Higher Education-Oriented Recommendation Algorithm for Personalized Learning Resource

https://doi.org/10.3991/ijet.v17i16.33179

Liang Zhang[1], Xiao Zeng[2,3] and Ping Lv[1](✉)
[1] School of Economics and Management, Guizhou Normal University, Guiyang, China
[2] Business school, Guizhou University of Finance and Economics, Guiyang, China
[3] King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand
100458229@gznu.edu.com

Abstract—As smart education is continuously deepened in higher education, personalized learning resource recommendation has developed into a significant research field of smart learning. Although the prediction accuracy has been improved by knowledge tracing models established on the basis of students' historical learning data, how to design and apply personalized learning recommendation by combining classroom teaching of higher education is a great difficulty. To recommend personalized learning resources meeting the teaching requirements in higher education, the Q-LRDP-D (Learning Resource Difficulty Prediction and Dijkstra based on Q matrix) algorithm was proposed in this study. First, learning resources were modeled in accordance with Q matrix theory. Then, students' learning difficulty was predicted through the long short-term memory (LSTM) algorithm of a learning resource difficulty prediction module, followed by cyclic prediction through combining the to-be-learned knowledge points as required by teaching units to form a directed path diagram of learning resources. Next, the least learning resources conforming to students' learning levels were recommended using the shortest path algorithm to complete learning tasks. Lastly, the accuracy and effectiveness of the established model were verified through undergraduate teaching experiments. Results demonstrate that the modeling of learning resources in higher education on the basis of Q matrix theory is applicable to the LSTM algorithm. In comparison with the benchmark algorithm, its precision is obviously improved. With the recommendation algorithm, the average return on learning resources in the experimental class is 6.31, which is considerably higher than that in the control group. This study provides a certain reference for improving students' learning efficiency through recommendation algorithms in higher education and teaching.

Keywords—LSTM, Q matrix, personalized learning, learning efficiency, higher education

1 Introduction

With the continuous development of network technologies, online education platforms, such as intelligent tutoring system and massive open online courses, have welcomed rapid development and popularization in higher education. E-learning is of increasing importance in higher education and learning communities as standard components in many courses [1]. Compared with the traditional off-line education, the greatest advantage of online learning systems lies in the fact that they keep the detailed learning trajectories of learners and provide conditions for investigating learners' behavioral effectiveness under different trajectories [2]. Mining the potential learning laws from students' learning trajectories to provide personalized teaching services has become an issue closely concerned by scholars [3].

For a personalized learning system in higher education, personalized learning resources catering to the current learning tasks with a moderate difficulty level must be recommended. Thus, personalized learning recommendation algorithms have been extensively investigated. Knowledge tracing (KT), a powerful tool realizing artificial intelligence (AI)-aided education [4], refers to tracking students' knowledge state (KS) in accordance with their past performance and formulating the corresponding exercises to enhance the learning efficiency. Despite the great success achieved in the KT field, the Bayesian KT model [5], one of the most popular KT models, is subjected to great problems. For example, the variable-KC correspondence is fuzzy, and the binary variables set fail to accord with the realistic learning process. Deep learning, which arouses wide attention from scholars by virtue of its powerful feature extraction ability, has been extensively applied to the KT field. In 2015, a groundbreaking deep KT (DKT) model was proposed [6]. DKT elevates the area under the curve of ROC by 20% without needing any manually annotated data [7] and can capture and utilize students' knowledge at a deeper layer for characterization [8]. Therefore, it is highly suitable for learning-centered teaching evaluation systems [9]. In spite of its better prediction performance than that of existing classical methods, the practicability of DKT in educational applications remains to be enhanced [10]. Different students encounter diverse difficulties in learning the same learning resource, given that they vary in the mastery degree of the corresponding knowledge concepts. Hence, learning is inseparable from individual learning state [11]. **With the sustained promotion of blended teaching in higher education, closely integrating personalized learning recommendation systems with classroom teaching in higher education is crucial. How to design and apply personalized learning recommendation by combining classroom teaching of higher education is a key problem.**

On the basis of the above analysis, the learning resources in higher education were first modeled through Q matrix theory, and the difficulty levels of learning resources to students were then predicted using the DKT algorithm in this study. Next, the least personalized learning resource sequence was recommended by combining knowledge points (KPs) as required by the current teaching to complete learning objects and strengthen college students' learning efficiency further.

2 State of the art

The extensively applied AI and big data technologies have introduced new momentum into intelligent education [12]. Learning recommendation systems, which constitute an important research direction of the educational data mining field, have been widely applied to all kinds of intelligent learning systems. Learner modeling is a prerequisite for constructing accurate, high-quality, and personalized learning recommendation systems [13]. An explicit modeling method extracts obvious learners' features or preference description data from systems or documents to embody learners' uniqueness or similarities [14,15]. From the perspective of learning path recommendation, Nabizaddeh, A. H. et al. set learners' personalized parameters as learning objectives, skill learning, knowledge background, time limitation, and learning style [16]. The learning recommendation model proposed by Oliveira, A. et al. directly performs learner modeling through acquiring learning styles and then develops a set of online questionnaire tools related to the learning styles; moreover, 'learners' learning style characteristics are described with a learning style classification method [17]. Mark, T. W. et al. constructed statistical models that would be used to identify students who are at risk of failing the course and to identify assessment tasks that students in our course find challenging, as a guide for the design of future interventional activities. Every constructed predictive model had an excellent capacity to discriminate between students who passed the course and those who failed [18].

Explicit learner modeling, which is efficient and intuitive, reserves learners' characteristics required by recommendation systems so that learner models will be of favorable interpretability. When lacking learners' characteristics, however, the explicit learner modeling method is usually invalid, so explicit learner modeling is the current research hotspot of learning recommendation systems. Zhu, T. et al. proposed a personalized exercise recommendation method of predicting learners' knowledge mastery state on the basis of a cognitive diagnosis model. In accordance with students' question answering records, this method forms a cognitive diagnosis model to express their knowledge mastery state, then uses PMF to predict the question answering results, and finally recommends exercises based on the prediction results [19]. In the exercise recommendation system raised by Gong, T. et al., two sets of identically structured stacked denoising autoencoders are used to generate an implicit representation of learners and exercises [20]. On the basis of learners' question answering records, Wu, Z. et al. predicted learners' probability to correctly answer KPs by using a long short-term memory (LSTM, a variant of RNNs)-based KT model. Given that each exercise may contain one or multiple KPs, the vector representation for the KP mastery level can be formed on this basis to establish learner models [21]. Wu, Z and Li, M et al. proposed an exercise recommendation method based on mapping knowledge domain representation in the experimental part. Regarding learners and exercises as entities, this method regards students' exercise performance results as relationships [22]. Pereira, C. K. et al. introduced a method of extracting learners' social data through their Facebook accounts to seek for the information defining education interests, access time preferences, language, and media preferences and then instantiating learner ontologies on the basis of such information [23]. Bourkoukou, O., and El Bachari, E. proposed a recommendation method based on machine learning technique. Based on this tool, a learning approach was designed to achieve personalized learning experiences by selecting the most appropriate learning activities [24].

The commodity recommendation methods in the e-commerce field are referenced by many personalized learning recommendation methods. Rosewelt, L. A. et al. first extracted the representation characteristics of learning resources through a feature selection model and then classified them in accordance with learners' understanding level, thus identifying the exact learning resource contents from mass data and recommending on the basis of learners' understanding level [25]. Ibrahim, M. E. proposed a hybrid method based on content recommendation and collaborative filtering recommendation for course recommendation. This method overcomes the information overload problem by using ontologies, i.e., mapping the attribute structure of courses and the feature structure of learners by using ontologies with similar structural layers. In addition, it integrates collaborative filtering recommendation and content-based recommendation and configures the optimal system parameters by using a genetic algorithm [26]. Wan, S. and Niu, Z. used the explicit learner modeling method to enrich learners' feature description as much as possible, clustered learners through a self-organizing recommendation strategy, and finally sorted and recommended learning resources by mining sequential patterns, thus enhancing the individualization and diversity of recommendation [27].

In recent years, the KT research based on deep learning has been increasingly profound. Ding et al. solved the poor interpretability by using the uncertainty evaluation method [28]. Yang et al. thought that the manual analysis of characteristics failed to discretize extremely large features and would cause manual errors, which were solved using the proposed DKT-DT model. Moreover, one-hot coded additional features (title, question answering frequency, and time) were preprocessed using classification and regression trees [29]. Huo et al. incorporated a title coding method carrying contextual information into the LSTMCQ model they proposed. To be specific, a Q matrix was manually created by an expert in the field, which saved the correspondence between exercises and KC [30]. Gan et al. came up with a modeling method integrating such features as learners' ability, the difficulty of cognitive items, learning, and forgetting [31]. Yeung et al. combined the prediction results of the DKT+ model with the student features extracted from datasets to predict whether students would be occupied in STEM professions [32]. Wu et al. introduced a KT-based test paper generation method [33].

In the abovementioned studies, learner modeling, learning resource modeling, and recommendation methods have been deeply explored. Deep learning algorithms have also been combined to establish and apply learning state models using historical exercise records. However, most research scenarios have been set on online learning platforms, while personalized learning recommendation specific to conventional classroom teaching has been scarcely involved. In this study, Q matrix theory was used to establish a learning resource model and construct an exercise library. Then, a learning resource difficulty prediction (LRDP) module was designed via the DTK model. Next, an alternative set was established in accordance with the KPs involved in the current classroom teaching tasks, and the LSTM algorithm of the DTK model was employed to select items meeting the current learning level of students to form a directed network graph. Lastly, a learning resource sequence with a moderate difficulty level and the least number was recommended to students by using the shortest path algorithm to complete the current learning tasks and enhance students' learning efficiency.

The remainder of this study is organized as follows: In Section 3, the idea of modeling learning resources on the basis of Q matrix theory is expounded, and the design idea of an LRDP module designed using the LSTM algorithm and a learning resource sequence generation (LRSG) module of the shortest path algorithm is followed. In Section 4, the experimental design and evaluation metrices are described, and the algorithm accuracy and effectiveness are analyzed. In the final section, the whole study is summarized, and conclusions are drawn.

3 Methodology

After completing the classroom teaching of one unit, a student needs to understand and master it further by learning the learning resources of related KPs. The Q-LRDP-D recommendation algorithm proposed in this study was mainly divided into three modules (Figure 1). First, Q matrix theory was used to establish an exercise model and construct an exercise library. Then, the probability of each student to master learning resources was predicted using the LSTM algorithm in the LRDP module in accordance with their learning history. Lastly, a directed graph for recommending learning resource sequences was generated by combining the KP requirements of teaching units and the LSTM algorithm, and the most efficient learning resource sequence was found through the Dijkstra algorithm for recommendation.

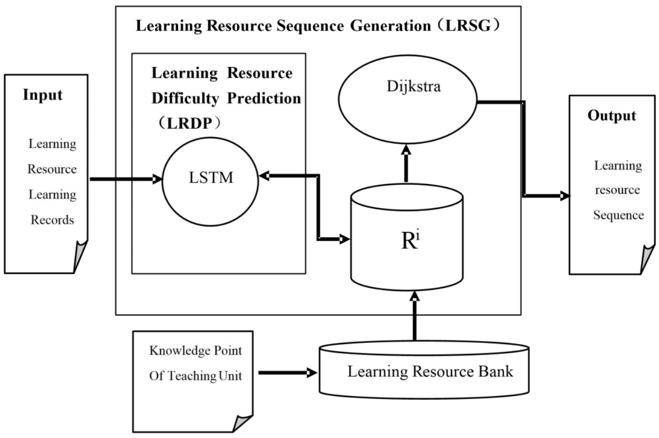

Fig. 1. Framework of Q-LRDP-D

3.1 Establishment of a learning resource library in accordance with Q matrix theory

Q matrix, first proposed by Embresin, develops into Q matrix theory after being perfected by Tasuka [34]. Q matrix theory mainly aims to determine the unobservable KPs tested by exercises and transform them into an observable item response mode to infer students' cognitive state. The Q matrix, which describes the relationship between learning resources and KPs, is generally a 0–1 matrix containing *m* rows (*m* is the number of learning resources) and *n* columns (*n* denotes the number of KPs). If $Q_{mn}=1$, then the learning resource *m* contains the KP n.

Given the layered relationship between KPs, the direct relations between KPs are reflected using an adjacency matrix (A matrix), a *K*-row and *K*-column 0–1 matrix. If two KPs are directly related, the corresponding value is 1; otherwise, it is 0. The accessible matrix (R matrix) reflects the direct, indirect, and self-relations between KPs. Only if two KPs present any of the above relations, the corresponding value is 0; otherwise, it is 0. To acquire the R matrix conveniently, Equation (1) can be applied to calculation, where *I* is a unit matrix. With the increase in *n*, the R matrix is obtained when the value is stably unchanged.

$$R=(A+I)^n \tag{1}$$

An ideal mastery mode refers to a logically reasonable mastery mode (also called KS) acquired in accordance with the hierarchical relations between attributes. KS is a *K*-column 0–1 matrix, where 1 means that the attribute is mastered, and 0 denotes the opposite meaning. The KS matrix can be obtained from the R matrix through an extension algorithm [35]. On the basis of such an ideal mastery mode, a learning resource library for higher education can be designed to acquire the corresponding Q matrix, and each column in KS will appear repeatedly in the Q matrix; *i.e.*, the same KS is involved in multiple learning resources. Each row of the Q matrix represents all KPs involved in this learning resource. If this learning resource is learned by a student, he/she will master all KPs involved in it.

3.2 LRDP

The difficulty level of a learning resource to a student can be measured by estimating his/her difficulty in learning this learning resource. In this study, the probability for this student to master the next learning resource at time $t+1$ was predicted in accordance with his/her learning resource mastery records from 0 to *t*. According to Q matrix theory, however, whether a student can master a learning resource depends on his/her mastery of the KPs involved in the learning resource. Here, $P(K^a)$ denotes the probability for student *a* to master all KPs, which was calculated through Equation (3). In general, the student was assumed to be able to master the learning resource when $P_m^a \geq 0.5$. Then, the difficulty level could be expressed by Equation (4).

$$P(K^a)=(P(k_1^a),P(k_2^a),...,P(k_n^a)) \tag{2}$$

$$P_m^a = \sqrt[n]{\prod_{i=1}^{n} P(k_i^a) \bullet Q_{mi}} \tag{3}$$

$$D_m^a = 1 - P_m^a \tag{4}$$

The DKT model, which uses an LSTM network, can be used to estimate P_m^a (Figure 2). The input X_t consists of (Q_m, A_m), where A_m is a 0–1 variable; 1 means that the student already masters the learning resource; otherwise, the value is 0. This variable determines whether the column vector of the Q matrix at time *t* and the answer are correct. The output Y_t is a vector with a length of *n*, where each component represents the probability to correctly answer the corresponding knowledge conceptions. The output Y_t of the well-trained LSTM network denotes the probability $P(K^a)$ for the student *a* to master all KPs in the curriculum.

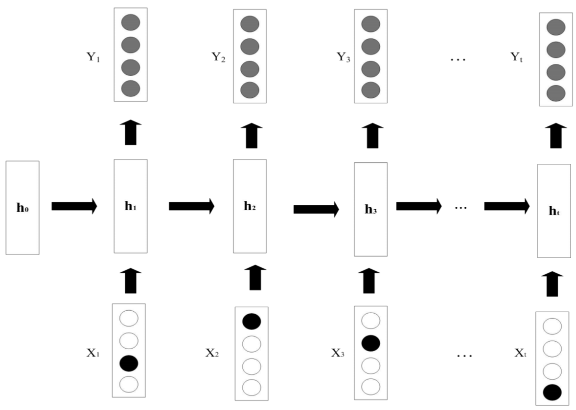

Fig. 2. DKT model structure

According to the DKT optimization method, the training objective lies in the negative logarithmic likelihood of students' answer observation sequence under the model. Here, $\phi(K^{t+1})$ denotes the KP vector at t+1, and a^{t+1} stands for the result obtained when the student learns the next learning resource. The loss function was constructed through a binary cross entropy, which, for a single student, could be expressed by Equation (5), where L_b is the binary cross entropy. Then, the optimized cost function for learning resource difficulty estimation could be denoted by Equation (6), where *S* is the number of students.

$$L_m = \sum_{t=0}^{T} L(y^t \bullet \phi(K^{t+1}), a^{t+1}) \tag{5}$$

$$L_{LRDP} = \frac{1}{\sum_{i=1}^{S} T_i} \sum_{i=1}^{S} \sum_{t=0}^{T_i} L(y^t \bullet \phi(K^{t+1}), a^{t+1}) \tag{6}$$

3.3 LRSG

Essentially, learning resource recommendation refers to a KS column vector catering to students' ability through their present KS and then seeking for the corresponding learning resources from the Q matrix on the basis of the row vector of KS. In this study, the recommendation mainly aimed to guide a student to master all KPs stipulated by a teaching unit with the least learning resources. Therefore, the difficulty level of recommended learning resources was required to conform to the student's present cognitive state. Moreover, the difficulty level was elevated step by step until the student could complete the learning resource containing the most stipulated KPs. First, the stipulated KPs were converted into an ideal mastery mode KS^U in accordance with Q matrix theory. Second, all column vectors of KS^U were found from the KS matrix as the recommended alternative set *U*. Third, all *U*-containing P_U^a were predicted using the LRDP module, and that meeting $P_{U(i)}^a \geq 0.5$ was regarded as the recommendation set $P_{U(i)}^a \geq 0.5$ in the first step. Fourth, the student was assumed to have smoothly completed the learning resources corresponding to R^1, and R^2 was generated by following the same idea. The rest was done in the same manner until the learning resource containing all stipulated KPs appeared in the recommendation set. Such R^n sets formed a directed graph, and the LRSG was turned into the shortest path problem, which could be solved through the Dijkstra algorithm (Figure 3).

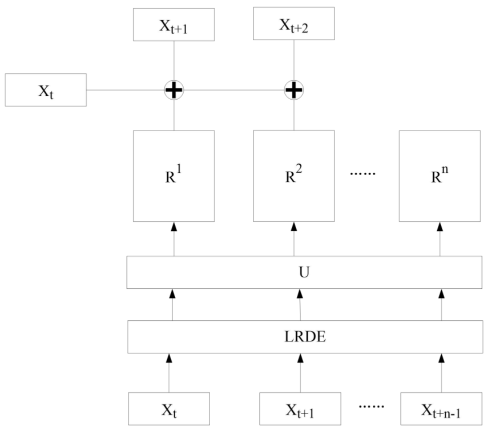

Fig. 3. LRSG graph

4 Result analysis and discussion

4.1 Experimental design

A personalized learning recommendation algorithm realizes recommendation by processing learner models and learning resource models. However, unlike other personalized recommendation systems, learning resources cannot only be such single learning resources as courseware, multimedia, and exercises but can also be a learning path composed of several associated learning resource combinations [17]. In this experiment, exercises were mainly used as learning resources, and data were derived from the exercise data of 625 undergraduates in 8 classes at 6 universities regarding 32 KPs of 4 teaching units in the course e-Marketing, as shown in Table 1.

Table 1. Overview of datasets

Course	KPs	Students	Exercise	Records
e-Marketing	32	625	457	84,254

To accelerate the LSTM training process in the LRDP module, the loss function was minimized in this study by using a small-batch (batch size: 64) random gradient descent. The model was trained at a learning rate of 0.01. Next, the LSTM network was trained in the LRDP module by using 70% of the datasets. Meanwhile, 10% were used as the verification datasets, and the remaining 20% served as the test datasets.

Student-based collaborative filtering (SBCF) and content-based filtering (CBR) were used as the benchmark algorithms. First, the SBCF algorithm constructed a student–student similarity matrix in accordance with the exercise answer records. Then, the first 10 students doing similar exercises were determined, and the exercise with the difficulty tag value closest to the expected value was found for recommendation. For the CBF algorithm, the distance from the difficulty tag of each exercise to the expected difficulty of all exercise weights was first calculated. Next, the similarities of all exercises to the exercises answered by the target students at time *t* were calculated. Lastly, the weighted sum was calculated, and the first 10 exercises were sorted in descending order and then recommended. The difficulty D_m of exercise *m* is defined in Equation (7), where A_m^i is the correctness of student *I* in answering exercise *m*, and A_m denotes the set of students answering exercise *m*.

$$D_m = \frac{\sum_{i=1}^{S} A_m^i}{|A_m|} \tag{7}$$

4.2 Experimental metrics

As for the LRDP module, the LSTM algorithm was mainly used to predict whether a student could correctly answer exercises, so precision and recall were regarded as the main verification metrics. To recommend the exercise sequence, first, exercises should be correctly answered by students; namely, the difficulty level should cater to students' KS. Second, better performance could be achieved using a small number of exercises. Given the two points, exercise sequences were measured using the correct response rate and the average return in exercise performance, as shown in Equations (8) and (9), respectively. In the equations, *Acc* is the accuracy rate of the exercise recommendation sequence, *S* stands for the number of students using the exercise recommendation sequence, R_j denotes the exercise recommendation sequence of student *j*, $Test_j$ is the test performance of student *j*, and *Eff* represents the average return on exercise performance.

$$Acc = \frac{1}{S}\sum_{j=1}^{S} \frac{\sum_{i \in R_j} A_i^j}{|R_j|} \tag{8}$$

$$Eff = \frac{1}{S}\sum_{j=1}^{S} \frac{Test_j}{|R_j|} \tag{9}$$

4.3 Analysis of experimental results

From Table 2, all metrics of LSTM were significantly better than those of benchmark algorithms CBR and SBCF. With increasing number of recommendations, the accuracy rate of LSTM was improved to a greater extent than that of the benchmark algorithms, indicating that the LSTM algorithm can more accurately predict whether a student can correctly answer exercises. The experimental results in Table 2 were acquired by assuming that the probability for students to answer exercises correctly satisfied $P_m^a \geq 0.5$. Figure 4 presents the changes in the MAE of the LSTM algorithm induced by the change in this threshold. The MAE value slightly increased when the threshold decreased, and it then declined slowly. Nevertheless, when the threshold was elevated, the MAE value declined fast, which was probably associated with our course. For liberal arts, such as network marketing, featured by a high degree of substitution between KPs, even a KP of one exercise poorly mastered by a student could still be correctly answered. Thus, the MAE value could be increased on the contrary if the threshold was reduced.

To verify the accuracy and effectiveness of the algorithm in recommending learning resource sequences, a total of 213 students were invited into the comparative experiment and largely divided into 4 groups. This experiment mainly aimed to verify whether the recommendation algorithm could effectively improve students' learning performance and efficiency, and there should be no differences among the four groups in pretest performance. First, a pretest was performed among the 213 participatory students before completing the first 2 teaching units. Then, they were grouped in accordance with the pretest performance through the following criterion: No significant differences existed among the four groups in the pretest performance. The one-way analysis of variance (ANOVA) of pretest performance of the four groups is displayed in Table 3. The grouping information and adopted exercise recommendation algorithms are exhibited in Table 4. Forty exercises were randomly recommended in class A, exercises were recommended by the two benchmark algorithms in classes B and C, and those in class D were recommended *via* the Q-LRDP-D algorithm with the number depending on this algorithm.

Table 2. Metric values of LSTM and benchmark algorithms

Top-N	Recommendation algorithm	Precision	Recall	F
5	CBR	0.6752	0.6592	0.6671
	SBCF	0.7102	0.6911	0.7005
	LSTM	0.8001	0.7245	0.7604
10	CBR	0.6812	0.6743	0.6777
	SBCF	0.7217	0.7101	0.7159
	LSTM	0.8101	0.7398	0.7734
15	CBR	0.6945	0.6882	0.6913
	SBCF	0.7381	0.7198	0.7288
	LSTM	0.8203	0.7549	0.7862
20	CBR	0.7134	0.7011	0.7072
	SBCF	0.7523	0.7321	0.7421
	LSTM	0.8374	0.7942	0.8152

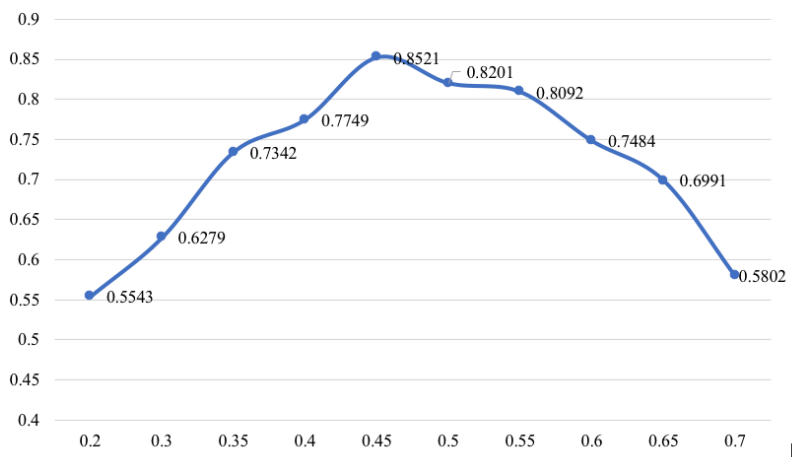

Fig. 4. Threshold-dependent change in the MAE value of the LSTM algorithm

Table 3. One-way ANOVA of pretest performance

	Quadratic sum	Degree of freedom	Mean square	*F*	Significance
Intergroup	1.929	3	0.942	0.021	0.973
Intragroup	6,231.613	129	50.420		
Total	6,233.542	212			

* The significance level of mean value difference is 0.05.

Table 4. Grouping information and adopted recommendation algorithms

Group	Number of students	Recommendation algorithm
A	53	/
B	54	CBR
C	52	SBCF
D	54	Q-LRDP-D

A test was carried out immediately after the teaching of the third unit was completed. The mean value and standard deviation of the performance and the number of exercises in the four groups are presented in Table 5. Table 6 displays the *t* test data of group D and the three other groups. The best average performance appeared in group C using the SBCF algorithm, followed by groups B, D, and A. From the *t* test results, however, the average performance of group D was insignificantly different from that of groups B and C. As for the performance distribution, the variance of group D was the minimum, that of group A was the maximum, and no difference was observed between groups C and D in variance. Therefore, the Q-LRDP-D algorithm did not obviously differ from the benchmark algorithms when influencing the test performance. Nonetheless, the performance was still obviously improved in comparison with that in randomly recommended exercises, and student performance was more approximate to the mean value without serious polarization. In addition, the average number of exercises was the maximum in group A, and that in group D was evidently smaller than that in the three other groups, manifesting that the Q-LRDP-D algorithm can effectively reduce the number of exercises to be done by students. The variance of group D was also apparently higher than those of groups B and C, revealing that the Q-LRDP-D algorithm can more effectively distinguish the number of exercises to be done by students compared with the benchmark algorithms.

Table 5. Statistical data of the performance and the number of exercises to be done by respondents in different groups

Group	Average performance	Variance	Average number of exercises	Variance
A	75.5	15.4	21.2	8.81
B	80.4	11.3	18.4	4.51
C	81.1	10.4	16.5	5.07
D	80.5	10.3	12.6	6.02

Table 6. T test of the performance and the number of exercises in different groups

Comparison	*T* test of average performance	*T* test of the average number of exercises
A-D	-3.24*	-2.59*
B-D	-0.45	-3.22*
C-D	-0.59	-4.01*

* The significance level of mean value difference is 0.05.

From the metrices *Acc* and *Eff* (Table 7), the accuracy rate of group D was higher than that of group A but slightly lower than those of groups B and C. Hence, the algorithm accuracy declined fast during the cyclic prediction. The change relation between the average accuracy rate and the number of cycles during the cyclic prediction of LSTM is displayed in Figure 5. With the increase in the number of cycles, the prediction accuracy declined rapidly, but the calculation was generally completed within 20 cycles, so the overall algorithm accuracy did not decline evidently. The *Eff* value of group D was obviously better than those of groups B and C, so the personalized learning recommendation algorithm could effectively improve students' learning efficiency. The benchmark algorithms performed the recommendation rightly by using the respondents' learning experience, while the Q-LRDP-D algorithm could more effectively identify students' current learning level, thus making the recommended exercises more pertinent.

Table 7. Experimental results of the exercise recommendation sequence module

Group	*Acc*	*Eff*
A	69.5	3.61
B	73.3	4.49
C	76.4	5.01
D	80.2	6.31

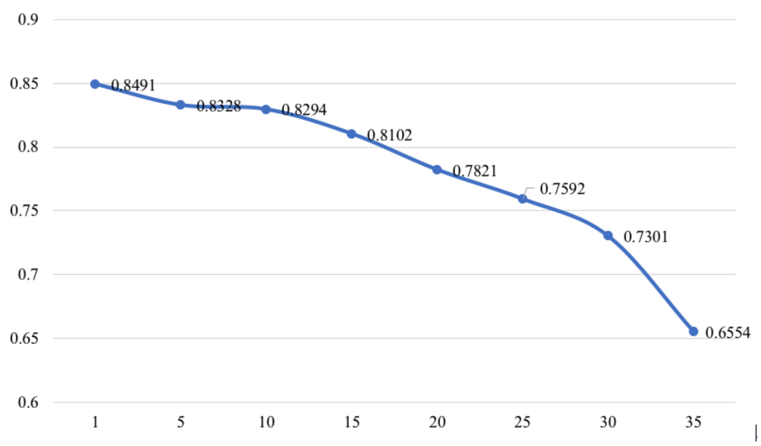

Fig. 5. Change in cyclic prediction accuracy of LSTM with the number of cycles

5 Conclusions

To meet the personalized learning needs of higher education, the current learning tasks were completed by recommending learning resource sequences meeting students' learning levels. First, learning resources were modeled in accordance with Q matrix theory. Then, whether a student would learn a learning resource was predicted using LSTM in the LRDP module. Moreover, the KPs of teaching units were cyclically predicted and filtered through the LRDP module to form a directed graph. Lastly, the learning resource sequences were recommended through the shortest path algorithm. The following conclusions could be drawn:

1. Q matrix theory can effectively model the learning resources of higher education and be applied to the DTK model.
2. Compared with the benchmark algorithms, the LSTM algorithm of the DTK model can markedly enhance the estimation accuracy for learning resource difficulties.
3. With the to-be-learned KPs combined, the LRDP module can produce a direct graph by cyclically predicting learning resource difficulties. In addition, the learning resource sequences recommended by the shortest path algorithm can effectively strengthen college students' learning efficiency.

In this study, the accuracy of the DTK model was verified through practical undergraduate teaching data. With the proposed Q-LRDP-D algorithm, students could accomplish learning tasks by completing a small quantity of learning resources, which would be of certain practical significance for enhancing college students' learning efficiency. Owing to the lack of real learning state data of college students in long-term teaching, the diversified learning state data of college students will be further combined in the follow-up study to correct the proposed algorithm so that the recommended learning resource sequences can meet practical teaching needs to a greater extent. Furthermore, the Q matrix vector will be automatically generated for new learning resources in the construction of learning resource banks.

6 Acknowledgment

This work was supported by the Teaching Reform Project of Guizhou Normal University (program number: 2021XJG20), Innovation and Entrepreneurship Project of Guizhou Normal University, 2018 Major Education Reform Project of Guizhou Normal University (Project Name: Innovation and Practice of E-commerce Talent Cultivation Mode with Deep Integration of Industry and Education in Guizhou Province, China), and 2019 National First-class Major in Marketing in China.

7 References

[1] Alrikabi, H. T., Jasim, N. A., Majeed, B. H., Zkear, A. A., & ALRubeei, I. R. N. Smart Learning based on Moodle E-learning Platform and Digital Skills for University Students. International Journal of Recent Contributions from Engineering, Science & IT, 2022, vol.10(1), pp. 109-120. https://doi.org/10.3991/ijes.v10i01.28995

[2] Li, F., Ye, Y., Li, X., et al. Application of knowledge tracing model in education: A review from 2008 to 2017. *Distance Education in China*, 2019, vol.7, pp.86-91.

[3] Wan, X., Anma, F., Ninomiya, T., et al. CollaboeNOTE: A Hybrid Recommender System for Group Learning Support. *International Journal of Computer Science and Network Security*, 2008, vol.8(9), pp.166-171.

[4] Ha, H., Hwang, U., Hong, Y., et al. Deep trustworthy knowledge tracing. *arXiv preprint, 2018, arXiv*: 1805. 10768.

[5] Corbet, A. T., Anderson, J. R. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, 1994, vol.4(4), pp.253-278. https://doi.org/10.1007/BF01099821
[6] Piech, C., Basen, J., Huang, J., et al. Deep knowledge tracing. In: Proceedings of the 28th Int Conference on Neural Information Processing System (NeurIPS), Cambridge, MA: MIT, 2015, pp.505-513.
[7] Xiong, X., Zhao, S., Inwegen, E. V., et al. Going deeper with deep knowledge tracing. *In: Proceedings of the 9th Int Conference on Educational Data Mining (EDM)*, Worcester, MA: IEDMS, 2016, pp.545-550.
[8] Mao, Y., Lin, C., Chi, M. Deep learning Bayesian knowledge tracing: Student models for interventions. *Journal of Educational Data Mining*, 2018, vol.10(2), pp.28-54.
[9] King, D. R. Production implementation of recurrent neural networks in adaptive instructional systems. In: *Proceeding of the 22nd Int Conference on Human-Computer Interaction (HCI). Berlin, Springer*, 2020, pp.350-361. https://doi.org/10.1007/978-3-030-50788-6_25
[10] Doleck, T., Lemay, D. J., Basnet, R.B., et al. Predictive analytics in education: A comparison of deep learning frameworks. *Education and Information Technologies, 2022, vol.*25 (3), pp.1951-1963. https://doi.org/10.1007/s10639-019-10068-4
[11] Lv, P., Wang, X., Xu, J., Wang, J. Utilizing knowledge graph and student testing behavior data for personalized exercise recommendation. *In: Proceedings of ACM Turing Celebration Conference, ACM,* 2018, pp.53–59. https://doi.org/10.1145/3210713.3210728
[12] Kekang, H. E. Significant Influence of Emerging Information Technology on Deepening Reformation of Education in the 21st Century. *e-Education Research,* 2019, vol.311(3), pp.7-14.
[13] Glrdpn, M., Emerson, A., Rowe, J. P., et al. Predictive student modeling in educational games with multi-task learning. In: Proceedings of the Thirty Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020. AAAI Press, 2020, pp.654-661. https://doi.org/10.1609/aaai.v34i01.5406
[14] Shi, D., Wang, T., Xing, H., et al. A learning path recommendation model based on a multidimensional knowledge graph framework for e-learning. *Knowledge-Based Systems*, 2020, vol.195(2), pp.105618. https://doi.org/10.1016/j.knosys.2020.105618
[15] Dias, A. D. S., Wives, L. K. Recommender system for learning objects based in the fusion of social signals, interests, and preferences of learner users in ubiquitous e-learning systems. *Personal and Ubiquitous Computing,* 2019, vol.23(2), pp.249-268. https://doi.org/10.1007/s00779-018-01197-7
[16] Nabizadeh, A. H., Leal, J. P., Rafsanjani, H. N., et al. Learning path personalization and recommendation methods: A survey of the state-of-the-art. *Expert Systems with Applications*, 2020, vol.159(5), 113596. https://doi.org/10.1016/j.eswa.2020.113596
[17] Oliveira, A., Teixeira, M., Neto, C. Recommendation of educational content to improve student performance: An approach based on learning styles. In: *Proceedings of the 12th International Conference on Computer Supported Education, Prague, Czech Republic,* May 2-4, 2020. SciTe Press, 2000, pp.359-365. https://doi.org/10.5220/0009436303590365
[18] Williams, M. T., Lluka, L. J., & Chunduri, P. Redesigning a First Year Physiology Course using Learning Analytics to Improve Student Performance. International Journal of Learn-

ing Analytics and Artificial Intelligence for Education, 2021, vol.3(1), pp. 4–19. https://doi.org/10.3991/ijai.v3i1.21799
[19] Zhu, T. Y., Huang, Z. Y., Chen, E. H., et al. Cognitive diagnosis based personalized question recommendation. *Chinese Journal of Computers, 2017, vol.*1, pp.176-191.
[20] Gong, T., Yao, X. Deep exercise recommendation model. *International Journal of Modeling and Optimization*, 2019, vol.9(1), pp.18-23. https://doi.org/10.7763/IJMO.2019.V9.677
[21] Wu, Z., He, T., Mao, C., et al. Exam paper generation based on performance prediction of student group. *Information Sciences*, 2020, vol.532(9), pp.72-90. https://doi.org/10.1016/j.ins.2020.04.043
[22] Wu, Z., Li, M., Tang, Y., et al. Exercise recommendation based on knowledge concept prediction. *knowledge-Based Systems*, 2020, vol.210(12),106481. https://doi.org/10.1016/j.knosys.2020.106481
[23] Pereira, C. K., Campos, F., Strele, V., et al. BROAD-RSI: educational recommender system using social networks interactions and linked data. *Journal of Internet Services and Applications, 2018, vol.9(1), 7:1-28*. https://doi.org/10.1186/s13174-018-0076-5
[24] Bourkoukou, O., & El Bachari, E. A Big-Data Oriented Recommendation Method in E-learning Environment. *International Journal of Emerging Technologies in Learning, 2022, vol.17(10), pp. 74–84*. https://doi.org/10.3991/ijet.v17i10.27861
[25] Rosewelt, L. A., Renjit, J. A. A content recommendation system for effective e-learning using embedded feature selection and fuzzy DT based CNN. *Journal of Intelligent Fuzzy Systems*, 2020, vol.39(1), pp.795-808. https://doi.org/10.3233/JIFS-191721
[26] Ibrahhim, M. E., Yang, Y., Ndzi, D. L., et al. Ontology based personalized course recommendation framework. *IEEE Access*, 2019, vol.7(1), pp.5180-5199. https://doi.org/10.1109/ACCESS.2018.2889635
[27] Wan, S., Niu, Z. A hybrid e-learning recommendation approach based on learners' influence propagation. *IEEE Transactions on Knowledge and Data Engineering*, 2020, vol.32(5), pp.827-840. https://doi.org/10.1109/TKDE.2019.2895033
[28] Ding, X., Larson, E. C. Incorporating uncertainties in student response modeling by loss function regularization. *Neurocomputing*, 2020, vol.409, pp.74-82. https://doi.org/10.1016/j.neucom.2020.05.035
[29] Yang, H., Cheung, L. P. Implicit heterogeneous features embedding in deep knowledge tracing. Cognitive Computation,2018, vol.10(1), 314. https://doi.org/10.1007/s12559-017-9522-0
[30] Huo, Y., Wong, D. F., Ni, L. M., et al. Knowledge modeling via contextualized representations for LSTM-based personalized exercise recommendation. *Information Sciences, 2020, vol.*523, pp.266-278. https://doi.org/10.1016/j.ins.2020.03.014
[31] Gan, W.， Sun, Y.， Peng, X., et al. Modeling learner's dynamic knowledge construction procedure and cognitive item difficulty for knowledge tracing. *Applied Intelligence, 2020, vol.*50(11), pp.3894-3912. https://doi.org/10.1007/s10489-020-01756-7
[32] Yeung, C. K., Yeung, D. Y., et al. Incorporating features learned by an enhanced deep knowledge tracing model for stem/non-stem job prediction. *International Journal of Artificial Intelligence in Education,* 2019, vol.29(3), pp.317-341. https://doi.org/10.1007/s40593-019-00175-1
[33] Wu, Z., He, T., Mao, C., et al. Exam paper generation based on performance prediction of student group. *Information Sciences*, 2020, vol.532: 7290. https://doi.org/10.1016/j.ins.2020.04.043
[34] Tatsuoka, K. K. Architecture of knowledge structure and cognitive diagnosis: A statistical pattern recognition and classification approach. In: P. D. Nichols, S. F. Chipman, and R. L.

Brennan (Eds.), Cognitively Diagnostic Assessment. Hillsdale, NJ: Lawrence Erlbaum Associates,1995.
[35] Tu, D. B., Cai, Y., Ding, S. L., Cognitive Diagnosis Theory, Beijing Normal University Press, 2021.

8 Authors

Ling Zhang is associate professor in the School of Economics and Management, Guizhou Normal University, China.

Xiao Zeng is teaching assistant in the Business school of Guizhou University of Finance and Economics, China and PhD student King Mongkut's Institute of Technology Ladkrabang, Thailand.

Ping Lv is professor in the School of Economics and Management, Guizhou Normal University, China (email: 100458229@gznu.edu.com).

Article submitted 2022-06-10. Resubmitted 2022-07-11. Final acceptance 2022-07-12. Final version published as submitted by the authors.