important components to construct the framework and are defined as a continuous path $N_1 \xrightarrow{r_{1\to2}} N_2 \xrightarrow{r_{2\to3}} \cdots \xrightarrow{r_{l-1\to l}} N_l$, that represents all relationships, denoted as $R = \mathrm{r}_1 \circ r_2 \circ \cdots \circ r_l$ to N_1. For example, if one learner's click behavior regarding a certain knowledge concept is propagated to another learner, the meta-path can be described as $Learner_i \xrightarrow{click} Knowledge\ \mathrm{Graph}_k \xrightarrow{click'} Learner_j$. Similarly, we can represent the meta-paths corresponding to Table 1 using the same approach.

Step 3. The attention mechanism and the multi-layer graph convolutional neural network are fused. Due to the strong autonomy and personalization of STEM online learning processes in MOOCs, it is necessary to input the content features and context features into a suitable graph convolutional neural network (GCN) to predict and explore potential entity information. Therefore, in the analysis process of GCN, it is important to achieve effective fusion of key features and track the complete temporal sequences of the learning process.

Given the heterogeneous information network G of STEM learning behaviors and the set of meta-paths MPs along with their corresponding adjacency matrix A, the related analysis process requires a multi-layer graph convolutional neural network that can fuse personalized non-linear features and propagate in a hierarchical manner. The corresponding rules are represented as: $h^{l+1} = \sigma\left(Ph^lW^l\right)(l = 0, 1, \cdots)$ (Formula 1), where h^{l+1} represents the new feature representation of relevant entities after propagation through the graph convolutional layers, and it is dependent on h^l. h^0 represents the initial features of learners or course contents. P is the adjacency matrix representing the meta-paths and the self-connected matrix, which is related to the learning behavior instances. $\sigma(\cdot)$ represents a non-linear activation function of neural network. In order to effectively identify a large amount of noises presented in the learning behavior instances, we adopt the Noisy ReLU activation function to handle the uncertainty caused by the learning behavior distribution towards a normal distribution, which is defined as n*oisy* ReLU $= \max(0, \mathrm{x} + \mathrm{Y})$, where x is the input neuron value, and $Y \sim N(0, Random(x))$ is a random variable.

To fuse the global features of learning process and effectively capture the relationships between learners and courses, as well as to deeply explore learners' preferences regarding knowledge concepts while considering the personalization and autonomy of learning behaviors, we adopt an feature fusion method based on attention mechanism, that is described as $\mathrm{a}TT\left(e_{MP_i}\right) = \frac{\exp\left(\lambda\left(AMe_{MP_i}\right)\right)}{\sum_{j \in |MP|} \exp\left(\lambda\left(AMe_{MP_j}\right)\right)}$ (Formula 2), e_{MP_i} represents the entity representation according to the target meth-paths, e_{MP_j} represents the entity representation of other meta-paths, AM is the trainable attention matrix, and λ is a non-linear aggregation function. In the process of feature fusion, the correlations between meta-paths are calculated by a feed-forward neural network.

Procedure

Based on the analysis results and corresponding definitions of above three parts, the hybrid recommendation approach can be derived from a massive amount of learning behavior instances, which is supported by the explicit features and implicit features of learners' effective knowledge concepts. This approach tracks the learning behavior tendency changes, analyzes the potential relationships, and explores potential learning interests. It constructs the trajectory of the learning process and establishes an effective order between learners and knowledge concepts (Kim & Tawfik, 2023). The analysis process of ReRo model is shown in Figure 3. The meta-path tracing enables better inference of the entities corresponding to knowledge concepts associated with deep learning routing. The related algorithm is described as Algorithm ERKC (ERKC means Entities Related to Knowledge Concepts).

Algorithm ERKC
Input: MPs//Meta path set;
A // It represents the adjacency matrix that is used to describe the relationships
F //It represents the feature matrix of potential target entities
d //It represents the dimension of entity representation
Output: TE //it represent the target entities
Begin
Initialize $TE \longleftarrow null$

$temp_paths \longleftarrow null$

Do while each $MP_i \in MPs$
$\sim A = A + I$ // I is the identity matrix; $\sim A$ is also one adjacency matrix related to a specific meta-path, that has has self connection.
$\sim D = diag(\sim A)$ // $diag(\sim A)$ is used to find the Diagonal matrix, $\sim D$ is the all-one vector.
$P^1 = P^2 = P^3 = \cdots P = \frac{1}{\sqrt{(\sim D)}} \cdot (\sim A) \cdot \frac{1}{\sqrt{(\sim D)}}$ // $P \in R^{y\times y}$, it describes the propagation process of content features or context features and converges to a stationary distribution, z is the spreading likelihood from knowledge concepts.

$\mathrm{h}^1 \longleftarrow F$

l^2

Do while (True)
{Calculate h^l by Formula 1

$\mathrm{l} = l + 1$

If there are some features that have not been correlated and analyzed
Then Continue
Else Break}
Return l and h^l

$temp\_paths \longleftarrow h^l$

$TE_{MP} = \mathrm{h}^l$

End while

$$TE = \sum_{l=1}^{|MPs|} \frac{\exp\left(\sigma\left(aTE_{MP_l}\right)\right)}{\sum_{k \in |MP|} \exp\left(\sigma\left(aTE_{MP_k}\right)\right)} \cdot TE_{MP_l} TE_{MP} = \mathrm{h}^l \quad \text{(Formula 3)//}$$

$\sum_{l=1}^{|MPs|} \frac{\exp\left(\sigma\left(aTE_{MP_l}\right)\right)}{\sum_{k \in |MP|} \exp\left(\sigma\left(aTE_{MP_k}\right)\right)}$ is the weight of attention mechanism, a is a regulatory factor.
Return TE
End.

ReRo model might consider and calculate the complete learning process, aiming to fully track the massive, sparse, and noisy natures of learning behaviors (Xia, 2020b). Additionally, it ensures the accuracy and reliability of learning behavior attributes and parameter calculations (Xia, 2021c). ReRo model incorporates the attention mechanism and iterative operation of multi-layer graph convolutional neural networks. The analytical and procedural steps guarantee the accuracy of uncovering potential interests in multidisciplinary and multi-course knowledge concepts. Moreover, it reliably predicts the meta-paths of deep learning routing.

Data (experiment) analysis

The training and testing experiments for ReRo model are designed and implemented based on the massive STEM learning

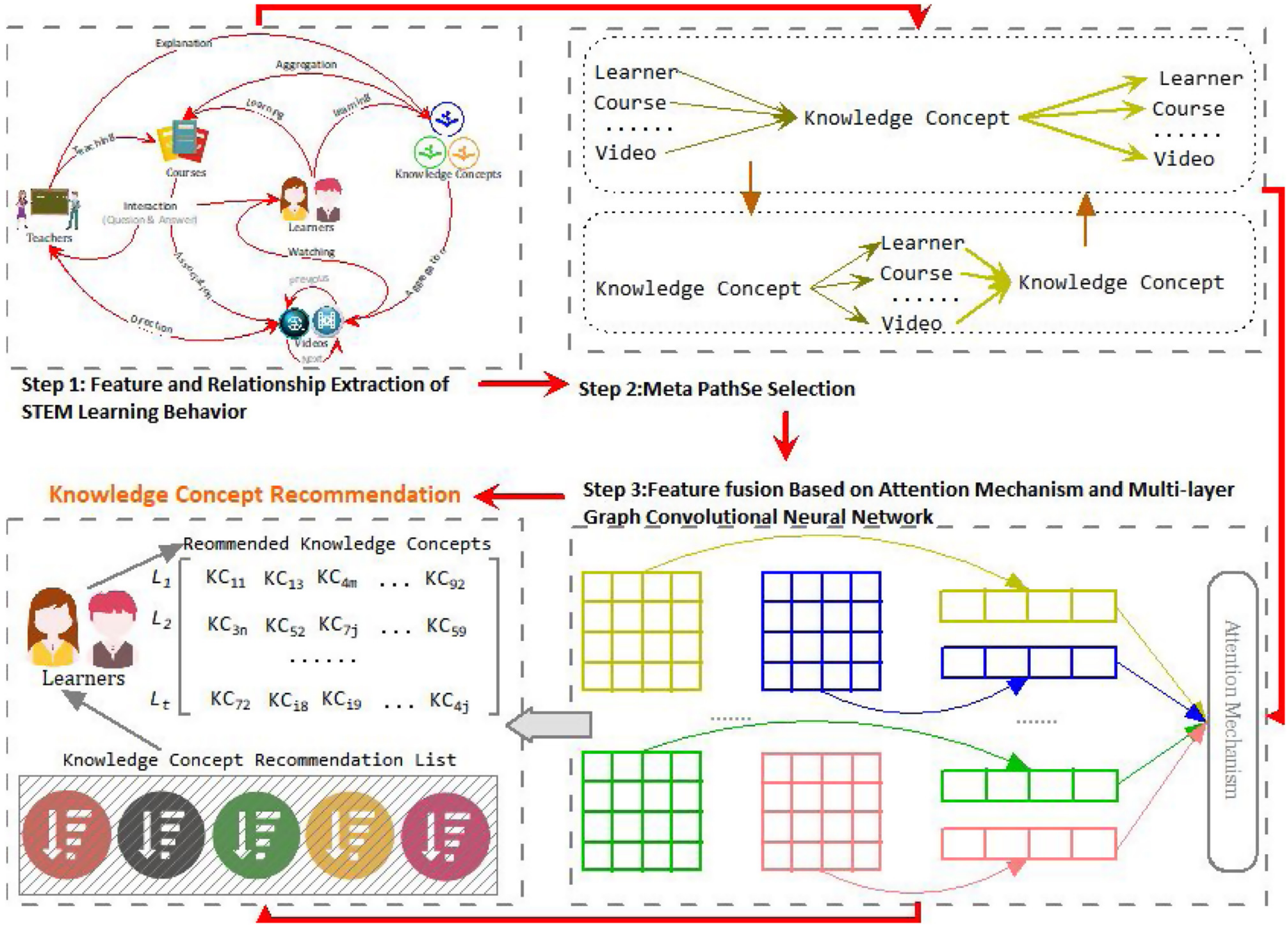

Figure 3. The analysis process of ReRo model.

behavior instances. We sort all the data in ascending order according to the temporal sequences and divide it into a training set and a testing set in 8:2. The learning behavior instances record the traces that learners participate in the learning process. During the training process, the knowledge concepts that the learners last clicked are used as the predictive target, while the remaining data represents the historical behavioral trajectory. A negative example is randomly generated to replace the target knowledge concepts. In the testing process, the entire testing set serves as the potential knowledge concepts to be learned, while within the training set, the knowledge concepts of same learner represent the historical learning sequences. To evaluate the hybrid recommendation effect of potential knowledge concept interests, each positive example in the testing set is matched with m negative examples to calculate the prediction results for $1 + m$ learning behavior instances.

To evaluate the effectiveness of knowledge concept recommendation, we will use Hit Ratio (HR@K) and Normalized Discounted Cumulative Gain (NDCG@K) as the metrics. K will be set to 5, 10, 20, and 30 respectively, and the metrics will be calculated for every $1 + m$ instances. Given a recommended knowledge concept item list T_l for learners, N represents the subset of learning behavior instances in the testing set, and N represents the total number of learners in the testing set, $HR@K = \frac{1}{N}\sum_{l} I\left(\left|R_l \cap T_1\right|\right)$ (Formula 4) is calculated using the indicator function $I(\cdot)$. A higher HR@K indicates the better model performance. $NDCG@K = \frac{1}{Z}\sum_{k=1}^{K} \frac{2^{I\left(\left\{r_l^j\right\} \cap Tl\right)} - 1}{\log(j+1)^{/>}}$ (Formula 5) is calculated using a normalization constant Z. Similarly, a higher NDCG@K indicates the better performance. Based on the results of HR@K and NDCG@K, the performances will be further evaluated using Mean Reciprocal Rank (MRR) and Area Under Curve (AUC), which consider the average reciprocal values of positive ranking positions and the overall ranking performances.

Firstly, the performances are evaluated on the combination of different meta-paths. During the experiment, we analyze the impact of multi meta-path combinations on ReRo model. The effective meta-paths can improve the efficiency and effectiveness of learning behaviors. About the comparative experiments, the performances of each individual meta-path are also evaluated. Based on Table 1, we select various types of meta-paths starting with learners and teachers. The results of the four performance metrics are shown in Tables 4 and 5.

Each individual meta-path has different performances, mainly due to the scale of heterogeneous nodes, which can affect the correlation of different entities. Since there are relatively few Prerequisite-dependency between learners and teachers for related STEM courses, the evaluation metrics for each meta-path may not be high. However, as the combination of meta-paths increases, the evaluation metrics of ReRo model significantly improve. When the number of combined meta-paths reaches its maximum, all four metrics achieve better values. Different meta-paths demonstrate different associations around knowledge concepts, and a stronger association leads to more reliable and comprehensive description of related entities.

Secondly, the parameter of ReRo model are evaluated. We will evaluate the parameters of ReRo model. The number of potential factors in the learning behavior matrix decomposition is a cru-

Table 4
Performances of meta-paths combination of learners centered on knowledge concepts

Learner Meta-path Combination	HR@20	NDCG@20	MRR	AUC
MP_{l1}	59.15%	40.17%	39.44%	88.19%
MP_{l2}	52.33%	38.21%	36.75%	85.53%
MP_{l3}	62.42%	47.39%	42.52%	88.94%
MP_{l4}	50.36%	34.37%	31.90%	82.98%
$MP_{l1\&\&}MP_{l2}$	61.60%	42.44%	40.72%	89.05%
$MP_{l1\&\&}MP_{l3}$	66.57%	45.75%	43.58%	92.02%
$MP_{l1\&\&}MP_{l4}$	65.02%	47.15%	44.81%	89.99%
$MP_{l2\&\&}MP_{l3}$	68.54%	50.27%	47.38%	91.92%
$MP_{l2\&\&}MP_{l4}$	49.54%	37.21%	36.15%	87.46%
$MP_{l3\&\&}MP_{l4}$	64.01%	46.43%	43.95%	90.87%
$MP_{l1\&\&}MP_{l2\&\&}MP_{l3}$	69.50%	51.43%	48.04%	93.11%
$MP_{l1\&\&}MP_{l3\&\&}MP_{l4}$	64.29%	47.47%	45.62%	91.85%
$MP_{l2\&\&}MP_{l3\&\&}MP_{l4}$	6737%	48.89%	46.53%	92.16%
$MP_{l1\&\&}MP_{l2\&\&}MP_{l3\&\&}MP_{l4}$	73.67%	52.62%	50.35%	94.83%

Table 5
Performances of meta-paths combination of teachers centered on knowledge concepts

Teacher Meta-path combination	HR@20	NDCG@20	MRR	AUC
MP_{t1}	62.30%	45.32%	40.72%	85.39%
MP_{t2}	68.19%	48.61%	45.62%	86.17%
MP_{t3}	72.91%	51.83%	48.42%	84.09%
$MP_{t1\&\&}MP_{t2}$	75.71%	53.25%	46.40%	89.74%
$MP_{t1\&\&}MP_{t3}$	77.37%	55.63%	46.66%	90.81%
$MP_{t2\&\&}MP_{t3}$	82.26%	50.27%	54.52%	91.04%
$MP_{t1\&\&}MP_{t2\&\&}MP_{t3}$	85.00%	55.59%	58.60%	92.88%

cial parameter. Different numbers of potential factors are obtained from content features and context features. We will use HR@K, NDCG@K, MRR, and AUC to verify the number of potential factors on ReRo model. K is set to 20, and the number of potential factors is set to 5, 15, 25, 35, 45, and 55 respectively. The evaluation results are shown in Figure 4. Data analysis reveals that when the number of potential factors is set to 35, the performance of ReRo model is more optimal.

Then, we further verify the influence of entity representation on ReRo model, with the number of potential factors set at 35. The dimension of entity representation is set to 40, 60, 80, 100, 120, 140, and 160 respectively. The evaluation results are shown in Figure 5. Data analysis reveals that when the dimension of entity representation is set to 100, ReRo model achieves superior performance across four metrics, that indicates that when both learners and knowledge concepts are represented as vectors whose dimension is set 100, the processing capabilities of heterogeneous information network improve the recommendation effectiveness. Therefore, setting the dimension of entity representation to 100 might be a suitable choice.

Thirdly, the explicit and implicit feature ablation experiment of ReRo model is implemented. To further validate the effectiveness of ReRo model in analyzing the explicit and implicit features of STEM learning behavior instances, we need to complete the ablation experiments based on the above two aspects of experimental analysis. These experiments are used to recommend potential knowledge concepts, that are involved three main steps. Step 1, for the selection of knowledge concepts, we employ an implicit feature-based recommendation mechanism using the rank of recommended items; Step 2, for the correlations of knowledge concepts and the correlations between courses, we use an explicit feature-based recommendation mechanism relying on predicted concept scores; Step 3, by integrating the multi entity attributes centered around knowledge concepts, along with their associations with Papers, Prerequisite-dependency, and Taxonomy, we develop the hybrid recommendation mechanism that leverages both explicit and implicit features, then analyze and calculate the four metrics. The experimental results are shown in Table 6.

From the data distribution in Table 6, the four metrics of hybrid recommendation mechanism combines the explicit and implicit features that surpass those of the individual features. The recommendation process associated with ReRo model can more comprehensively utilize the features when they are acquiring STEM knowledge concepts, effectively ensuring the robustness and reliability of entire experimental process.

Finally, ReRo model and other related approximate models are comparatively tested. Based on the experimental analysis and verification results of above three aspects, ReRo model and related approximate models are analyzed and compared in the experiments, the corresponding comparative models are as follows:

(1) BPR (Bayesian Personalized Ranking), it is a model based on learners' implicit features, provides learners with key recommended items. It is a pairwise method that based on matrix factorization, and might optimize the pairwise ranking loss of the recommended task in a Bayesian way.
(2) MLP (Multi-Layer Perceptron), it is a type of feedforward artificial neural network that maps a set of input vectors to a set of output vectors. MLP can be seen as one directed graph consisting of multi layers of nodes, where each layer is fully connected to the next layer. Apart from the input nodes, each node is a neuron (or processing unit) with a nonlinear activation function.
(3) FM (Factorization Machines), it is a machine learning algorithm based on matrix factorization, that is widely used in advertisement estimation models and has greatly improved performances. FM combines the advantages of SVM (Support Vector Machine) and factorization models. Similar to SVM, FM is a versatile predictor that can handle arbitrary feature vectors.
(4) FISM (Factored Item Similarity Models), it is a project-based method uses for generating top-N recommendation, which learns an item-item similarity matrix constructed by multiplying two low-dimensional potential factor matrices and models by using structural equations.
(5) NAIS (Neural Attentive Item Similarity model), it is a model that uses the attention mechanisms to analyze the historical learner

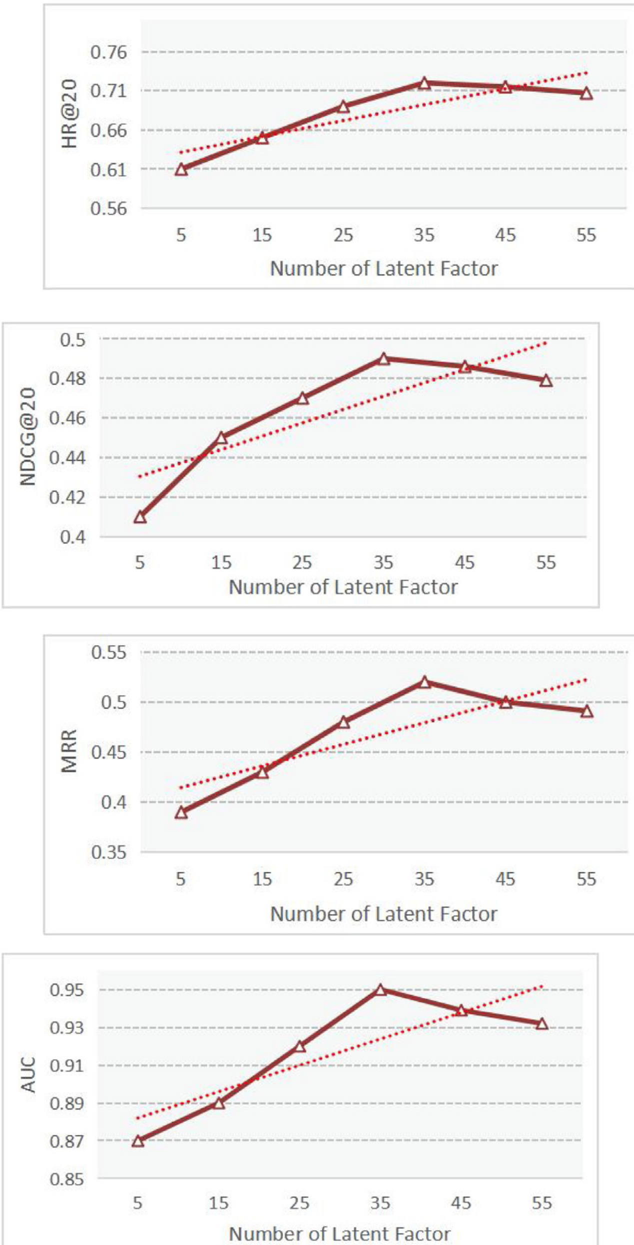

Figure 4. The impact of potential factors on ReRo model.

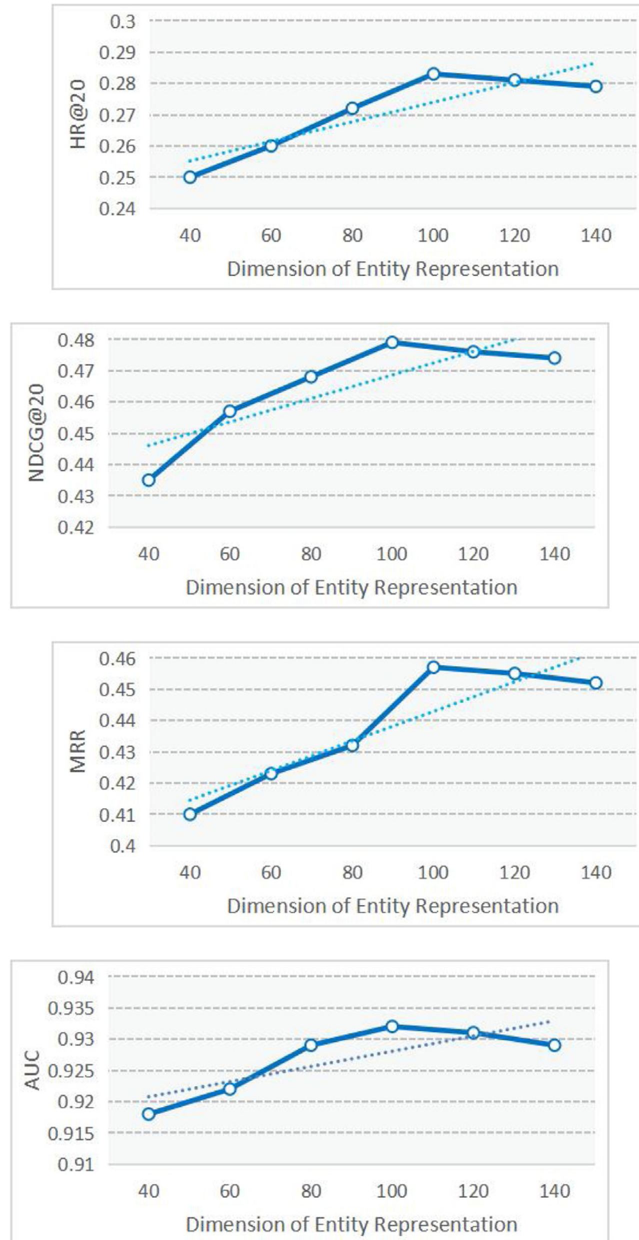

Figure 5. The impact of entity representation on ReRo model.

behaviors that are important for future predictions. It has been proven to be more effective than traditional item-based collaborative filtering, with stronger expressive power.

(6) NASR (Neural Attentive Session-based Recommendation), it is a model that implements GRU (Gated Recurrent Unit) for sequential modeling. It improves the recurrent neural network and handles long-distance problems in RNN well.

(7) Ackrec (Attentional Heterogeneous Graph Convolutional Deep Knowledge Recommender), it is an attention heterogeneous graph convolutional deep knowledge recommendation method based on end-to-end graph neural networks, including multi variations, Ackrec has been proven to achieve adaptive recommendation of knowledge concepts. Like other recommendation methods, it also has sparsity issues.

This study selects a processing flow that combines entity-based heterogeneous context features and content features to describe entity information networks. However, for meta-paths, the interest propagation rules with three layers are fixed, so we implement the multi-layer scheduling for optimal decision propagation. All models calculate four metrics when dealing with STEM learning behavior instances, with K set to 20 for HR@K and NDCG@K. The experimental results are shown in Figure 6. From the linear distribution of each performance metric, ReRo model demonstrates the adaptability in handling heterogeneous information, meta-path combination, and knowledge concept recommendation by implementing multi-layer propagation rules. It fully considers the content features and context features, making it superior to other seven compared models. The experimental analysis and verification results of the above aspects indicate that the learning behavior instances might accurately meet the potential interest recommendation needs for STEM knowledge concepts in MOOCs, ReRo model achieves the fusion of heterogeneous information network, attention mechanism, and multi-layer graph convolutional neural

Table 6
Metrics of ablation experiment

Procedure	HR@20	NDCG@20	MRR	AUC
Step 1	79.28%	55.84%	51.02%	88.69%
Step 2	77.30%	57.46%	49.51%	87.65%
Step 3	81.89%	59.94%	55.32%	90.07%

network, and also adaptively track and evaluate the experimental results. The data analysis results and conclusions are reliable.

Results

The interdisciplinary nature and course connections of STEM primarily come from the ability to contextualize knowledge concepts. It plays a crucial role in the integration of different subjects and courses, as there are potential and inherent relationships between knowledge concepts. These relationships dictate the distribution of knowledge concepts during the learning process, establishing a certain sequential order. Different knowledge concepts also enable learners to explore and construct applicable learning behaviors, they will also be related to the entire learning process of learners, and their online learning participation and learning ability will also be accompanied by changes in the representation of learning behavior characteristics and updates in attribute values, making learning state, psychological process, and learning behaviors integrated together (Gilligan-Lee et al., 2022). Driven by knowledge concepts, learners can also focus on related courses and even derive interdisciplinary topics, thus facilitating their self-growth and the self-propagation of knowledge concepts in STEM education. During the learning process, it is important for learners to autonomously explore positive learning behaviors, implement more efficient learning modes, and achieve the positive and sustainable learning attitude towards the learning process (Ioannou & Gravel, 2024). So the learning process involves the exploration, association, or propagation of personal interests, gradually uncovering effective learning methods and achieving efficient learning behaviors, resulting in high learning outcomes. The sequences of such learning behaviors gradually promotes deep learning routing (Solomon et al., 2022).

To some extent, the extensive utilization of online learning has enabled the STEM education and deep learning strategies. Extensive data analysis has discovered that STEM learners exhibit higher enthusiasm for online learning in MOOCs, generating a vast number of learning behavior instances, that provide critical data for describing, analyzing, and predicting the complete STEM learning process, promoting the tracking and description of effective learning process for knowledge concepts. It becomes the significant factor to influence the formation of potential interests, guide and recommend the feasible learning tendencies accurately (Edelsbrunner et al., 2023). Therefore, the research on the potential interest recommendation of STEM knowledge concepts and deep learning routing has strong practical implications for the study and implementation of deep learning in STEM education.

This study focuses on the massive STEM learning behavior instances in MOOCs, thoroughly analyzing and demonstrating the potential interest recommendation of knowledge concepts and deep learning routing. It presents corresponding research problems, designs the analysis and prediction model capable of tracking the complete learning process, and verifies the reliability and accuracy through extensive experiments and comparative analysis. This section combines the four problems to analyze and infer regular patterns based on the analysis results of ReRo model. So we select the corresponding courses related to four STEM domains and learning behavior instances to argue and analyze the four problems. The related courses are Science→Data Structure, Technology→Java, Engineering→Software Engineering, and Mathematics→Advanced Mathematics. Drawing on the related entities and items in Tables 2 and 3, as well as the experimental results of ReRo model, the thorough analysis and pattern inference will be conducted for the four problems.

Analysis of key knowledge concepts

According to the key items in Table 3, the knowledge concepts of four domains' corresponding courses are analyzed and extracted. ReRo model is applied to explore and calculate the potential interests of meta-paths and knowledge concepts, as well as to construct their potential relationships, that results in the formation of knowledge concept paths, as shown in Figure 7, each domain's course selects the top ten knowledge concepts, which are classified into two categories: core knowledge concepts and associated knowledge concepts, The former plays an important guidance in learning other knowledge concepts within the course, the latter is closely related to the core concepts, and it also plays an important role in the formation of course knowledge system by learners. It can be observed from Figure 7 that all ten concepts of Advanced Mathematics are classified as the core knowledge concepts. By tracking and predicting the related descriptive items and learning behavior tendencies, the potential relationships between knowledge concepts are derived, forming the distribution pattern with relatively strict learning sequences. In order to deepen the understanding and application, learners need to follow their inherent relationships of knowledge concepts (Daker et al., 2021). On this basis, the relationships of knowledge concepts in each domain

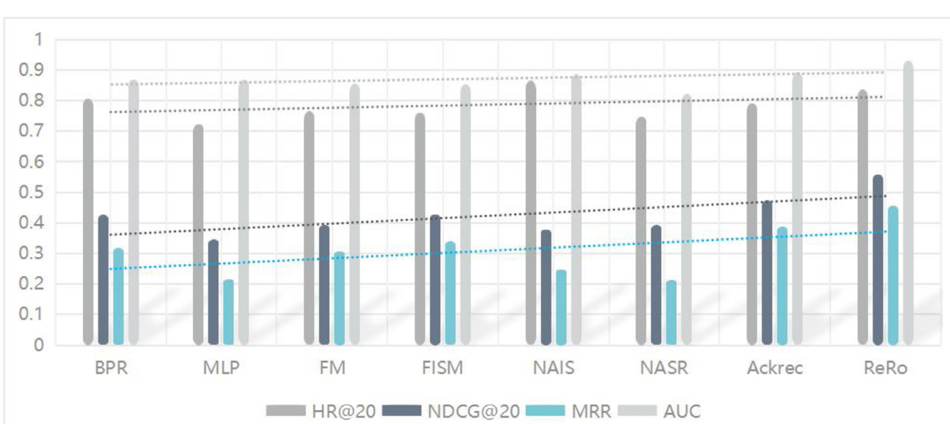

Figure 6. Performances of comparative models.

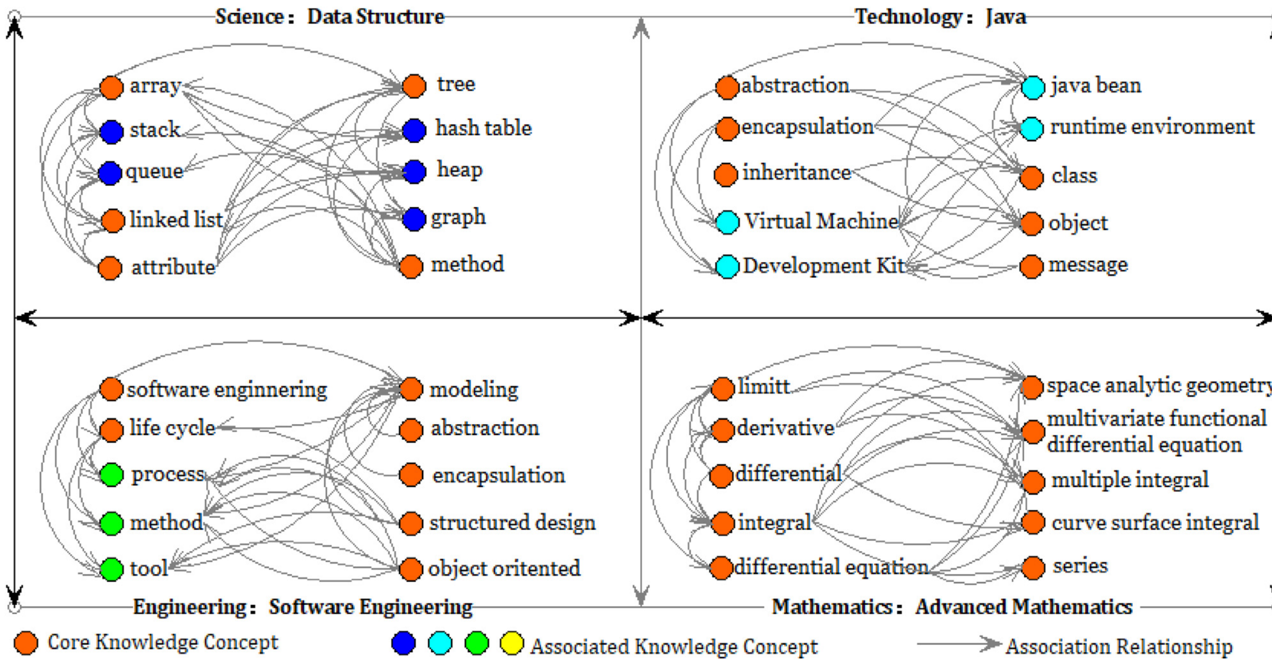

Figure 7. Distribution and relationships of key knowledge concepts.

Table 7
Data statistics and test results

Learning behavior	Course	Knowledge concept		relationship	
		Overlap	Significance	Overlap	Significance
Efficient	Data Structure	96.30%	0.000***	98.3%	0.000***
	Java	80.44%	0.004**	85.48%	0.000***
	Software Engineering	88.6%	0.009**	89.30%	0.002**
	Advanced Mathematics	99.52%	0.000***	95.42%	0.000***
Inefficient	Data Structure	92.93%	0.075	53.03%	0.114
	Java	79.35%	0.707	40.22%	0.894
	Software Engineering	84.65%	0.809	46.00%	0.955
	Advanced Mathematics	93.48%	0.067	55.09%	0.108
Invalid	Data Structure	30.05%	–	4.45%	–
	Java	21.17%	–	2.32%	–
	Software Engineering	25.04%	–	4.55%	–
	Advanced Mathematics	36.63%	–	6.60%	–

***p < .001, **p < .01, *p < .05.

are trained. The overlap degree between them is compared, that refers to the proportion of identical knowledge concepts or relationships in different domains, described as a percentage (Xia & Qi, 2023b). Learners' assessment results are divided into three types: Distinction and Pass as the first type, corresponding to the efficient learning behaviors; Fail as the second type, corresponding to the inefficient learning behaviors; and the third type defines as learners who discontinued the learning process without assessment results, corresponding to the ineffective learning behaviors. The overlap degree is calculated in two parts: the consistency of relevant knowledge concepts and the consistency of related relationships. Furthermore, using the overlap degree of knowledge concepts and relationships as independent variables and learners' assessment results as the observation variable, it is tested whether the relationships of knowledge concepts significantly influence the learning behavior routes in a given domain. After data statistics and analysis, Table 7 is obtained.

From Table 7, it can be observed that the three different types of learning behaviors correspond to the four domains' respective courses. When the learning behavior is efficient, there is a high overlap degree in knowledge concepts and their relationships across the four courses. Particularly for Data Structures and Advanced Mathematics, the learning behavior routes exhibit the strong significance. When the learning behaviors are inefficient, there is a high overlap degree of knowledge concepts across the four courses, but the overlap degree in related relationships decreases significantly, this indicates that no significant learning behavior routes are formed. For ineffective learning behaviors such as dropout, learners have low engagement with knowledge concepts. Most learners have not covered all the knowledge concepts, resulting in very low overlap in both knowledge concepts and their relationships. Consequently, the significant tests cannot be conducted. Therefore, the potential relationships of knowledge concepts significantly influence the routing tendencies of efficient learning behaviors.

Analysis of knowledge concept clusters

Based on the core knowledge concepts and associated knowledge concepts obtained from P1, as well as the potential

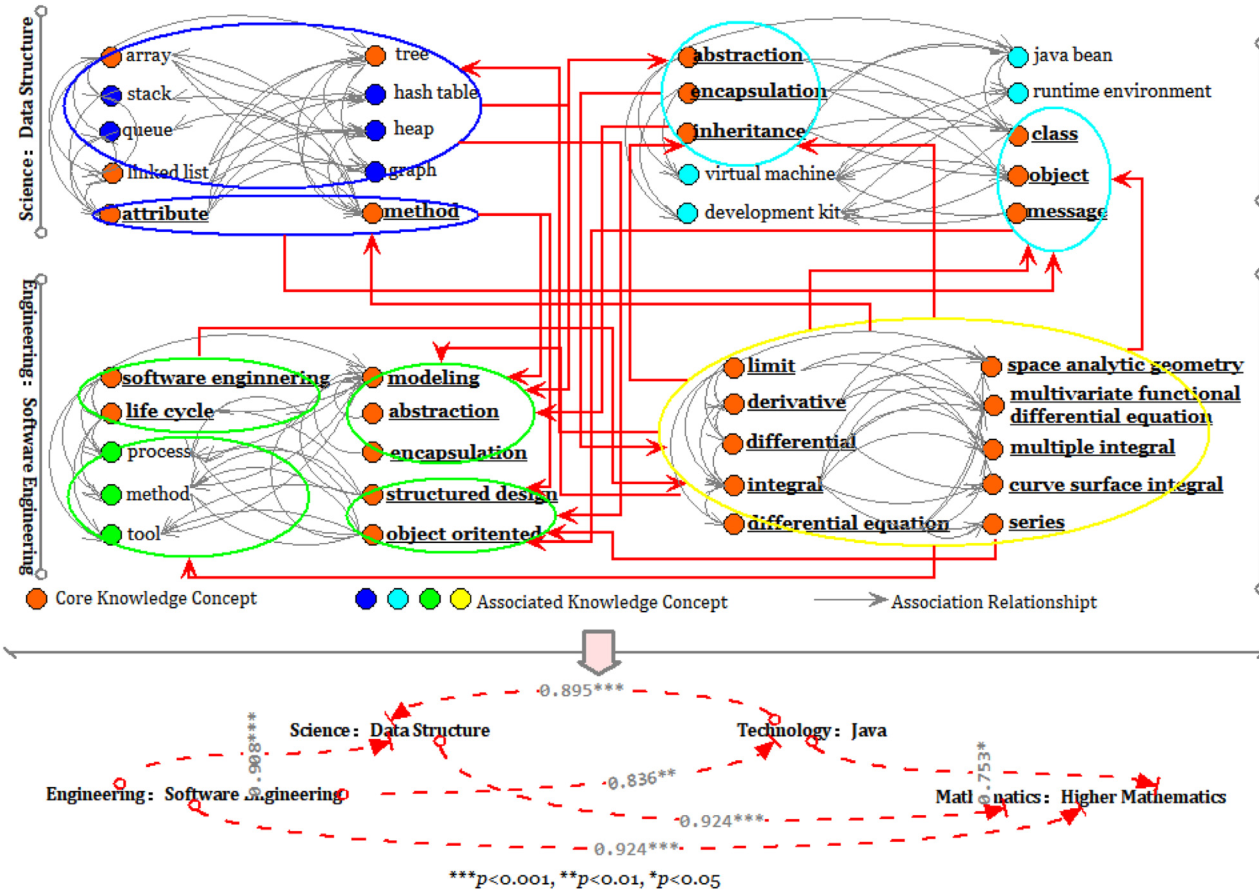

Figure 8. Clustering of knowledge concepts and potential correlation.

relationships between them, the clustering of key knowledge concepts in relevant courses of four domains is described in Table 3. The distribution of associated items and meta-paths based on ReRo model are used to analyze the clustering. As shown in Figure 8, it can be observed that each core knowledge concept in Data Structure and Advanced Mathematics is highly important for the other domains, resulting in two main concept clusters. For the corresponding courses in the other domains, their core knowledge concepts are divided into two or four clusters. The core knowledge concepts of these four domains form clear association tendencies, with Advanced Mathematics serving as an important prerequisite course for other three domains. The core knowledge concepts of Data Structure form the foundation for Java and Software Engineering, with Java driving the effective learning of Software Engineering. The distinct links between the clusters of knowledge concepts in different domains provide a clear knowledge system during learning these courses.

According to the analysis results of efficient learning behavior instances for each course from P1, the data analysis results reveal the cross-domain interest tendencies in the learning process. By utilizing the meta-path routing and propagation strategies of ReRo model, the potential learning behavior tendencies are tracked and predicted among four domains. Through associated calculation and sufficient statistics, the correlation and significance of learners' attention to other domains are determined. The lower part of Figure 8 illustrates the potential relationships driven by knowledge concepts among four domains. Specifically, the efficient learning behaviors in Software Engineering show significant influences on the relevant knowledge concepts of Advanced Mathematics, Data Structures, and Java. Similarly, the efficient learning behaviors in Java exhibit the personalized and autonomous enthusiasm for studying Data Structure. However, the efficient learning behaviors in Advanced Mathematics or Data Structure does not guide strong attention or substantial intention towards Java and Software Engineering, the motivation for learners to learn these two courses in association is weak, and there is no psychological motivation or tendency to turn to multiple fields and courses.

Therefore, when learners engage with practical and applied learning contents and demonstrate efficient learning behaviors, they actively establish the multi course connections and engage in self-directed learning driven by the principles and rules of related STEM domains. This significantly influences their potential interests in other domains.

Analysis of specific knowledge concept learning paths

The knowledge concepts are clustered based on the test results of P2, as well as the potential relationships are associated between different domains, that can be used to analyze the interrelationships of knowledge concepts in four domains. By employing ReRo model and its corresponding meta-paths, efficient learning behaviors for the four domains are identified. With the knowledge concepts as nodes, the potential learning behavior routing with strong correlations between knowledge concepts are constructed. During tracking learning behavior routing it is discovered that associated knowledge concepts in Data Structure and Software Engineering have become core knowledge concepts, the knowl-

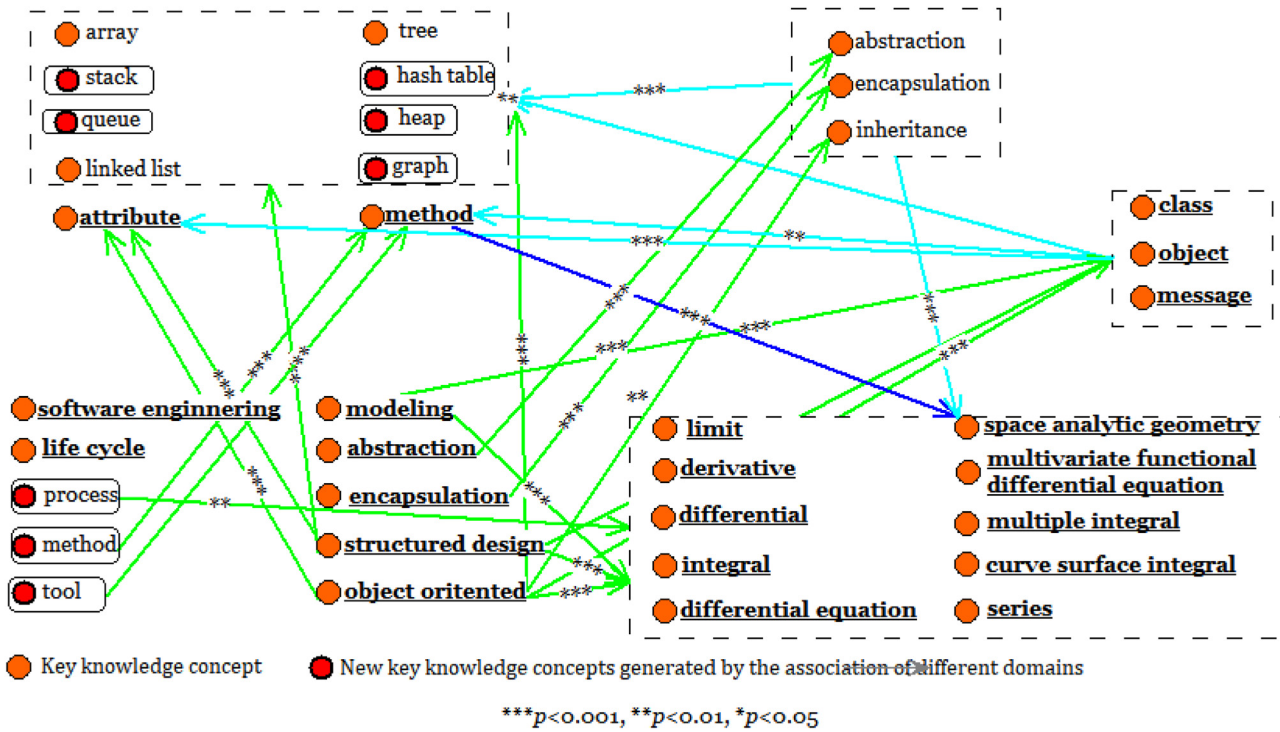

Figure 9. Specific knowledge concept learning paths for different domains.

edge concept labels Figure 9 are significantly different from those in Figure 8. As a result, all ten knowledge concepts in Data Structure have become important prerequisites for learners to learn Java and Software Engineering. The knowledge concepts of "process," m̈ethod,änd ẗoolïn Software Engineering have transformed into core knowledge concepts during the formation of their potential learning tendencies. The learning behavior routing formed between knowledge concepts are consistent with Figure 8. Specifically, the courses with stronger application and practice tend to prompt learners to initiate the learning of related principles and knowledge concepts from other courses more consciously. By selecting and learning the knowledge concepts with problem-solving objectives, learners can improve their learning motivation and effectiveness. The potential learning behavior routing from Software Engineering to Data Structure and Java is significant. In this way, the potential learning behavior routing from Software Engineering to Data Structure or Java is closely related to knowledge concepts, indicating that Data Structure and Java are important precursor courses for learning Software Engineering.

Therefore, when the efficient learning behaviors are applied to study practical and applied courses, learners tend to focus on and participate in specific knowledge concepts, using them as the enabled conditions to explore related, similar, and associated knowledge concepts, thereby forming the strong learning behavior routing. Software Engineering clearly drives the learning enthusiasm of Data Structure and Java, and also drives learners to systematically construct deep learning routing related to multi courses and knowledge concepts, and promotes the definition and use of more core knowledge concepts. It also significantly affects learning behavior routing related to other STEM domains.

Analysis of deep learning behavior routing

P4. Based on the validation results of P3, we verify the efficient deep learning behavior routing formed by learners among knowledge systems of different domains. The test results of P3 reveal that learners autonomously focus on and reinforce the foundational and theoretical knowledge concepts when engaging in the learning of applied and practical courses. Moreover, learners might establish strong correlation between knowledge concepts across different STEM domains. To further explore the efficient deep learning routing among these four STEM domains, ReRo model generates corresponding meta-paths and core knowledge concepts. Additionally, new core knowledge concepts that emerge during the inter-domain associations in the testing process of P3 are identified. Using the efficient learning behavior instances for training and testing, a thorough analysis of P4 is achieved to construct a feasible and efficient deep learning routing, As shown in Figure 10, this implies that learners need to develop a solid and stable learning effect across four STEM domains, that effectively follows the learning process from theory to application. Furthermore, learners should establish a strong knowledge system among core knowledge concepts in related courses to fully comprehend the principles, practical requirements, and application rules. This will also enable more effective associations with other STEM domains to provide efficient accumulation of prerequisite knowledge concepts. So the descriptive sequence of the deep learning routing is: (Mathematics→Data Structure)→Java)→Software Engineering.

In the case presented in this study, it is advisable to first systematically learn Advanced Mathematics, followed by a deeper understanding of knowledge concepts and relationships in Data Structure, then proceed to Java with the goal of engineering development. Software Engineering exhibits more significant asso-

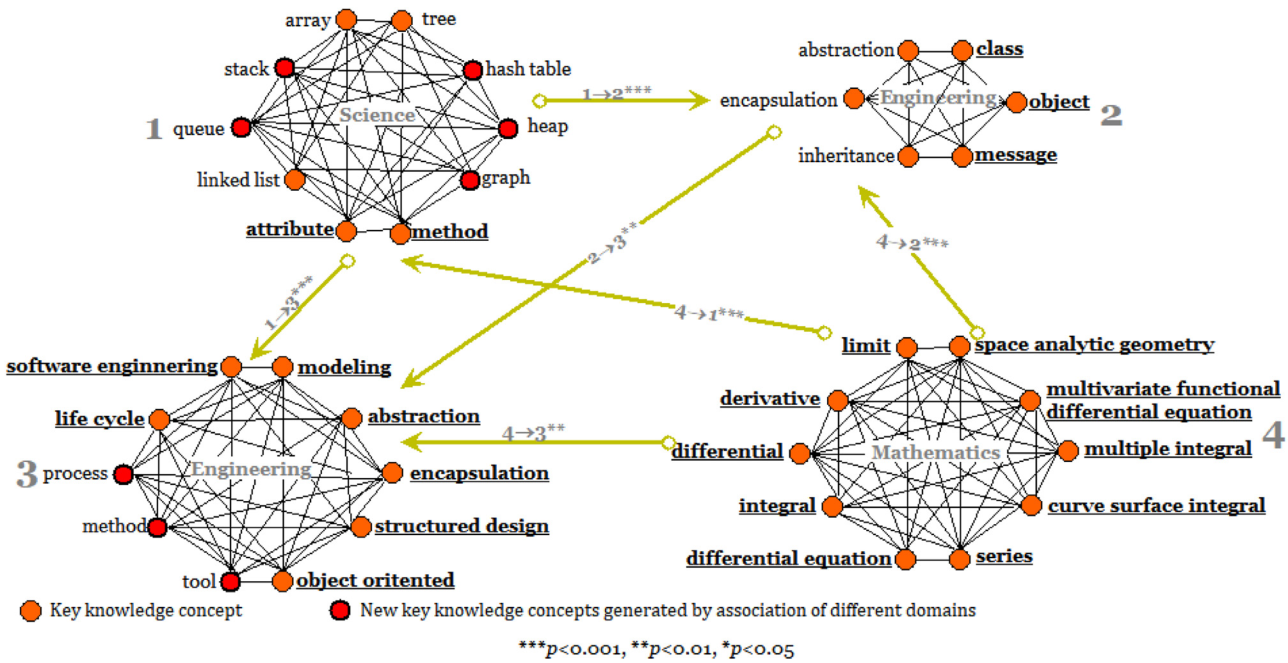

Figure 10. Deep learning routing in different domains.

ciation and intersection with other knowledge concepts. This establishes the effective routing strategies for deep learning within four STEM domains. From the analysis results, it can be inferred that the deep learning of related STEM courses requires learners to respect the regularities of learning contents and develop a coherent knowledge system for core knowledge concepts (Chu et al., 2022). Furthermore, based on the data analysis of the four corresponding courses, it can be observed that an exceptional computer manager, particularly a software engineer (Architect, algorithmist), must possess a strong mathematical background, theoretical learning background, and research and development experience.

Through the analysis and test of above four problems, it has been found that there are the strong correlations among the knowledge concepts in the STEM domains. Additionally, a knowledge system has been formed in each domain. The efficient learning behaviors occur within these correlations, leading to the potential learning tendencies and forming corresponding learning interests. This facilitates the systematic learning and enables the establishment of knowledge systems. Regarding the online learning behaviors of achieving positive and proactive learning psychological awareness, the solid deep learning routing might be constructed and propagated.

Discussion

With the emergence of online learning platforms such as MOOCs, the autonomy and personalization of online collaborative learning have been well demonstrated, especially in STEM education. Meanwhile, The psychological changes and learning attitudes of learners in online learning process will also be displayed through the correlation analysis of learning behavior instances, and positive learning behavior routing will be derived through the analysis of the complete learning process. Key influencing factors, potential relationships, and change characteristics will be explored for negative learning outcomes or dropout phenomena, and timely tracking, warning, and intervention will be implemented to sustainably guide and improve learners' initiative and positive enthusiasm (Van Hoe et al., 2024). Teachers are no longer direct guides in the learning process, as learners can engage in course learning anytime in MOOCs. The entire learning process is initiated and explored by the learners themselves (Arizmendi et al., 2023). In learners' self-directed engagement, the generated data in MOOCs becomes the direct description of learning behaviors. However, despite learners having ample agency, online learning has also brought about various challenges. One key issue is that learners struggle to effectively analyze the relevant knowledge concepts and understand there relationships. Faced with vast learning contents, learners are prone to getting lost with massive negative psychological experiences, that may even prematurely discontinue the learning process (Borrella et al., 2022). This inability to achieve effective learning outcomes not only hampers the full utilization of MOOCs resources, but also impacts the scheduling mechanisms of the learning process. This drawback is particularly detrimental in STEM education, which involves the association and integration of knowledge concepts across multi subjects and courses (Guo et al., 2023).

Adequate data analysis and problem verification indicate that interest-driven learning of STEM knowledge concepts in MOOCs is active, that might enable the effective learning behavior routing, contribute to enhancing learner participation and gradually constructing deep learning routing with interrelated knowledge concepts from multi courses in different domains. This study focuses on the exploring and prediction of potential interests during learning the knowledge concepts. By extracting the multi features, defining the meta-paths, as well as completing the feature fusion helped by the attention mechanism and multi-layer graph convolutional neural network, we have achieved the correlation analysis and efficient prediction of STEM learning behavior instances. The whole analysis and prediction process has been proven to reliable and effective.

STEM involves four domains that have many connections and different associations, forming various subjects, even hot topics and issues. The widespread application of MOOCs has promoted the sharing and propagation of high-quality STEM courses, facilitated the formation of course knowledge framework, and enabled the integration and fusion of STEM-related course concepts. According to the four domains of STEM, this study analyzes and demonstrates the relationships between different courses, proves the logical correlation of knowledge concepts and their drives self-directed learning. Case studies have found that the applied and practical courses make it easier for learners to develop the strong interests in related foundational or theoretical knowledge concepts, and derive systematic and solid mathematical and scientific learning outcomes, guiding learners to gradually build effective deep learning routing. Two main decision recommendation are deduced from this study:

(1) Construct a multi-layer knowledge concept learning method with an application-oriented practice background. During the STEM educational process, the courses are often categorized based on their designations. For example, many programs define mathematics as a core prerequisite course, which is then further divided into different courses that can be shared across different learning periods. Learners need to complete the required credits before moving on to subsequent learning contents. Within the definition of professional courses, some are classified as core courses, which directly teach principles, methods, and rules related to the professional skills, while others are elective courses that are still related to different majors but offer different research directions, allowing learners to make personalized choices based on their professional interests. Regardless of different courses, if learners only focus on the internal knowledge concepts and relationships, they may find it difficult to understand the meanings and applications of those complex mathematical formulas, theorems, and lemmas, making it challenging to their learning objectives or research directions. Learners struggle to ensure the applied fields and practical ways of knowledge concepts, resulting in a passive learning process.

Therefore, it is more beneficial for a teacher with a systematic professional background to explore and catalyze learners' potential interests and stimulate their active learning motivation, in order to develop the psychological awareness of continuous learning (Silva et al., 2023). To some extent, learners might strengthen their higher requirements on the exposition of knowledge concepts and the process of video explanations. Through sufficient heuristic case teaching, learners are given appropriate guidance to drive their self-discovery and problem-solving abilities.

(2) Deepen the understanding and expansion of foundational and theoretical knowledge concepts in engineering development. Regarding the whole analysis of interest mining in knowledge concepts and deep learning routing involves P2 and P3, this study has demonstrated that when learners study the knowledge concepts of Software Engineering, they actively associate them with Advanced Mathematics, Data Structure, and Java, these courses they may have studied before. To deepen their understanding and expand their current course's knowledge concepts, learners are likely to choose and consolidate the related contents from other related courses. The propagation of learning interests might enable to form the multi course knowledge system of some particular domain. Conversely, when learners start with foundational and theoretical knowledge concepts, they do not necessarily explore the associated applied and practical courses. Furthermore, the foundational and theoretical knowledge concepts should be introduced earlier than engineering development courses, but if learners focus only on foundational and theoretical courses without the guidance and implementation strategies of engineering practice cases, the learning outcomes may be significant worse. Learners can develop negative emotions about course learning due to the complexity of knowledge concepts and reasoning processes, and even discontinue their learning process midway. This further confirms the needs to create ample engineering development contexts to facilitate the effective deep learning routing.

Therefore, the STEM course learning should follow the engineering development processes and realize the linearity and coherence of knowledge concepts. In engineering practice, we might deepen the understanding and expansion of foundational and theoretical knowledge concepts by combining theory with practice. During the explanation of engineering courses, the relevant knowledge concepts of core courses should be appropriately linked, providing learners with assistance and instructions to establish the multi-layer iterative and correlated knowledge system.

Conclusion

STEM education aims to promote the interdisciplinary integration of multi subjects and courses, creating the closely interconnected knowledge system, driving learners to engage in interdisciplinary learning with proactive actions and positive attitudes. The learning process often involves the interrelationships and constraints among various concepts, where the understanding of prerequisite concepts might be very essential for the study of particular courses, moreover, some courses may serves as the foundation for subsequent related courses. About the traditional teaching models, the limitations and topological order between concepts are determined by teachers' syllabus and teaching task, this is one common passive learning approach that often fails to provide learners with a comprehensive understanding of related learning contents, leading to the disengagement and disconnect between knowledge acquisition and practical application.

In this study, we analyze and demonstrate the potential interest recommendation strategies and deep learning routing schemes for STEM knowledge concepts in MOOCs. We have designed relevant models and methods that integrate the selection and recommendation of learning contents. Nonetheless, it remains a challenge to empower learners to autonomously establish effective relationships to better connect and correlate the relevant knowledge concepts based on their learning backgrounds and current study materials. To address this challenge, our study first enhances the natural fusion approach based on the end-to-end graph convolutional neural network, that is used to convert the rich heterogeneous context features and content features of STEM learning behaviors into an effective recommendation process for knowledge concepts. This approach establishes a heterogeneous information network that centers around knowledge concepts, supporting multi entities, features, and courses in a more natural and intuitive manner. Secondly, we propose the learning method based on graph convolutional neural network and attention mechanism, that might explore and represent the meta-paths through which different entities can successfully propagate the knowledge concepts each other. This study enables the effective association and clustering of knowledge concepts and the formation of deep learning routing.

Extensive experiments analyze and verify a wealth of STEM learning behavior instances in MOOCs. The results indicate that our proposed methods can effectively and reliably identify learners' potential interests in knowledge concepts and provide accurate and suitable recommendation, facilitating precise deep learning routing. However, it is important to note that this study focuses on STEM learning behavior instances, and further experiments and validation are needed to assess the generality and robustness of our methods, particularly when applied to new STEM learning behavior instances, it is argued that the number of convolutional layers may require further adjustment and optimization. Meanwhile, our research focuses on the four domains of STEM. As we all know, with the increasing interdependence between disciplines, STEM has expanded to STEAM (Science, Technology, Engineering, Art, and Mathematics). Art has also been proven to have a certain potential relationship with the four domains of STEM. STEAM has expanded the breadth of knowledge concept propagation, and the related features, attributes, and relationships have become more complex, the interdisciplinary connections and intersections are closer and more comprehensive. For the larger interdisciplinary network formed by the association between the five domains of STEAM or more social sciences and natural sciences (DeLuca et al., 2024), further data analysis and method testing are needed, and more comprehensive and sufficient iterative prediction and calculation will also be needed for learners. The correlation and intersection between disciplines are closer and more comprehensive. As learners relate their learning goals to more disciplines, courses, knowledge, and even specific sub goals and sub needs, this poses greater challenges for sustainable online learning processes and also leads to more complex problems. As the key participants of entire learning process, learners need to achieve organic integration and adaptive scheduling of more diverse disciplines, courses, knowledge concepts, attributes, features, and resources, this is because learners are more prone to burnout and boredom during multi associated learning processes, and even completely abandon online learning. Accurately describing the learner's state and tracking psychological changes throughout the entire teaching and learning process is crucial. Cognition will also be accompanied by greater expansion challenges and deep reinforcement, and the exploration of related learning behavior tracking and deep learning routing will be more difficult. However, the method design and problem testing of this study provide a feasible analytical approach for more interdisciplinary and related studies, and also achieve the innovative design and analysis of key parts for sufficient argument about STEAM, that should potentially involve the definition of new problems, as well as the optimization or redesign of new methods and patterns. Our work might verify the feasibility of knowledge concept propagation and deep learning routing formation through massive and real online learning behavior instances, and derives relevant implementation decisions and rule, which could have more practical significance, and provide a certain decision basis for the construction and reliability testing of subsequent STEAM related issues (Guimerans-Sanchez et al., 2024). We have be aware of this issue and will implement the corresponding research in a phased manner in subsequent research work.

Hence, in the future study, it is suggested to expand the categories of features and types of meta-paths for STEM or STEAM knowledge concepts, explore more factors influencing learners' potential interests, guide learners in constructing more rational deep learning modes, drive effective learning behavior routing, improve learning outcomes, optimize learning state, build psychological awareness of self-directed learning, explore adaptive learning strategies, and achieve more effective data analysis and decision recommendation for online STEM or STEAM education.

Availability of supporting data

The datasets used or analyzed during the current study are available from the corresponding author on reasonable request.

Conflict of interest

The authors declare that there are no conflicts of interest.

Compliance with ethical standards

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled. The whole research does not involve human participants and/or animals No conflict of interest exits in the submission of this manuscript, and manuscript is approved by all authors for publication. I would like to declare on behalf of my co-authors that the work described was original research that has not been published previously, and not under consideration for publication elsewhere, in whole or in part. All the authors listed have approved the manuscript that is enclosed.

Funding

This study is supported by Natural Science Foundation of Shandong Province (Grant NO. ZR2023MF099).

Acknowledgements

Thanks for the technical support provided by University of Oxford (UK), University of Pennsylvania (USA) and Tsinghua University (China), as well as the theoretical guidance and practical reference of University of Agder (Norway) and Qufu Normal University (China).

References

Aldowah, H., Al-Samarraie, H., Alzahrani, A. I., & Alalwan, N. (2020). Factors affecting student dropout in MOOCs: a cause and effect decision-making model. *Journal of Computing in High Education*, *32*(2), 429–454. https://doi.org/10.1007/s12528-019-09241-y

Anttila, S., Lindfors, H., Hirvonen, R., Määttä, S., & Kiuru, N. (2023). Dropout intentions in secondary education: Student temperament and achievement motivation as antecedents. *Journal of Adolescence*, *95*(2), 248–263. https://doi.org/10.1002/jad.12110

Arizmendi, C. J., Bernacki, M. L., Rakovic, M., Plumley, R. D., Urban, C. J., Panter, A. T., Greene, J. A., & Gates, K. M. (2023). Predicting student outcomes using digital logs of learning behaviors: Review, current standards, and suggestions for future work. *Behavior Research Methods*, *55*(6), 3026–3054. https://doi.org/10.3758/s13428-022-01939-9

Bañeres, D., Rodríguez-González, D. M., Guerrero-Roldán, A. E., & Cortadas, P. (2023). An early warning system to identify and intervene online dropout learners. *International Journal of Educational Technology in Higher Education*, *20*(1), 3. https://doi.org/10.1186/s41239-022-00371-5

Borrella, I., Caballero-Caballero, S., & Ponce-Cueto, E. (2022). Taking action to reduce dropout in MOOCs: Tested interventions. *Computers & Education*, *179*(1), Article 104412. https://doi.org/10.1016/j.compedu.2021.104412

Buckley, J., Gumaelius, L., Nyangweso, M., Hyland, T., Seery, N., & Pears, A. (2023). The impact of country of schooling and gender on secondary school students' conceptions of and interest in becoming an engineer in Ireland, Kenya and Sweden. *International Journal of STEM Education*, *10*(1), 28. https://doi.org/10.1186/s40594-023-00416-9

Calvera-Isabal, M., Santos, P., & Hernandez-Leo, D. (2023). Towards citizen science-inspired learning activities: The co-design of an exploration tool for teachers following a human-centred design approach. *International Journal of Human-Computer Interaction*, *2023*(4), 1–13. https://doi.org/10.1080/10447318.2023.2201554

Cetron, J. S., Connolly, A. C., Diamond, S. G., May, V. V., Haxby, J. V., & Kraemer, D. J. M. (2020). Using the force: STEM knowledge and experience construct shared neural representations of engineering concepts. *NPJ Science of Learning*, *5*(1), 6. https://doi.org/10.1038/s41539-020-0065-x

Chen, P., Teo, D. W. H., Foo, D. X. Y., Derry, H. A., Hayward, B. T., Schulz, K. W., Hayward, C., McKay, T. A., & Ong, D. C. (2022). Real-world effectiveness of a social-psychological intervention translated from controlled trials to classrooms. *NPJ Science of Learning*, *7*(1), 20. https://doi.org/10.1038/s41539-022-00135-w

Chu, H. C., Hwang, G. H., Tu, Y. F., & Yang, K. H. (2022). Roles and research trends of artificial intelligence in higher education: A systematic review of the top 50 most-cited articles. *Australasian Journal of Educational Technology*, *38*(3), 22–42. https://doi.org/10.14742/ajet.7526

Costello, R. A., Salehi, S., Ballen, C. J., & Burkholder, E. (2023). Pathways of opportunity in stem: Comparative investigation of degree attainment across different demographic groups at a large research institution. *International Journal of STEM Education*, *10*(1), 46. https://doi.org/10.1186/s40594-023-00436-5

Daker, R. J., Gattas, S. U., Sokolowski, H. M., Green, A. E., & Lyons, I. M. (2021). First-year students' math anxiety predicts STEM avoidance and underperformance throughout university independently of math ability. *NPJ Science of Learning*, *6*(1), 17. https://doi.org/10.1038/s41539-021-00095-7

Dash, R., Ranjan, K. R., & Rossmann, A. (2022). Dropout management in online learning systems. *Behaviour and Information Technology*, *41*(9), 1973–1987. https://doi.org/10.1080/0144929X.2021.1910730

DeLuca, C., Dubek, M., & Dubek, M. (2024). How hermeneutics can guide grading in integrated STEAM education: An evidence-informed perspective. *British Educational Research Journal*, (1), 1–18. https://doi.org/10.1002/berj.3979

Edelsbrunner, P. A., Malone, S., Hofer, S. I., Küchemann, S., Kuhn, J., Schmid, R., Altmeyer, K., Brünken, R., & Lichtenberger, A. (2023). The relation of representational competence and conceptual knowledge in female and male undergraduates. *International Journal of STEM Education*, *10*(1), 44. https://doi.org/10.1186/s40594-023-00435-6

Evenhouse, D., Lee, Y., Berger, E., Rhoads, J. F., & DeBoer, J. (2023). Engineering student experience and self-direction in implementations of blended learning: A cross-institutional analysis. *International Journal of STEM Education*, *10*(1), 19. https://doi.org/10.1186/s40594-023-00406-x

Flegr, S., Kuhn, J., & Scheiter, K. (2023). How to foster STEM learning during Covid-19 remote schooling: Combining virtual and video experiments. *Learning and Instruction*, *86*(1), Article 101778. https://doi.org/10.1016/j.learninstruc.2023.101778

Gijsen, M., Catrysse, L., De Maeyer, S., & Gijbels, D. (2024). Mapping cognitive processes in video-based learning by combining trace and think-aloud data. *Learning and Instruction*, *90*(1), Article 101851. https://doi.org/10.1016/j.learninstruc.2023.101851

Gilligan-Lee, K. A., Hawes, Z. C. K., & Mix, K. S. (2022). Spatial thinking as the missing piece in mathematics curricula. *NPJ Science of Learning*, *7*(1), 10. https://doi.org/10.1038/s41539-022-00128-9

Gomes, S., Costa, L., Martinho, C., Dias, J., Xexeo, G., & Santos, A. M. (2023). Modeling students' behavioral engagement through different in-class behavior styles. *International Journal of STEM Education*, *10*(1), 21. https://doi.org/10.1186/s40594-023-00407-w

Guimerans-Sanchez, P., Alonso-Ferreiro, A., Zabalza-Cerdeirina, M. A., & Monreal-Guerrero, I. M. (2024). E-textiles for STEAM education in primary and middle school: A systematic review. *Revista Iberoamericana de Educación a Distancia*, *27*(1), 1–15. https://doi.org/10.5944/ried.27.1.37645

Guo, L., Du, J., & Zheng, Q. (2023). Understanding the evolution of cognitive engagement with interaction levels in online learning environments: Insights from learning analytics and epistemic network analysis. *Journal of Computer Assisted Learning*, *39*(3), 984–1001. https://doi.org/10.1111/jcal.12781

Gupta, A., Garg, D., & Kumar, P. (2022). Mining sequential learning trajectories with hidden Markov models for early prediction of at-risk students in E-learning environments. *IEEE Transactions on Learning Technologies*, *15*(6), 783–797. https://doi.org/10.1109/TLT.2022.3197486

Hsu, L. (2023). EFL learners' self-determination and acceptance of LMOOCs: The UTAUT model. *Computer Assisted Language Learning*, *36*(7), 1177–1205. https://doi.org/10.1080/09588221.2021.1976210

Ioannou, A., & Gravel, B. E. (2024). Trends, tensions, and futures of maker education research: A 2025 vision for STEM plus disciplinary and transdisciplinary spaces for learning through making. *Educational Technology Research and Development*, *72*(1), 1–14. https://doi.org/10.1007/s11423-023-10334-w

Khor, E. T., & Dave, D. (2022). A learning analytics approach using social network analysis and binary classifiers on virtual resource interactions for learner performance prediction. *The International Review of Research in Open and Distributed Learning*, *23*(4), 123–146. https://doi.org/10.19173/irrodl.v23i4.6445

Khoushehgir, F., & Sulaimany, S. (2023). Negative link prediction to reduce dropout in Massive Open Online Courses. *Education and Information Technologies*, *28*(8), 10385–10404. https://doi.org/10.1007/s10639-023-11597-9

Kim, K., & Tawfik, A. A. (2023). Different approaches to collaborative problem solving between successful versus less successful problem solvers: Tracking changes of knowledge structure. *Journal of Research on Technology in Education*, *55*(4), 628–645. https://doi.org/10.1080/15391523.2021.2014374

Kleinschmit, A. J., Rosenwald, A., Ryder, E. F., Donovan, S., Murdoch, B., Grandgenett, N. F., Pauley, M., Triplett, E., Tapprich, W., & Morgan, W. (2023). Accelerating STEM education reform: Linked communities of practice promote creation of open educational resources and sustainable professional development. *International Journal of STEM Education*, *10*(1), 16. https://doi.org/10.1186/s40594-023-00405-y

Lee, H. Y., Cheng, Y. P., Wang, W. S., Lin, C. J., & Huang, Y. M. (2023). Exploring the learning process and effectiveness of STEM education via learning behavior analysis and the interactive-constructive-active-passive framework. *Journal of Educational Computing Research*, *61*(5), 951–976. https://doi.org/10.1177/07356331221136888

Maric, D., Fore, G. A., Nyarko, S. C., & Varma-Nelson, P. (2023). Measurement in STEM education research: A systematic literature review of trends in the psychometric evidence of scales. *International Journal of STEM Education*, *10*(1), 39. https://doi.org/10.1186/s40594-023-00430-x

Mccarthy, S., Rowan, W., Kahma, N., Lynch, L., & Ertiö, T. P. (2021). Open e-learning platforms and the design–reality gap: An affordance theory perspective. *Information Technology & People*, *35*(8), 74–98. https://doi.org/10.1108/ITP-06-2021-0501

Mourdi, Y., Sadgal, M., Elabdallaoui, H. E., El Kabtane, H., & Allioui, H. (2023). A recurrent neural networks based framework for at-risk learners' early prediction and MOOC tutor's decision support. *Computer Applications in Engineering Education*, *31*(2), 270–284. https://doi.org/10.1002/cae.22582

Mubarak, A. A., Cao, H., Hezam, I. M., & Hao, F. (2022). Modeling students' performance using graph convolutional networks. *Complex & Intelligent Systems*, *8*(3), 2183–2201. https://doi.org/10.1007/s40747-022-00647-3

Norris, C. M., Taylor, T. A., & Lummis, G. W. (2023). Fostering collaboration and creative thinking through extra-curricular challenges with primary and secondary students. *Thinking Skills and Creativity*, *48*(1), Article 101296. https://doi.org/10.1016/j.tsc.2023.101296

Parviainen, M., Aunola, K., Torppa, M., Poikkeus, A. M., & Vasalampi, K. (2020). Symptoms of psychological ill-being and school dropout intentions among upper secondary education students: A person-centered approach. *Learning and Individual Differences*, *80*(1), Article 101853. https://doi.org/10.1016/j.lindif.2020.101853

Pattison, S. A., Gontan, I., Ramos-Montaez, S., Shagott, T., Francisco, M., & Dierking, L. (2020). The identity-frame model: a framework to describe situated identity negotiation for adolescent youth participating in an informal engineering education program. *Journal of the Learning Sciences*, *29*(4-5), 550–597. https://doi.org/10.1080/10508406.2020.1770762

Silva, M. P. R. I. R., Rupasingha, R. A. H. M., & Kumara, B. T. G. S. (2024). Identifying complex causal patterns in students' performance using machine learning. *Technology Pedagogy and Education*, *12*, 1–17. https://doi.org/10.1080/1475939X.2023.2288015

Solomon, F., Champion, D., Steele, M., & Wright, T. (2022). Embodied physics: Utilizing dance resources for learning and engagement in stem. *Journal of the Learning Sciences*, *31*(1), 71–106. https://doi.org/10.1080/10508406.2021.2023543

Van Hoe, A., Wiebe, J., Rotsaert, T., & Schellens, T. (2024). The implementation of peer assessment as a scaffold during computer-supported collaborative inquiry learning in secondary STEM education. *International Journal of STEM Education*, *11*(1), 3. https://doi.org/10.1186/s40594-024-00465-8

Weston, T. J., Laursen, S. L., & Hayward, C. N. (2023). Measures of success: Characterizing teaching and teaching change with segmented and holistic observation data. *International Journal of STEM Education*, *10*(1), 24. https://doi.org/10.1186/s40594-023-00413-y

Wu, J., & Uttal, D. H. (2024). Diversifying computer science: An examination of the potential influences of women-in-computing groups. *Science Education*, *1*, 1–24. https://doi.org/10.1002/sce.21861

Xia, X. (2022). Application technology on collaborative training of interactive learning activities and tendency preference diversion. *SAGE Open*, *12*(2), 1–15. https://doi.org/10.1177/21582440221093368

Xia, X. (2020). Learning behavior mining and decision recommendation based on association rules in interactive learning environment. *Interactive Learning Environments*, *8*, 1–16. https://doi.org/10.1080/10494820.2020.1799028

Xia, X. (2021). Sparse learning strategy and key feature selection in interactive learning environment. *Interactive Learning Environments*, *11*, 1–25. https://doi.org/10.1080/10494820.2021.1998913

Xia, X. (2020). Random field design and collaborative inference strategies for learning interaction activities. *Interactive Learning Environments*, *12*, 1–25. https://doi.org/10.1080/10494820.2020.1863236

Xia, X. (2021b). Decision application mechanism of regression analysis of multi-category learning behaviors in interactive learning environment. *Interactive Learning Environments*, *4*, 1–14. https://doi.org/10.1080/10494820.2021.1916767

Xia, X. (2021c). Interaction recognition and intervention based on context feature fusion of learning behaviors in interactive learning environments. *Interactive Learning Environments*, *1*, 1–19. https://doi.org/10.1080/10494820.2021.1871632

Xia, X., & Qi, W. (2024). Driving STEM learning effectiveness: dropout prediction and intervention in MOOCs based on one novel behavioral data analysis approach. *Humanities and Social Sciences Communications*, *11*(1), 430. https://doi.org/10.1057/s41599-024-02882-0

Xia, X., & Qi, W. (2022). Early warning mechanism of interactive learning process based on temporal memory enhancement model. *Education and Information Technologies*, *7*, 1–22. https://doi.org/10.1007/s10639-022-11206-1

Xia, X., & Qi, W. (2023a). Dropout Prediction and Decision Feedback Supported by Multi Temporal Sequences of Learning Behavior in MOOCs. *International Journal of Educational Technology in Higher Education*, *6*, 1–24. https://doi.org/10.1186/s41239-023-00400-x

Xia, X., & Qi, W. (2023b). Learning behavior interest propagation strategy of MOOCs based on multi entity knowledge graph. *Education and Information Technologies*, *3*, 1–29. https://doi.org/10.1007/s10639-023-11719-3
Xia, X., & Qi, W. (2023c). Interpretable early warning recommendations in interactive learning environments: A deep-neural network approach based on learning behavior knowledge graph. *Humanities & Social Sciences Communications*, *10*. https://doi.org/10.1057/s41599-023-01739-2
Xia, X., & Wang, T. (2022). Multi objective evaluation between learning behavior and learning achievement. *Asia-Pacific Education Researcher*, *12*, 1–15. https://doi.org/10.1007/s40299-022-00703-z