Online-Journals.org

iJET International Journal of Emerging Technologies in Learning

iJET | eISSN: 1863-0383 | Vol. 18 No. 18 (2023) | OPEN ACCESS

https://doi.org/10.3991/ijet.v18i18.43503

PAPER

Data-Driven Prediction of Students' Online Learning Needs and Optimization of Knowledge Library Management

Jiajia Li(✉)

School of Electronic Information Engineering, Hebi Polytechnic, Hebi, China

hbljj123456@163.com

ABSTRACT

Thanks to the advancement of information technology, online learning has become a crucial tool of modern education, and the management of modern education is facing the challenges of how to effectively predict students' learning needs and how to optimize the management of the knowledge library to support these needs. However, existing data-driven prediction approaches are flawed in handling complex learning environments and timely adapting to changes, so this study attempts to solve these questions by exploring the correlation between learning needs, the extraction of knowledge linkages, and the optimization of knowledge libraries based on rule updates. In our work, a new method was proposed for extracting learning needs and knowledge linkages to more accurately identify and predict students' learning needs, and a rule-based knowledge library management optimization method was introduced to allow the knowledge library to more flexibly adapt to students' learning needs and the changes in educational resources. In this way, this study offers a comprehensive and flexible solution for education management via the combination of these two aspects, which not only increases student satisfaction and improves teaching quality but also reduces resource waste and gives students a more personalized and efficient learning experience. Furthermore, the methods and findings of this study could also be used as references for data-driven decision-making in other fields.

KEYWORDS

data-driven, learn online, learning needs, knowledge library management

1 INTRODUCTION

In this era of information technology, constant reform and progress are seen in modern education. In this context, students' learning styles and needs are changing dramatically these days, and online learning is regarded as a common trend in modern education [1] [2]. Such big changes have posed great challenges for

Li, J. (2023). Data-Driven Prediction of Students' Online Learning Needs and Optimization of Knowledge Library Management. *International Journal of Emerging Technologies in Learning (iJET)*, 18(18), pp. 150–164. https://doi.org/10.3991/ijet.v18i18.43503

Article submitted 2023-05-27. Revision uploaded 2023-07-21. Final acceptance 2023-07-28.

education management [3–6], as educators must master students' learning needs in a timely manner and support and satisfy these needs through effective knowledge library management. However, as student diversity and learning content complexity increase constantly, traditional educational management methods can no longer meet the new learning needs [7] [8].

To cope with these matters, now that a number of educational institutions and researchers have begun to pay attention to the application of data-driven methods in education management, it's been found that students' learning needs can be predicted more accurately through the analysis of students' online learning behavior and feedback [9–11], and educational resources could be updated and optimized effectively via data analysis and management of knowledge libraries, so as to meet students' personalized learning needs [12–16]. However, existing studies mostly talk about every single analysis and prediction method, and few of them have been concerned about how to effectively combine these methods with knowledge library management.

Although data-driven prediction approaches have great potential in education management, there are a series of shortcomings with currently available methods. For instance, existing prediction methods of learning needs cannot deal with complex and dynamic learning environments, so they may not be able to accurately reflect the real needs of students [17–20]. The existing knowledge library management methods are not flexible enough to timely adapt to changes in students' learning needs and the new educational resources. Additionally, little research has considered how to optimize learning need prediction through knowledge library management or how to optimize knowledge library management according to the prediction of learning needs [21–24].

Regarding the blank in the related research field mentioned above, this study attempts to explore the correlation between learning needs, the extraction of knowledge linkages, and the optimization of the knowledge library based on rule updates. At first, a new method was proposed for extracting learning needs and knowledge linkages to more accurately identify and predict students' learning needs. Then, a rule-based knowledge library management optimization method was introduced to allow the knowledge library to more flexibly adapt to students' learning needs and the changes in educational resources. This method not only updates the content of the knowledge library based on real-time feedback and behavior data from students but also adjusts according to predefined rules to improve the response speed and accuracy of the knowledge library. By combining these two aspects, this study provides a comprehensive and flexible solution for education management to provide more effective support for students' online learning.

The significance of this study lies in that it increases the accuracy of learning need prediction and enhances the adaptability and pertinence of educational resources via knowledge library management optimization. For educational institutions, the findings of this study can improve teaching quality and student satisfaction and reduce the waste of educational resources, while students can get more personalized and efficient learning experiences. Moreover, the methods and findings of this study could provide references and insights for data-driven decision-making in other fields and have very good application prospects and values.

2 EXTRACTION OF LEARNING NEEDS AND KNOWLEDGE LINKAGES

In the research topic of this study, relationship extraction is a core component that helps to reveal students' online learning needs and the internal associations between knowledge in the knowledge library. By figuring out these relationships, educational institutions can make a more accurate judgment on students' learning needs, thereby making wiser and more effective decisions about knowledge library management.

The adopted relation extraction model in this study offers evident advantages in relation extraction. Firstly, the introduced *Attention* mechanism can focus on the part that is most relevant to the prediction task in the study, through which the model can place its focus on the relational extraction of the most important words and phrases, thereby improving prediction accuracy. Secondly, the introduced *BiGRU*, namely the Bi-Gated Recurrent Unit, can capture richer semantic information through forward and backward text processing, which is very important for understanding complex learning needs and knowledge linkages. Lastly, the *Gumbel Tree* gate control unit was introduced to capture hierarchical information. Since the relationship between the knowledge required by students' online learning and the knowledge linkages in the knowledge library are usually complicated and hierarchical, the *Gumbel Tree* gate can help the model better understand these hierarchies. Figure 1 gives a diagram showing the architecture of text sentence embedding in the model.

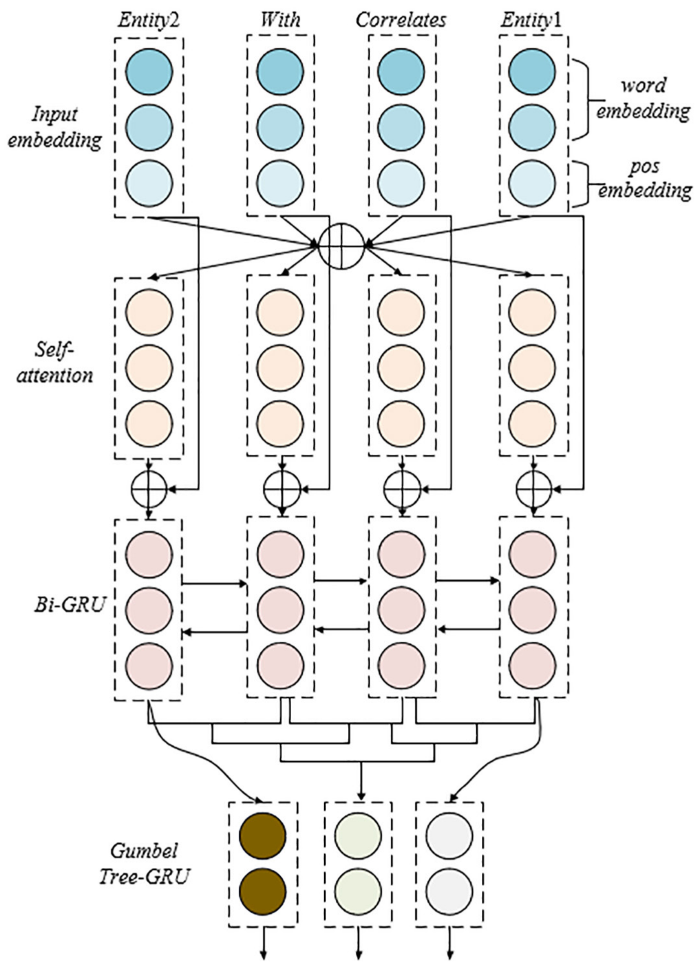

Fig. 1. Text sentence embedding in the model

When building the relation extraction model, text-based data samples were required for training and validation. These data samples usually contain sentences, text segments, tagged entities, and the relations between them. Here are two examples of the data samples: Text: "Many students say that they are confused when learning linear algebra"; Entities: student, linear algebra; Relation: student confused by linear algebra; another example: Text: "To help beginners, the knowledge library of the programming course has been updated to the *Python* Tutorial"; Entities: programming course, *Python* Tutorial; Relation: programming course updated knowledge library *Python* Tutorial.

The self-attention mechanism is the mapping of a *Query* to a series of key-value pairs <*Key*, *Value*>. The self-attention layer processes the input word vectors, allowing the model to focus on relevant information about the current task. It helps to capture long-term dependencies in sentences, even if they are far from each other in the text. For complex texts such as learning materials or educational resources, such long-term dependencies might be very important. The self-attention mechanism enables the model to capture such relations, thereby improving the accuracy of relation extraction.

Based on the *Query* values corresponding to sample data, the weight coefficients corresponding to each *Key* value can be obtained by calculating the similarity between *Query* and *Key* values. Assuming: *Mz* represents the length of *Source* consisting of <*Key*, *Value*> data pairs, *Q* represents the value of given elements, then the value of *Attention* is the weighted sum of these weight coefficients, namely:

$$AT\left(Q,Origin\right)=\sum_{u=1}^{Mz}SIM\left(Q,K_u\right)*V_u \tag{1}$$

In this study, the *Attention* mechanism has been summarized into three stages. First, the model determines the focus of its attention, namely to decide which parts of the input data it should pay attention to, and this process is usually accomplished by generating a query vector; each part of the input data is expressed as a key value pair, and the model calculates the match score between the query vector and each key, and a higher score means more attention is required for this part of the data by the model. Second, if the scores are normalized so that the sum of them is 1, then the scores can be interpreted as a probability distribution. Third, according to this probability distribution, the model carries out weighted averaging on the values and generates the final output. In the first stage, the similarity between the *Query* and each *Key* can be obtained by calculating the vector dot product of the two:

$$SIM\left(Issue,Crux_u\right)=Issue\cdot Crux_u \tag{2}$$

The vector *Cosine* similarity of *Query* and each *Key* can be solved by the following formula:

$$SIM\left(Issue,Crux_u\right)=\frac{Issue\cdot Crux_u}{\left\|Issue\right\|\cdot\left\|Crux_u\right\|} \tag{3}$$

or can be calculated based on the *MLP* neural network.

$$SIM\left(Issue, Crux_u\right) = SJWL\left(Issue, Crux_u\right) \tag{4}$$

In the second stage, in order to find out the probability distribution corresponding to the *Value* so as to get weight coefficients, this study chose to normalize the similarity between *Query* and *Key* based on the *Softmax* function. Assuming β_u represents the attention distribution, ZYL_u represents the scoring mechanism of attention, denoted as $a(z_u, w)$, that is:

$$\beta_u = \text{softmax}\left(ZYL_u\right) = \frac{e^{ZYL_u}}{\sum_{k=1}^{M_Z} e^{ZYL_k}} \tag{5}$$

To get the *Attention* value, the weighted sum of the calculation results of attention distribution was attained based on the following formula, that is:

$$AT\left(Issue, Oirgin\right) = \sum_{u=1}^{M_Z} \beta_u \cdot Worth_u \tag{6}$$

The *BiGRU* extracts word features by bi-directional text processing, it takes into account both the preceding and subsequent text information and generates a rich feature representation for each word. Such bi-directional processing allows the model to get a more comprehensive understanding of the text, which is conducive to understanding the complex relations between entities.

Assuming: x_y and s_y respectively represent the update gate and reset gate of BiGRU; $\tilde{g}$ represents the current candidate hidden state, g_y represents the output, f represents the nonlinear activation function *tanh*, then the formulas for calculating each state of *GRU* are:

$$\begin{aligned} X_y &= \delta\left(q_x \cdot \left[g_{y-1}, z_y\right] + n_x\right) \\ s_y &= \delta\left(q_e \cdot \left[g_{y-1}, z_y\right] + n_e\right) \\ \tilde{g} &= f\left(Q \cdot \left[s_y * g_{y-1}, z_y\right] + n_y\right) \\ g_y &= \left(1 - x_y\right) * g_{y-1} + x_y * \tilde{g} \end{aligned} \tag{7}$$

Figure 2 shows the architecture of text sentence aggregation in the model. The *Gumbel Tree-GRU* layer set by the relation extraction model was used to search for the optimal path combination. It received the output of the *BiGRU* layer and found the optimal path combination in all feasible solutions through the greedy strategy. This structure can effectively process structured and hierarchical information, which is very important for extracting the complex relations between the knowledge required by students' online learning and the knowledge linkages in the knowledge library. At last, the text features of entities were embedded in the sentences, which helps to emphasize the part of the sentences that is related to the entities so that the relations between entities can be found more accurately. In this way, the model can better focus on the interaction between entities, thereby more accurately identifying the relations between them.

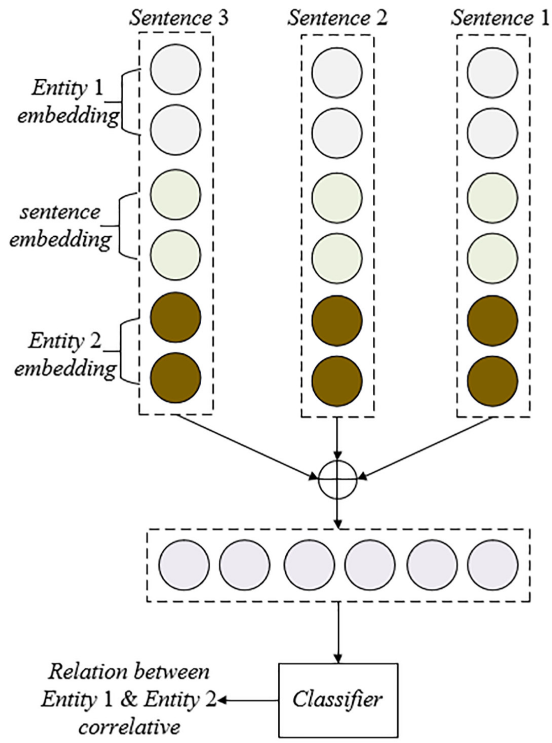

Fig. 2. Text sentence aggregation in the model

3 OPTIMIZATION OF KNOWLEDGE LIBRARY MANAGEMENT BASED ON RULE UPDATE

Since short texts generally have insufficient context information, it may result in low accuracy for the relation extraction model. Through the twin network, the text expression segments of students' online learning needs and the rule segments in the knowledge library can be compared to capture the similarity between them, which facilitates providing richer context information and improving model accuracy. By matching students' online learning needs to the rules in the knowledge library, the rules that need to be updated can be effectively identified and updated. This is because the model can understand the similarities and differences between students' needs and the content of the knowledge library, thereby giving accurate and targeted updates and references for knowledge library rules. Figure 3 shows the updated scheme of knowledge library rules.

The main task of the input layer and encoding layer was to convert raw texts into a form that is more suitable for the processing of machine learning models, the task usually involves the generation of word vectors, and capturing the semantic relations of words and their contexts. Assuming: $S = \{v_1, v_2, ..., v_L\}$ $D = \{c_1, c_2, ..., c_M\}$ represent encoded vectors of a short text containing M words, wherein $v_i \in R^d c_u \in E^f$ represents the df-dimensional vector expression of the iu-th word, and $S \in R^{L \times d} A \in E^{M \times f}$ represents the word vector sequence of the input short text.

The purpose of the feature extraction layer was to extract useful features from the coded text. These features can help the model understand and compare the similarity between two short texts. In this study, a multi-scale convolution was set up to use one-dimensional convolutions of different scales to perform feature extraction on the text matrix. Through splicing features of different scales, the multi-scale feature representation of the texts was attained. Assuming: $c_{u:u+g-1}$ represents the sentence vector consisting of u to $u + g - 1$ words, g represents the size of the one-dimensional convolution window, q_v represents the convolution kernel, and n represents the bias term, then the feature extraction formula is:

$$v_u = f\left(q_v c_{u:u+g-1} + n\right) \tag{8}$$

After going through the convolution layer, the feature matrix V was attained, and the generated multi-scale feature matrix can be written as:

$$D = CT\left(V_1, V_2, V_3\right) \tag{9}$$

The output layer was responsible for making the final prediction based on extracted features, namely the similarity between the two texts. With the fully connected layer as the discriminant layer, the feature representation extracted by the multi-attention mechanism fusion module was mapped to the space of marked samples.

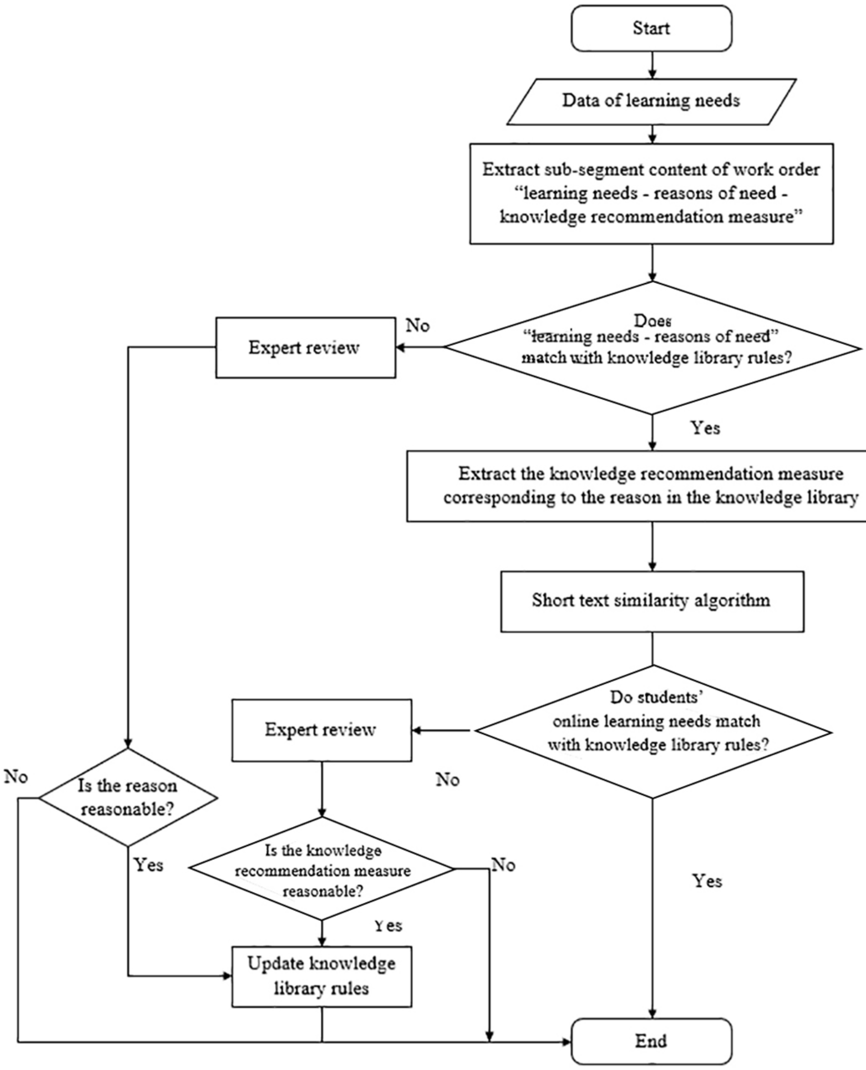

Fig. 3. Update scheme of knowledge library rules

It's of great significance to introduce the multi-attention mechanism fusion module into the short text similarity recognition model established based on the twin network. In case where there are many model parameters, the model may have to process large amounts of input information; therefore, the module needs to comprehensively consider multiple attention mechanisms so that the model can focus on key text feature information from different angles and hierarchies. This design is critical to improving the efficiency and accuracy of the model.

The multi-head self-attention mechanism was composed of multiple self-attention mechanisms; it focused on word dependencies within the text so that the model could give different levels of attention to information inside the text to realize the transformation operation of the text vectors of students' online learning needs.

Assuming, $B \times F$ represents the dimension of the input text word vector matrix Z, B represents sentence length, f represents word vector dimension, the text word vector matrix Z was multiplied by three weight matrices Q^W, Q^J, and Q^C to be transformed into *Query*, *Key*, and *Value*; $f \times f_w$, $f \times f_j$, and $f \times f_c$ respectively represent the dimensions of weight matrices; $B \times f_w$, $B \times f_j$, and $B \times f_c$ respectively represent the dimensions of *Query*, *Key*, and *Value*; in the self-attention mechanism, the dimensions of *Query* and *Key* were the same, then there are,

$$W = Z \cdot Q^W \tag{10}$$

$$J = Z \cdot Q^J \tag{11}$$

$$C = Z \cdot Q^C \tag{12}$$

To avoid too large WJ^Y, a penalty factor f_j was introduced, and the *softmax* activation function was adopted to normalize the attention score.

$$ATT(W, J) = \text{softmax}\left(\frac{WJ^Y}{\sqrt{f_j}}\right) \tag{13}$$

The attention weight attained from above formula was multiplied by matrix C to get the final attention score, that is,

$$AT(Q, J, C) = \text{softmax}\left(\frac{QJ^Y}{\sqrt{f_j}}\right)C \tag{14}$$

The multi-head self-attention mechanism generated $W = (W_0, QW_1, \ldots, W_u, \ldots, W_{g-1})$, $J = (J_0, J_1, \ldots, J_u, \ldots, J_{g-1})$ and $C = (C_0, C_1, \ldots, C_u, \ldots, C_{g-1})$ through g-times of different linear transformations, and the result of the u-th attention mechanism was calculated to be,

$$X_u = AT(W_u, J_u, C_u) \tag{15}$$

The result of the multi-head attention mechanism can be attained by performing f_c-dimensional splicing and linear transformation on the g-th attention mechanism calculation result:

$$MH(W, J, C) = CT(X_0, X_1 \ldots\ldots, X_{g-1})Q^0 \tag{16}$$

Using spatial attention to acquire the feature space relation between text pairs enables the model to pay attention to the similarities and differences between two texts in the feature space. Assuming: D_1 and D_2 represent feature maps extracted by multi-scale convolution and they satisfy $D_1 \in E^{1\times G\times Q}$ and $D_2 \in E^{1\times G\times Q}$, $d^{1\times 1}$ represents a 1×1 convolution layer, $L_A(D)$ represents the weight of spatial attention, δ represents the *Sigmoid* activation function, then the formula for calculating spatial attention is,

$$L_a\left(D\right) = \delta\left(f^{1\times 1}\left(\left[D_1 ; D_2\right]\right)\right) \tag{17}$$

Interactive attention can capture the interactive information between text pairs and focus on the semantic interaction between them. Assuming: L and B represent two short text feature maps extracted by the convolution layer with each row containing some higher-level semantic features of the text, denoted as $L = (l_1, l_2,...,l_u,...,l_j)$ and $B = (b_1, b_2,..., b_u,...,n_j)$. In order to get the attention of m_i to feature map N, the degree of correlation between l_u and each row in B was first calculated based on the following formula:

$$AT\left(l_u, b_k\right) = l_u b_k^Y \tag{18}$$

The degree of correlation between each row in L and the u-th row in B can be calculated as follows:

$$AT\left(L, b_u\right) = l_o b_u^Y \tag{19}$$

The attention weight of each row in B can be calculated by the following formula:

$$ATT\left(L, B\right) = \frac{e^{\sum_{o=1}^{j} l_o b_u^Y}}{\sum_{k}^{J} e^{\sum_{o=1}^{j} l_o b_u^Y}} \tag{20}$$

Each row in B was weighted based on the weight of attention, and finally, the interactive information between the text features of the knowledge required by students' online learning and the knowledge of the knowledge library was attained.

4 EXPERIMENTAL RESULTS AND ANALYSIS

Table 1 lists the performance of each model in the relation extraction task. In the table, P, R, and $F1$ respectively represent precision, recall rate, and $F1$ score. *CNN* performed poorly in the relation extraction task, with a precision of 61.2%, a recall rate of 63.8%, and an $F1$ score of 69.4%. This is because *CNN* mainly captures local features, while the relation extraction task requires a deeper understanding of the context. The *Bi-LSTM* performed well in terms of precision, reaching as high as 74.1%, but its recall rate was low, only 61.9%, and its $F1$ score was 72.4%. This is because *Bi-LSTM* can capture long-term dependencies of sequences but is limited in capturing complex relations. The precision of the proposed model was 70.9%, its recall rate was 64.2%, and its $F1$ score was 71.8%. This is because the model has integrated the attention mechanism, the *BiGRU*, and the *Gumbel Tree* gate control unit, so it has certain advantages in capturing complex relations and long-term dependencies.

In case the *Attention* mechanism was not introduced, the precision was 74.2%, the recall rate was 67.5%, and the $F1$ score was 73.6%, indicating that introducing the

Attention mechanism was helpful since it could help the model focus on more important information. In case the *BiGRU* was not introduced, the precision was 77.8%, the recall rate was 63.8%, and the F1 score was 79.5%, indicating that the *BiGRU* layer had a significant effect on the performance of the model. In case the *Gumbel Tree* was not introduced, the precision was 70.5%, the recall rate was 69.4%, and the F1 score was 71.4%, indicating that the *Gumbel Tree* gate control unit played an important role in the model. The experimental results show that the proposed model has some advantages in balancing precision and recall rate and is more suitable for application scenarios that need to comprehensively consider the two indicators.

Table 1. Comparison of experimental results of relation extraction

Model	*P*(%)	*R*(%)	*F*1(%)
CNN	61.2	63.8	69.4
Bi-LSTM	74.1	61.9	72.4
The proposed model	70.9	64.2	71.8
Without *Attention* mechanism	74.2	67.5	73.6
Without *BiGRU*	77.8	63.8	79.5
Without *Gumbel Tree*	70.5	69.4	71.4

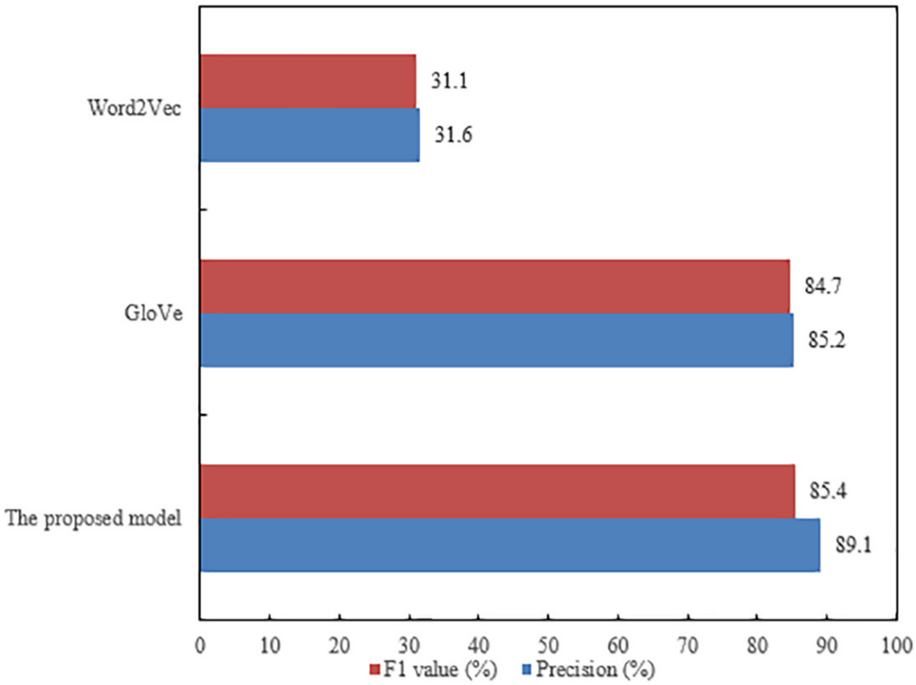

Fig. 4. Comparison of precision and F1 value of different word vector models

Figure 4 shows the performance of three different word vector models in terms of the precision and F1 value of a certain task. In this task, the performance of the proposed model was the best; the precision was 89.1%, and the F1 value was 85.4%. The performance of the *GloVe* was slightly worse than the proposed model; its precision was 85.2% and its F1 value was 84.7%. The *GloVe* is a word vector model established based on global word frequency statistics, and it attempts to map global statistical information into word vectors. Compared with the proposed model, its ability to capture co-occurrence information is stronger, but in terms of the said task, the proposed model performed better. The precision and F1 value of *Word2Vec* were far lower than those of *GloVe* and the proposed model, which were 31.6% and 31.1%, respectively, indicating that the model failed to capture enough semantic information about words effectively when generating the word vectors.

Figure 5 shows the training results before and after introducing the multi-attention mechanism fusion module. The data were indicators under different iteration epochs, and they should be some kind of loss or error indicator, and the smaller the value, the better the performance of the model. After the multi-attention mechanism fusion module had been introduced into the constructed model, starting from iteration epoch 0, the indicator of the model began with 1.85 and decreased gradually as the iteration epoch grew. When the iteration epoch reached 500, the indicator value dropped to 0.35, indicating that the model's performance had been constantly improved during the training. Before introducing the multi-attention mechanism fusion module into the constructed model, for the model without this module, starting from iteration epoch 0, the indicator of the model began at 21 and decreased as the iteration epoch grew. When the iteration epoch reached 500, the indicator value was 1.65, which was still higher than that of the model after introducing the multi-attention mechanism fusion module. According to the above experimental results, introducing the multi-attention mechanism fusion module can significantly improve model performance. This is because the multi-attention mechanism enables the model to focus more on key information, thereby improving the accuracy and efficiency of the model. During the training process, the performance of the model with the multi-attention mechanism fusion module improved constantly. Under the same iteration epoch, its performance was always better than that without the said module.

Table 2 lists the accuracy of three different models in the text similarity recognition task. *Max*1, *Max*2, *Min*1, *Min*2, *Aver*1, and *Aver*2 are different aspects or subtasks of the evaluation indicators. The performance of the *TF-IDF* + cosine similarity model was lower in terms of most evaluation indicators, especially on *Min*1 (91.34%) and *Aver*2 (91.27%); its performance was the worst. The *BERT* model performed better than the *TF-IDF* + cosine similarity model in terms of almost all evaluation indicators, and its performance on *Min*1 (92.25%) and *Aver*2 (95.81%) was especially outstanding. The proposed model performed better than other models in terms of all indicators, suggesting that the proposed model can provide richer and more accurate representations when dealing with text similarity recognition tasks.

Table 3 gives the *TPR* (true rate), *FPR* (false rate), and *AUC* (area under the subject work feature curve) of four differently configured models in the text similarity recognition task. In case the twin network was not introduced, the model's *TPR* was 92.58%, its *FPR* was 2.51%, and its *AUC* was 91.41%. In case the multi-scale convolution was not introduced, the model's *TPR* was 91.24%, its *FPR* was 1.62%, and its *AUC* was 94.57%. In case the multi-attention mechanism fusion module was not introduced, the model's *TPR* was 90.63%, *FPR* was 2.48%, and *AUC* was 92.85%, which were lower than those of the model without twin networks and without multi-scale convolution. The model adopted in this study included configurations containing twin networks, multi-scale convolution, or multi-attention mechanism fusion modules. In terms of *TPR*, the proposed model significantly outperformed the other three configurations, reaching 98.57%. In the meantime, its *FPR* was 1.2%, which was the lowest among the four configurations. Then it's known that introducing a twin network, multi-scale convolution, or multi-attention mechanism fusion module significantly improved the *TPR* and reduced the *FPR* of the model, indicating that the proposed model showed good performance in correctly recognizing positive samples and reducing wrongly recognized positive samples.

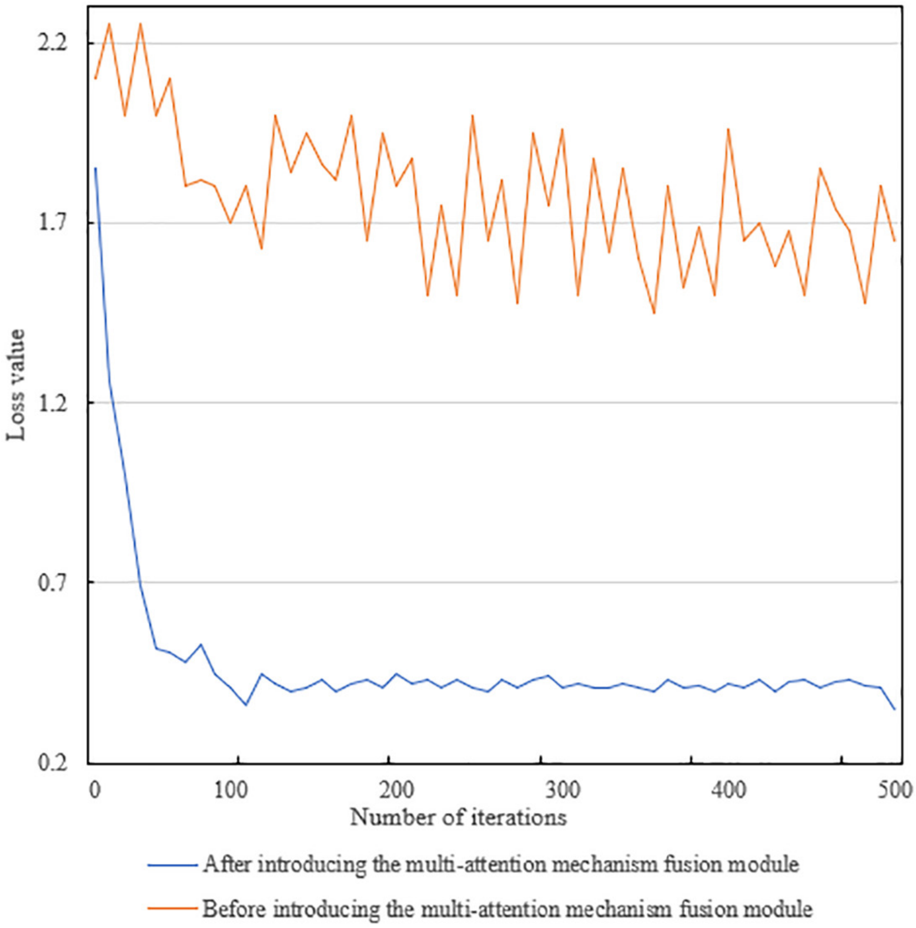

Fig. 5. Comparison of training results before and after introducing the multi-attention mechanism fusion module

Table 2. Similarity recognition accuracy of different models

Model	*Max*1	*Max*2	*Min*1	*Min*2	*Aver*1	*Aver*2
TF-IDF + Cosine similarity model	92.41%	91.52%	91.34%	96.48%	92.35%	91.27%
BERT	95.84%	94.57%	92.25%	95.74%	94.58%	93.81%
The proposed model	96.35%	95.85%	94.52%	97.16%	96.32%	95.38%

Table 3. Similarity recognition accuracy of the model before and after improvement

Model	TPR	FPR	AUC
Without the twin network	92.58%	2.51%	91.41%
Without multi-scale convolution	91.24%	1.62%	94.57%
Without multi-attention mechanism fusion module	90.63%	2.48%	92.85%
The proposed model	98.57%	1.2%	96.61%

5 CONCLUSION

This study proposes a short text similarity recognition model that is based on a twin network. This model solves the low accuracy of conventional relation extraction models caused by insufficient context information when processing short texts. By combining the multi-scale convolution and multi-attention mechanism fusion modules, the model can capture the semantic relations between texts more accurately and effectively match students' online learning needs with the rule fields in the knowledge library. Based on experimental results, the following conclusions could be drawn:

1. Specific solutions were proposed for challenges in extracting relations from short texts in this study. With the help of the twin network, the proposed model could process paired texts and enhance its expression ability by learning the similarity between text pairs.
2. The feature extraction had been enhanced: Multi-scale convolution played a key role in the feature extraction stage as it enabled the model to capture different levels of information and analyze multiple aspects of texts.
3. Fusion of attention mechanisms: By fusing three mechanisms of multi-head self-attention, spatial attention, and interactive attention, the proposed model can focus more comprehensively on the key features of texts. Such a fusion strategy has enhanced the model's ability in to pay attention to different-type information.
4. Accuracy and efficiency had been improved: The proposed model outperformed the baseline model in terms of multiple performance indicators, suggesting that it can process the similarity of short texts and complete relation extraction tasks more accurately and efficiently.
5. Knowledge library management was optimized: The proposed model has great value for knowledge library management since it can assist rule updates by accurately identifying and matching relevant information, which is very important for maintaining the quality and reliability of the knowledge library and offering personalized services.

6 REFERENCES

[1] X. Ma, "Influence study of learners' independent learning ability on learning performance in online learning," *International Journal of Emerging Technologies in Learning*, vol. 17, no. 9, pp. 201–213, 2022. https://doi.org/10.3991/ijet.v17i09.30925

[2] Y. Sun and R.Q. Chai, "An early-warning model for online learners based on user portrait," *Ingénierie des Systèmes d'Information*, vol. 25, no. 4, pp. 535–541, 2020. https://doi.org/10.18280/isi.250418

[3] L. Zhao, "Use of a deep learning approach for the evaluation of students' online learning cognitive ability," *International Journal of Emerging Technologies in Learning*, vol. 18, no. 12, pp. 58–74, 2023. https://doi.org/10.3991/ijet.v18i12.41093

[4] S.Y. Wang, "Online learning behavior analysis based on image emotion recognition," *Traitement du Signal*, vol. 38, no. 3, pp. 865–873, 2021. https://doi.org/10.18280/ts.380333

[5] K. Krpálková Krelová, K. Berková, P. Krpálek, and A. Kubišová, "Perception of selected aspects of online learning by Czech higher education students," *International Journal of Engineering Pedagogy*, vol. 12, no. 5, pp. 4–25, 2022. https://doi.org/10.3991/ijep.v12i5.32243

[6] Y. Dong, S. Zhang, and Y. Huang, "Construction of educational quality evaluation index system based on project management," *International Journal of Information Systems in the Service Sector (IJISSS),* vol. 12, no. 2, pp. 65–75, 2020. https://doi.org/10.4018/IJISSS.2020040105
[7] S. Afifi, H.B. Santoso, and L.M. Hasani, "Investigating students' online self-regulated learning skills and their e-learning experience in a prophetic communication course," *Ingénierie des Systèmes d'Information*, vol. 27, no. 3, pp. 387–397, 2022. https://doi.org/10.18280/isi.270304
[8] S.B. Seryakova, X. Zhang, O.V. Galustyan, N.N. Askhadullina, T.V. Pushkareva, and E.V. Zvonova, "Application of online learning within education of future engineers during the Covid-19 pandemic," *International Journal of Engineering Pedagogy*, vol. 12, no. 1, pp. 95–103, 2022. https://doi.org/10.3991/ijep.v12i1.25009
[9] Y. Yan, J.C. Lee, and E.W. Cooper, "Inference of student needs in an online learning environment based on facial expression," In *2022 10th International Conference on Information and Education Technology (ICIET)*, pp. 113–117, 2022. https://doi.org/10.1109/ICIET55102.2022.9779022
[10] N. Nonthamand, "Development of an online learning lesson in the educational technology equipment operation course for pre-service teachers," *International Journal of Interactive Mobile Technologies*, vol. 17, no. 10, pp. 30–46, 2023. https://doi.org/10.3991/ijim.v17i10.37775
[11] K. Saddhono, B.W. Setyawan, Y.M. Raharjo, and R. Devilito, "The diagnosis of online game addiction on Indonesian adolescent using certainty factor method," *Ingénierie des Systèmes d'Information*, vol. 25, no. 2, pp. 191–197, 2020. https://doi.org/10.18280/isi.250206
[12] R.K. Chenoori and R. Kavuri, "Online transaction fraud detection using efficient dimensionality reduction and machine learning techniques," *Revue d'Intelligence Artificielle*, vol. 36, no. 4, pp. 621–628, 2022. https://doi.org/10.18280/ria.360415
[13] M. Ashar, W. Kamdi, and D.T. Kurniawan, "Professional skills development through the network learning community using an online learning platform," *International Journal of Interactive Mobile Technologies*, vol. 15, no. 12, pp. 202–210, 2021. https://doi.org/10.3991/ijim.v15i12.21587
[14] M. Hamse, S. Lotfi, and M. Talbi, "Identification of student-teachers groups' needs in physical education and sport for designing an open distance learning on the model of small private online courses," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 10, pp. 97–105, 2020. https://doi.org/10.14569/IJACSA.2020.0111013
[15] S. Ivan, G. Maria, and D. Debor, "The impact of online learning toward the independence of primary school students during the pandemic in Jakarta, Indonesia," *International Journal of Online and Biomedical Engineering*, vol. 18, no. 8, pp. 23–33, 2022. https://doi.org/10.3991/ijoe.v18i08.31965
[16] F.Y. Cao, Y.T. Feng, and B. Wei, "Practical research of online teaching platform on reform of computer network with flipped classroom," *Review of Computer Engineering Studies*, vol. 9, no. 2, pp. 67–70, 2022. https://doi.org/10.18280/rces.090204
[17] S. Shreyas, S. Bhat, and S.H. Prasad, "Effective crop prediction considering water and moisture needs using machine learning techniques," In *2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon)*, pp. 1–5, 2022. https://doi.org/10.1109/NKCon56289.2022.10127028

[18] M. Nakatsugawa, Z. Cheng, A. Kiess, A. Choflet, M. Bowers, K. Utsunomiya, S. Sugiyama, J. Wong, H. Quon, and T. McNutt, "The needs and benefits of continuous model updates on the accuracy of RT-induced toxicity prediction models within a learning health system," *International Journal of Radiation Oncology Biology Physics*, vol. 103, no. 2, pp. 460–467, 2019. https://doi.org/10.1016/j.ijrobp.2018.09.038
[19] N. Jalinus, Z. Ganefri, M.A. Wulansari, R.A. Nabawi, and H. Hidayat, "Hybrid and collaborative networks approach: Online learning integrated project and kolb learning style in mechanical engineering courses," *International Journal of Online and Biomedical Engineering*, vol. 18, no. 15, pp. 4–16, 2022. https://doi.org/10.3991/ijoe.v18i15.34333
[20] J. Choudhury, F.B. Ashraf, A. Shakil, and N. Raonak, "Predicting priority needs for rehabilitation of refugees based on machine learning techniques from monitoring data of rohingya refugees in Bangladesh," In *2020 IEEE Region 10 Symposium (TENSYMP)*, pp. 210–213, 2020. https://doi.org/10.1109/TENSYMP50017.2020.9230867
[21] C. Wu, X. Li, R. Jiang, Y. Guo, J. Wang, and Z. Yang, "Graph-based deep learning model for knowledge base completion in constraint management of construction projects," *Computer-Aided Civil and Infrastructure Engineering*, vol. 38, no. 6, pp. 702–719, 2023. https://doi.org/10.1111/mice.12904
[22] R. He, M. Sandoval-Reyes, I. Scott, R. Semeano, P. Ferrao, S. Matthews, and M.J. Small, "Global knowledge base for municipal solid waste management: Framework development and application in waste generation prediction," *Journal of Cleaner Production*, vol. 377, p. 134501, 2022. https://doi.org/10.1016/j.jclepro.2022.134501
[23] C. Wu, P. Wu, J. Wang, R. Jiang, M. Chen, and X. Wang, "Ontological knowledge base for concrete bridge rehabilitation project management," *Automation in construction*, vol. 121, p. 103428, 2021. https://doi.org/10.1016/j.autcon.2020.103428
[24] F.N. Fote, A. Roukh, S. Mahmoudi, S.A. Mahmoudi, and O. Debauche, "Toward a big data knowledge-base management system for precision livestock farming," *Procedia Computer Science*, vol. 177, pp. 136–142, 2020. https://doi.org/10.1016/j.procs.2020.10.021

7 AUTHOR

Jiajia Li graduated from Henan Normal University in 2017 and currently works at Hebi Polytechnic. Jiajia's research focuses on political education and student management. Since 2015, Jiajia has published a total of 13 papers, 4 works, and obtained 2 patents. In 2014, Jiajia was appointed as a business mentor by the Hebi Municipal People's Social Bureau. (E-mail: hbljj123456@163.com; Orcid: https://orcid.org/0009-0005-5620-2626).