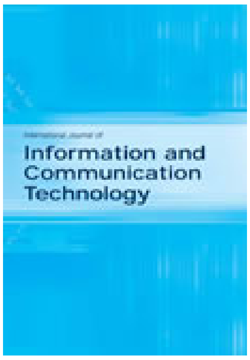


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Personalised foreign language learning path recommendation strategy based on disciplinary knowledge graph

Jia Guo, Zhe Zhou

Article History:

Received:	10 February 2025
Last revised:	20 February 2025
Accepted:	21 February 2025
Published online:	16 April 2025

Personalised foreign language learning path recommendation strategy based on disciplinary knowledge graph

Jia Guo*

School of Foreign Languages and International Trade,
Wuhan Polytechnic,
Wuhan 430074, China
Email: guojia19882021@163.com
*Corresponding author

Zhe Zhou

School of Resource and Environmental Sciences,
Wuhan University,
Wuhan 430079, China
Email: zz87710@126.com

Abstract: With an aim to increase learners' learning efficiency and experience, this paper suggests a personalised foreign language learning path suggestion technique based on subject knowledge graph. Using a knowledge graph in the domain of foreign language studies, integrating multidimensional knowledge points such as vocabulary, grammar, and pragmatics, and revealing their correlations and hierarchical structures, a fusion recommendation algorithm based on multiple learning factors was designed. The method improves learning path recommendation by thoroughly considering the sequence and similarity of knowledge items. According to the trial data, this approach can help students master information points rather successfully, optimise the length of learning paths, and significantly enhance learning interest and enthusiasm, providing important references for the design of foreign language learning systems.

Keywords: knowledge graph; recommended learning paths; similarity of knowledge points; cognitive diagnosis; collaborative filtering.

Reference to this paper should be made as follows: Guo, J. and Zhou, Z. (2025) 'Personalised foreign language learning path recommendation strategy based on disciplinary knowledge graph', *Int. J. Information and Communication Technology*, Vol. 26, No. 8, pp.49–69.

Biographical notes: Jia Guo received her PhD from Wuhan University in 2020. Currently, she is a Lecturer at Wuhan Polytechnic. Her research interests are foreign language education and teaching, vocational education and so on.

Zhe Zhou received his Master's degree from China Geo University in 2013. Currently, he is a doctoral student at Wuhan University. His research interests are spatial cognition and spatio-temporal knowledge map under deep learning.

1 Introduction

In today's rapidly developing information technology, the education sector is undergoing profound changes. Personalised learning, as a learner centred educational model, is gradually receiving widespread attention. Foreign language learning is an important component of the education system, and its learning process involves a large number of complex knowledge points, such as vocabulary, grammar, pragmatics, and other multidimensional content (Golonka et al., 2014). However, traditional foreign language teaching mainly focuses on fixed curriculum design, ignoring individual differences among different learners (Andreyeva et al., 2019). This' one size fits all 'teaching method not only fails to meet the diverse needs of learners, but also easily leads to problems such as decreased learning interest and low learning efficiency. Therefore, how to support the design of personalised foreign language learning paths through technological means has become an important issue that urgently needs to be addressed.

Knowledge graph is a technology that has emerged in the field of artificial intelligence in recent years. Due to its advantages in structured information representation and knowledge association mining, it is widely used in intelligent recommendation, medical diagnosis, and educational analysis (Hao et al., 2021). In foreign language learning, knowledge graphs can be used to visually represent subject knowledge points and their associated structures, providing solid theoretical and data support for personalised learning path recommendations. By combining knowledge graphs with learners' behavioural data, a dynamically adaptive personalised learning path can be constructed, effectively addressing the shortcomings of traditional teaching models (Huang and Zhu, 2021).

Research already conducted in the domains of personalised recommendation systems, knowledge tracking, and path optimisation yields several outcomes. For instance, educational systems have extensively applied collaborative filtering-based recommendation systems to suggest learning materials to students (Gong, 2010); behavioural modelling techniques based on learning analysis can help understand learners' knowledge states (Park, 2009). However, there are still several shortcomings in the specific application of these studies in foreign language learning scenarios

The deep modelling of knowledge graphs is insufficient, and the knowledge associations in previous foreign language learning are mainly based on simple hierarchical or linear relationships, which fail to fully explore the multidimensional associations between complex knowledge points and make it difficult to fully reflect the knowledge structure of foreign language disciplines. Insufficient consideration of learners' dynamic needs, significant differences exist in learners' needs at different stages of learning, and existing recommendation methods lack dynamic analysis and adaptation of learning behaviours and states, making it difficult to generate learning paths that match their actual needs. The optimisation problem of personalised path generation, how to balance the relationship between learner knowledge point coverage and path length when recommending learning paths, and improve the efficiency of path generation and recommendation effectiveness, is still an urgent technical challenge to be solved.

In the sphere of education, research on individualised learning path suggestion and knowledge graph application, a large number of achievements have explored personalised recommendation systems, learning path optimisation, and the construction of subject knowledge graphs.

One of the main areas of research in the field of educational technology is personalised recommendation system. Traditional recommendation algorithms are mostly based on collaborative filtering methods, aiming to recommend suitable learning resources for learners by analysing their historical behaviour and interest preferences. For example, Burke et al. (2011) proposed a collaborative filtering-based recommendation framework applied in online learning environments to provide personalised course recommendations for learners. However, existing recommendation methods focus more on resource recommendation and less on personalised design of learning paths.

Knowledge tracking technology aims to dynamically model learners' knowledge mastery status to provide personalised learning support. Abdelrahman et al. (2023) proposed a deep learning-based knowledge tracking method that predicts students' mastery of knowledge points through long short-term memory networks (LSTM). This approach offers significant basis for the planning of learning paths and may dynamically assess the knowledge level of students.

Knowledge graph, as a technique for representing and processing knowledge, has been widely applied in multiple fields. In the field of education, subject knowledge graphs can intuitively represent the relationships between various knowledge points, providing learners with a systematic knowledge framework. For example, Zhang et al. (2023) proposed a graph convolutional network-based approach for modelling high-order correlations on knowledge graphs to more accurately capture learners' preferences. This method constructs a knowledge graph of the disciplinary field, revealing the hierarchical and relational relationships between knowledge points, and providing learners with a systematic learning path planning.

Learning path optimisation is one of the key technologies in personalised learning recommendation. Hidalgo-Paniagua et al. (2017) proposed a multi-objective firefly algorithm to solve path optimisation problems, this addresses three different goals in order to provide precise and workable answers. This approach can raise learners' interest and passion of studying in addition to their efficiency of learning.

In recent years, an increasing number of studies have attempted to combine knowledge graphs with personalised recommendations to enhance recommendation effectiveness. Based on knowledge graphs, Liu et al. (2023) suggested a personalised learning recommendation system that uses their structural representation of topic knowledge and shapes a learning path recommendation system based on knowledge graphs. This system uses node concentration and node weights to extend the knowledge graph system, thereby better expressing the structural connections between knowledge. The great possibilities of knowledge graphs in customised suggestions are shown in this work.

The emotional state of learners plays a crucial role in the learning process. Pooja and Bhalla (2022) studied the application of sentiment analysis in education and proposed to improve learners' engagement and learning outcomes by analysing students' emotional changes, adjusting learning paths and resource recommendations in real time. Emotion analysis technology provides new ideas for dynamically adjusting personalised learning paths.

Based on learners' knowledge mastery, Nabizadeh et al. (2020) suggested an intelligent learning route suggestion approach, which can intelligently adjust learning paths according to students' progress in mastering knowledge points and achieve personalised education. Through dynamic adjustment of the learning path, this approach increases the relevance and efficiency of education.

Application study of personalised recommendation technology is progressively becoming more important in the field of foreign language acquisition. Based on deep learning, Cheng and Di (2024) suggested a customised foreign language learning recommendation system that analyses student behaviour and feedback to create customised learning recommendations. The system can maximise the learning path and automatically modify the learning content and progress based on the real state of learners.

Based on reinforcement learning, Lin et al. (2022) suggested a dynamic learning path recommendation approach that automatically modifies the learning path to fit the always shifting learning demands of learners by means of real-time analysis of feedback. This method can effectively improve learning effectiveness and reduce the recommendation of ineffective learning content.

Introducing the NSGA-II evolutionary method for handling multi-objective optimisation problems, Qu and Li (2022) suggested a path optimisation algorithm using big data analytic approaches. Under a novel mode, the improved genetic algorithm was coupled with the multi-objective path optimisation model to derive the Pareto optimal solution set for the research topic in this work.

In response to the foregoing problems, this paper suggests a personalised foreign language learning path recommendation technique grounded on disciplinary knowledge graph. Based on a knowledge graph, this approach suggests individualised learning paths to students across the pre-existing and dependent connections in the graph. Next, we designed a fusion recommendation algorithm based on multiple learning factors, which targets learners' personalised weak knowledge points and learning interests, comprehensively considers the sequence and similarity of knowledge points, enhances the accuracy, coherence, and pertinence of learning path recommendation, and can better help learners fill knowledge gaps, improve learning efficiency, learning experience, and user satisfaction.

This paper carried experimental study in real-world foreign language learning environments in order to confirm the efficiency of the suggested approach. The proposed recommendation strategy in this research can considerably increase learners' mastery of knowledge points, maximise the length of learning routes, and increase learning interest and efficiency, according to the experimental results when compared with conventional learning path design methods. This study not only provides important references for the design of foreign language learning systems, but also offers new ideas for the theory and practice of personalised learning path recommendations.

2 Relevant technologies

2.1 Overview of knowledge graph

2.1.1 Construction of knowledge graph.

Knowledge graph (KG) is a graph structure for representing knowledge and its connections, where nodes represent entities (such as knowledge points, courses, learners, etc.) and edges represent relationships between entities (such as dependencies between knowledge points, mastery levels between learners and knowledge points, etc.). In the field of education, building subject knowledge graphs and transforming them into forms that can be processed by computers through embedding technology has become an important means of improving the performance of personalised learning recommendation systems.

Firstly, extract key entities from textbooks, reference books, online courses, and other resources in the field of the discipline, which can be individual knowledge points, themes, or concepts. For example, in foreign language learning, entities can be vocabulary, grammar rules, syntactic structures, etc.

The core of a knowledge graph is the relationships between entities. Relationships can be dependency relationships, similarity relationships between different knowledge points, or the degree of mastery between learners and knowledge points. The method of relationship extraction can be achieved through natural language processing (NLP) techniques such as dependency parsing, named entity recognition, etc., to automatically identify and extract semantic associations between entities from text. For example, the relationship between 'verb' and 'tense' can be linked through 'being a part of the tense change'.

Through the process of entity recognition and relationship extraction, a preliminary disciplinary knowledge graph can be constructed. Nodes in a graph represent knowledge points or entities, while edges represent semantic or dependency relationships between knowledge points. In a graph, the attributes of nodes can include descriptions and types of knowledge points, while the attributes of edges can include types and weights of relationships.

After constructing a preliminary knowledge graph, it is usually necessary to optimise the graph through expert knowledge or automated tools. For example, certain knowledge points may be missing in the graph and need to be manually supplemented; in addition, some relationships may be ambiguous or inaccurate and require algorithmic adjustment or correction.

2.1.2 Embedding knowledge graph

Knowledge graph embedding is the process of transforming entities and relationships in a graph into dense vector representations in a low dimensional vector space, enabling computers to process and infer information in the graph more efficiently (Wang et al., 2017). The embedding technique turns graph items and connections into vector representations, allowing the structure and semantics of the knowledge graph to be captured and transmitted, thereby promoting the application of the graph in practical tasks such as personalised learning path recommendation and problem solving.

Knowledge graph embedding aims mostly to learn low dimensional vector representations so preserving the structure and semantic information in the graph. The perfect embedding technique should allow vector space's similarity to reflect the similarity between entities.

Representing the SPO triplet of the knowledge graph as (h, r, t), i.e., (head, relation, tail), where h and t represent the head entity and tail entity respectively, and r represents the relationship between the head and tail entities is the main step of knowledge graph embedding. Firstly, randomly initialise each named entity and relationship into an n-dimensional vector format (requiring a custom vector dimension of n), which is the initial feature expression vector; then split the several triplets in proportionate training, testing, and validation sets. Furthermore separate the training set's positive and negative samples using the score functions matching each model to determine the loss values of the positive and negative samples correspondingly; calculate the objective function then depending on the concept of reducing the positive sample loss value and maximising the negative sample loss value; at last, reverse based on the objective function values optimise the feature expression vectors of every entity and connection. Embedding results of the knowledge graph are the feature expression vectors of every item and relationship acquired by iterative training.

The TransE model (Le et al., 2023) has few parameters, is simple and efficient to use, and has received widespread attention from many scholars.

For each triplet (h, r, t), the TransE model takes the relationship vector r as the connection vector between the head and tail entity vectors h and t, that is, the model wants $h + r \approx t$. Therefore, the scoring function of the model is expressed as follows:

$$f(h,t) = \| h + r - t \|_2^2 \tag{1}$$

The TransE model's objective function may be found from the aforementioned scoring function as follows:

$$L = \sum_{(h,r,t)\in S} \sum_{(h',r,t')\in S'} \max\left(0, f(h,t) + \gamma - f\left(h',t'\right)\right) \tag{2}$$

where S is a positive sample composed of positive triplets, and S' is a negative sample composed of positive triplet replacements; $f(h, t)$ is the score for positive samples, and $f(h', t')$ is the score for negative samples. When the score difference between positive and negative samples is greater than γ, the objective function can be minimised.

However, due to the fact that the TransE model is only embedded in one plane, it cannot handle the 'one to many' and 'many to many' relationships between entities well.

To compensate for the shortcomings of the TransE model, the TransH model (Ebisu and Ichise, 2019) has emerged. In each (h, r, t) triplet, the relationship r corresponds to a hyperplane, and the model projects the entities h and t onto the hyperplane corresponding to the relationship r.

The specific projection method of the TransH model is as follows, where W_r represents the projection matrix of the relationship r.

$$h_\perp = h - W_r^T h W_r \tag{3}$$

$$t_\perp = t - W_r^T t W_r \tag{4}$$

The representation of relationship r on its hyperplane is d_r, and the scoring system runs like this:

$$f_r(h,t) = \left\| h_\perp + d_r - t_\perp \right\|_2^2 \tag{5}$$

Based on the scoring function, the objective function of the TransH model can be given, where S and S' are positive and negative samples composed of positive and negative triplets, respectively; $f_r(h, t)$ and $f_r(h', t')$ represents the scores of positive and negative samples, respectively. The objective function can be minimised when the score difference between positive and negative samples exceeds γ.

$$L = \sum_{(h,r,t)\in S} \sum_{(h',r,t')\in S'} \max\left(0, f_r\left(h,t\right) + \gamma - f_r\left(h',t'\right)\right) \tag{6}$$

Nevertheless, the TransH model limits the useful use of the model since it projects both the head and tail entities into the relational space, so they have the same dimension.

Multiple projection matrices can be employed for head and tail entities, therefore addressing the issues of the TransH model; the TransR model (Wang et al., 2023) projects entities and relationships into multiple low dimensional regions. The TransR model's head and tail entity mapping technique is demonstrated here.

$$h_r = hM_r \tag{7}$$

$$t_r = tM_r \tag{8}$$

For the mapped entities h_r, t_r and relationship r, during the TransR model's training program, it is necessary to make $h_r + r \approx t_r$ as much as possible, so the scoring function is as follows:

$$f_r(h, t) = \|h_r + r - t_r\|_2^2 \tag{9}$$

The TransR model's objective function, derived from the scoring function, is exactly that of the TransH model. It will not be mentioned here as the fundamental principle is that the score difference between positive and negative samples is more than γ to minimise the objective function.

2.2 Overview of personalised recommendation system

The personalised recommendation system aims to provide users with customised content or services based on their historical behaviour, interests, needs, and other relevant information. In the field of education, personalised recommendation systems recommend courses, knowledge points, or routes of instruction fit for their educational requirements by means of analysis of their learning behaviour, learning preferences, and other related factors. In order to achieve personalised recommendation, researchers have proposed various recommendation algorithms, among which the most typical are collaborative filtering algorithm and knowledge graph-based recommendation algorithm.

2.2.1 Collaborative filtering

One of the most often used recommendation systems, collaborative filtering generates recommendations by means of past behaviour analysis of the user and the behaviour of like users. Two important forms of CF are user-based collaborative filtering and object-based collaborative filtering.

Finding additional users (neighbours) with like habits or preferences to the target user is the basis of user-based collaborative filtering; subsequently, it recommends products that these neighbours like. The basic presumption is that user A might like additional goods that user B loves if user A and user B have similar ratings or historical behaviour toward some items.

Based on user collaborative filtering, the similarity between user u and other users v is computed and k neighbours with great similarity are chosen for suggestion based on u's recommendation. Usually, cosine similarity or Pearson correlation coefficient measures similarity; the often used equation is:

$$Sim(u, v) = \frac{\sum_{i \in I} r_{ui} r_{vi}}{\sqrt{\sum_{i \in I} r_{ui}^2} \sqrt{\sum_{i \in I} r_{vi}^2}} \tag{10}$$

Among them, r_{ui} and r_{vi} are the ratings or behaviours of user u and user v towards item i, respectively. I is a collection of objects co-owned by users u and v. Calculation of similarity helps one to derive the similarity matrix between users.

Based on user collaborative filtering, recommend items that user u has not encountered before:

$$\hat{r}_{u,i} = \frac{\sum_{v \in N_u} Sim(u, v) r_{v,i}}{\sum_{v \in N_u} Sim(u, v)} \tag{11}$$

Among them, N_u is the neighbour set of user u, and $\hat{r}_{u,i}$ is the predicted rating for user u and item i.

Analysing users' past evaluations of the products helps collaborative filtering based on items to determine item similarity. Assuming that item i and item j have resemblance, it is advised that user u loves item j if the user has comparable rating behaviour for both items. Usually, cosine similarity or Pearson correlation coefficient finds the similarity between objects. The ratio is:

$$Sim(i, j) = \frac{\sum_{u \in U} r_{ui} r_{uj}}{\sqrt{\sum_{u \in U} r_{ui}^2} \sqrt{\sum_{u \in U} r_{uj}^2}} \tag{12}$$

where r_{ui} is the rating of item i by user u, and U is the set of all users. Collaborative filtering based on items recommends item j to user u:

$$\hat{r}_{u,j} = \frac{\sum_{i \in N_u} Sim(i, j) r_{u,i}}{\sum_{i \in N_u} Sim(i, j)} \tag{13}$$

where N_u is the set of items that user u has rated, and $\hat{r}_{u,j}$ is the predicted rating for user u's item j.

Even though personalised recommendation systems heavily rely on collaborative filtering techniques, they also have some issues, such as cold start problems (lack of sufficient data for new users or items) and sparsity problems (sparse user rating matrices).

2.2.2 Recommendation algorithm based on knowledge graph.

Knowledge graphs have emerged as a significant development path in the field of personalised recommendation in recent years since they provide basis for knowledge-based recommendation algorithms. Unlike traditional collaborative filtering algorithms, knowledge graph-based recommendation methods not only consider users' historical behaviour and similarity between items, but also fully utilise the semantic relationships between entities in the graph, which can more accurately capture users' needs and interests. Two main categories define knowledge graph-based recommendation systems: graph embedding-based recommendations and graph inference-based recommendations.

Semantic information in the knowledge graph can be used by the knowledge graph recommendation algorithm grounded on graph embedding to enhance the representation of users or objects. This paper presents several knowledge graph embedding models, which can be used to embed knowledge graphs from high-dimensional, sparse graph form into low dimensional, dense vector form. Thus obtaining semantic vector representations of users, recommended items, and the relationships between items. Then, using similarity measurement methods, calculate similarity based on semantic vector representation, and using rate prediction, suggest things users might find interesting.

The recommendation algorithm based on graph embedding can be combined with other traditional recommendation algorithms to optimise the feature modelling of users and items with its rich semantic information, in order to achieve higher recommendation accuracy. Nevertheless, the graph embedding technique misses the direct high-dimensional interactions between items in the graph since it embeds entities and relationships in the graph as low dimensional vectors.

The method based on graph reasoning utilises semantic relationships in knowledge graphs to automatically derive new knowledge points or recommendation paths. For instance, inferring the user's possible points of interest in the knowledge graph or the next learning topic the user might be interested in depending on the knowledge points they presently possess.

Assuming that the knowledge graph contains a 'subset' relationship between knowledge points, if a user masters a certain knowledge point A, it can be inferred that the user may also master a subset B of A (i.e., the dependency relationship between knowledge points). This kind of reasoning allows the recommendation system to dynamically suggest the next knowledge points to be acquired depending on the learning development of the user and the knowledge points already acquired.

In this method, the inference process can be achieved through deep learning models such as graph convolutional networks (GCN), which utilise the adjacency relationships of graph structures for node representation learning.

3 The construction method of foreign language subject knowledge graph and its application in personalised learning path recommendation

The main components of this work are the building of a foreign language subject knowledge graph and its implementation in a personal learning path suggestion. Foreign language learning is a complex task that involves multiple levels of knowledge structures, such as vocabulary, grammar, listening, speaking, etc. Therefore, it is crucial to construct an efficient and accurate knowledge graph of foreign language subjects. This graph not only helps to systematically organise and manage foreign language knowledge, but also provides support for personalised learning path recommendations. The construction approach of foreign language subject knowledge graph will be thoroughly introduced in the following together with discussion of its use in recommendations for tailored learning paths.

3.1 Overall framework

This paper proposes the DKG-PFLR personalised learning path recommendation algorithm by merging knowledge graph, cognitive diagnosis, and collaborative filtering in order to solve the problems of ignoring the sequential relationship, semantic similarity relationship, and rating similarity relationship between knowledge points in current learning path recommendation methods, which cannot form a coherent learning path, and ignoring learners' personalised weak knowledge points, which cannot help learners identify and improve learning efficiency.

Figure 1 DKG-PFLR algorithm framework (see online version for colours)

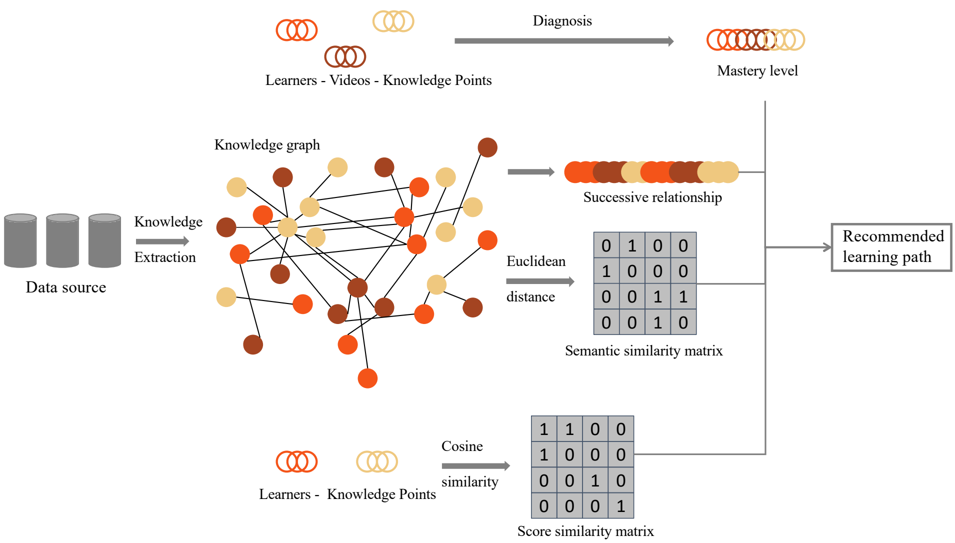


Figure 1 shows the general framework of the method. First, tailored weak knowledge points of learners are found and the learner's mastery level of knowledge points is diagnosed based on learner video knowledge point related data depending on an improved cognitive diagnostic model. Then, thorough consideration of the semantic information in the knowledge graph and the user's historical learning records was taken into account to get more accurate similarity information of knowledge points. Based on the embedding findings of the knowledge graph, the semantic similarity of knowledge points was computed; conversely, the collaborative filtering technique computed the score similarity of knowledge points. The two similarities were fused and normalised to provide the last similarity between knowledge points. This article derives the knowledge point order based on the knowledge graph and recommends personalised learning paths to learners by combining the knowledge point order, knowledge point mastery level, and knowledge point similarity since many times learners are unable to grasp the current knowledge points because they have not mastered the basic knowledge points that follow the current knowledge point level. This approach mostly targets learners' personalised weak knowledge points, recommends learning paths to learners in the sequence of knowledge points, helps learners identify and make up for personalised weak knowledge points, and increases learning efficiency based on consideration of learners' interests.

3.2 Construction method of knowledge map of foreign language discipline

The building of knowledge map of foreign language discipline is a difficult work which entails the methodical and ordered processing of knowledge in the field of foreign language discipline. In this process, it is necessary to form a knowledge map that can effectively support personalised learning path recommendation through multiple steps such as accurate knowledge point extraction, entity definition, relationship mining and map optimisation. This map not only provides a comprehensive knowledge framework for foreign language teaching, but also provides accurate guidance for learners' learning path.

First of all, the construction of knowledge map of foreign language discipline needs to start from the extraction of knowledge points. Foreign language learning involves many aspects, such as vocabulary, grammar, listening, speaking, reading and writing skills. The extraction of knowledge points is the basis for the construction of the whole map, and these knowledge points include not only specific language units (such as words, phrases, sentence patterns, tenses, grammar rules, etc.), but also some more abstract concepts (such as language learning strategies, listening skills, oral expression methods, etc.). In order to achieve full coverage of knowledge points, knowledge points extraction usually depends on a variety of methods, including textbook analysis, expert experience and automated text analysis. By analysing the course contents of foreign language textbooks, reference books and online learning platforms, we can identify the systematic knowledge framework and teaching objectives, and then extract various knowledge points.

After knowledge points are extracted, the entities in the map need to be defined next. Entities in the knowledge map represent important learning elements in foreign language disciplines. These entities include not only nominal concepts in the traditional sense (such as 'English tense', 'word pronunciation', etc.), but also entities related to learners, such as students' learning status, learning progress, learning style, etc. In order to ensure the efficient use of knowledge mapping, these entities should not only cover all aspects of foreign language disciplines, but also be closely related to learners' learning process and personalised needs. For example, 'student a', as an entity, may have a mastery relationship with 'English verb tense', indicating that the student has mastered this grammatical knowledge. In this way, learners' personal information and knowledge points of foreign language disciplines are organically combined to form a multi-dimensional knowledge map.

Knowledge mapping is fundamentally based on the definition of the relationships among entities. The knowledge map of foreign language education mostly consists on entity relationships, which directly influences the operability and usefulness of the map. Semantic links, dependencies, and pre-repair ties between items can all find expression in relationships. For instance, knowledge points frequently depend on one another; 'English verb tense' depends on 'verb form'. When students grasp a particular knowledge point, for instance, they can have a 'master' relationship with it, like 'student a has mastered the English tense'. Furthermore quite significant is the interaction among knowledge points. Students must, for instance, become proficient in the ordinary present tense before mastering the present perfect tense. The formulation of these links not only clarifies the hierarchical structure among knowledge points but also offers the foundation for the recommendation for a tailored learning route.

The construction of the map is not only a simple combination of knowledge points and relationships, but also needs to optimise the map. The purpose of optimisation is to remove redundancy, correct errors and improve the integrity of the map. Because in the construction process, there may be repeated knowledge points or some inaccurate relationships, it is necessary to correct the map through expert approval, algorithm optimisation and other methods. For example, some knowledge points may be expressed by multiple teachers or textbooks in different ways, and it is necessary to ensure the simplicity and accuracy of the atlas by de recalculation. In addition, knowledge mapping is a dynamic system. With the development of foreign language knowledge and the change of learners' needs, new knowledge points and relationships will continue to emerge. Therefore, it is also very important to update and expand the knowledge map. Through the continuous revision of automated tools and experts, the map can continue to develop and maintain its effectiveness and timeliness.

In the construction process, the representation of knowledge map is usually in the form of graph database. Graph database can effectively store and manage complex entity relationship data, and provide flexible query ability. Common graph databases, such as neo4j, can support efficient storage, query and update of nodes and relationships in the graph. This structured storage method not only facilitates the management of knowledge, but also enables the atlas to support rapid reasoning and query, providing technical support for personalised learning path recommendation system.

The construction of foreign language subject knowledge map is a multi-level and multi-dimensional process. We combine the domain knowledge, educational objectives and students' needs, and ultimately form a comprehensive, accurate and dynamic subject knowledge map through accurate knowledge extraction, reasonable entity definition, detailed relationship construction and continuous map optimisation. This map cannot only help foreign language teaching provide a systematic knowledge framework, but also play an important role in personalised learning path recommendation, helping learners obtain the most suitable learning content according to their own learning progress and needs, thus as to increase the learning experience and effect.

3.3 Personalised learning path recommendation algorithm

The foundation of personalised learning path recommendation algorithm is individual traits of the learners, learning status and needs, and intelligently recommends learning content that conforms to their learning progress and interests. Its core goal is to customise personalised learning path for each learner, optimise learning effect and improve learning efficiency by analysing learners' behaviour data, mastery, interest preferences and other factors. The implementation of personalised learning path recommendation algorithm involves multiple levels of technologies and methods. The most critical is how to use the combination of subject knowledge map, learner model and recommendation algorithm to form personalised learning recommendation through effective reasoning and prediction mechanism.

Let us now recommend a learning path for learner P_u course C_n. First initialise the recommended path *path_all*; then, according to the constructed foreign language knowledge map, the set of knowledge points of course C_n is deduced, and the score prediction of knowledge points is calculated according to the similarity of knowledge points. The method is as follows:

$$SP_{ui} = \frac{\sum_{j \in N_u} Sim(i, j) \times C_{uj}}{\sum_{j \in N_u} Sim(i, j)} \tag{14}$$

where N_u represents the knowledge points contained in all the courses that learner P_u has learned (having learned a course does not mean having completed all the knowledge points in the course), $Sim(i, j)$ represents the similarity between knowledge points S_i and S_j, C_{uj} represents whether learner P_u has learned a knowledge point. Higher knowledge point scores indicate more student interest in learning the knowledge point and more likelihood of later understanding the information point.

Considering the mastery degree of knowledge points obtained from the score prediction and cognitive diagnosis model, the initial recommendation set is calculated, so that the lower the mastery degree of learner P_u, the more interested knowledge points are ranked higher in the set; then take out the knowledge points from the set in turn, after choosing the 'most similar + weakest' previous knowledge point of the past, choose the 'most similar + weakest' prior knowledge point of the current knowledge points... And so on, until there are no prior knowledge points or the maximum number of prior knowledge points of a single knowledge point is reached, the learning path of the current knowledge point is formed, and the path is added to *path_all*; if the recommended *path_all* reaches the maximum length, the resources for learning connected to every knowledge point in *path_all* are obtained from the database, including knowledge point explanation, knowledge point related videos and slides, and returned to learner P_u together with the knowledge point *path_all*, which is the learning path of course C_n recommended for the learner.

4 Experimental verification and analysis

This part confirms that the DKG-PFLR personalised learning route recommendation algorithm is more in accordance with learners' learning behaviours from shallow to deep and can assist learners enhance learning efficiency. First, the experimental data – including the knowledge map in the fields of MySQL database and computer science – are presented. Three experiments are then set up, and the corresponding experimental techniques, evaluation indices, experimental findings, and analysis of each three experiment is introduced separately.

4.1 Experimental data

In the experimental stage, the knowledge map MySQL database created by the mooccube dataset forms the data used in this paper (Liu et al., 2024). The knowledge map mainly contains semantic information. The data volume of entities and entity related attributes in the knowledge map is shown in Table 1.

Based on the above entities, the knowledge map also constructs a large number of relationships between entities, and Table 2 shows the data volume.

Table 1 Amount of entity and related attributes data

Entity name	*Entity attribute*	*Number*
Concept	Knowledge point name, knowledge point explanation	3,200
Course	Course name, advance course, and course brief introduction	100
Video	Video name	9,000
School	School name and school brief introduction	100
Teacher	Teacher's name and teacher's brief introduction	350
Field	Domain name	5

Table 2 Relationship data

Relation	*Inverse relation*	*Related entities*	*Number*
FieldHasCourse	IsCourseOfField	Domain-course	300
CoursehasConcept	IsConceptOfCourse	Curriculum-knowledge points	40,000
SchoolHasCourse	IsCourseOfSchool	School-course	300
FieldhasConcept	IsConceptOfField	Domain-knowledge points	7,000
TeacherhasCourse	IsCourseOfTeacher	Teacher-course	800
SchoolhasTeacher	IsTeacherOfSchool	School-teacher	650

MySQL database mainly contains information about learners. It includes 15,547 learners in total, 84,213 records of learners' course selection, 11,842,711 records of learners' learning of knowledge points, and 2,226,418 records of learners' watching videos.

4.2 Experimental procedure

The first part of the experiment is to determine the fusion factor, which directly affects the quality of the final knowledge point similarity. We design experiments so that when the values of the fusion factor are 0, 0.1, 0.2 … and 1.0 accordingly, we simply provide suggestions depending on the similarity of knowledge points; so, the chance that the suggested outcomes satisfy the user's need is acquired. When this probability reaches the optimal value, it is the final fusion factor.

The indicators to measure whether the recommendation results meet the user interest include the user interest rate P_1 and the user preference coverage rate P_2. Among them, the rate of conforming to user interest is the probability that the recommendation result conforms to user interest (recommending knowledge points that users like, or not recommending knowledge points that users do not like); user preference coverage is the proportion of knowledge points recommended by algorithm among all knowledge points of interest to users. Table 3 shows the calculation parameters of P_1 and P_2.

Table 3 The calculation parameters of user interest rate P_1 and user preference coverage P_2

Real results	*User interested*	*User not interested*
Actual recommendation	TP	FP
Not recommended	FN	TN

The calculation method of user interest rate P_1 is as follows:

$$P_1 = \frac{TP + TN}{TP + FN + FP + TN} \tag{15}$$

The user preference coverage P_2 is calculated as follows:

$$P_2 = \frac{TP}{TP + FN} \tag{16}$$

The second section of the experiment is a comparison one arranged according to knowledge point path. The computation equation of the knowledge point order of the path is as follows assuming a recommendation algorithm suggests a learning path for the course of learners:

$$\text{Order degree} = \frac{\sum_{i=1}^{n-1} f(a_i, a_{i+1})}{n-1} \tag{17}$$

where $f(a_i, a_{i+1})$ indicates whether there is a succession relationship between two adjacent knowledge points in the path. The greater the order of knowledge points, the more consistent with learners' daily learning habits, that is, to learn the basic knowledge points first, and then to learn the subsequent knowledge points from shallow to deep.

A comparative study of learning efficiency forms the third component of the experiment. This work evaluates the learning efficiency by means of a comparison between the paths suggested by learners across several algorithms and the enhancement of the mastery of knowledge points in the same learning period. Trans CF recommendation algorithm (Zhou et al., 2021) and Ontology-CF recommendation method (Agarwal et al., 2022) are the two comparison studies of the later two experiments.

4.3 Experimental results and analysis

The fusion factors were 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0. Because the value of fusion factor directly affects the quality of the final knowledge point similarity, and then affects whether the recommendation results meet user preferences. Thus, under each value of the fusion factor, this study successively forecasts the score based on the similarity of knowledge points and provides the recommendation results; counts the recommended user interest rate P_1 and user preference coverage rate P_2. The value of the fusion factor makes the horizontal axis; the curve shows the user preference coverage rate P_2 and the user interest rate P_1 accordingly.

It can be concluded that the overall trend of the curves in line with user interest rate P_1 and user preference coverage rate P_2 is basically similar, both rising first and then falling. When $\alpha = 0$, the similarity of knowledge points just addresses the user score similarity, which is the similarity calculation method of traditional collaborative filtering algorithm. At this time, the values of P_1 and P_2 are not ideal. When the value of α gradually increases, the proportion of semantic similarity in knowledge point similarity gradually increases, and P_1 and P_2 show a tortuous upward trend. Until $\alpha = 0.7$ o'clock, the two indicators reached the maximum. Then continue to increase the value of α, and the values of P_1 and P_2 decrease rapidly. When $\alpha = 1.0$, the similarity of knowledge points only considers the semantic similarity, which is the traditional similarity calculation method based on the knowledge map recommendation algorithm. At this time, P_1 and P_2 both decline. To sum up, only considering the score similarity or semantic similarity cannot get a good recommendation result. It is necessary to integrate the two. When the fusion factor is 0.7, the most accurate similarity of knowledge points can be obtained, and the optimal user interest rate P_1 and user preference coverage rate P_2 can be achieved. Therefore, in the following two comparative experiments, this work presents a 0.7 fusion factor of DKG-PFLR learning path recommendation method.

Figure 2 Meet the user coverage rate P_1 (see online version for colours)

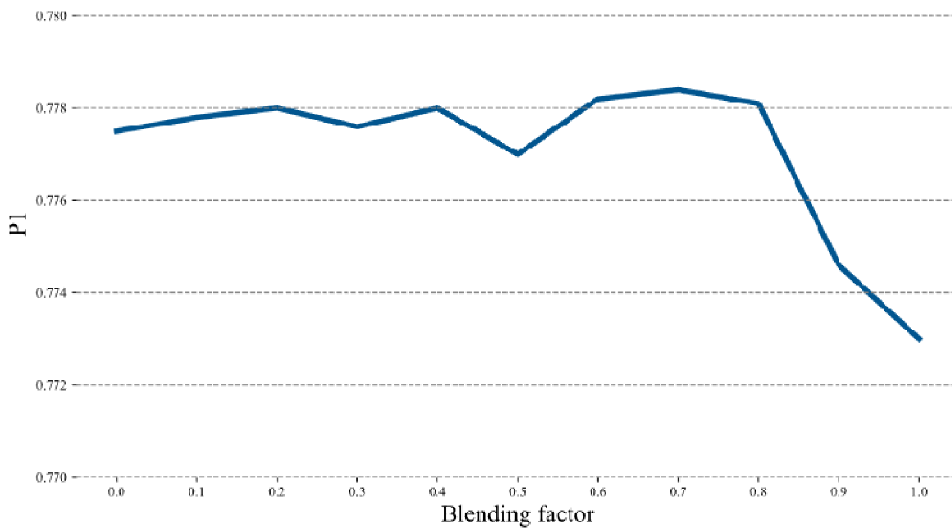


Figure 3 User preference coverage rate P_2 (see online version for colours)

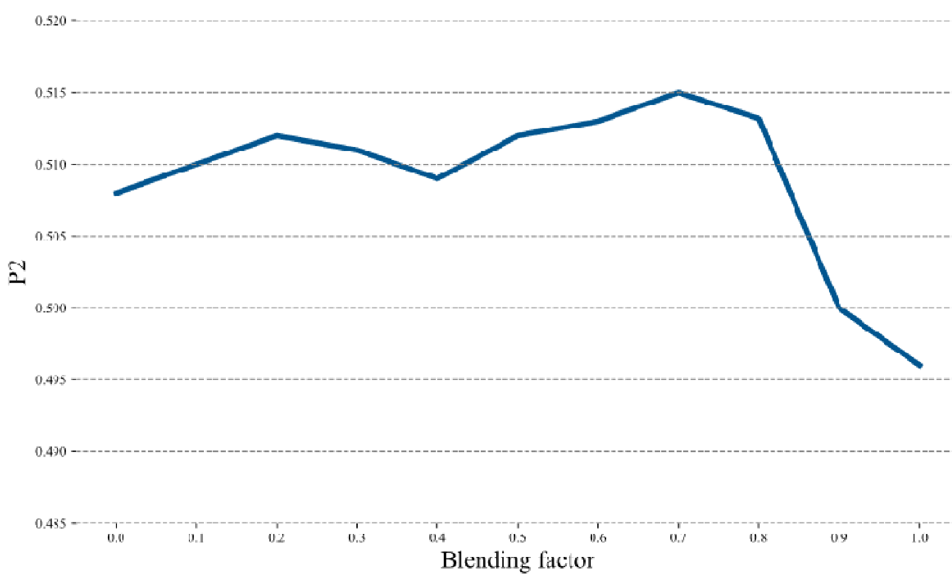
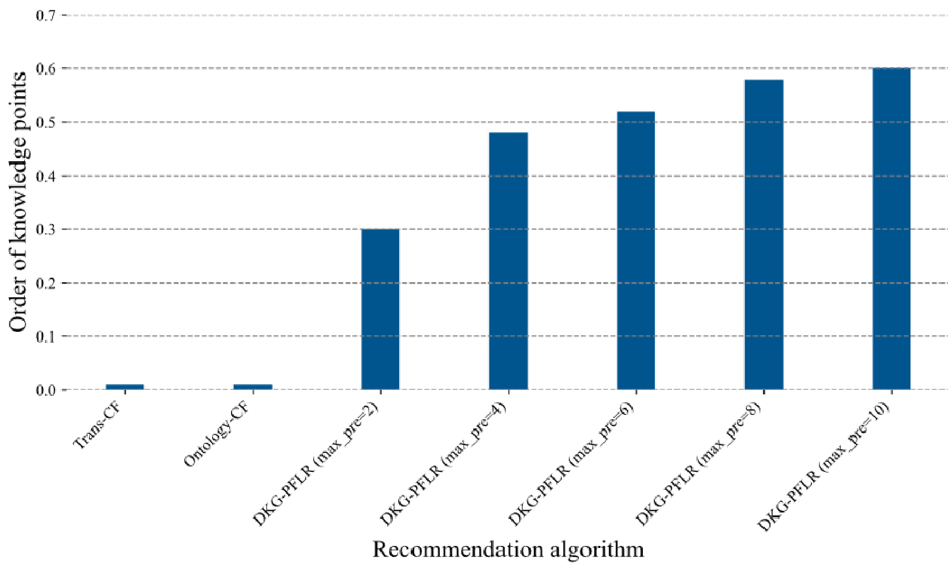


Figure 4 Comparative experiment of knowledge point order (see online version for colours)



In this work, the maximum number of prior knowledge points in the learning path max_pre is chosen as a crucial parameter influencing the performance of DKG-PFLR algorithm in the comparative experiment of the order of knowledge points in the path. While suggesting the path of knowledge points in the first set of knowledge points to be recommended, max_pre is the maximum number of prior knowledge points to query. Since the succession of knowledge points is not considered in the Trans-CF algorithm and Ontology-CF algorithm, the value of max_pre will not affect their recommendation results. When max_pre is 2, 4, 6, 8 and 10, the order degree of knowledge points of the recommended path of DKG-PFLR algorithm in this paper is calculated, and compared with the order degree of knowledge points of the recommended path of Trans-CF algorithm and Ontology-CF algorithm.

Figure 4 displays the experimental findings; the knowledge point order index of every suggested algorithm is displayed by the blue column chart. The analytical results reveal that:

1 The learning path advised by Trans-CF algorithm and Ontology-CF method approaches zero in the order index, which reflects the lack of logic and systematicness of the sequence of knowledge points generated, which is not conducive to the construction of a complete knowledge system for learners. In contrast, the DKG-PFLR algorithm proposed in this paper shows significantly higher order values under different max_pre parameter settings, indicating that it can effectively improve the systematisation, explicability and learners' satisfaction of the learning path.

2 The order index of DKG-PFLR algorithm is positively correlated with the maximum number of prior knowledge points max_pre. When the value of max_pre is too small, the logical association between knowledge points in the learning path is weak, which may lead to learners' subsequent learning without fully mastering the prerequisite

knowledge, affecting the construction of knowledge system; however, when the value of max_pre is too large, learners may over concentrate on the prerequisite knowledge learning of a single knowledge point and reduce the overall learning efficiency.

The above analysis guides this work to set the max_pre value to 4, which not only ensures that learners fully grasp the prerequisite knowledge of the current weak knowledge points, but also takes into account the learning progress of other knowledge points, and achieves the balance between learning efficiency and knowledge mastery.

This study is based on the learners' learning records of knowledge points of a course, which are arranged in chronological order, and K is used to indicate their present learning development in the comparative experiment of learning efficiency. This paper makes a comparison between the average mastery degree of the course knowledge points obtained in the same follow-up learning time after using the DKG-PFLR algorithm, Trans-CF algorithm, Ontology-CF algorithm and the learning path selected by learners, when there is no follow-up learning.

Figure 5 Comparative experiment of learning efficiency (see online version for colours)

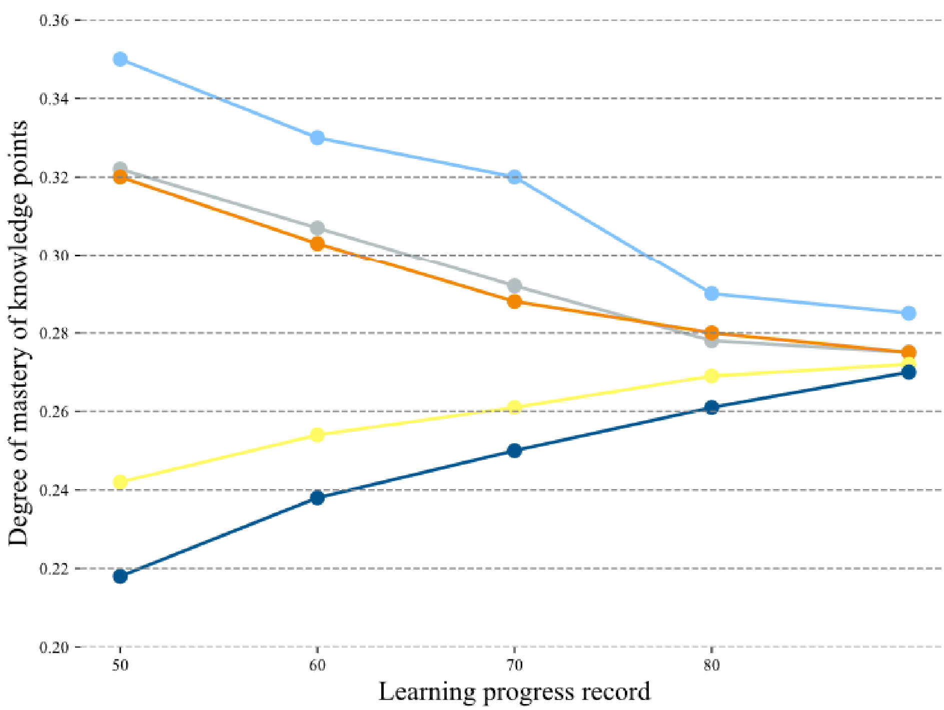


Figure 5 displays the average of all course recommendations in the dataset, where the vertical axis denotes the mastery of knowledge points under several learning routes and the horizontal axis reflects the value of the current learning progress *K*. It is clear from comparing the dark blue dotted line (without follow-up learning) with the other four curves (different recommended paths) that the four recommended paths can increase the mastery of knowledge points; but, the improvement effect progressively weakens as *K* value increases – that is, the remaining learning time reduces. Further comparison of the light blue (DKG-PFLR), orange (Trans-CF), grey (Ontology-CF) curve and yellow curve (learner autonomy path) reveals that the DKG-PFLR algorithm, Trans-CF algorithm and Ontology-CF algorithm suggested in this paper are better than learners' autonomous learning path, so indicating that these three recommended algorithms can effectively improve learning efficiency. Among them, the DKG-PFLR method in this work has the most notable enhancement effect; Trans-CF algorithm performs essentially the same as Ontology-CF algorithm. This result reveals that the DKG-PFLR algorithm suggested in this work can greatly increase learning efficiency and optimise learners' mastery of knowledge points in the same period.

5 Conclusions

Based on the knowledge map of foreign language disciplines, this study suggests a personalised learning path recommendation algorithm and demonstrates the efficiency of the method in enhancing learners' learning effect with pragmatic experiments. First, we create the knowledge map of foreign language discipline, methodically arrange the link between knowledge areas, and suggest, via the requirements and dependencies in the map, individualised learning routes for students. We then developed a fusion recommendation system using several learning considerations. Learners' individual weak knowledge points and learning interests let the algorithm fully consider the succession and similarity of knowledge points, increase the accuracy, consistency and pertinence of learning path recommendation, and could help learners make up for knowledge gaps, so improving learning efficiency, learning experience and user satisfaction.

Three tests are intended in this work to confirm the efficiency of the suggested approach: the determination of fusion factor, the comparison of the order of knowledge points in the path and the comparison of learning efficiency. In many respects, experimental results reveal that the suggested algorithm outperforms the conventional recommendation system. First of all, the most precise similarity of knowledge points can be obtained when the fusion factor is 0.7. Second, the order degree of recommendation path is higher than the conventional recommendation method based on time order; the order degree of knowledge points reveals that the recommendation method based on subject knowledge map can better ensure the order rationality between knowledge points. At last, in the comparison experiment of learning efficiency, our suggested approach's learning efficiency score is much higher than that of the comparative technique, thereby indicating that personalised learning path recommendation may greatly raise the learning efficiency of learners.

This study has made some successes, but further research and direction of improvement are still needed. From the following points, further studies can be deepened and enlarged.

Based on the stationary knowledge point structure, the subject knowledge map developed in this work. With the development of the subject and the update of the learning content, the content of the knowledge map needs to be continuously expanded and updated. In the future, we can consider introducing an automatic update mechanism to dynamically adjust and optimise the knowledge map through the update of learners' behaviour data and course content.

Although we make personalised recommendations through historical behaviour, mastery of the situation and interest preferences, these factors cannot fully cover all the needs of learners. In the future, more learning data, such as learners' emotional state and learning habits, can be combined to further improve the personalised degree of recommendation algorithm and create a more accurate personalised learning path.

Declarations

All authors declare that they have no conflicts of interest.

References

Abdelrahman, G., Wang, Q. and Nunes, B. (2023) 'Knowledge tracing: a survey', *ACM Computing Surveys*, Vol. 55, No. 11, pp.1–37.

Agarwal, A., Mishra, D.S. and Kolekar, S.V. (2022) 'Knowledge-based recommendation system using semantic web rules based on learning styles for MOOCs', *Cogent Engineering*, Vol. 9, No. 1, p.2022568.

Andreyeva, Y.A., Fakhrutdinova, A.V., Korneva, I.G. et al. (2019) 'The effectiveness of the use of computer technology compared to the traditional methods in the process of a foreign language teaching', *Universal Journal of Educational Research*, Vol. 7, No. 10, pp.21–26.

Burke, R., Felfernig, A. and Göker, M.H. (2011) 'Recommender systems: an overview', *AI Magazine*, Vol. 32, No. 3, pp.13–18.

Cheng, H. and Di, W. (2024) 'Construction of personalized learning content recommendation system based on recommendation algorithm in English learning', *Journal of Electrical Systems*, Vol. 20, No. 7s, pp.3047–3060.

Ebisu, T. and Ichise, R. (2019) 'Generalized translation-based embedding of knowledge graph', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 32, No. 5, pp.941–951.

Golonka, E.M., Bowles, A.R., Frank, V.M. et al. (2014) 'Technologies for foreign language learning: a review of technology types and their effectiveness', *Computer Assisted Language Learning*, Vol. 27, No. 1, pp.70–105.

Gong, S. (2010) 'A collaborative filtering recommendation algorithm based on user clustering and item clustering', *Journal of Software*, Vol. 5, No. 7, pp.745–752.

Hao, X., Ji, Z., Li, X. et al. (2021) 'Construction and application of a knowledge graph', *Remote Sensing*, Vol. 13, No. 13, p.2511.

Hidalgo-Paniagua, A., Vega-Rodríguez, M.A., Ferruz, J. et al. (2017) 'Solving the multi-objective path planning problem in mobile robotics with a firefly-based approach', *Soft Computing*, Vol. 21, pp.949–964.

Huang, Y. and Zhu, J. (2021) 'A personalized English learning material recommendation system based on knowledge graph', *International Journal of Emerging Technologies in Learning (Online)*, Vol. 16, No. 11, p.160.

Le, T., Le, N. and Le, B. (2023) 'Knowledge graph embedding by relational rotation and complex convolution for link prediction', *Expert Systems with Applications*, Vol. 214, p.119122.

Lin, Y., Lin, F., Zeng, W. et al. (2022) 'Hierarchical reinforcement learning with dynamic recurrent mechanism for course recommendation', *Knowledge-Based Systems*, Vol. 244, p.108546.

Liu, C., Zhang, H., Zhang, J. et al. (2023) 'Design of a learning path recommendation system based on a knowledge graph', *International Journal of Information and Communication Technology Education (IJICTE)*, Vol. 19, No. 1, pp.1–18.

Liu, H., Chen, X. and Zhao, F. (2024) 'Learning behavior feature fused deep learning network model for MOOC dropout prediction', *Education and Information Technologies*, Vol. 29, No. 3, pp.3257–3278.

Nabizadeh, A.H., Leal, J.P., Rafsanjani, H.N. et al. (2020) 'Learning path personalization and recommendation methods: a survey of the state-of-the-art', *Expert Systems with Applications*, Vol. 159, p.113596.

Park, S.Y. (2009) 'An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning', *Journal of Educational Technology & Society*, Vol. 12, No. 3, pp.150–162.

Pooja and Bhalla, R. (2022) 'A review paper on the role of sentiment analysis in quality education', *SN Computer Science*, Vol. 3, No. 6, p.469.

Qu, L. and Li, H. (2022) 'Analysis of distribution path optimization algorithm based on big data technology', *Journal of King Saud University-Science*, Vol. 34, No. 5, p.102019.

Wang, Q., Mao, Z., Wang, B. et al. (2017) 'Knowledge graph embedding: a survey of approaches and applications', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 29, No. 12, pp.2724–2743.

Wang, X., Chen, G., Qian, G. et al. (2023) 'Large-scale multi-modal pre-trained models: a comprehensive survey', *Machine Intelligence Research*, Vol. 20, No. 4, pp.447–482.

Zhang, X., Liu, S. and Wang, H. (2023) 'Personalized learning path recommendation for e-learning based on knowledge graph and graph convolutional network', *International Journal of Software Engineering and Knowledge Engineering*, Vol. 33, No. 1, pp.109–131.

Zhou, K., Yang, C., Li, L. et al. (2021) 'A folksonomy-based collaborative filtering method for crowdsourcing knowledge-sharing communities', *Kybernetes*, Vol. 52, No. 1, pp.328–343.