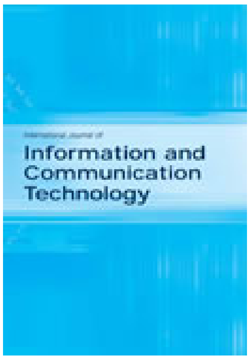


International Journal of Information and Communication Technology

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

Adaptive content recommendation for distance education based on fuzzy logic and knowledge graph

Chunmei Du, Donghai Xu

DOI: 10.1504/IJICT.2025.10071553

Article History:

Received:	10 February 2025
Last revised:	20 February 2025
Accepted:	20 February 2025
Published online:	18 June 2025

Adaptive content recommendation for distance education based on fuzzy logic and knowledge graph

Chunmei Du* and Donghai Xu

Tangshan Open University,
Tangshan, 063000, China
Email: duchunmei2024@163.com
Email: tslbtbt@163.com
*Corresponding author

Abstract: Intending to the issue that existing adaptive content recommendation methods for distance education ignore the dynamic uncertainty of learners' cognitive level, the top-down approach is first used to construct the distance education KG (DEKG), and the TransR model is utilised to vectorise the representation of the DEKG. Secondly, based on fuzzy logic, the cognitive level of the learners is determined, and the matching degree and cognitive level are combined to calculate the similarity of knowledge points. Then, the degree of learner preference was measured using fuzzy logic to represent the knowledge point similarity as a vector over the content labels. Subsequently, a corresponding rating prediction formula is designed to realise more effective and accurate mining of distance education content that meets learners' characteristics for recommendation. The experimental results show that the proposed method improves the recall and F1 by at least 3.21%.

Keywords: adaptive content recommendation; fuzzy logic; knowledge graph; knowledge point similarity; rating prediction.

Reference to this paper should be made as follows: Du, C. and Xu, D. (2025) 'Adaptive content recommendation for distance education based on fuzzy logic and knowledge graph', *Int. J. Information and Communication Technology*, Vol. 26, No. 20, pp.41–55.

Biographical notes: Chunmei Du obtained her Master's degree from the Hebei Normal University, China in 2020. She is currently an Associate Professor at Tangshan Open University. Her research interests include modern educational technology and distance education.

Donghai Xu received his Master's degree from the Changsha University of Science and Technology, China in 2012. He is currently an Associate Professor at Tangshan Open University. His research interests include artificial intelligence.

1 Introduction

The concept of lifelong education has been gradually accepted and recognised in today's society, and distance education as an important means of implementing lifelong education has made rapid development in recent years (Attr, 2012). Based on the development of science and technology, the transmission and sharing of resources has become more convenient and fast, and the application of computer network technology in the field of distance education breaks the boundaries of time and space, so that high-quality educational resources circulate quickly, and to a certain extent, solves the problem of uneven distribution of educational resources and geographic differentiation of the teacher’s strength (Mahmudi et al., 2023). However, the widening of information flow channels has also brought about the problem of information overload, and learners are often overwhelmed by the vast amount of information resources available to them. Adaptive distance learning system can personalise and push the distance learning content that fits the learners’ characteristics, indicate the navigated learning path for the learners, stimulate the learning motivation and accomplish the learning goals (Chen et al., 2018). Thus, how to accurately realise the adaptive content recommendation of distance education has far-reaching significance for improving the quality and efficiency of distance education.

Madani et al. (2020) obtained and analysed the data of learners in the process of distance learning, obtained the learning preferences of learners, and used genetic hierarchical recommendation algorithms to recommend appropriate courses for learners, which can improve the learning efficiency. Okubo et al. (2022) developed a content-based personalised adaptive learning recommendation model that can be used for ubiquitous learning, where learning resources are analysed through semantic descriptions of that learning, KNS networks, and learning activities. Venkatesh and Sathyalaksmi (2022) proposed an adaptive content recommendation system based on deep belief networks by combining the user and course feature vectors of a distance education platform to mine the user’s interest in the course, but the system suffers from the cold-start problem.

Knowledge graphs (KGs) can provide rich semantic relationships and contextual information to help distance education adaptive content recommendation systems understand users’ needs and interests more deeply, thus improving the accuracy of recommendations. Ma et al. (2023) built a distance education recommendation model for learning paths using a multidimensional KG framework, but it could not fully utilise the relationships between knowledge. Liu et al. (2022) built an adaptive learning system using KGs combining multiple disciplines to analyse the entire learning performance of a learner, plan a learning path for the learner, and make resource recommendations. Wu et al. (2022) recommended the most suitable learning resources from the distance learning resources database through the processes of KG construction, vector generation of knowledge points to be learned, and similarity iteration algorithm design. Huo et al. (2020) proposed an adaptive distance learning content recommendation method based on joint KG and temporal characteristics, but the accuracy of recommendation is not high.

In practice, learners are sometimes unable to describe their perception of a content with a precise score, which is less influenced by user subjectivity than the exact score. Therefore, the use of fuzzy feelings helps the algorithm to calculate user similarity more accurately, thus reducing the recommendation error. Then, how to deal with fuzzy feelings appropriately have become a difficult problem in similarity algorithms, and the application of fuzzy logic can solve this problem effectively. Fuzzy logic is mainly used to manage some concepts, objects or information, etc. which cannot be represented in a precise way in the real world (Klir and Yuan, 1996) and works mainly on fuzzy sets (Peng and Selvachandran, 2019) and affiliation functions (Povolotskaya and Mach, 2013). Wu et al. (2020) used trapezoidal fuzzy numbers to express personalised information containing user preferences and rating patterns, and constructed a rating fuzzy similarity model to improve recommendation accuracy. Alagarsamy et al. (2021) used a fuzzy tree to match learning content. However, due to the structural characteristics of the tree, it is not possible to correlate the learning contents of different paths, thus affecting the results of content recommendation. Abolghasemi et al. (2022) for the user's fuzzy preferences for items, and subsequently replace the ratings with fuzzy preferences for subsequent calculations, but relying only on a single surface rating is highly susceptible to user subjectivity, which biases some of the calculations.

To summarise, existing studies fail to fully consider learners' preferences and the dynamic uncertainty of learners' cognitive level, resulting in unsatisfactory recommendation accuracy, so this paper proposes an adaptive content recommendation method for distance education based on fuzzy logic and KG. Firstly, a top-down approach is adopted to construct the distance education KG (DEKG), design the KG ontology model, extract the data from the data source and perform knowledge fusion, and utilise the TransR model to perform the vectorised representation of the KG on the fused DEKG. Then, based on fuzzy logic, the learners are diagnosed and analysed to determine their cognitive level, and the Euclidean distance between each two knowledge points is calculated as the semantic similarity of knowledge points. The collaborative filtering algorithm is used to calculate the similarity of knowledge points according to the relationship between learners and knowledge points, and at the same time, the matching degree and cognitive level are integrated to calculate the final similarity of knowledge points. Secondly the degree of learner preference was measured using fuzzy logic to represent the knowledge point similarity as a vector over the content labels. Subsequently, based on the relationship between this vector and the predicted content labels, a corresponding score prediction formula is designed to realise more accurate mining of distance education content that meets the learner's own characteristics for recommendation. The experimental outcome indicates that the suggested approach not only improves the recommendation accuracy, but also reduces the number of neighbours needed for the optimal prediction of each metric and exhibits excellent recommendation performance.

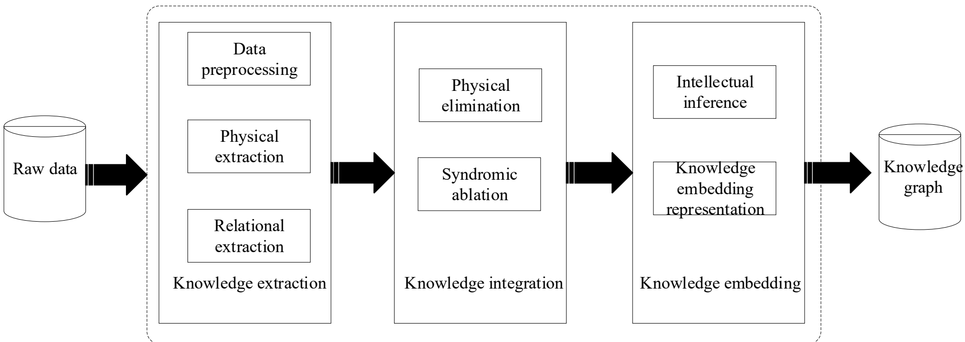
2 Relevant theoretical foundations

2.1 Knowledge graph

KG is a structured semantic knowledge base for describing concepts and their interrelationships in the physical world. It mainly consists of 'entity-relationship-entity' triples and entity-related attribute-value pairs. A KG is formed by combining entities and relations into a mesh structure (Guan et al., 2019). KG can increase the diversity of recommendations by connecting different kinds of relationships, connecting the user's interest history and recommendation results, making the recommendation results richer and more diverse. In addition, the semantic relations of KG make the recommendation results more interpretable, and learners can more easily understand the rationale and justification of the recommendations.

In generalised KG construction, there are three types of construction methods: fully automatic, semi-automatic, and fully manual. The fully automated approach relies on high-quality and large-volume data sources, and the accuracy of knowledge extraction is not satisfactory; the manual approach is of high quality, but there are some limitations in the scale of the map, which may not be able to satisfy the demand, and the cost of labour and time is immeasurable. From a technical point of view, the construction of KG involves various technologies such as knowledge representation, information extraction, knowledge fusion and knowledge storage (Zhong et al., 2023). The overall construction process of knowledge graph generalisation is shown in Figure 1.

Figure 1 The overall construction process of KG



2.2 Fuzzy logic theory

The theory of fuzzy logic (Mouzouris and Mendel, 1997) is a science based on multi-valued logic that uses the method of fuzzy sets to study fuzzy thinking, linguistic forms and their laws. Due to the fuzzy nature of the criteria, it is difficult to determine when an object belongs to this category. This leads to the concept of fuzzy sets as follows. For a given domain U, one can define a fuzzy subset H that is determined by a function $\mu_H(u)$ whose range is [0, 1]. This function reflects the membership degree from the element u on the domain to the fuzzy set H.

Fuzzy sets are used in information retrieval. First, all the content sets $D = \{d_1, d_2, d_3, \ldots, d_n\}$ and all the feature sets $T = \{t_1, t_2, t_3, \ldots, t_k\}$ are listed and the degree of correlation between different feature items is calculated. Then, with D as the domain and each feature term as a fuzzy set, this defines the degree of affiliation of a single content d in D to a fuzzy set as follows.

$$\mu_H(d_i) = 1 - \prod_{t_k \in d_i} (1 - c_{HK}) \tag{1}$$

where the fuzzy set H corresponds to the feature item t_h, the fuzzy set K corresponds to the feature item t_k, and c_{HK} is the correlation between the feature items t_h and t_k. The value of c_{HK} ranges from 0 to 1.

If fuzzy subset C and D have the same thesis U, where $u_1, u_2, u_3, \ldots, u_n$ is an element in the thesis. Let $\mu_c(u_i)$ and $\mu_D(u_i)$ be the affiliation functions of the elements in U for the fuzzy subsets C, D respectively, then the correlation relation ρ_{CD} between C and D is as follows.

$$\rho_{CD} = \frac{\sum_{i=1}^{n} \left(\mu_C(u_i) - \overline{\mu_C}\right)\left(\mu_D(u_i) - \overline{\mu_D}\right) / (n-1)}{S_c{}^{*} S_D} \tag{2}$$

$$\overline{\mu_C} = \frac{\sum_{i=1}^{n} \mu_C(u_i)}{n}, \overline{\mu_D} = \frac{\sum_{i=1}^{n} \mu_D(u_i)}{n} \tag{3}$$

$$S_C = \sqrt{\frac{\sum_{i=1}^{n} \left(\mu_C(u_i) - \overline{\mu_C}\right)^2}{n-1}}, S_D = \sqrt{\frac{\sum_{i=1}^{n} \left(\mu_D(u_i) - \overline{\mu_D}\right)^2}{n-1}} \tag{4}$$

where $\overline{\mu_C}, \overline{\mu_D}$, S_C, and S_D are the mean affiliation of C, the mean affiliation of D, the sample standardised variance of C, and the sample standardised variance of D, respectively.

In the recommender system, the content of interest provided by the user is used as the training set, a fuzzy set is established for the two categories, and then the content to be recommended is expressed as a fuzzy set, and the degree of relevance is calculated to determine whether to recommend the content or not.

3 Fine-grained distance education knowledge graph construction

From the introduction above, it is clear that a complete KG with high accuracy is crucial for realising adaptive content recommendation for distance education. In this paper, we adopt a top-down approach to construct a fine-grained DEKG oriented to knowledge points. First, the DEKG ontology model is designed, which mainly includes the analysis and determination of conceptual classes, data attributes, and object attributes. Then extract the data from the data source and perform knowledge fusion to form a knowledge graph.

1 DEKG ontology construction. The constructed DEKG includes conceptual classes, data attributes, and object attributes. The conceptual class is the entity, the data attributes are the attributes of the entity, and the object attributes are the relationships between the entities. Firstly, the goal of constructing DEKG is clarified and the conceptual classes in the domain are extracted accordingly, then the data attributes and object attributes of the conceptual classes are analysed and defined, and finally the ontology modelling tool is used to implement the DEKG ontology, as shown in Figure 2.

2 Knowledge Extraction. In the knowledge extraction task, many deep learning models are applied and greatly improve the effectiveness of the task and do not require the design of complex features. In this paper, we adopt the common BiLSTM-CRF model (Meng et al., 2022), which has been proved to have high extraction efficiency by previous research experiments. First, the sentences are converted into word vectors, and the word vectors of each word are fed into the BiLSTM layer, and the probability of each word corresponding to each label is obtained through the learning as described above and below. All the probabilities output from the previous layer will be used as inputs to the CRF layer, and the label order and other constraints learned in the training set will be used to get the prediction result of each word. The overall model structure is shown in Figure 3.

Figure 2 DEKG construction (see online version for colours)

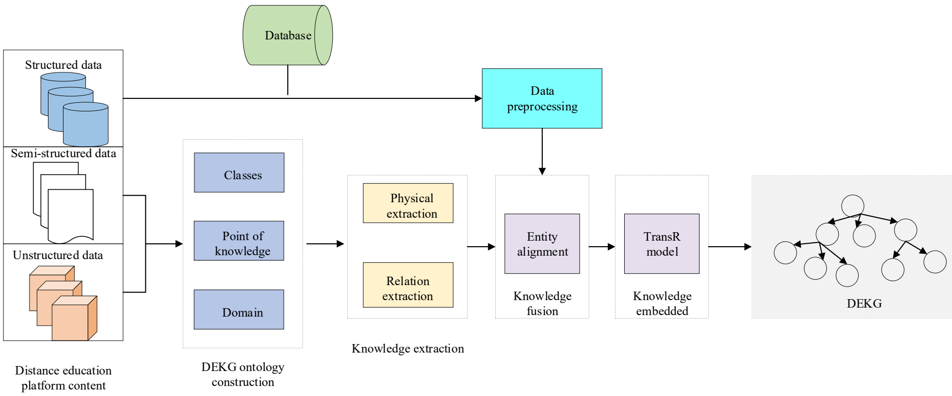
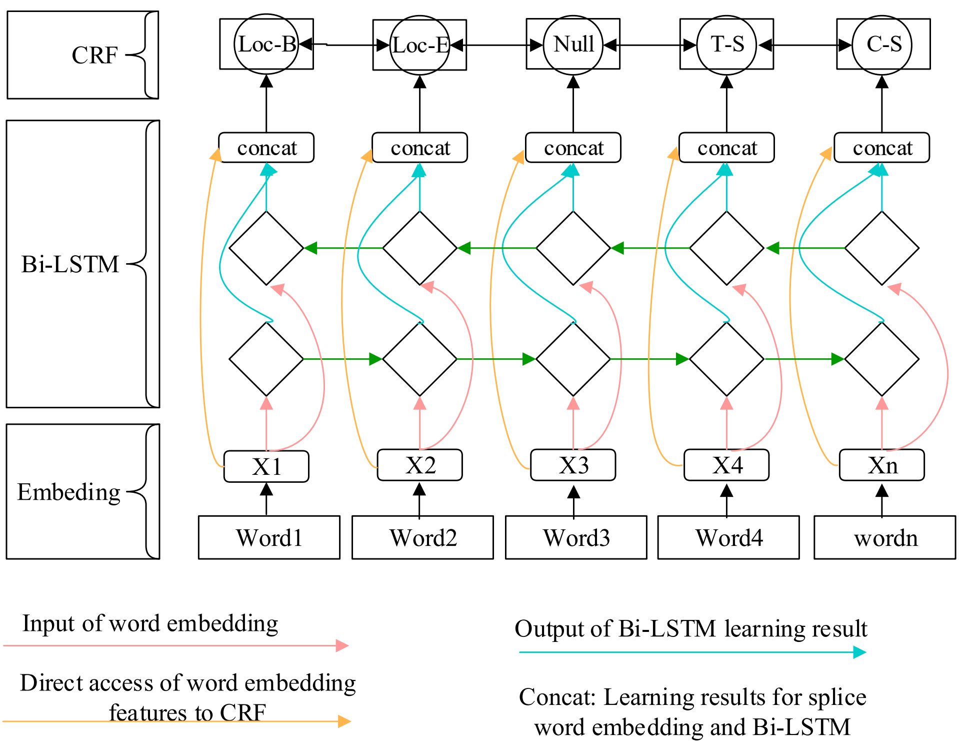


Figure 3 The structure of BiLSTM-CRF (see online version for colours)



3 Knowledge fusion. After the knowledge has been extracted, the data obtained is in a more heterogeneous form, and there may be entities with duplicated meanings and confusing structures. Therefore, there is a need for knowledge fusion, which takes a large number of named entities extracted from knowledge and performs steps such as data cleansing and entity alignment to improve the quality of DEKG. Firstly, redundant, incorrect, and useless data in KG are removed, and then entities with

repeated meanings in DEKG are uniformly named, for example, 'data structure' and 'data structure course' are named 'data structure'.

4 KG embedding representation. In order to vectorise the representation of entities and their relationships in the fused DEKG, this paper adopts the TransR model, which is currently the most efficient, as the knowledge embedding tool. For each triad (h, r, t), the entities h, t originally belonging to the entity space are mapped into the relation space by the projection matrix M_r of the relation r, forming h_r, t_r. This projection brings entities with relationship r closer to each other, while keeping entities without relationship r away from each other. The mapping of the head and tail entities of the TransR model is shown below.

$$h_r = hM_r \tag{5}$$

$$t_r = tM_r \tag{6}$$

For the mapped entities h_r, t_r and relation r, during the training process of the TransR model, it is necessary to make $h_r + r \approx t_r$ as much as possible, so the score function is as follows

$$f_r(h, t) = \|h_r + r - t_r\|_2^2 \tag{7}$$

4 Adaptive content recommendation for distance education based on fuzzy logic and knowledge graph

4.1 Fuzzy logic-based diagnosis of learners' cognitive level

To address the problem that existing recommendation methods fail to fully consider the learner's learning purpose and the dynamic uncertainty of the learner's cognitive level, firstly, according to the DEKG constructed above, diagnose and analyse the learner based on fuzzy logic, determine the learner's cognitive level, and then integrate the matching degree and cognitive level to compute the similarity between the knowledge points in the content. After cognitive diagnosis, the degree of learner preference was measured using fuzzy logic to represent knowledge point similarity as a vector over content labels. Subsequently, based on the relationship between the vector and the label of the predicted target item, the corresponding similarity calculation and score prediction formulas are designed to realise more effective and accurate mining of distance education content that meets the learner's own characteristics for recommendation. The recommended method flow is shown in Figure 4.

For each knowledge point k in the constructed DEKG there exists a fuzzy set (J, μ_k) corresponding to it, μ_k: $J \rightarrow [0, 1]$ is the corresponding affiliation function, and for any learner j in J, the learner's level of knowledge α_{jk} is equal to the affiliation in the corresponding fuzzy set $\mu_k(j)$.

$$\alpha_{jk} = \mu_k(j) = \frac{1}{1 + \exp\left[-1.7a_{jk}\left(\theta_j - b_{jk}\right)\right]} \tag{8}$$

where α_{jk} is learner j's mastery of knowledge point k, which is equal to $\mu_k(j)$, a is the knowledge point differentiation, b is the knowledge point difficulty, θ_j is the learner's latent traits, and –1.7 is an empirical constant.

In the designed cognitive diagnosis of fuzzy logic, the knowledge points are compensated for each other and the objective content is considered to contain all the knowledge points examined; the better mastered knowledge points in the subjective content are considered to compensate for the less mastered knowledge points. According to the definitions of fuzzy intersection and fuzzy concatenation in equations (9) and (10), the learners' mastery of the knowledge points on objective content and subjective content are obtained as shown in equations (11) and (12), respectively.

$$(A \cap B)(x) = MIN\{A(x), B(x)\} \tag{9}$$

$$(A \cup B)(x) = MAX\{A(x), B(x)\} \tag{10}$$

$$(Obj)\eta_{ji} = MIN\left(\alpha_{jk}\right),\ 1 \le k \le K \tag{11}$$

$$(Sub)\eta_{ji} = MAX\left(\alpha_{jk}\right),\ 1 \le k \le K \tag{12}$$

For mastery of subjective and objective content, the proposed model uses Bernoulli and Gaussian distributions for objective and subjective content, respectively, as follows.

$$(Obj)P\left(R_{ji} \middle| \eta_{ji}, s_i, g_i\right) = (1 - s_i)\eta_{ji} + g_i\left(1 - \eta_{ji}\right) \tag{13}$$

$$(Sub)P\left(R_{ji} \middle| \eta_{ji}, s_i, g_i\right) = N\left(R_{ji} \middle\| \left[(1 - s_i)\eta_{ji} + g_i\left(1 - \eta_{ji}\right)\right] \middle\| \sigma^2\right) \tag{14}$$

where R_{ji} is learner j's score on point i, and s_i and g_i are i's failure rate and guessing rate, respectively. $(1 - s_i)\eta_{ji}$ is when the student masters the objective content correctly without errors, and $g_i(1 - \eta_{ji})$ is when the student fails to master the objective content correctly. The probability of correctly mastering subjective content obeys a Gaussian distribution with $(1 - s_i)\eta_{ji} + g_i(1 - \eta_{ji})$ as the mean and σ^2 as the variance.

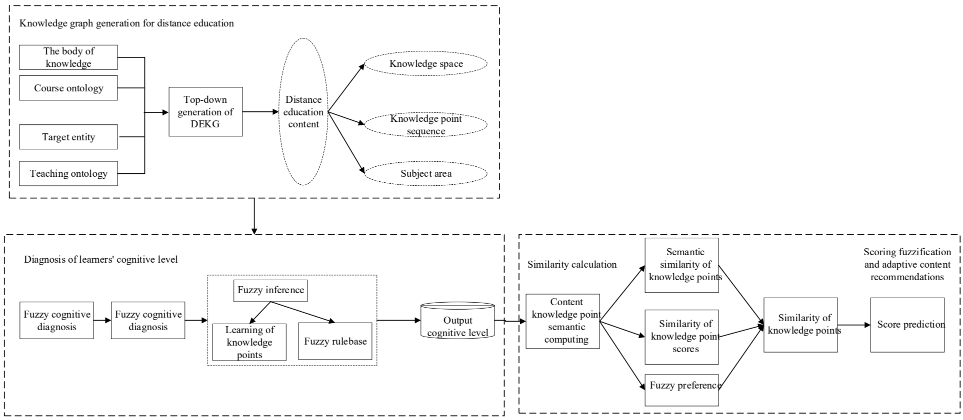
4.2 Content knowledge point semantics and scoring similarity calculations

After the diagnostic analysis of learners' cognitive levels, the matching and cognitive levels are fused to compute the semantic similarity between the knowledge points. All the entities and relations in DEKG are embedded as d-dimensional semantic vectors. Take any two knowledge points S_i and S_j and let their semantic vectors in the low-dimensional space be $\vec{S_i} = (E_{1i}, E_{2i}, \ldots, E_{di})^T$, $\overrightarrow{S_j} = (E_{1j}, E_{2j}, \ldots, E_{dj})^T$. The similarity between any knowledge points can be calculated using Euclidean distance as shown in equation (15).

$$sim_{sg}\left(S_i, S_j\right) = \frac{1}{1 + D_e\left(\vec{S_i}, \overline{S_j}\right)} = \frac{1}{1 + \sqrt{\sum_{p=1}^{d}\left(E_{pi} - E_{pj}\right)^2}} \tag{15}$$

where $D_e(\vec{S_i}, \overline{S_j})$ is the Manhattan distance, $D_e(\vec{S_i}, \vec{S_j}) = \sqrt{\sum_{p=1}^{d}(E_{pi} - E_{pj})^2}$.

Figure 4 The recommended method flow



From equation (15), if two knowledge points are similar, then whether or not they are learned by different learners should also be similar. Then the similarity of scores between knowledge points S_i and S_j is calculated based on the cosine of the angle between the two knowledge point learning vectors. When the two vectors are in the same direction, the cosine value is 1, and the similarity between the ratings of the two knowledge points is maximised; when the two vectors are orthogonal, the cosine value is 0, and the similarity between the ratings of the two knowledge points is minimised, as shown in equation (16).

$$sim_{cf}\left(S_i, S_j\right) = \cos\left(S_i, S_j\right) = \frac{S_i \cdot S_j}{\left\|S_i\right\|} = \frac{\sum_{u=1}^{U} S_{ui} \cdot S_{uj}}{\sqrt{\sum_{u=1}^{U} S_{ui}^2} \cdot \sqrt{\sum_{u=1}^{U} S_{uj}^2}} \tag{16}$$

Because of the different hierarchical structure of knowledge points, even if the matching and cognitive levels of and are identical, it does not mean that they are the same. In the paper, the quality, quantity and hierarchical structure of the knowledge points are taken into consideration, and the final similarity of the knowledge points is obtained by combining $sim_{sg}(S_i, S_j)$ and $sim_{cf}(S_i, S_j)$ and setting the fusion factor, as shown below.

$$sim_{ij} = \alpha \times sim_{sg}\left(S_i, S_j\right) + \beta \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{S_i S_j}{\mu_{ij}} + (1 - \alpha - \beta) \times sim_{cf}\left(S_i, S_j\right) \tag{17}$$

where μ_{ij} is the student's cognitive level, α and β are fusion factors.

4.3 Scoring fuzzification and adaptive content recommendations

In the above calculation of the similarity of knowledge points, the learners' preference is not considered, for this reason, this paper firstly analyses the learners in the scoring matrix one by one to derive the learners' fuzzy preference for each knowledge point. The corresponding liking index λ_1 or disliking index λ_2 is subsequently calculated by means of the affiliation function in fuzzy logic. The application of the fuzzy affiliation function makes the calculation of sim_{ij} more objective, logical and accurate, and helps to reduce the final score prediction error.

When a learner's rating of a knowledge point in the content is greater than its mean rating, it indicates that the learner's preference is like and vice versa. The use of the mean of each learner's individual ratings instead of the mean of all learners' ratings is intended to better integrate the ratings with the user's scoring habits in order to reduce the preference error. After obtaining the user's preference for the content, the learners' fuzzy preference index for the content can be obtained by applying equations (18) and (19).

$$\lambda_1\left(r_{qi}\right)=\begin{cases}1, & r_{qi}=r_q^{\max}\\ \dfrac{r_{qi}-\overline{r}_q}{r_q^{\max}-\overline{r}_q}, & \overline{r}_q\le r_{qi}<r_q^{\max}\\ 0, & r_{qi}<\overline{r}_q\end{cases} \tag{18}$$

$$\lambda_2\left(r_{qi}\right)=\begin{cases}1, & r_{qi}=r_q^{\min}\\ \dfrac{\overline{r}_q-r_{qi}}{\overline{r}_q-r_q^{\min}}, & r_q^{\min}<r_{qi}\le\overline{r}_q\\ 0, & r_{qi}>\overline{r}_q\end{cases} \tag{19}$$

where r_{qi} is learner q's rating of knowledge point i; r_q is the average rating; $r_q^{\min}$ and $r_q^{\max}$ are learner q's lowest and highest ratings in history, respectively; $\lambda_1(r_{qi})$ is q's fuzzy preference like index for knowledge point i; and $\lambda_2(r_{qi})$ is the dislike index.

The learners' evaluation of content knowledge points is transformed from subjective scoring values to a pair of fuzzy preference indices. Obviously, the global similarity between learners is also based on these two fuzzy preference indices, so the similarity of the liking index and the similarity of the disliking index between learners need to be computed as shown in equations (20) to (21). Finally, the preference index similarity is incorporated into equation (15) to calculate the global similarity between learners, as shown in equation (22).

$$sim_{like}=T'^{\circ}\sum_{i=1}^{n}\left[g_{qi}g_{wi}\left[1-\left|\lambda_1\left(r_{qi}\right)-\lambda_1\left(r_{wi}\right)\right|\right]\cdot L_i\right] \tag{20}$$

$$sim_{dislike}=T'^{\circ}\sum_{i=1}^{n}\left[g_{qi}g_{wi}\left[1-\left|\lambda_2\left(r_{qi}\right)-\lambda_2\left(r_{wi}\right)\right|\right]\cdot L_i\right] \tag{21}$$

$$sim_{ij}^{'}=sim_{ij}+\tau sim_{like}+(1-\tau)sim_{dislike} \tag{22}$$

where n is the number of content knowledge points jointly rated by learners q and w. g_{qi} and g_{wi} are whether q and w have rated knowledge point i. L_i is the label vector of knowledge point i. The symbols ° denotes the Hadamard product of the vectors.

In the rating prediction, it is necessary to base on the ratings of the neighbour learners on the content knowledge points, in order to improve the efficiency of the recommender system and the speed of the algorithm, take the top N neighbour learners with the highest similarity to the target learner to form the predicted user set, and in the case that the number of neighbour users is less than N, then all the neighbour users will be selected as the predicted user set. When rating prediction is performed, the similarity between the target user and its neighbouring users for the item to be predicted varies depending on the tags contained in the item. Since the global similarity A between two users is a vector, it is transformed into a scalar by equation (23) to be computed, followed by equation (24) for score prediction. The specific formula is as follows.

$$SIM_{ij} = \sqrt{\sum_{l=1}^{k} \left(sim'_{ijl} \right)^2} \tag{23}$$

$$r_{qi} = \overline{r}_q + \frac{\sum_{w=1}^{N} \left[\left(r_{wi} - \overline{r}_w \right) SIM_{ij} \right]}{SIM_{ij}} \tag{24}$$

where SIM_{ij} is the weight of each neighbour of the target learner in the final prediction value; N is the total number of neighbours eventually taken by the target learner q, $\overline{r}_w$ is the average rating of the learner w, r_{wi} is the rating of the learner w on the knowledge point i, and r_{qi} is the final prediction rating.

5 Experimental results and analyses

To evaluate the effectiveness of the proposed recommendation method FLKG in practice, the course dataset of a distance education platform collected by Romero and Ventura (2017), which contains 26,387 learning records and evaluation information of 3,205 students and 109 course learning resources, is used. In the experiment, the dataset was randomly divided into ten copies for the experiment through ten-fold cross-validation, where any one of the copies could be used as the test set for the remaining nine copies, and the average of the ten experiments was taken as the final result of the experiment. For all methods, the variables used in this paper are the same, and the experiments were conducted with the number of neighbours from 1 to 200. The experimental environment is based on Python 3.7, Tensorflow framework, Pycharm2022, Intel(R) Core(TM) i7-12700H processor, and NVIDIA GeForce RTX 3070 Ti Laptop GPU.

Recommender system as a kind of information prediction tools, its output recommendation results contain a lot of information, so when evaluating the recommendation effect of the model needs to be considered from multiple perspectives. In this paper, mean absolute error (MAE), half-life utility (HLU), AUC, Recall@N, F1@N, and NDCG@N are selected as evaluation metrics, where N denotes the number of course content. The comparison methods were selected as LSTM-KG (Huo et al., 2020), CFR-IV (Wu et al., 2020), FCRSA (Alagarsamy et al., 2021), and PA-GRS (Abolghasemi et al., 2022), and the comparisons of MAE and HLU for different methods are shown in Figure 5.

From Figure 5(a), it can be seen that when the number of neighbours is less than 22, the MAE value of FLKG proposed in this paper is always the lowest and the best. Meanwhile, in the range of the number of neighbours less than 40, FLKG improves the MAE value by 12.38% at most and 2.59% on average compared with the comparison method with the same number of neighbours. In addition, when the number of neighbours of the optimal value of FLKG is taken, the MAE value is improved by 1.57%. From Figure 5(b), it can be seen that the FLKG proposed in this paper achieves the global HLU maximum of 1.26 at the number of neighbours of 20 compared to other methods. FLKG outperforms the other four algorithms in terms of HLU for all values of the number of neighbours, with a maximum improvement of 13.30% and an average improvement of 4.45% for the same number of neighbours, and an improvement of 7.02% for the optimal number of neighbours for this algorithm.

Figure 5 The comparison of MAE and HLU for different methods, (a) comparison of MAE for different (b) comparison of HLU for different (see online version for colours)

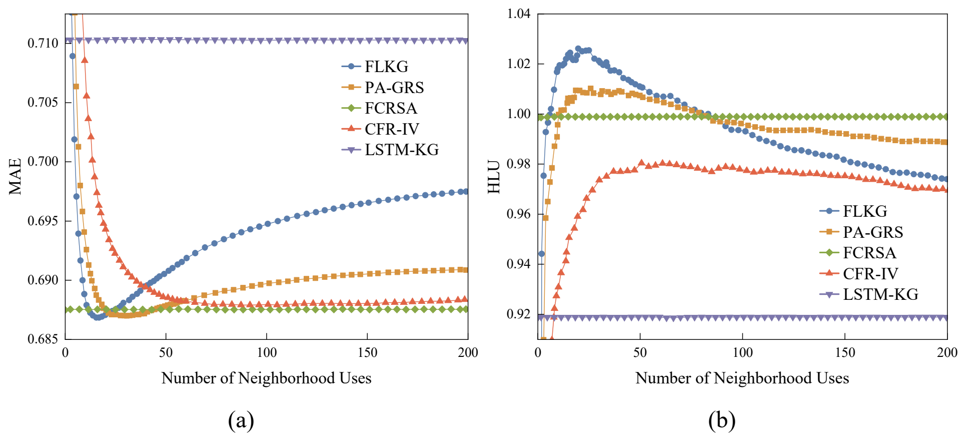


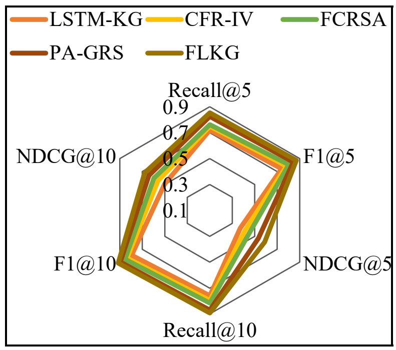
Table 1 Comparison of AUC under different training set ratios

Method	*20%*	*40%*	*60%*	*80%*	*100%*
LSTM-KG	0.781	0.745	0.768	0.715	0.804
CFR-IV	0.824	0.811	0.831	0.802	0.847
FCRSA	0.859	0.846	0.854	0.838	0.872
PA-GRS	0.863	0.878	0.897	0.872	0.907
FLKG	0.892	0.904	0.937	0.915	0.942

To verify that using KG as auxiliary information can effectively improve the accuracy of recommendation, the validation set and the test set are kept unchanged, and the ratio of the training set is adjusted to compare the AUC, as shown in Table 1. The new training data are randomly selected from the original training set at a ratio of 20% to 1. As the ratio of the training set increases, the AUC also increases. However, FLKG's recommendation outperforms the comparison method for any proportion of training data. When the training set ratio is 60%, the AUC of FLKG is 0.937, which is 4.46%, 9.72%, 12.76% and 22.01% higher than that of LSTM-KG, CFR-IV, FCRSA and PA-GRS, respectively. In the case of sparse training data, the AUC of FLKG is 0.937. The FLKG method with KG was the most effective, which verified the effectiveness of the introduction of KG.

When the number of contents N is taken as 5 and 10, the recommended performance comparison of different methods is shown in Figure 6. When N is taken as 5, FLKG's Recall and F1 are improved by at least 3.21% and NDCG is improved by at least 10.74% compared to the other four methods. When N is taken as 10, recall, F1 and NDCG of FLKG are 0.8926, 0.9139 and 0.6789 respectively, which are higher than the comparison methods. LSTM-KG does not take into account the fuzzy preferences of learners, although it considers the introduction of KG to adaptive content recommendation. CFR-IV, FCRSA, and PA-GRS are all fuzzy logic-based recommendation methods, and PA-GRS takes into account learners' fuzzy preferences, but does not model course knowledge points to construct the corresponding KG. CFR-IV and FCRSA, on the other hand, only utilise fuzzy logic to build similarity models, so the recommendation performance of CFR-IV and FCRSA is not as good as that of PA-GRS.FLKG not only builds DEKGs, but also takes the learner's cognitive level and fuzzy preference into account, which greatly improves the recommendation effect.

Figure 6 Comparison of recommended performance indicators for different methods (see online version for colours)



Although the model proposed in this paper achieves better recommendation performance, there are still problems such as wasted vector space and large subjective influence of learners, which will be investigated in the following two aspects in the future.

1 Simplify or downsize the content label vectors to reduce the waste of vector space to increase the scalability of the algorithm; and will try to expand the fuzzy set and deepen the degree of fuzzy to further reduce the subjective influence of the learner and improve the accuracy of similarity calculation.

2 Comparative experiments of the proposed model on multiple datasets are conducted to verify the generalisation and robustness of the proposed model under different situations.

6 Conclusions

With the explosive growth of distance education learning content, learners are faced with the problem of information overload, which leads to the low efficiency of resource searching and the difficulty of information screening in the process of adaptive learning. Therefore, this paper proposes an adaptive content recommendation method for distance education based on fuzzy logic and KG. The DEKG is first constructed using a top-down approach, and the TransR model is utilised to provide a vectorised representation of entities and relationships for the DEKG. Then, based on fuzzy logic, the learners are diagnosed and analysed to determine their cognitive level, and the Euclidean distance between each two knowledge points is calculated as the semantic similarity of knowledge points. The collaborative filtering algorithm is used to calculate the similarity of knowledge points according to the relationship between learners and knowledge points, and at the same time, the matching degree and cognitive level are integrated to calculate the final similarity of knowledge points. Secondly, fuzzy logic is used to measure the degree of learner preference, the similarity of knowledge points is expressed as a vector on the content label, and the content label is combined with its corresponding fuzzy preference index, constituting a fuzzy preference label vector and calculating the scoring prediction formula, to realise a more accurate distance education adaptive content for recommendation. The experimental outcome indicates that the MAE and AUC values of the suggested method are both improved to a large extent compared to the comparison model, and have a high recommendation accuracy. In the subsequent work will try to simplify or downscale the content label vectors to reduce the waste of vector space to increase the scalability of the algorithm; and will try to expand the fuzzy set and deepen the degree of fuzzy, to further alleviate the learner's subjective influence and to improve the accuracy of similarity calculation.

Declarations

All authors declare that they have no conflicts of interest.

References

Abolghasemi, R., Engelstad, P., Herrera-Viedma, E. et al. (2022) 'A personality-aware group recommendation system based on pairwise preferences', *Information Sciences*, Vol. 595, pp.1–17.

Alagarsamy, R., Arunpraksh, R., Ganapathy, S. et al. (2021) 'A fuzzy content recommendation system using similarity analysis, content ranking and clustering', *Journal of Intelligent & Fuzzy Systems*, Vol. 41, No. 6, pp.6429–6441.

Attr, A.K. (2012) 'Distance education: problems and solution', *International Journal of Behavioral Social and Movement Sciences*, Vol. 1, No. 4, pp.42–58.

Chen, Y., Li, X., Liu, J. et al. (2018) 'Recommendation system for adaptive learning', *Applied Psychological Measurement*, Vol. 42, No. 1, pp.24–41.

Guan, N., Song, D. and Liao, L. (2019) 'Knowledge graph embedding with concepts', *Knowledge-Based Systems*, Vol. 164, pp.38–44.

Huo, Y., Wong, D.F., Ni, L.M. et al. (2020) 'Knowledge modeling via contextualized representations for LSTM-based personalized exercise recommendation', *Information Sciences*, Vol. 523, pp.266–278.

Klir, G.J. and Yuan, B. (1996) 'Fuzzy sets and fuzzy logic: theory and applications', *Possibility Theory versus Probab. Theory*, Vol. 32, No. 2, pp.207–208.

Liu, Y., Gu, F., Gu, X. et al. (2022) 'Resource recommendation based on industrial knowledge graph in low-resource conditions', *International Journal of Computational Intelligence Systems*, Vol. 15, No. 1, p.42.

Ma, Y., Zhang, X., Gao, C. et al. (2023) 'Enhancing recommendations with contrastive learning from collaborative knowledge graph', *Neurocomputing*, Vol. 523, pp.103–115.

Madani, Y., Ezzikouri, H., Erritali, M. et al. (2020) 'Finding optimal pedagogical content in an adaptive e-learning platform using a new recommendation approach and reinforcement learning', *Journal of Ambient Intelligence and Humanized Computing*, Vol. 11, pp.3921–3936.

Mahmudi, A.A., Fionasari, R., Mardikawati, B. et al. (2023) 'Integration of artificial intelligence technology in distance learning in higher education', *Journal of Social Science Utilizing Technology*, Vol. 1, No. 4, pp.190–201.

Meng, F., Yang, S., Wang, J. et al. (2022) 'Creating knowledge graph of electric power equipment faults based on BERT-BiLSTM-CRF model', *Journal of Electrical Engineering & Technology*, Vol. 17, No. 4, pp.2507–2516.

Mouzouris, G.C. and Mendel, J.M. (1997) 'Nonsingleton fuzzy logic systems: theory and application', *IEEE Transactions on Fuzzy Systems*, Vol. 5, No. 1, pp.56–71.

Okubo, F., Shiino, T., Minematsu, T. et al. (2022) 'Adaptive learning support system based on automatic recommendation of personalized review materials', *IEEE Transactions on Learning Technologies*, Vol. 16, No. 1, pp.92–105.

Peng, X. and Selvachandran, G. (2019) 'Pythagorean fuzzy set: state of the art and future directions', *Artificial Intelligence Review*, Vol. 52, pp.1873–1927.

Povolotskaya, E. and Mach, P. (2013) 'Affinity diagram as a risk analysis tool with support of fuzzy logic', *Advanced Science Letters*, Vol. 19, No. 8, pp.2391–2394.

Romero, C. and Ventura, S. (2017) 'Educational data science in massive open online courses', *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol. 7, No. 1, p.e1187.

Venkatesh, M. and Sathyalaksmi, S. (2022) 'Memetic swarm clustering with deep belief network model for e-learning recommendation system to improve learning performance', *Concurrency and Computation: Practice and Experience*, Vol. 34, No. 18, p.e7010.

Wu, C., Liu, S., Zeng, Z. et al. (2022) 'Knowledge graph-based multi-context-aware recommendation algorithm', *Information Sciences*, Vol. 595, pp.179–194.

Wu, Y., ZHao, Y. and Wei, S. (2020) 'Collaborative filtering recommendation algorithm based on interval-valued fuzzy numbers', *Applied Intelligence*, Vol. 50, No. 9, pp.2663–2675.

Zhong, L., Wu, J., Li, Q. et al. (2023) 'A comprehensive survey on automatic knowledge graph construction', *ACM Computing Surveys*, Vol. 56, No. 4, pp.1–62.