Research Article

Personalized Learning Path Recommendation Based on Weak Concept Mining

Xiuli Diao,1 Qingtian Zeng,1 Lei Li,1 Hua Duan,2 Hua Zhao,1 and Zhengguo Song1

1College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao, Shandong, China
2College of Mathematics and Systems, Shandong University of Science and Technology, Qingdao, China

Correspondence should be addressed to Qingtian Zeng; qtzeng@163.com

Received 23 November 2021; Revised 24 December 2021; Accepted 15 March 2022; Published 14 May 2022

Copyright © 2022 Xiuli Diao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Discovering valuable learning path patterns from learner online learning data can provide follow-up learners with effective learning path reference and improve their learning experience and learning effects. In this paper, a personalized learning path recommendation method based on weak concept mining is proposed. Firstly, according to the degree of concept mastery of historical learners, concept maps of different types of learners are generated by clustering and association rule mining algorithms. A set of weak concept learning paths are then automatically generated through topological sorting algorithm. Secondly, the long short-term memory neural network based on the attention mechanism (LSTM+attention) is trained to predict the learning effect of the weak concept learning path. Finally, the personalized weak concept path that meets the expected learning effect is selected from the path prediction results. In the experiment, the proposed method is not only compared with the traditional recommendation method, but also, a comparative experiment on the impact of different learning effect prediction models is carried out. The experiment results show that our proposed method has obvious advantages in recommendation performance.

1. Introduction

At present, due to the advantages of time and space flexibility, as well as convenience and timeliness, many learners have met their learning needs through e-learning. At the same time, the e-learning platforms have also accumulated a big data from the learner’s learning process. Learning path rules of history learners and implicit learning path patterns are discovered in the learning process of different individuals or groups. These may include the sequence of concepts and resources that the learners access. It can provide effective learning path reference for follow-up learners [1] and improve the learning experience as well as learning effect of online learners [2, 3], thus realizing personalized and accurate e-learning [4] which has become one of the research hotspots of personalized learning service.

The answer record is one of the important learning behavior sets generated in the learner’s e-learning process. Through the historical answer records, the learner’s concept mastery can be obtained, and the knowledge loopholes in the learning process are accurately found. For example, a learner’s historical answer records may show that the exercise error rate of a certain concept is higher, meaning that the learner has a weak grasp of that concept. Due to the inseparable relationship between concepts, it is necessary to plan the follow-up learning path to help the learner to completely grasp the weak concepts [5]. However, it is difficult for learners to understand their own mastery of concepts in the process of doing exercises in fact [6], and they may be also unable to review and consolidate relevant and targeted concepts based on the results of the exercises that have been done. Therefore, from the learner’s answer records, relationship between concepts can be automatically discovered, and the weak concepts are diagnosed to provide targeted learning path guidance in time [7]. It can also help learners fill up knowledge gaps and complete the learning goal of proficiency in weak concepts as soon as possible while providing follow-up learners with an effective learning experience of weak concepts. It can be seen that the research faces two important challenges. One is how to automatically
mine the weak concepts according to the learner’s answer records and discover the relationship between the concepts in order to lay the foundation for the recommendation path. The second is how to ensure the expected learning effect while recommending a personalized weak concept learning path for the target learner.

From the above analysis, our research involves two stages of learning path discovery and recommendation based on learning process data. We discover the relationship between concepts judging by the historical learners’ answer records to construct concept maps and then use the topological sorting algorithm to generate learning paths for different learners. Finally, we recommend a personalized weak concept learning path that meets the learning needs of the target learner. Currently, most of the existing related work pay attention to the rules of the existing learning path data and then find the path with cooccurrence, sequence, and other group characteristics in the learning path group. As an example, in [8], lag sequence analysis is used to discover the knowledge construction behavior pattern of college students’ collaborative translation activities. As another example, the course sequence and grades visited by learners are obtained from historical learning process data, and then, similar users are found for a new user. A variable-length genetic algorithm can then be used to recommend personalized learning paths for new users in [2]. At present, the research of learning path recommendation based on learning path discovery has mostly recommended personalized and highly similar optimal learning paths for target learners [2, 3, 9]. In [10], considering the different learning path preferences of learners in different learning situations, personalized learning paths were recommended such as the shortest learning path and key learning path. However, none of the above works involves mining the weak concepts of learners in the learning process. In addition, most of them do not fully consider the final learning effect after the learners accept the recommended path, resulting in low recommendation accuracy (less than 50%). Furthermore, since learners have different knowledge base, the optimal learning path recommended in existing research may not be suitable for every learner. For instance, if a learner’s usual test score is 60 points. Rapidly improving to 90 points in a short period of time before the exam may be burdensome. Thus, it is of great practical significance to let a teacher or a learner set the expected learning effect and provide them with personalized path guidance.

In view of the shortcomings of existing related work, our research first uses association rules to mine concept map from historical learner’s answer records in the path discovery stage [11]. Concept map reflects the association relationship between concepts, mainly the prerequisite relationship. Since the concept map is a directed acyclic graph, topological sorting can be used to generate the concept sequence, that is, the learning path. Combined with the learner’s concept mastery, the weak concept learning path can be diagnosed [7]. The follow-up recommendation phase researches for the weak concept path recommendation which not only meets the expected learning effect of the target learner but also aims to help the target learner master the weak concept. Specifically, we first discover the full learning paths of different learner groups based on historical learner answer records. The weak concepts of the learners are then diagnosed according to their concept mastery, and the association relationships between weak concepts generate the learning path of weak concepts. Finally, the LSTM+attention neural network is used to predict the learning path and learning effect. A personalized weak concept learning path is then recommended for the target learner. This learning path should meet the learner’s personalized learning need and realize personalized and accurate learning support services at the same time.

Furthermore, in order to meet the challenge that it is difficult to ensure the quality of learning effect in personalized learning recommendation, literature [3] uses LSTM neural network to predict learning effect before recommending the optimal full learning path which is at the level of learning resource. However, our research objective is different from [3]. We recommend the learning path which is based on the concept level and does not involve the recommendation of subsequent specific learning resources. The purpose is to recommend a weak concept learning path that meets the expected learning effect, not necessarily the optimal learning path with the best learning effect. To this end, we have improved the method. We are based on the LSTM neural network and introduce the attention mechanism to distinguish the contribution of different weak concepts and key concepts in the learning path to the learning effect prediction, so as to effectively improve the accuracy of the learning effect prediction.

In summary, it is different from the existing research. We employ a LSTM+attention neural network model to predict the learning effect to build a new learning path recommendation model in an e-learning environment. The technical contributions of our research include (1) a new framework for a personalized weak concept learning path recommendation model and (2) LSTM+attention trained based on the learner’s concept mastery and the weak concept learning paths to predict the learning path and learning effect.

The rest of this paper is organized as follows. Section 2 introduces the related work of learning path recommendation, concept map, and weak concept diagnosis. Section 3 defines the research question. Section 4 introduces the learning path recommendation model framework and the specific research content of each module in detail. Section 5 introduces the experimental process and verifies the recommended quality, while Section 6 discusses the research content and looks forward to further research work that can be carried out in the future.

2. Related Works

This section mainly reviews the related works of learning path recommendation, concept map, and weak concept diagnosis, as follows.

2.1. Learning Path Recommendation. With the application of recommendation technology in the field of e-learning, the development of the research field of personalized learning
recommendation has been promoted. Personalized learning recommendation is a key link in achieving personalized learning service. Its objective is to provide individual learners with learning resources that meet their learning needs to enhance learning experience and effects. In the current related research, the objects of personalized learning recommendation involve learning resources such as exercises [12–14], courses [15, 16], friends [17], and learning paths [18, 19]. Among them, learning path recommendation is one of the hotspots of personalized learning recommendation research.

The purpose of learning path recommendation is to help individual learners achieve their learning goals; just as the following suggest to the learners’ characteristics and needs, personalized learning path is recommended. Some learning content sequence (learning object, concept, and course) that matches their prerequisites can then be called a learning path [20, 21]. At present, some researchers have produced many different personalized learning path recommendation methods through various knowledge models, algorithms, and technologies.

Research on learning path recommendation is mostly based on the idea of constructing a knowledge model from a graph [18, 19]. The graph could be a concept map [9, 21, 22], knowledge map [10, 23], ontology [24], topic map [25], or knowledge graph [26–28]. The nodes represent the learning content, while the directed edges represent the relationship between the nodes. The objective is to generate a path that meets the user’s characteristics and requirements. Among them, the knowledge model can be constructed by the expert himself or through educational data mining technology. The concept map constructed by the expert is limited in that it mainly reflects the explicit relationship between concepts, which is highly subjective, time-consuming, and labor-intensive. In contrast with the one constructed by data mining technology which in addition to containing the explicit relationship between the concepts, saving time, and effort, it reflects the implicit relationship between the concepts [7].

In the research of learning path recommendation, parameters are generally used to determine the personalized characteristics and needs of users. These parameters also describe different learning scenarios, besides describing the learning needs of users and their different characteristics, such as learning style and knowledge background [29]. Through literature research, it has been found that the researchers consider different personalized parameters to recommend personalized learning path. The main parameters are as follows: concept mastery, knowledge background, learning goal, learning style, and learning time limit. Among them, concept mastery occurs most frequently. Time limit [30] appears least which could be attributed to the more difficult controllability of the time factor. In [21], based on the educational concept map (ECM), individual learners first selected a set of topics from the ECM to determine their knowledge background. After the user selecting the initial theme and the target theme, the expert deleted the known themes in the concept map. Finally, an algorithm was used to linearize the concept map and generate a path. In [31], a learning path composed of a sequence of learning objects was then generated according to learning goals and knowledge levels. Based on graph theory, after determining the user’s knowledge level and needs, the shortest learning path was then generated and recommended to the user in [32]. Hence, a successful learning path was recommended according to the learner’s available time and knowledge background in [33].

With the application of big data technology in the field of education, collecting a learner’s learning behavior and learning result data has increasingly become convenient. A rising numbers of researchers are paying attention to recommending learning path for target learner by considering the learning paths of historical learners to improve the accuracy of path recommendation. Data mining technology can assist new learners in choosing learning paths based on historical learners’ learning paths, such as Tutorlet agent based on hidden Markov model [34], and cross-layer frequent pattern mining based on FP-Tree in e-teaching systems [35]. These systems make use of statistics to summarize the learning paths adopted by most previous history learners and then recommended them to current new learners. However, some researchers believe that the most frequent learning paths summarized from data mining may not be suitable for target learners. Two factors should also be considered when choosing a learning path: (1) the characteristics of the history learner and (2) the learning effect of the history learner after completing the path in [36, 37].

Moreover, some existing data mining methods use clustering to combine the characteristics of historical learners. Some studies use intelligence and optimization algorithms to combine the learning effect of historical learners. These include genetic algorithm [2], ant colony algorithm (ACO) [38], and neural network [3] which have been used by researchers to solve learning path generation and recommendation. In [39], ACO-based learning path planning technology treated a learner as ant-like agent. It first calculated the pheromone in the entire content network and then recommended a learning path to a new student. Research by [40] used the concept familiarity of all former learners in the same cluster in the learning path discovery algorithm, together with ant colony algorithm (ACO) to generate learning paths and recommend learning paths to new learners.

In recent years, neural networks have been the most used method for learning path prediction in the field of learning path recommendation. In [3], a personalized learning full-path recommendation model was proposed based on LSTM neural network, which relied on clustering and machine learning technology. First, a group of learners was clustered based on their feature similarity measurement, and a long short-term memory (LSTM) model was trained to predict the learning path and learning effect. The personalized learning full path was then selected from the path prediction result. A method of constructing a learning path recommendation system was proposed using ability graphs. The ability graphs were used to display user’s scores. Recursive neural networks were used to construct sequence-based prediction models to predict and recommend next questions that should be learned by the user [41].
In summary, the current research on learning path recommendation has achieved heaps results, but none of them involves the recommendation of weak concept learning path. To this end, in view of the above existing research, our research is focused on historical learners' answer records and existing learning experience to recommend personalized weak concept learning path for the target learner. Among them, the knowledge model is a concept map mined by association rules, and the learner's characteristics are calculated according to the learner's concept mastery. The weak concept path of each historical learner is the experience for the target learner to learn. Before recommending learning paths for a target learner, we combine the learning characteristics of historical learners to classify target learner into corresponding categories through similarity calculations. The weak concept learning paths for similar learners are found, and then, deep neural networks are used to predict the learning effect before learning path recommendation.

2.2. Concept Map and Weak Concept Diagnosis. As an effective knowledge visualization tool, concept map is used to structurally present concepts and the dependencies between concepts by map [42]. It is widely used in the field of education. Considering the learner's concept mastery, the learner's knowledge loopholes can be accurately located based on the concept map, and the weak concepts and their associated relationships can be mined. This helps teachers provide personalized guidance for students [7, 43]. Therefore, it has become a hot research topic of constructing a concept map based on the learner's answer records and mining the weak concepts.

In [44], the Apriori algorithm was used to automatically construct the concept map from the students' answer records. Based on the generated concept map, following the students' academic performance, their weak concepts were then diagnosed. Through the establishment of a remedial teaching path, all students were given unified learning guidance. [11] also used data mining technology to automatically generate dynamic concept maps from the answer records. This method could update the concept map in time following the learner's answer records, which could match the requirements of the adaptive learning system. In order to visually present the learning situation of students, concept maps were automatically excavated to provide personalized guidance for different students in [45]. [46] used direct hashing and pruning algorithms to construct concept maps from student answer data. Guided by the constructed concept map, the weak concepts were excavated according to the test situation of the students, and the learning path of the weak concepts was generated to guide the students' follow-up learning. In [7], an association rule mining algorithm was applied to generate concept maps for different categories of learners. By the topological sorting algorithm, the learning paths of different types of learners were automatically generated. Different learning paths of weak knowledge points were generated for each learner considering the learner's conceptual mastery to provide the learner with personalized learning guidance. In [47], learners were first grouped considering the degree of concept mastery through clustering. The concept map was automatically constructed using direct hashing and pruning algorithms. Furthermore, concepts which learners had weak grasp were unearthed, and learning problems are automatically diagnosed.

The above research automatically constructs concept maps by mining the association rules between concepts based on student answer records and diagnoses the concepts that learners have weakly grasp. It can provide targeted learning guidance for individual learners or groups. Therefore, the method in [7] is used to automatically discover the relationship between concepts from historical learner's answer records in this paper. Then, a concept map is constructed, and topological sorting algorithm is used to generate learning paths for different learners. The weak knowledge points are then diagnosed according to the concept mastery of each learner. The weak concept learning path of the historical learner is generated, which lays the foundation for the target learner to recommend personalized learning path.

3. Problem Statement

The research includes two stages of weak concept learning path discovery and recommendation. Specifically, the history learner’s answer records and the matrix of the relationship between exercises and concepts can be used to calculate the learner's concept mastery in the first stage. Furthermore, the weak concepts of the learner can be diagnosed, and the implicit relationship between the weak concepts can be discovered. Finally, the weak concept learning path of each learner is obtained. The second stage of the path recommendation is to recommend the weak concept learning path of similar history learners for the target learner while ensuring the learning effect of the recommended path.

Therefore, two types of learners are involved in this research. One is the target learner who is the object of the research’s personalized learning service. The other is historical learners who have learned the same concept. Their answer records are the data set of the research, which provides a data basis for the smooth development of the research. In order to better carry out the research work, the following two basic assumptions are made.

_Hypothesis 1._ Learners with similar concept mastery follow a similar learning path when learning the same concepts, while learners with different concept mastery may have different learning paths. Furthermore, learners can be divided into numerous of different groups in accordance with the concept mastery. Different concept maps are generated for different groups of learners in this paper.

_Hypothesis 2._ Since different learners have different levels of concept mastery, the learners or teachers expect different learning outcomes. In the paper, personalized learning path is recommended considering the learning effect expected by the teacher or learner.
To complete the final purpose of the research, we split the research content into the following subcontents. These subcontents are the following:

1. Mining concept map: considering the historical learner answer records (R) and the relationship between the question and the concept (QC), the learner concept error rate (F) is calculated. As a feature, we divide many historical learners into different clusters G. The associated rules are then utilized to excavate each cluster concept map M, where the concept map is a directed acyclic graph, which embodies the prerequisite relationship between the concepts. An example of a conceptual drawing is shown in Figure 1. The circles in the figure indicate conceptual nodes, while the lines with arrows indicate the prerequisite relationships between concepts.

2. Generating weak learning path: we use the topology sort algorithm based on the concept map M to generate the full learning path (FLP) of each cluster. Then, considering the concept error rate F of each learner, the full learning path is simplified, and the weak concept learning path set (WLP) is obtained. This lays a data foundation for training the learning effect prediction model and path recommendation.

3. Training learning effect prediction model: the learning path in this research is a set sequence of concepts which is sequential. As an improved recurrent neural network, LSTM is an efficient tool for processing sequence data. Therefore, considering the historical learner's concept mastery F and the weak concept learning path set WLP that has been constructed. The LSTM+attention neural network is trained to predict the learning effect and path for a target learner.

4. Recommending a personalized weak concept learning path: for the answer records of the target learner, we initially apply the clustering algorithm both to find out the cluster to which it belongs and find the top-k similar users in the cluster through the feature similarity measure. We then take the target learner's concept mastery and the weak concept path of each similar user as input. The trained learning effect prediction model is applied to predict the learning effect in turn. Finally, the weak concept path with the most similar user and the learning effect in line with the expected is recommended to a target learner. After analyzing the learner's historical learning effect, the expected learning effect can be set by the learner himself or the teacher.

5. Evaluating the proposed method: in the experiment, the proposed method is compared with the traditional recommendation algorithm for performance evaluation. We then compare the impact of the learning effect prediction model using different model parameters and data embedding methods (such as random embedding and graph embedding) on the final learning path recommendation performance.

4. The Proposed Method

The proposed personalized learning path recommendation method in this research is composed of three components: concept map mining and weak concept learning path generation, training LSTM+attention learning effect prediction model, and personalized weak concept learning path recommendation. Among them, the personalized weak concept learning path recommendation for target learner includes three steps: learning path recall, learning effect prediction, and learning path conversion. The overall research framework of the proposed method is shown in Figure 2. The function of each component is described in detail below.

4.1. Concept Map Mining and Weak Concept Learning Path Generation

This module contains two tasks. One is to classify historical learners through clustering while mining different types of concept maps M. Topological sorting is then used to generate a full learning path $\text{flp}_k$ that includes all concepts for learners in different clusters. The other is judging by the full learning path and learner concept error rate to diagnose each historical learner’s weak concept learning path and construct the weak concept learning path set WLP.

Inspired by the literature [7], based on the historical learner’s answer records R and the relationship between the question and the concept QC, the learner’s grasp of the concept is calculated, that is, the concept error rate F. The concept error rate F is divided into three levels (high, medium, and low) and learner characteristics $F^G$ are obtained. We then use a clustering algorithm to divide learners into different clusters $G$ according to the learner’s characteristics $F^G$. For each cluster, the association rule mining algorithm is applied to generate the directed acyclic concept map $M_k$. Then, we use the topological sorting algorithm to generate a concept sequence for each concept map. It is to generate the full learning path $\text{flp}_k$ for each cluster. Finally, the full learning path of each cluster is simplified according to the concept error rate F of each learner, and the weak concept learning path set WLP is constructed. Among them, the framework diagram of the automatic generation of the full learning path is shown in Figure 3.

For a learner, some of the concepts in the learning path generated above may reach a high level of mastery; in that case, no further learning is required. In such a case, the follow-up learning link only needs to learn the concepts with a poor mastery. Thus, the full learning path of each cluster can be broken down and the weak concept learning path $\text{wlp}_m$ of
each history learner diagnosed based on the learner's concept error rate \( F \). A set of weak concept learning paths of all history learners WLP is then constructed. Algorithm 1 describes in detail the generation process of the weak concept learning path set.

4.2. Training LSTM+Attention Learning Effect Prediction Model. In recent years, using machine learning to predict learner performance has attracted the attention of many researchers. This includes techniques like learner performance prediction [48] and course failure risk prediction [49]. In view of the advantages of neural networks in data prediction, deep neural networks have been used in the education field to carry out a lot of learner performance prediction work. Some were based on the learner’s emotion analysis [50], while others centered the learner’s behavior analysis [51]. In [52], LSTM neural network was applied to predict risk learners from the clickstream information generated when the learner interacted with the e-learning platform. There are also studies that predict the learner’s performance in answering questions. The deep knowledge tracing (DKT) model proposed by Piech et al. used exercise answer sequence data as training data and applied LSTM neural network to predict whether learners could answer specific exercises correctly in [53]. In [54], Liu et al.’s research team designed a recurrent neural network to track and model the learning status of learners in the practice process and combined this network with the question content information to predict the learner’s answer score.

In summary, considering the advantages of LSTM neural network in sequence data prediction, and the outstanding performance of the attention mechanism in solving the long input sequence of LSTM/RNN model, the interpretability of the model has been improved. It is then explanatory and widely used in image processing [55], natural language processing [56], speech recognition [57], and other fields. Since the learning process data also reflect the time sequence, the LSTM neural network based on the attention mechanism is applied to predict the learning effect in this paper, which lays the foundation for the personalized weak concept learning path recommendation.

In this paper, the concept error rate \( F_l \) of a learner \( l \) and the weak concept learning path \( \text{wlp}_l \) of similar learners \( i \) recommended for his or her are given. \( F_l = \{ f_{l1}, f_{l2}, \ldots, f_{ln} \} \), \( f_{ln} \in [0, 1] \), and \( n \) is the number of concepts in the full path; \( \text{wlp}_l = < C_1, C_2, \ldots, C_p > \), and \( p \) is the number of concepts included in the weak concept path. For this learner, the learning effect is \( y_o \), and the learning effect category set is \( y = \{ y_1, y_2, y_3, y_4, y_5 \} \). The goal of the learning effect prediction task is to predict the learner’s learning effect category \( y_e \) according to the given learner’s concept error rate \( F_l \) and the recommended weak concept path \( \text{wlp}_l \). The model
structure is shown in Figure 4, including three parts: the input coding layer, the middle attention layer, and the label prediction layer. The structure and principle of each layer are described in detail below.

4.2.1. Input Coding Layer. This layer mainly vectorizes the learner’s characteristics and the weak concept path and stitches the two as the input of the model. It can be seen from the above that the lengths of the two are inconsistent. For example, if the concept error rate of the learner $l$ is $F_l = (f_{l1}, f_{l2}, \ldots, f_{ln})$, where $n$ is the number of concepts in the full path; it can be seen that the dimensions of $wlpm$ and $F_l$ are different. Although LSTM can handle variable-length sequences when training the model in batches, the data is input into the model in the form of tensors, and the length of the data needs to be consistent. Because each element of the learner’s learning feature is meaningful, if $f_{l}^{x} \in [0, 1]$ after $wlpm$ according to the ordinary operation, some feature elements will be discarded while splicing with the learning feature. Therefore, we choose to fill in the middle of $wlpm$ and replace the concepts that are discarded in the weak concept learning path with the label 0. That is, $wlpm = \langle C_1, C_2, \ldots, C_p, \ldots, C_n \rangle$. After embedding, it is spliced with the concept error rate $F_l$ as the input of the model.

As shown in the input layer of Figure 4, each time step will input a node in the learning path and a node in the concept error rate. After embedding, each node in the learning path becomes a $q$ dimensional vector, where $q$ is the

---

**Algorithm 1: Weak concept learning path set WLP generation.**

**Input:** Learners’ answer records $R$, Question-Concept matrix $QC$, Number of clusters $k$.

**Output:** Learner clusters $G = \{G_1, G_2, \ldots, G_p, \ldots, G_k\}$, Weak concept learning path set $WLP$.

1: $R$ is the right or wrong status of each question of the student, 1 is right and 0 is wrong. $QC$ is the relationship between the question and the concept (The concept can only belong to one concept, and the concept can contain multiple questions);

2: Based on $R$ and $QC$, calculate the concept error rate $F = \{f_1, f_2, \ldots, f_n\}$, where $f_i \in [0, 1]$, $n$ is the number of concepts;

3: Convert $F$ to $F'$, if $0 \leq f_i \leq 0.3$, then $f'_i = 0$, else if $0.3 \leq f_i \leq 0.7$, then $f'_i = 0.5$, else $f'_i = 1$;

4: Use the K-Means algorithm to divide the learner into $k$ clusters according to the feature $F'$, and get $k$ learner clusters $G = \{G_1, G_2, \ldots, G_p, \ldots, G_k\}$;

5: Use the concept map mining algorithm in [11] to dig out the concept map for each cluster of learners, and get the concept map set $M = \{M_1, M_2, \ldots, M_k\}$, the concept map is a directed acyclic graph and each learner cluster has different concept map;

6: Use the topological sorting algorithm to generate learning paths for $k$ concept maps, and get the full learning path set $FLP = \{flp_1, flp_2, \ldots, flp_k\}$;

7: According to the concept error rate $F$ of each learner, set the threshold, simplify the learning path $LP$ respectively, and construct a set of weak concept learning paths $WLP = \{wlpm, wlpm, \ldots, wlpm\}$; // $m$ is the number of history learners;

8: return $G$, $WLP$. 

---

![Figure 4: Structure diagram of learning effect prediction model.](image-url)
specified embedding dimension. Next, the concept error rate is sorted according to the order of the concepts in the learning path, and the concatenation performs sequentially. Each node is then a \((q+1)\) dimensional vector. There is a total of \(n\) nodes, so the dimension of the input vector of the middle layer of the model is \(n \times (q+1)\).

4.2.2. Intermediate Attention Layer. Considering the different influence degrees of different concepts on the learning effect, the model introduces the attention mechanism to distinguish the contribution of different concepts in the learning path in the prediction of learning effect. It is used to assign appropriate attention weights to the concepts in the learning path to further enhance the learning effect prediction ability of the model.

The first layer of the middle layer is the LSTM layer. Its input is the splicing of the embedding learning path and the concept error rate, that is, \(x = \text{wlp}_m \oplus \text{Flc}\), where \(m\) is the number of students, and the hidden vector of the previous time step. For example, \(x_i\) is the input of the \(i\)-th LSTM time step. It is a \((q+1)\)-dimensional vector. The total length of the LSTM time step is equal to the number of concepts \(n\). The hidden vector \(h = (h_1, h_2, \ldots, h_t), t \in [1, n]\) and each gate vector of the LSTM layer can be calculated iteratively using

\[
\begin{align*}
i_t &= \sigma(W_i x_t + U_i h_{t-1}), \\
f_t &= \sigma(W_f x_t + U_f h_{t-1}), \\
o_t &= \sigma(W_o x_t + U_o h_{t-1}), \\
\tilde{c}_t &= \tanh(W_c x_t + U_c h_{t-1}), \\
c_t &= i_t \odot \tilde{c}_t + f_t \odot c_{t-1}, \\
h_t &= o_t \odot (\tanh(c_t)),
\end{align*}
\]

(1)

where \(i_t\) is the input gate at time step \(t\), \(f_t\) is the forget gate at time step \(t\), \(o_t\) is the output gate at time step \(t\), \(c_t\) is the candidate cell state at time step \(t\), \(\tilde{c}_t\) is the cell state at time step \(t\), and \(h_t\) is the hidden vector at time step \(t\) and the output of the LSTM layer. Meanwhile, \(h_t\) is also input to the LSTM cell at time step \(t+1\) to continue the calculation until \(t\) is equal to the maximum time step \(n\).

The second layer of the middle layer is the attention layer. The attention structure is shown in Figure 5. The input is the latent vector of the LSTM layer \(h = (h_1, h_2, \ldots, h_t), t \in [1, n]\). First, formula (2) is used to input the hidden vector \(h_t\) into a single-layer neural network to obtain \(k_t\).

\[
k_t = \tanh(W_d h_t + b_k).
\]

(2)

After that, formula (3) is used to calculate the similarity between \(k_t\) and query vector \(q\) through vector dot multiplication. The normalized weight \(\alpha_t\) is then obtained by using softmax function.

\[
\alpha_t = \frac{\exp(k_t^T q)}{\sum \exp(k_i^T q)}
\]

(3)

To sum it up, the most influential concepts in the learning path are focused on to improve the effect of the model by calculating similarity and weighted summation methods.

4.2.3. Label Prediction Layer. The task of this layer is to predict the learning effect category of the target learner based on the learner’s concept error rate \(F_l\) and the weak concept learning path \(\text{wlp}_m\) code that are stitched together in the attention layer. We correspond the learning effect with the learning score interval. The 5 learning score intervals correspond to 5 learning effect categories, as shown in Table 1.

The output code of the attention layer is passed through a fully connected layer to construct a learning effect prediction model. The specific calculation method is shown in

\[
a_f = W a + b_a,
\]

(5)

where \(a_f\) which is a 5-dimensional vector is the learning effect label. \(W\) and \(b_a\) are the model parameters of the fully connected network. The final model output layer employs the softmax function to obtain the prediction score \(s_i\) of each
Table 1: Learning score and learning effects.

<table>
<thead>
<tr>
<th>The learning score intervals</th>
<th>Learning effect categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,60]</td>
<td>( y_1 = 0 )</td>
</tr>
<tr>
<td>(60,70]</td>
<td>( y_2 = 1 )</td>
</tr>
<tr>
<td>(70,80]</td>
<td>( y_3 = 2 )</td>
</tr>
<tr>
<td>(80,90]</td>
<td>( y_4 = 3 )</td>
</tr>
<tr>
<td>(90,100]</td>
<td>( y_5 = 4 )</td>
</tr>
</tbody>
</table>

learning effect category, and the specific calculation method is shown in

\[
s_j = \frac{\exp \left( a'_i \right)}{\sum \exp \left( a'_j \right), i = 1, 2, \ldots, 5,}
\]

where \( s_j \) is the probability of the corresponding label, \( s_j \in [0, 1] \), and \( \sum s_j = 1 \). The label corresponding to the maximum value of \( s_j \) is the final output of the model, that is, the learning effect category of the model prediction.

4.3. Personalized Weak Concept Learning Path Recommendation. For a given target learner’s answer records, the objective of personalized learning path recommendation is to find the appropriate learning path in the weak concept learning path set and to ensure that the learning effect after learning is as much expected as possible. The expected learning effect which should generally be higher than the learner’s historical learning effect is set by the learner or teacher. The recommendation process mainly includes learning path recall, learning effect prediction, and learning path conversion.

4.3.1. Learning Path Recall. For the target learner \( l_{\text{new}} \), the task of path recall is to find the weak concept learning path TLP of top-\( k \) similar users in the weak concept learning path set WLP, which is recommended as a preselected path.

In the learning path automatic generation Algorithm 1, the historical learners are clustered into 5 clusters through the \( K \)-means clustering algorithm. To this end, it is first necessary to determine the cluster of the learner \( l_{\text{new}} \) and then calculate the Euclidean value of \( l_{\text{new}} \) and the center vector of each cluster. The cluster with the smallest distance is the cluster which \( l_{\text{new}} \) belongs. In the path recall stage, the cosine similarity between \( l_{\text{new}} \) and each learner in the cluster is calculated and sorted according to the similarity. The weak concept learning path of the top-\( k \) learners with the most significant similarity is used as the preselected recommended learning path TLP. The specific algorithm is shown in Algorithm 2. We then find the \( k \) learning paths from the larger weak concept learning path set to \( l_{\text{new}} \), which lay the foundation for the next predicted path to the learning effect of \( k \) learning paths.

4.3.2. Learning Effect Prediction. In order to recommend a weak concept learning path that meets the expected learning effect for a learner \( l_{\text{new}} \), it is necessary to combine his/her concept error rate \( F_{\text{new}} \) to sequentially predict the learning effect of top-\( k \) recommended weak concept learning paths TLP. The specific algorithm is shown in Algorithm 3. The concept error rate \( F_{\text{new}} \) of \( l_{\text{new}} \) and the recalled top-\( k \) learning path TLP are inputted. The LSTM+attention learning effect prediction model can then get \( \text{Effect} = \{ \text{effect}_1, \text{effect}_2, \ldots, \text{effect}_k \} \).

4.3.3. Learning Path Conversion. The learning path conversion algorithm is shown in Algorithm 4, and the specific description is as follows.

The \( k \) predicted learning effects are traversed, and the corresponding learning path will be discarded if it is less than the expected learning effect \( \sigma_{\text{new}} \) of \( l_{\text{new}} \). Among them, \( \sigma_{\text{new}} \) can be specified by the learner or the teacher. If a learning path greater than or equal to the expected learning effect \( \sigma_{\text{new}} \) of \( l_{\text{new}} \) is found, it is recommended to \( l_{\text{new}} \). If the \( k \) predicted learning effects meet the conditions, the learning path corresponding to the learning effect with the most significant similarity is then recommended. This learning path not only meets the expected learning effect but also has the highest similarity with \( l_{\text{new}} \), which is most suitable for \( l_{\text{new}} \) to learn. If the \( k \) learning effects do not meet the expected learning effect \( \sigma_{\text{new}} \) of \( l_{\text{new}} \), the learning path corresponding to the learner with the highest similarity to \( l_{\text{new}} \) is then recommended.

5. Experiment and Evaluation

In this section, we conduct a series of experiments to verify the performance of the proposed method. We first elaborate data sets and the related experimental preparations. Then, the learning path generation and learning effect prediction model training are introduced. Finally, the proposed method is compared to the traditional recommendation method. We also evaluated the multifaceted impact of the proposed method from different architectures, different data embedded methods, and attention mechanism.

5.1. Data Set and Experiment Preparation. The experimental data is got from a public data set on GitHub. The data contains 617,940 test records of 6866 learners in large-scale computer culture basic course experiments, including 90 exercises covering 29 concepts [7]. Among them, each exercise examines one concept, and each concept can be examined by multiple exercises. In this way, the relationship matrix QC between exercises and concepts and the learner’s answer record matrix R can be obtained, where QC is given by authoritative experts in related fields to ensure correctness and authority.

Since the original data set only has the learner’s historical question data, there is no learning effect (i.e., academic performance) after the personalized learning path is recommended. Inspired by the literature [3], the method of simulating data is applied to generate the learning effect. From a generally statistical point of view, a learner’s academic performance approximates to a Gaussian distribution, with a mean of 75 and a variance of 8. The academic performance is the manifestation of the learner’s learning ability. For this reason, firstly, formula (7) is used to generate the ability value \text{Learning} \_ \text{ability} that obeys the Gaussian distribution for each learner.
Considering the learner’s concept error rate $F_i$, combined with the learner’s ability, formula (8) is used to generate the learning effect.

$$\text{Effect}_i = (1 - F_i + \sigma) \times \text{Learning_ability}_i, \quad (8)$$

where $\sigma \sim N(0.35, 1)$ represents the difficulty deviation of the exercises to the learner. $\text{Effect}_i$ represents the learning effect of the learner, where $\text{Effect}_i \in [0, 100]$.

In the experiment, 6866 learners are divided into training set and test set in 8:2. The number of learners included, respectively, is shown in Table 2.

The experimental running environment is Windows 10 operating system. The programming language is Python 3.6. The software development environment is VS Code and Jupyter. The deep learning framework is Pytorch 1.10, and the hardware environment is GeForce GTX 1050 Ti and 8G memory.

5.2 Concept Map Mining and Learning Path Generation Experiment. The experiment first uses the answer records and the relationship QC between the exercises and the concept to generate the learner’s error rate $F$ about the concept. $F$ is then divided into three levels of high, medium, and low to obtain student characteristics $F'$. Next, learner clustering operation is performed by $K$ -means. The experimental results show that when the number of clusters is equal to 5, the effect is best better. Therefore, the learners are gathered into 5 clusters. Figure 6 shows the two-dimensional visualization of the clustering results. It can be seen that the learners are well separated, which is consistent with the research results in [7].

Using the concept map mining algorithm based on association rules [11], 5 concept maps can be generated for 5 types of learners, respectively. The 5 full learning paths are then generated through topological sorting, respectively. Based on the full learning path, the learning path is simplified for each learner in the 5 categories and the concepts with the concept error rate less than the threshold are removed. Thus, the weak concept learning path for each learner can be obtained. The generated concept map, full learning path, and simplified weak concept learning path are detailed in [7]. They will not be shown in this paper.
Table 2: The number of learners included in the training set and test set.

<table>
<thead>
<tr>
<th>Data set</th>
<th>The number of learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training set</td>
<td>5,492</td>
</tr>
<tr>
<td>Test set</td>
<td>1,374</td>
</tr>
</tbody>
</table>

5.3. Training Experiment of Learning Effect Prediction Model. Before training the learning effect prediction model, the relationship between the data set, clusters, and the simulated learning effects are visualized to verify the rationality of the data set division and the simulated learning effect. Through the Sankey diagram shown in Figure 7, it can be clearly seen that the learners in the training set and the test set are evenly divided into 5 clusters. It can also be observed that the distribution of simulated learning effects obeys the Gaussian distribution, which is in line with the actual situation of teaching. From the perspective of the data flow of clusters and learning effects, the learners of the third cluster have better learning effects than other clusters, while the learners of the first cluster have poorer learning effects. It is necessary to provide learning guidance for them.

We used the concept error rate of learners in the divided training set and the weak concept learning path to train the LSTM+attention learning effect prediction model. The data embedding method uses random embedding, and the single-layer LSTM network is selected. The different learning rate and embedding size are trained separately. Figure 8 shows how the accuracy of the model on the test set changes with the training period when the learning rate is 1e-1, 1e-2, and 1e-3. Figure 9 shows the performance of the model when embedding size is 8, 16, and 32, respectively.

In the model training process, the final hyperparameter settings of the LSTM+attention learning effect prediction model are shown in Table 3 by comparing the effects of different hyperparameters. The optimizer chooses to use Adam, and the loss function uses the cross-entropy loss function. At this time, the accuracy of the model training varies with the training period as shown in Figure 10.

5.4. Performance Evaluation of Personalized Learning Path Recommendation. This section first gives the experimental evaluation indicators and then compares the effects of the proposed method with the traditional recommendation method. The final recommended effect comparative experiment is to verify the impact of the learning effect prediction model with LSTM as the main body and different structures on the proposed method.

5.4.1. Evaluation Indicators. In the experiment, in order to accurately recommend the weak concept learning path, other nonweak concepts also need to be recommended together. But these concepts need to be set to 0, so that the concept sequence in the recommended path is consistent with the sequence of knowledge points in the desired path. Therefore, we begin by conducting the evaluation of the overall learning recommendation effect to verify the proposed method. Evaluation indicators include concept-level Macro-precision, Macro-recall (recall rate), and Macro-F1 score. The recommended learning path for learners is $L_{p_i} = (l_{p_i}^1, l_{p_i}^2, \ldots, l_{p_i}^n)$, where $n$ is the number of concepts, and the target learning path is $L_{pi} = (l_{pi}^1, l_{pi}^2, \ldots, l_{pi}^n)$. The concepts recommended by learners (that is, nonweak concepts) in the two paths do not need to be marked with a unified subscript. Only when the corresponding concepts in the two paths are the same, the concept prediction is considered correct. For example, if $l_{pi}^i = l_{p_i}^i$, the prediction of the $i$-th concept is correct, and if $l_{pi}^i \neq l_{p_i}^i$, the prediction of the $i$-th concept is wrong. Formulas (9)–(11) are used to calculate Precision, Recall, and $F_1$ score of the $i$-th concept.

\[
\text{Precision}^i = \frac{#l_{p_i}^i}{(#l_{p_i}^i + \#l_{pi}^i)},\tag{9}
\]

\[
\text{Recall}^i = \frac{#l_{pi}^i}{(#l_{p_i}^i + \#l_{pi}^i)},\tag{10}
\]

\[
F_1 \text{score}^i = \frac{2 \times \text{Precision}^i \times \text{Recall}^i}{\text{Precision}^i + \text{Recall}^i},\tag{11}
\]

where $#l_{p_i}^i$ is the recommended learning path contains the $i$-th concept, and the corresponding position of the target learning path is also the $i$-th concept. $#l_{pi}^i$ is the recommended learning path contains the $i$-th concept, and the corresponding position of the target learning path is other concept. $#l_{pi}^i$ is the corresponding position of the target learning path is the $i$-th concept, and the recommended learning path is other concept.

After calculating Precision, Recall, and $F_1$ score for each concept through the above formulas, formulas (12)–(14) are used to calculate Macro-precision, Macro-recall, and Macro-F1 score.

\[
\text{Macro-precision} = \frac{1}{n} \sum_{i=1}^{n} \text{Precision}^i,\tag{12}
\]

\[
\text{Macro-recall} = \frac{1}{n} \sum_{i=1}^{n} \text{Recall}^i,\tag{13}
\]

\[
\text{Macro-F1 score} = \frac{2 \times \text{Macro-precision} \times \text{Macro-recall}}{\text{Macro-precision} + \text{Macro-recall}},\tag{14}
\]

5.4.2. Comparison with Traditional Recommended Methods. Since none of the existing learning path recommendation research has the same goals as ours, the proposed method is first compared with three traditional recommendations, that is, cluster-based recommendation [58], content-based recommendation [59], and learner-based collaborative filtering recommendation method [60]. The following briefly introduces the recommended ideas of three traditional recommended methods.
(i) Clustering-based recommendation (CL-BR): we start by calculating the learners’ concept error rate based on their answer data and obtain learner characteristics. We then use K-means for clustering to obtain 5 learner clusters. When recommending learning path for a target learner, we first locate the cluster to which the target learner belongs and identify the most similar top 5 learners in the cluster. Finally, the concepts in the 5 learners whose concept error rate is lower than the threshold are merged based on the full learning path and added to the recommendation list.

(ii) Content-based recommendation (CO-BR): initially, we construct a learner-learner similarity matrix, which is calculated from learner characteristics. A learner who is most similar to the target learner is then selected, and his/her weak concept learning path is recommended to the target learner.

(iii) Learner-based collaborative filtering recommendation (LE-BCFR): we firstly calculate the cosine similarity between learners based on their characteristics and construct a learner-learner similarity matrix. We then identify the top 5 most similar to the target learner. Next, we compare the learner characteristics of these 5 learners and the target learner and find the concept with an error rate lower than the threshold. Finally, we carry out the concept merging operation based on the full learning path and add it to the recommendation list.

The comparison results between the proposed method and the above three traditional methods are shown in...
It can be observed that the path recommendation method using the LSTM+attention for learning effect and path prediction has the best effect, and the Macro-F1 score can reach 0.8255. It can be observed that the proposed method is 14% higher than the best traditional recommendation method. As there is no quantified learning effect, the effect of clustering-based recommendation and learner-based collaborative filtering recommendation is deviated, and there is a big gap between them and the proposed deep learning-based method. The gap between Macro − precision and Macro − recall based on content recommendation is very large, indicating that the false-positive rate \( \#lp_{TP} \) is relatively large, and many unrelated concepts are recommended.

### 5.4.3. Comparison of the Effects of Different Structural LSTM Models on the Proposed Methods

In addition to the comparison with traditional recommendation methods, we also compare their results on the test set with LSTM as the main body and different structure of the learning effect prediction model. The following four models with different structures are set up for training, where the data embedding method uses random embedding and graph embedding.

<table>
<thead>
<tr>
<th>Model Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Random embedding+single-layer unidirectional LSTM</td>
</tr>
<tr>
<td>(2) Random embedding+single-layer bidirectional LSTM</td>
</tr>
<tr>
<td>(3) Random embedding+single-layer one-way LSTM + attention</td>
</tr>
<tr>
<td>(4) Graph embedding+single-layer unidirectional LSTM + attention</td>
</tr>
</tbody>
</table>

Furthermore, we also set the evaluation indicator \( \text{Find}_{rate} \) to evaluate the accuracy of the proposed learning path recommendation method. The calculation method is shown in formula (15), where \( \text{RNum}_{s} \) and \( \text{RNum}_{t} \), respectively, refer to the number of learners and the total number of learners who have successfully recommended a learning path in the test set.

\[
\text{Find}_{rate} = \frac{\text{RNum}_{s}}{\text{RNum}_{t}} \quad (15)
\]

As shown in Figure 11, Most time is the number of similar learners selected when recommending learning resources. As the number of similar learners increases, the number of those reaching the expected learning effect increases, together with the corresponding path recommendation effect. As it improves, the best recommendation is attained when the number of similar learners strikes 50.

It should be noted that since the neural network is the only model used to predict the learning effect in this paper,
only the learning effect prediction model with LSTM as the main body and different structures is used to compare the performance of the learning path recommendation method on the index Find_rate.

From Table 5, it can be seen that the recommendation method based on a single-layer and random embedding LSTM+attention has the best effect. The Macro – F1 score can reach 0.8255, and the one-way LSTM in the prediction model is better than the two-way LSTM. It may be an inherent logical sequence between concepts, generally learning from front to back, from simple to complex. The two-way LSTM in the prediction model which does not conform to the learning law increases the back-to-front direction, so it will reduce the model’s effect. Furthermore, the Find_rate of the proposed method is also the best, which is better than the single-layer LSTM without attention mechanism. It verifies the effectiveness of the model after the introduction of attention mechanism. The deviation of the prediction effect of LSTM+attention based on graph embedding than random embedding may be related to the graph embedding method of learning path data.

6. Conclusions

Educational data mining and analysis have great potential in the design of education systems. The booming of e-learning websites has brought a lot of learning data to the education field. This provides a data basis for mining the implicit relationship between concepts and diagnosing the learner’s knowledge loopholes.

In this paper, personalized weak concept learning path recommendation method is proposed. Firstly, the concept error rate is calculated by the learner’s answer records, and clustering is performed accordingly. The concept map is
mined by the association rule method. The full learning path is then generated through topological sorting, and the learner’s weak concept learning path is diagnosed. The LSTM network based on the attention mechanism is then used for learning effect and path prediction, which solves the problem that the learning effect cannot be quantified in the traditional recommendation algorithm. Finally, experiments verify that the proposed method has strong advantages over traditional recommended methods. The learning path recommendation algorithm can use historical learners’ answer records to diagnose weak concept and their relationship, provide personalized learning guidance to target learner, and ensure the feasibility and effectiveness of the guidance.

It should be pointed out that the concept map is a directed acyclic graph. For this reason, we try to embed the concept map data by using graph embedding in the experiment and hope to get a better embedding effect. However, it is found that the applying of Deep Walk’s graph embedding method reduces the recommendation effect. Therefore, we could employ more graph embedding methods for experiments in the follow-up research.

Besides the fact that the proposed method has great advantages compared with the traditional recommendation method, it also achieves a better recommendation effect. However, due to the limitation of experimental conditions, we have only veriﬁed the effect of the method on an open-source data set. The next step can be verified on multiple real education big data, and more in-depth research can be carried out in combination with the learner’s behavioral characteristics such as the time of answering questions, the number of answering questions, and the number of reminders used.

Data Availability

We used the data set accessed via GitHub website. The data set website address is as follows: diligentlee/LPG-algorithm-datasets. The data sets of LPG algorithm is as follows: http://github.com/.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the Shandong Postgraduate Education Quality Improvement Plan (No. SDYJG19075), Shandong Education Teaching Research Key Project (No. 2021JXZ010), Qingdao City Philosophy and Social Science Planning Project (No. QDSSL2001128), National Statistical Science Research Project (No. 2021LY053), Shandong Province Educational Science “Thirteenth Five-Year Plan” Special Project (No. 2020ZBY008), and Shandong University of Science and Technology Education and Teaching Research “Stars Plan” Project (No. QX2021M29).

References

[16] A. Esteban, A. Zafra, and C. Romero, “Helping university students to choose elective courses by using a hybrid multi-


