Research Article

Wenyi Lu, Ting Wei, Zijun Guo, and Jianhong Ren*

Construction of GCNN-based intelligent recommendation model for answering teachers in online learning system

Abstract: In response to the limitations of the existing online learning system regarding the efficiency and accuracy of the question-and-answer (Q&A) teacher recommendation method, this research develops a Q&A teacher recommendation model utilizing a Graph Convolutional Neural Network. First, a time-sensitive online learning Q&A teacher recommendation model (A Time Sensitive Online Learning Q&A Teacher Recommendation Model; TSRM) is proposed to address the shortcomings that current recommendation methods ignore, i.e., the teacher's ability to answer questions with time changes. Then, a TSRM based on Short and Long Term Interest for Answering Questions (LSTR) is proposed to address the problem that the current recommendation methods ignore, i.e., the types of questions of student users' concerns can change. Finally, we combine the TSRM model and LSTR model to build an intelligent recommendation model for answering teachers. The conclusion is that the accuracy rate of TSRM model on the test set is 99.5%, and the recommendation success rate of LSTR model reaches 98.4%, which are better than the other two models. The above results can show that the LSTR model and TSRM model constructed by the study have high performance and can effectively perform the recommendation of answering teachers in the online learning system, thus improving the efficiency of solving students' problem, improving students' learning effect, and contributing to the development of university education informatization.

Keywords: graph convolutional neural network, online learning, intelligent recommendation, Q&A teacher

1 Introduction

With the rapid development of Internet technology, the traditional education model is gradually changing to the online learning system, and in this process, the intelligent recommendation of answering teachers has become the key link to improve the learning efficiency [1]. Effective teacher matching can not only enhance the learning experience, but also promote the rational distribution and utilization of educational resources [2]. Unfortunately, the current online learning platform still has many shortcomings in answering teacher recommendations, such as low accuracy and insufficient personalization, which limits the development potential of online education model [3]. In order to solve this problem, the aim of this study is to construct an intelligent recommendation model for question answering teachers based on graph convolutional neural network (GCNN). By mining the multi-dimensional interaction data between teachers and students, the model captures

---

* **Corresponding author: Jianhong Ren,** Educational Media Construction Division, Jiangxi Education Evaluation and Assessment Institute, Nanchang, 330038, China, e-mail: renjianhong2023@126.com
Wenyi Lu, **Zijun Guo:** School of Information Engineering, Jiangxi Vocational and Technical College of Communications, Nanchang, 330013, China
Ting Wei: Department of Student Affairs, Jiangxi Vocational and Technical College of Communications, Nanchang, 330013, China

the individual needs of learners and makes intelligent matching based on it. GCNNs provide an ideal tool for capturing complex teacher–student interaction patterns due to their strong representation ability on graph data [4]. GCNN is an emerging deep learning technique that extends the convolutional operations in traditional CNN to graph data, and is widely used in fields such as intelligent recommendation and traffic prediction [5,6]. The innovation of this study is that it integrates the advantages of GCNN in node classification and link prediction, and recommends the most suitable teacher for each learner. Through deep learning of node features and efficient encoding of edge relationships, the model can gain insight into the real needs of learners and the professional competence of teachers, so as to achieve high-precision personalized recommendation. The study offers a novel viewpoint and methodological approach for enhancing intelligent recommendation within online learning ecosystems, aiming to refine the precision of recommendations as well as to augment the level of personalization. The deployment of this model is anticipated to substantially elevate instructional quality and augment efficiency in the online learning milieu, thereby fostering innovation and progression in the domain of online education. Section 2 is a review of the research results related to GCNN and intelligent recommendation, Section 3 is the construction of two intelligent recommendation models based on GCNN, Section 4 is the performance analysis of the intelligent recommendation models for answering teachers, and Section 5 is the summary of the whole research.

2 Related works

GCNN is a new technique built on CNN, which can extend the convolutional operation in traditional CNN to graph data, and has wide applications in citation networks, intelligent recommendation, intelligent traffic prediction, and computer vision [7]. Yu et al. designed a GCNN-based ResGNet-C neural network and applied it to the detection of neocrown pneumonia, thus improving the efficiency and accuracy of neocrown pneumonia detection [8]. Yu et al. used GCNN to extract features and achieve intelligent prediction of road traffic speed considering the existence of regional spatio-temporal correlation of traffic roads, contributing to the construction of intelligent transportation [9]. Chen et al. designed a CNN model based on graph convolutional features and used the model to make intelligent prediction of stocks so as to provide data support for investors' investment decisions [10]. Louis et al. proposed a GCNN with an improved structure and constructed a machine learning model for material performance prediction based on this improved GCNN. The experiment showed that the model owns high accuracy [11]. Wang and Chen, aiming at the trend of intelligent traffic system, proposed a real-time video object detection structure with high compression rate based on GCNN, and predicted the future traffic safety trend according to the previous traffic flow. The experimental results on real traffic datasets prove the excellent performance of this method [12]. Mou et al. designed a nonlocal GCNN and applied it to the field of hyperspectral image classification. Experiments showed that the model has high classification accuracy and classification efficiency in the field of hyperspectral image classification [13]. Min et al. proposed a strategy to improve the GCNN in order to avoid the over-smoothness property in the GCNN that affects the performance of the model, thus designing a scattering GCNN. This proved that the model can effectively overcome the over-smoothness in the traditional GCNN [14]. Spinelli et al. proposed a GCNN model based on adversarial training and applied the model to missing data interpolation so as to improve the integrity and quality of the data, which verified that the model can effectively perform missing data interpolation [15].

At present, with the overall spurt of information resources, information overload has become a problem of the times, and it has become a challenge to dig out effective resources from the massive information center and how to improve the utilization of information. In this context, recommendation systems can effectively alleviate the problem of information overload and allow two-way selection between information and users, and even tap into the potential needs of users and the potential value of information [16]. Zaidan and Zaidan analyzed and reviewed the literature related to IoT-based smart home applications and explored the application path and effect of AI-based smart recommendation technology in it [17]. Milano et al. conducted an application of current recommendation systems, in-depth exploration, and analysis, and expressed their views

and discussions on their ethicality, which provided theoretical support for the research on the application development of recommendation systems [18]. Tahmasebi et al. designed a deep automatic coding network based on Twitter data and used the network to build a social movie recommendation system, so as to achieve personalized and intelligent social movie recommendations [19]. Wahab et al. proposed a trust-based joint learning method for the cold start problem in intelligent recommendation, so as to improve the recommendation accuracy of the recommendation system. The method showed that it has good results and can improve the performance of the recommender system [20]. Paleti et al. proposed an alternating least squares decomposition method based on community detection, thus solving the cold start problem in recommender systems and improving the recommendation efficiency and accuracy [21]. Brandão et al. designed a wavelet-based recommender system for cancer patients with intelligent drug recommendation to reduce patients' pain and prolong their survival [22]. Chen et al. designed an autoclave process formulation parameter recommendation system and conducted an empirical study to verify the effectiveness of the system for parameter recommendation. The use of the system was effective in improving the autoclave process and enhancing the production efficiency [23]. Jelodar et al. proposed an LP algorithm that integrates Latent Dirichlet Allocation (LDA) and PageRank algorithm to improve the recommendation system in the educational question-and-answer (Q&A) platform. Experimental results have proved the effectiveness and excellence of this method [24]. Lyu et al. proposed a Weighted Hypertext Induced Topic Search (WHITS) recommendation algorithm based on link analysis, aiming to identify and recommend high-quality Q&A content and authoritative experts on online education platforms, and realize accurate and effective recommendation by analyzing the interactive relationship between students and answer content, as well as professional mutual recognition among teachers [25].

As described above, GCNNs are widely used and the research results related to intelligent recommendation play an important role in various fields. However, there are still more problems in the current research results of intelligent recommendation, which affect the recommendation effect, and there are few studies related to the recommendation of answering teachers in online learning systems. To address these problems, the research is based on GCNN to obtain the feature vector of student users and then construct the intelligent recommendation model of answering teachers. The research results can help the online learning system to achieve intelligent and efficient teacher recommendation, so as to solve students' problems and improve their learning efficiency in a timely manner, which has certain guiding significance for the development of university education informatization.

3 GCNN-based teacher recommendation model construction for answering questions

3.1 Time-sensitive teacher recommendations based on Q&A

In the current online learning systems, Q&A teacher recommendations are generally based on the correlation between student and instructor users' concerns and historical responses, or the sociality between student and instructor users. However, this approach ignores the problem that the teacher's knowledge base and expertise grow dynamically over time [26]. Therefore, for a certain type of question, a certain teacher user cannot give an effective answer in the present, but in the future, that teacher user may be competent enough to answer that type of question. To address the above problem, the study proposes a time-sensitive online learning Q&A teacher recommendation model (A Time Sensitive Online Learning Q&A Teacher Recommendation Model, TSRM) based on the basic structure shown in Figure 1.

In the TSRM model shown in Figure 1, the number of likes of answers is taken as an important index to evaluate the quality of answers, and more likes indicate higher quality of answers and prove the higher level of Q&A teacher users. The relationship between users' co-response and co-concern is analyzed, and a multi-relationship co-response network is constructed based on the analysis results, and a sequence of Q&A text is constructed by time series. The semantic features of the constructed text sequences are extracted by
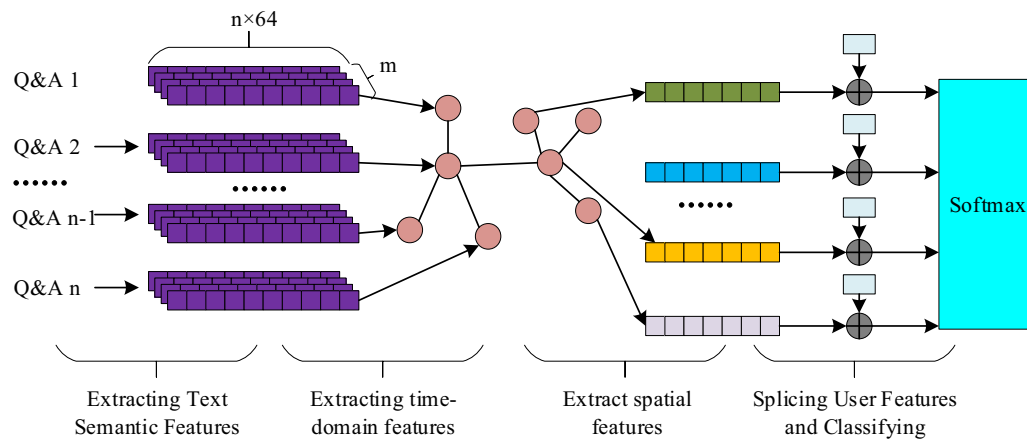
Figure 1: The basic structure of TSRM.

Bidirectional Encoder Representations from Transformers (BERT) model, the topology of the constructed answer network is learned to obtain the empty domain social features, and the time series on the quiz text The topology of the constructed response network is learned to obtain the spatial social features, and the change in the number of likes on the time series of Q&A texts is learned to obtain the time domain features. Finally, the extracted features are classified so that each node of the network can be predicted to determine whether teacher users are better at answering certain types of questions. Based on the GCNN idea, the undirected graph of the multi-relational co-answer network constructed in the study is noted as $G = (V, E)$, where E denotes the set of edges in the topological graph, and V denotes the set of all network nodes, i.e., the set of users. Based on GCNN, it can predict whether a user can answer a question well, as shown in formulas (1)–(3) [27].

$$p(Y_U|X_V) = \prod_{i \in V} p(y_i|X_V).$$ (1)

where $X \in \mathbb{R}^{n \cdot u}$ is the feature matrix of users in the network, where n is the total number of users and u is the feature dimension; Y_U denotes the untagged users; and $p(y_i|X_V)$ is the label distribution of users i, which are considered independent of each other in the study. The TSRM model predicts the user label distribution as $q(y_i|X_V)$, then the objective function is set as equation (2).

$$L(p, q) = -\sum_{i=1}^{n} p(y_i|X_V) \log(q(y_i|X_V)).$$ (2)

The greater the similarity between the predicted and actual distributions of the model, the greater the objective function. The study introduces the relative entropy method to calculate the similarity between the predicted and actual distributions of the model. In the study, the question-answering teacher recommendation problem is viewed as a binary classification problem of graph nodes. Therefore, the goal of the TSRM model is to minimize the error between the predicted value $\hat{y}_i$ and the true value y_i. Therefore, the study introduces cross-entropy loss to construct the objective function, which is given by equation (3).

$$L(y, \hat{y}) = -\sum_{i=1}^{n} [y_i \log y_i + (1 - y_i) \log(1 - y_i)].$$ (3)

In the above, the time-domain features are generally obtained by the Long Short-Term Memory (LSTM) artificial neural network model, whose basic structure is shown in Figure 2.

In Figure 2, we can see that the LSTM unit contains 3 internal gate structures: forgetting, input, and output gate. The LSTM is able to record long-term information through the gating mechanism, which has a good application in different scenarios. However, the structure of LSTM is relatively complex and the training time is long, so it is simplified to obtain Gated Recurrent Unit (GRU). Since GRU is simplified from LSTM, it has the
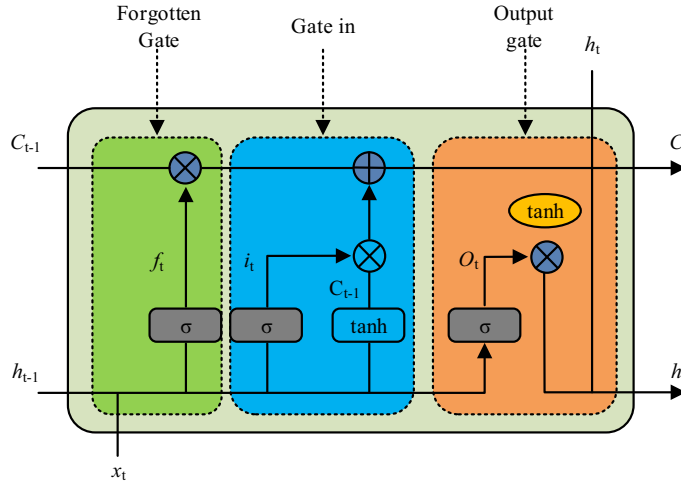
Figure 2: The basic structure of LSTM.

advantages of LSTM while its structure is simpler and involves fewer parameters, so the training efficiency is higher. Based on the above reasons, the study uses GRU to obtain time-domain features. In this process, four main parts are involved, which are reset gate, update gate, and output gate. Based on LSTM, its calculation is specifically shown in formulas (4)–(8) [28]. Among them, the structure of the update gate can be expressed as equation (4), which is able to control the extent of the state information generated during the last training when it enters this training.

$$u_t = \sigma(W_u[X_t, h_{t-1}] + b_u). \tag{4}$$

where σ is the nonlinear classification function, the Sigmoid function used in the study; W is the weight matrix in the training process; X_t denotes the Q&A feature at the moment of t; b is the bias vector in the training process; h_t denotes the hidden state at the moment of t, which can be calculated using equation (5).

$$h_t = u_t \cdot h_{t-1} + (1 - u_t) \cdot c_t. \tag{5}$$

where c_t denotes the stored state inside the model. The forgetting gate can be expressed as equation (6), which is able to control the degree of elimination of the state information generated during the last training.

$$r_t = \sigma(W_r x[X_t, h_{t-1}] + b_r). \tag{6}$$

At the moment of t, the storage state inside the model can be expressed as equation (7).

$$c_t = \tanh(W_c[X_t, (h_{t-1} \cdot r_t)] + b_c). \tag{7}$$

In equation (6), W is the weight matrix from the input layer to the implied layer. h_t can be expressed as equation (8).

$$h_t = (1 - z_t) \circ \vec{h}_t + z_t \circ h_{t-1}. \tag{8}$$

Finally, GCNN is used to obtain the social features in the null domain. The specific process is as follows: first, the text features of user i are extracted using GRU and input to the GCNN layer as the initial vector H_i^0; then, the feature vector matrix $H^0 = [H_1^0, H_2^0, ..., H_n^0]$ is put into the GCNN layer together with the adjacency matrix of the graph by repeating the above operation for all users, so as to extract the airspace social features from the multi-relational common network. The above process is shown in Figure 3.

The TSRM model is designed to provide students with the most suitable Q&A teacher. To do this, the model will take into account the multi-dimensional characteristics of students and teachers, and learn how to match teachers and students most effectively through intelligent algorithms. First, data about students and teachers need to be collected, and after the data are collected, it is pre-processed to clean, normalize, and encode it to ensure that the data are suitable for model processing. Then, natural language processing and machine
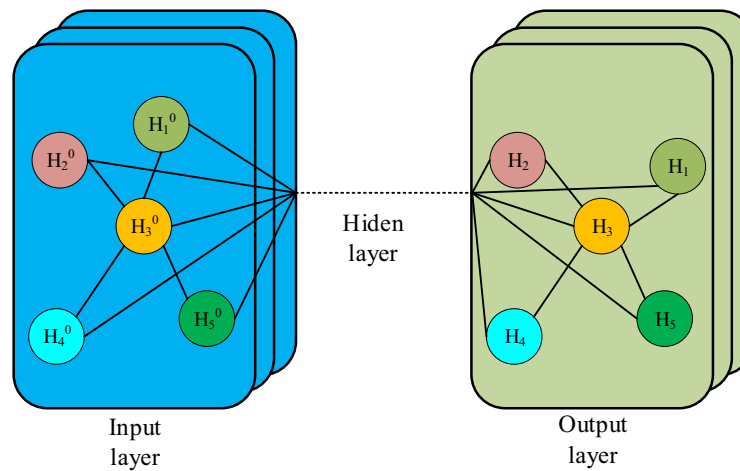
Figure 3: GCNN-based process for obtaining spatial social features.

learning techniques are used to extract feature vectors from students' and teachers' Q&A records and learning behavior logs. Next a graph structure is constructed, with nodes representing students and teachers and edges representing possible matching relationships between them, and a graph neural network is used to extract complex relationships and patterns between nodes. Finally, node features and graph structure are used to predict the matching score.

3.2 Teacher recommendation model based on long- and short-term interest in answering questions

In different periods of time, as time goes by, the learning content of student users will change, and therefore the types of problems that student users focus on will also shift to some extent. However, in current intelligent recommendation methods, this is often ignored, and the recommended teachers for answering questions are more fixed, which affects the efficiency of solving students' problem [29]. And with the change in time, the user's concern and good problem solving will also change somewhat, i.e., an interest shift occurs. If the user has a long time of interest shift, it is called long-term user shift; accordingly, if the user has a short time of interest shift, it is called short-term user shift. To remedy the shortcomings of traditional intelligent recommendation methods, a TSRM based on Short- and Long-Term Interest for Answering Questions (LSTR) is proposed. The model is divided into three parts: question vector feature extraction, user vector representation, and intelligent recommendation based on the extracted question vector features and user vector. In the problem vector feature extraction part, the study uses the BERT model to obtain the word embeddings of problem titles and labels. Based on BERT, the calculation process is shown in formulas (9)–(12) [30]. Let a question be q, the title text of the question is represented as q_t, and the label text of the question is represented as q_l, then the title word embedding of the question X_{Qt} can be calculated using equation (9).

$$X_{Qt} = \text{BERT}(q_t). \tag{9}$$

where $\text{BERT}(q_t)$ denotes the word embedding vector obtained by pre-training the title text of the problem q in the BERT model. The labeled word embedding of this problem X_{Ql} can be calculated using equation (10).

$$X_{Ql} = \text{BERT}(q_l). \tag{10}$$

where $\text{BERT}(q_l)$ denotes the word embedding vector obtained by pre-training the labeled text of the problem q in the BERT model. The feature vector of this problem X_Q can be calculated using equation (11).

$$X_Q = X_{Qt} \oplus X_{Ql}. \tag{11}$$

where $\oplus$ denotes vector splicing. The word embedding of the question title is stitched with the word embedding of the question label to obtain the vector representation of the question. The above process is represented in Figure 4.
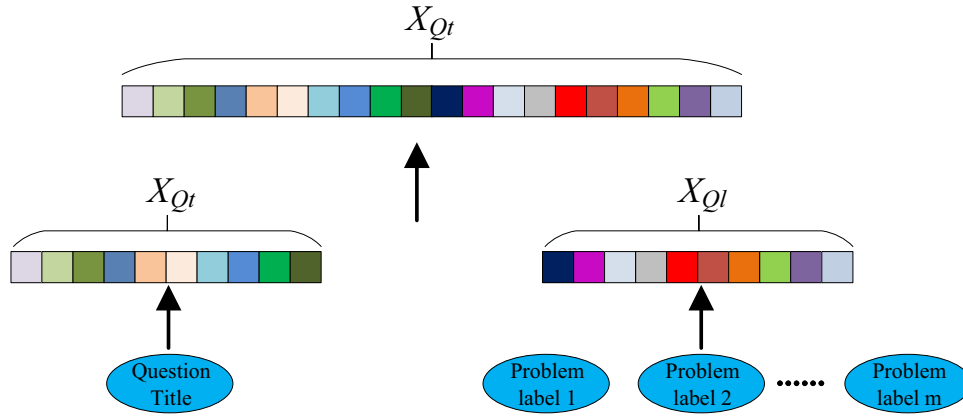

Figure 4: The process of obtaining vector representation of problems.

In addition, the social relationships of the users in the online learning system are analyzed. Users who have a relationship such as a common concern, a common answer, or a concern are considered to have a social relationship with each other. If a user is interested in a certain type of question, it is assumed that users with whom the user has a social relationship may also be interested in that type of question. Based on this theory, the list of questions that a user is interested in is considered as an attribute of the user, and each user is considered as a node, and the relationship between users in terms of concern or common answer is considered as an edge, thus building a multi-relationship common answer network with attributes. Based on GCNN, we extract the features of each node in the network, which can be regarded as the long-term interest of the user. Let the user's concern question text be p, and the word embedding features of the topic text are extracted by BERT model as shown in equation (12).

$$X_v = \text{BERT}(p). \tag{12}$$

Then, the GraphSage model is used to process and analyze the word embedding features to obtain the long-term interest features of users X_{Ul}. On the basis of the time series of user's responses, we can analyze the change pattern of user's interests and process the analysis results to obtain the short-term interests of users. The short-term interest features of users can be extracted by RNN, the LSTM and GRU are both based on the improvement of RNN. The GRU model has a simpler structure and is more efficient in training, as can be seen in Section 3.1. Therefore, the study uses the GRU model to extract the short-term interest features of users X_{Us}. A complete user feature vector is obtained by stitching the long-term interest features obtained grounded on GCNN with the short-term interest features obtained in line with RNN. The above process is shown in Figure 5.
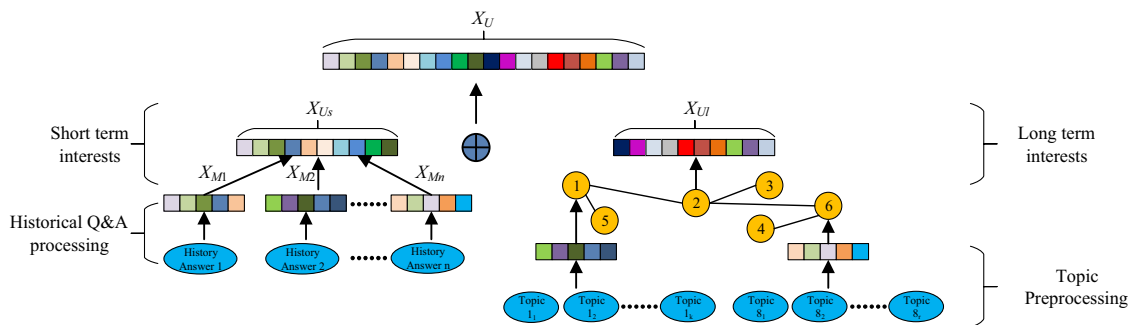

Figure 5: User vector representation process.

Based on the obtained question vector and user vector, the answer teacher recommendation is implemented. The study uses a Deep Structured Semantic Model (DSSM) for result prediction. The extracted user feature vector X_U and question vector X_Q are added into two fully connected layers, respectively, and mapped to the low-dimensional space to obtain the low-dimensional vectors. The cosine similarity method is utilized, the similarity between the user features and the problem features $S(X_Q, X_U)$ is able to be calculated. The teachers with the highest number of likes on the question are considered as a positive sample, and the teacher users who did not answer the question are considered as a negative sample. Based on the random sampling method with Softmax, it is able to convert the similarity between the question and the positive sample to the posterior probability of finding a positive sample for that question $P(X_U|X_Q)$. During the training process, the loss function in equation (13) needs to be minimized [31].

$$f_L = -\log \prod_{X_Q, X_U} P(X_U|X_Q). \tag{13}$$

The above contents can be represented as Figure 6. The LSTR model is completed by combining the above contents. Combining the LSTR model with the TSRM model, efficient and high-precision teacher recommendation for answering questions can be achieved. Compared with the existing learning-based networks, the model uses graph structure to represent and process data, which is particularly suitable for processing graph structure data and can capture complex relationships and patterns between nodes. At the same time, it is able to process non-Euclidean structured data, such as social network data, namely, the relationship between students and teachers. This capability makes it suitable for recommendation systems, especially in situations where interactions between users need to be captured and exploited. And the model can solve the problem of teacher recommendation in online learning system, which is not covered by the traditional learning network.
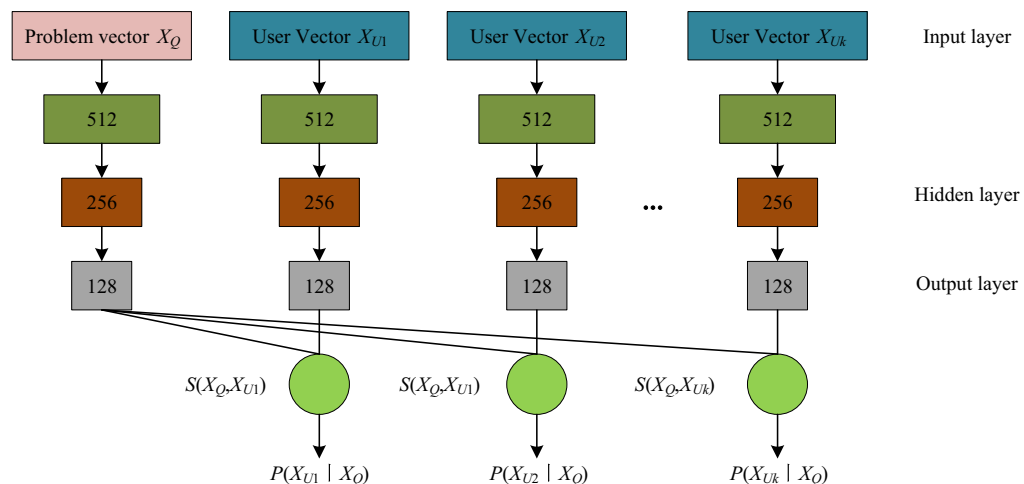

Figure 6: The basic structure of deep structured semantic model.

4 Performance analysis of LSTR model and TSRM model

4.1 Performance analysis of TSRM model

On the premise of obtaining consent, data were obtained from an online learning system of a university to construct a dataset, which included student dimension information, teacher dimension information, classroom content information, and interactive dimension information. Student dimension information includes students' basic information, learning behavior, and historical question answering request. Teacher dimension information includes teachers' basic information, historical evaluation, and historical Q&A information. The classroom content information includes the basic information of the course, the difficulty of the course, the

students' participation, and the degree of completion. Interactive dimension information includes historical question answering records between students and teachers. The dataset is divided into two subsets according to the ratio of 7:3, which are noted as the training set and the testing set, for the training and testing of the model. The detailed configuration parameters of hardware and software used in the study are shown in Table 1.

Table 1: Details about software and hardware configuration parameters

Hardware			Software		
Name	**Supplier**	**Detail**	**Name**	**Supplier**	**Detail**
GPU	NVIDIA	GeForce GTX 1660 Super	OS	Linux	Ubuntu 20.04 LTS
CPU	Intel	Core i5-10600K	Deep learning framework	CUDA	TensorFlow 2.6
RAM	Kingston	16GB-DDR4	GCNN library	Deep Graph Library	1.0
MEM	SUMSUNG	512GB-NVMe	Python environment	Anaconda	3.8

Among the current online learning systems, there are two main state-of-the-art methods for teacher recommendation, which are LDA-PageRank model (LP) and WHITS model. First, the performance of TSRM model, LP model, and WHITS model is compared. First, the TSRM, LP, and WHITS model are trained using the training set, and Figure 7 shows the error variation and loss value variation of the three models during the training process. In Figure 7, we can see that the error and loss values of the three models decrease with the increase in the number of model iterations, and after the number of model iterations reaches a certain level, the error and loss values of the models stabilize and do not change significantly. Among them, the TSRM model needs 99 iterations to reach the optimal state, which is 48 and 103 times less than the LP model and WHITS model, respectively. In Figure 7(a), the error values of LP and WHITS are 0.02 and 0.03 higher than those of TSRM (0.02). In Figure 7(b), the loss values of TSRM (0.3) are 0.4 and 0.8 lower than those of LP and WHITS, respectively.
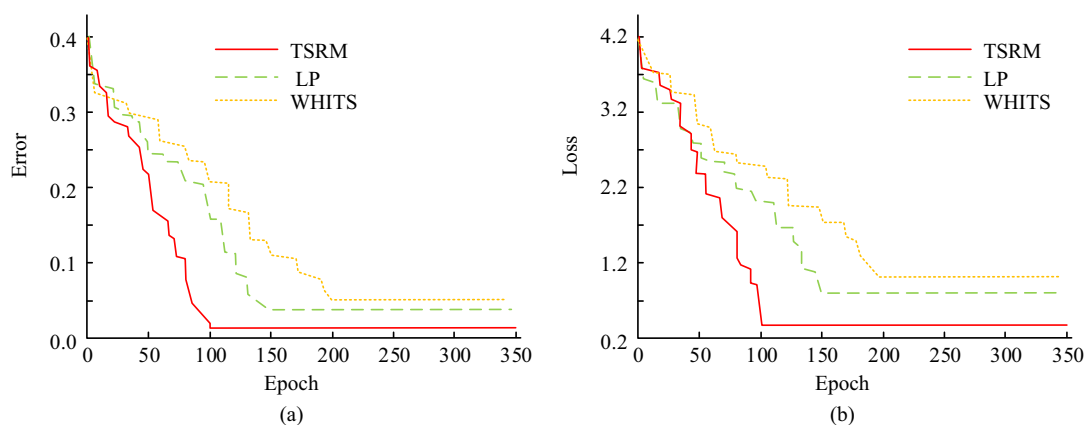

Figure 7: Error changes and loss value changes of three models during training. (a) Error. (b) Loss.

The performance of the TSRM model, LP model, and WHITS model was tested using the test set, and they are evaluated using $F1$ values and Recall values, as shown in Figure 8. In Figure 8, it can be seen that the $F1$ value and Recall value of the models increase as the number of iterations increases. In Figure 8(a), the TSRM's $F1$ value is 96.5%, which is 1.1 and 1.4% higher than that of the LP and the WHITS. In Figure 8(b), the TSRM's Recall value is 96.3%, which is 0.5 and 1.0% higher than the other two models.
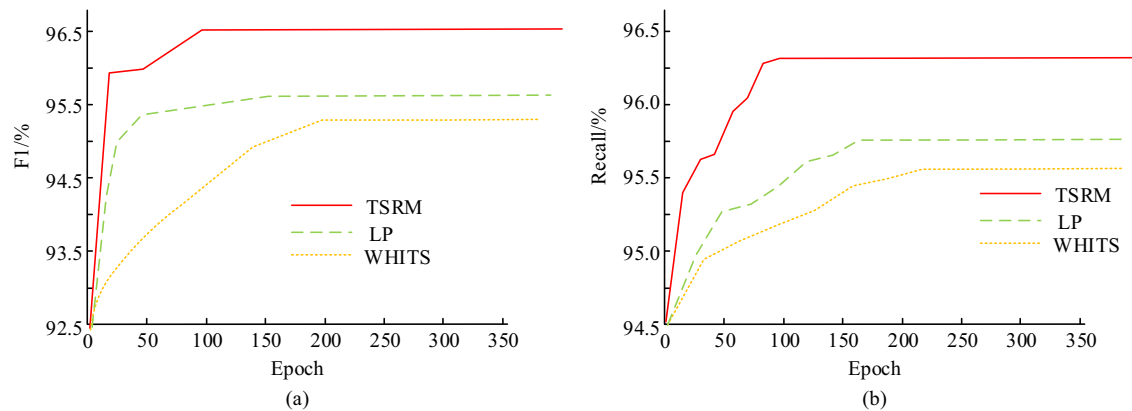
Figure 8: *F*1 and Recall values of TSRM model, LP model, and WHITS model. (a) *F*1. (b) Recall.

The recommendation accuracy of the TSRM, the LP, and the WHITS model are compared on the training set and the test set, respectively. Figure 9 is the recommendation accuracy of the three models. The models' accuracy values increase with the iterations on both the training and test sets. In Figure 9(a), the accuracy variation of the TSRM model, LP model, and WHITS model on the training set is shown. The TSRM accuracy is 98.5%, which is 0.5 and 1.2% higher than the LP and the WHITS. In Figure 9(b), the accuracy variation of the TSRM model, LP model, and WHITS model on the test set is shown. The accuracy of the TSRM is 99.5%. It is 0.4 and 0.9% beyond than those of other two. In summary, the TSRM model has better performance in the work of Q&A teacher recommendation for online learning systems.
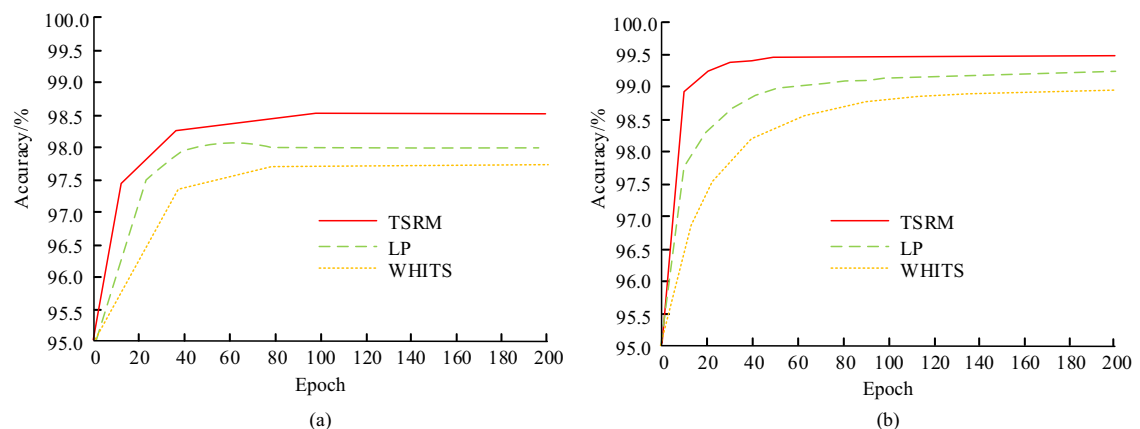

Figure 9: Change in accuracy of models on training and testing sets. (a) Training set. (b) Test set.

4.2 Performance analysis of LSTR model

The LSTR model was tested and compared with the LP model and the WHITS model using the training and test sets described above to demonstrate the LSTR model performance. The variation in MAE values and RMSE values of the LSTR model, LP model, and WHITS model on the training set are exhibited in Figure10. In Figure 10(a), the mean MAE of the LSTR is 3.7%. It is 0.9 and 1.4% lower than LP model and the WHITS, respectively. In Figure 10(b), the mean value of RMSE for the LSTR is 3.3%. It is 0.6 and 1.2% lower than that of the LP and the WHITS, respectively.

In the test set, the recommended success rate and computational complexity test of LSTR model, LP model, and WHITS model are shown in Figure 11. It can be seen that with the increase in the number of model
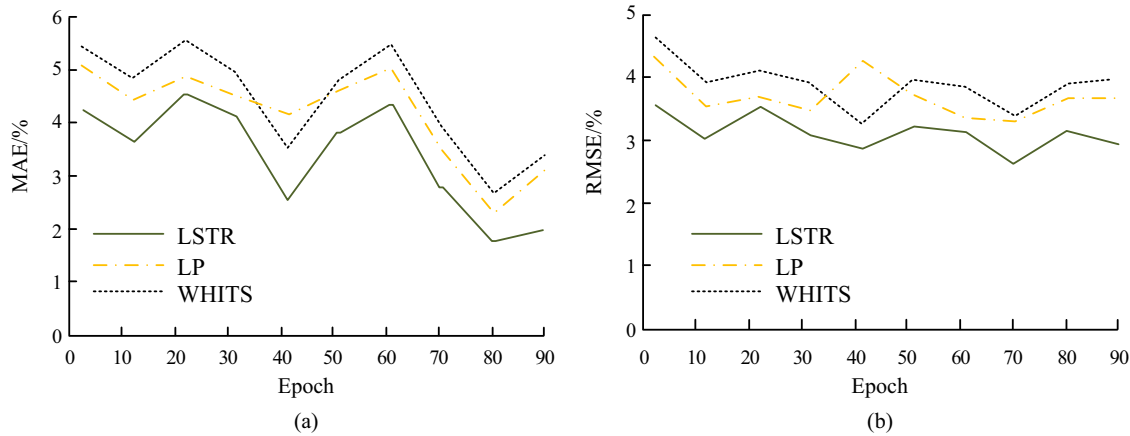
Figure 10: Changes in MAE and RMSE values of the model. (a) MAE. (b) RMSE.
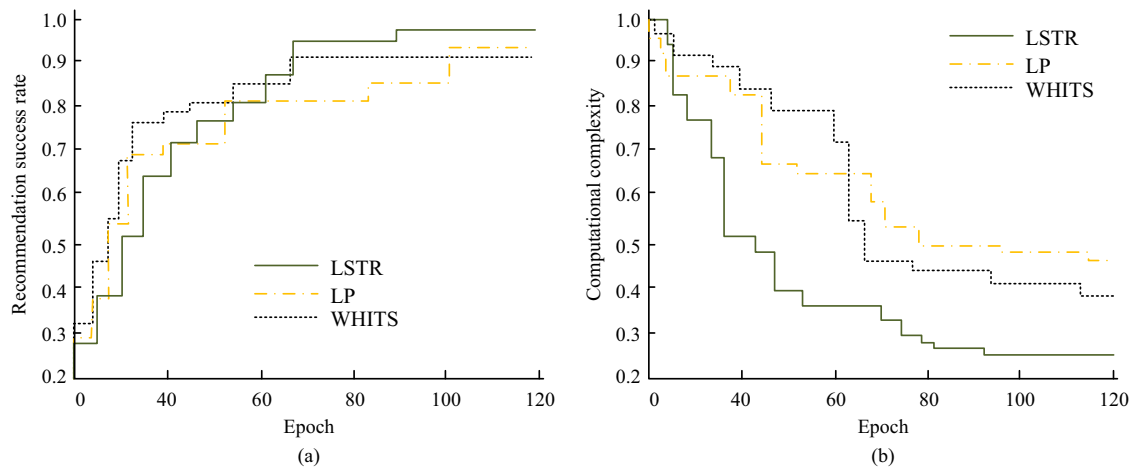

Figure 11: The recommendation success rate of LSTR model, LP model, and WHITS model. (a) Recommendation success rate. (b) Computational complexity.

iterations, the recommendation success rate of LSTR model, LP model, and WHITS model is also increasing. As can be seen in Figure 11(a), the LSTR model has a recommendation success rate of 98.4%, which is 4.2 and 4.9% higher than the LP model and WHITS model, respectively. As can be seen from Figure 11(b), the computational complexity of the LSTR model has been optimized very effectively, and it is 63.7 and 86.5% ahead of the LP model and WHITS model, respectively

The comprehensive performance of LSTR model with LP model and WHITS model was evaluated by ROC curve. Figure 12 displays the evaluation results. The AUC value of the LSTR model reaches 0.925, which is 0.014 and 0.020 higher than that of the LP model and the WHITS model.

The mean reciprocal ranking (MRR) and the normalized discounted cumulative gain (NDCG) of the three models were tested, and the test results are shown in Table 2. As can be seen from Table 2, the average reciprocal rank of LSTR model is 0.23, which is higher than LP model and WHITS model by 0.05 and 0.08, respectively. The cumulative normalized loss gain of LSTR model is 0.25, which is 0.08 and 0.09 higher than that of LP model and WHITS model, respectively.

In summary, the LSTR model and the TSRM model constructed by the study have high performance and can effectively carry out the recommendation of the answering teacher in the online learning system, thus improving the efficiency of solving students' problem, improving their learning effect, and contributing to the development of education informatization in higher education institutions.
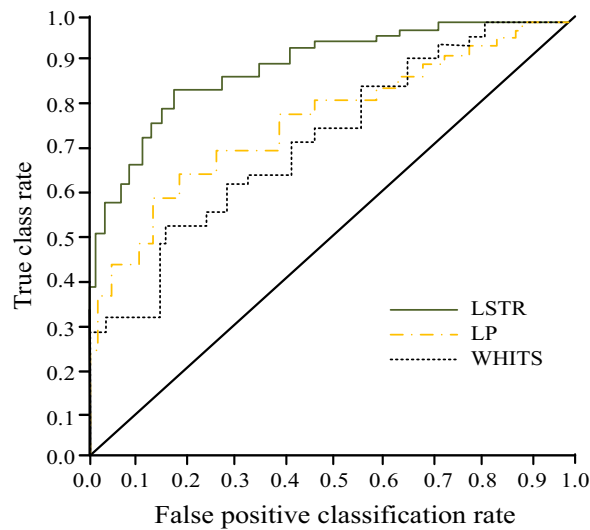
Figure 12: Comprehensive performance evaluation of LSTR model, LP model, and WHITS model.

Table 2: MRR and NDCG test results of three models

Times	LSTR		LP		WHITS	
	MRR	**NDCG**	**MRR**	**NDCG**	**MRR**	**NDCG**
1	0.24	0.26	0.19	0.21	0.16	0.17
2	0.22	0.25	0.18	0.20	0.15	0.18
3	0.23	0.26	0.18	0.19	0.15	0.17
4	0.22	0.24	0.19	0.21	0.14	0.16
5	0.24	0.26	0.18	0.18	0.14	0.17
Average	0.23	0.25	0.18	0.19	0.15	0.17

5 Conclusion

In the university online learning system, the intelligent recommendation of answering teachers can effectively improve the problem-solving effect of students and enhance their learning efficiency. Compared with the previous model of the same type, this model has better performance, and also has better comprehensive performance in the actual use test, with stronger practicability and usability. To this end, the LSTR model and TSRM model are constructed and combined to achieve efficient and intelligent Q&A teacher recommendation. The experiment demonstrated that the TSRM requires 99 iterations to reach the optimal state, which is 48 and 103 times less than the LP model and WHITS model, respectively; the error value is 0.02, which is 0.02 and 0.03 lower than the LP and WHITS, respectively; the Loss value is 0.3, which is 0.4 and 0.8 lower than the LP and WHITS, respectively; the $F1$ value is 96.5% which is 1.1 and 1.4% beyond the LP and WHITS, respectively; Recall value (96.3%) is 0.5 and 1.0% beyond the LP and WHITS, respectively; accuracy of the training set (98.5%) is 0.5 and 1.2% higher than the LP and WHITS, respectively; accuracy of the test set (99.5%) is 0.4% beyond the other two; The mean MAE of LSTR model is 3.7%, which is 0.9 and 1.4% lower than LP model and WHITS model, respectively; the mean RMSE is 3.3%, which is 0.6 and 1.2% lower than LP model and WHITS model, respectively; the recommendation success rate is 98.4%, which is 4.2 and 4.4% higher than LP model and WHITS model, respectively. In summary, the LSTR model and the TSRM model constructed in the study have high performance and can effectively perform the recommendation of answering teachers in online learning systems. The study used the number of likes to assess the quality of teachers' Q&A, but the index has some limitations, which led to the possible bias of the experimental results. Therefore, the index needs to be

improved in the follow-up study so that it can reflect the quality of Q&A more comprehensively and accurately and improve the credibility of the experimental results.

References

[1]   Castro MDB, Tumibay GM. A literature review: Efficacy of online learning courses for higher education institution using meta-analysis. Educ Inf Technol. 2021;26(2):1367–85.

[2]   Verawardina U, Asnur L, Lubis AL, Hendriyani Y. Reviewing online learning facing the Covid-19 outbreak. J Talent Dev Excell. 2020;12(3s):385–92.

[3]   Simamora RM. The challenges of online learning during the COVID-19 pandemic: An essay analysis of performing arts education students. Stud Learn Teach. 2020;1(2):86–103.

[4]   Mohammadi M, Sharifi A. Evaluation of convolutional neural networks for urban map** using satellite images. J Indian Soc Remote Sens. 2021;49(1):2125–31.

[5]   Nejad SMM, Abbasi-Moghadam D, Sharifi A, Farmonov N, Amankulova K, Lászlź M. Multispectral crop yield prediction using 3D-convolutional neural networks and attention convolutional LSTM approaches. IEEE J Sel Top Appl Earth Observ Remote Sens. 2022;16:254–66.

[6]   Esmaeili M, Abbasi-Moghadam D, Sharifi A, Tariq A, Li Q. Hyperspectral image band selection based on CNN embedded GA (CNNeGA). IEEE J Sel Top Appl Earth Observ Remote Sens. 2023;16(1):1927–50.

[7]   Ali A, Zhu Y, Zakarya M. Exploiting dynamic spatio-temporal graph convolutional neural networks for citywide traffic flows prediction. Neural Netw. 2022;145:233–47.

[8]   Yu X, Lu S, Guo L, Zhang Y. ResGNet-C: A graph convolutional neural network for detection of COVID-19. Neurocomputing. 2021;452:592–605.

[9]   Yu B, Lee Y, Sohn K. Forecasting road traffic speeds by considering area-wide spatio-temporal dependencies based on a graph convolutional neural network (GCN). Transp Res Part C: Emerg Technol. 2020;114:189–204.

[10]   Chen W, Jiang M, Zhang WG, Chen Z. A novel graph convolutional feature based convolutional neural network for stock trend prediction. Inf Sci. 2021;556:67–94.

[11]   Louis SY, Zhao Y, Nasiri A, Wang X, Song Y, Liu F, et al. Graph convolutional neural networks with global attention for improved materials property prediction. Phys Chem Chem Phys. 2020;22(32):18141–8.

[12]   Wang J, Chen Q. A traffic prediction model based on multiple factors. J Supercomputing. 2021;77(3):2928–60.

[13]   Mou L, Lu X, Li X, Zhu X. Nonlocal graph convolutional networks for hyperspectral image classification. IEEE Trans Geosci Remote Sens. 2020;58(12):8246–57.

[14]   Min Y, Wenkel F, Wolf G. Scattering GCN: Overcoming oversmoothness in graph convolutional networks. Adv Neural Inf Process Syst. 2020;33:14498–508.

[15]   Spinelli I, Scardapane S, Uncini A. Missing data imputation with adversarially-trained graph convolutional networks. Neural Netw. 2020;129:249–60.

[16]   Sturluson A, Raza A, McConachie GD, Siderius DW, Fern XZ, Simon CM. Recommendation system to predict missing adsorption properties of nanoporous materials. Chem Mater. 2021;33(18):7203–16.

[17]   Zaidan AA, Zaidan BB. A review on intelligent process for smart home applications based on IoT: Coherent taxonomy, motivation, open challenges, and recommendations. Artif Intell Rev. 2020;53(1):141–65.

[18]   Milano S, Taddeo M, Floridi L. Recommender systems and their ethical challenges. Ai Soc. 2020;35:957–67.

[19]   Tahmasebi H, Ravanmehr R, Mohamadrezaei R. Social movie recommender system based on deep autoencoder network using Twitter data. Neural Comput Appl. 2021;33:1607–23.

[20]   Wahab OA, Rjoub G, Bentahar J, Cohen R. Federated against the cold: A trust-based federated learning approach to counter the cold start problem in recommendation systems. Inf Sci. 2022;601:189–206.

[21] Paleti L, Radha Krishna P, Murthy JVR. Approaching the cold-start problem using community detection-based alternating least square factorization in recommendation systems. Evolut Intell. 2021;14:835–49.

[22] Brandão L, Belfo FP, Silva A. Wavelet-based cancer drug recommender system. Procedia Comput Sci. 2021;181:487–94.

[23] Chen YX, Wang LC, Chu PC. A recipe parameter recommendation system for an autoclave process and an empirical study. Procedia Manuf. 2020;51:1046–53.

[24] Jelodar H, Wang Y, Zhao G, Rabbain M, Zhao R, Ayobi S, et al. Recommendation system based on semantic scholar mining and topic modeling on conference publications. Soft Comput. 2021;25:3675–96.

[25] Lyu D, Chen L, Xu Z, Yu S. Weighted multi-information constrained matrix factorization for personalized travel location recommendation based on geo-tagged photos. Appl Intell. 2020;50(1):924–38.

[26] Caro-Martínez M, Jiménez-Díaz G, Recio-García JA. Conceptual modeling of explainable recommender systems: an ontological formalization to guide their design and development. J Artif Intell Res. 2021;71:557–89.

[27] Guan H, Shizhong HE, Qiuqiu LI, Yang Z, Qin C, Weichu HE. A review of convolutional neural networks in equipment wear particle recognition. Tribology. 2022;42(2):426–45.

[28] Liu J, Wang Z, Xu M. DeepMTT: A deep learning maneuvering target-tracking algorithm based on bidirectional LSTM network. Inf Fusion. 2020;53(1):289–304.

[29] Chalyi S, Leshchynskyi V. Method of constructing explanations for recommender systems based on the temporal dynamics of user preferences. EUREKA Phys Eng. 2020;3(3):43–50.

[30] Li M, Li W, Wang F, Jia X, Rui G. Applying BERT to analyze investor sentiment in stock market. Neural Comput Appl. 2021;33(10):4663–76.

[31] Kou J, Zhan T, Zhou D, Xie Y, Da Z, Gong M. Visual attention-based siamese CNN with SoftmaxFocal loss for laser-induced damage change detection of optical elements. Neurocomputing. 2023;517(14):173–87.