Journal of Network Intelligence ©2025 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **10**, Number **4**, November **2025**

Knowledge Sharing Behavior of English Learning Based on Knowledge Graph in Social Network Environment

Yuan-Yuan Huang*, Xiao-Fei Tong

Department of Public Course Teaching
Nanyang Vocational College of Agriculture
Nanyang 473000, P. R. China
cathyhuang9785@163.com, wuyan202409@126.com

Jian Liu

Krirk University
Bangkok 10220, Thailand
mg3528@163.com

*Corresponding author: Yuan-Yuan Huang
Received September 5, 2024, revised May 21, 2025, accepted September 22, 2025.

Abstract. *English knowledge sharing is the chief way for students to accumulate English knowledge, so the prediction of their sharing behavior is of great value. However, existing methods have a large number of incomplete relationships, leading to unsatisfactory prediction results. Based on this, this paper investigates a knowledge graph (KG)-based method for predicting English learning knowledge sharing behaviors in a social network environment. Firstly, the knowledge points of English learning are modelled as KGs, the mapping relationships of different entity structures are established, the TransD model is optimized based on the relationship path information and hierarchical relationship structure (ETransD), the representation of entity information is added to the path, and the information of hierarchical relationship structure is introduced in order to improve the ability of entity and relationship representation. Then Graph Convolutional Network (GCN) is used to capture the correlation between users and social relations, mine user features, and combine with ETransD to extract shared knowledge features. Meanwhile, the attention mechanism is used to aggregate user and shared knowledge features, and the aggregated vectors and user vectors are computed to get the vector similarity by inner product, and the final prediction probability is obtained by prediction function. The experimental results show that the average prediction accuracy of the proposed method reaches 91.8%, and the coefficient of determination R^2 reaches 0.962, which is better than the comparison method and greatly improves the prediction efficiency.*
Keywords: Social network; Knowledge graph; Graph convolutional network; Attention mechanism; Sharing behavior prediction

---

1. **Introduction.** Knowledge sharing refers to members' knowledge exchange and discussion with each other through various channels within an organization, with the aim of expanding the use value of knowledge and generating the effect of knowledge through the exchange of knowledge [1]. In today's big data era, knowledge sharing has become an important decision for national big data and one of the important supporting technologies for data economy. While English is a subject that involves a lot of knowledge, knowledge sharing is the main way for students to accumulate and integrate their knowledge

[2]. With the progress of the times, Internet technology has received wide attention and rapid development, a variety of network applications are emerging, always producing huge amount of English knowledge resources, the amount of data shows explosive rapid growth, how to more accurately and quickly predict the English knowledge sharing behavior, as well as efficiently use these data resources to facilitate people's lives, is the current issue that requires to be addressed urgently [3, 4].

1.1. **Related work.** Traditional studies have mainly focused on the analysis of user-level factors of knowledge-sharing behavior. Shih and Lou [5] showed that knowledge-sharing behavior can be influenced by individually oriented expectations, but more accurate predictions can be made through socially oriented expectations. Chen et al. [6] used individual perceived bridging capital as a variable and used SVM for knowledge sharing behavior prediction. Cai et al. [7] found that perceived ease of use was the factor that had the greatest impact on sharing behavior, and input it into a regression tree for behavior prediction. Chen et al. [8] categorized the distribution of topics of shared knowledge based on LDA, and used Logistic regression models to construct a prediction model.

The neural network structure captures the higher-order features of the input data, and these features enable it to process knowledge data efficiently, so many researchers have used deep learning techniques and auxiliary information such as social relationships to generate more effective and comprehensive feature representations to improve the prediction performance of knowledge sharing behaviors. Chang et al. [9] used CNN-LSTM to build a prediction model based on a search engine environment to study the trend of users' sharing behavior on different pages. Tong and Zhan [10] used a multilayer perceptron (MLP) network to predict knowledge sharing behaviors using data on English knowledge and user characteristics shared by users on a MOOC platform. Zhong et al. [11] considered knowledge features, social relationships between users and used RNN and attention mechanisms to predict user sharing behavior on Facebook.

Recently, knowledge graph (KG), as an emerging approach to knowledge representation and management, has gradually received attention in the field of knowledge sharing behavior research [12,13]. Mining and analyzing knowledge using KG technology can facilitate more efficient knowledge sharing among users. Rossi et al. [14] Ying and used Graph Neural Network (GNN) for feature extraction of English teaching KGs, and achieved the prediction of sharing behaviors by softmax, but the prediction effect is not satisfactory. Liang et al. [15] proposed an efficient graph convolutional neural network to analyze sentiment data thus improving the sharing behaviour prediction performance. Ramezani et al. [16] combined the knowledge graph representation technique TranSD and graph attention network to achieve the prediction of users' knowledge sharing behavior. To improve the accuracy of prediction, researchers and scholars have incorporated social

network data into their models [17]. Social relationship data belongs to graph data, which reflects the trust relationship of users, and GNN is extremely suitable for dealing with this kind of data. Tran et al. [18] used GCN analysis to extract social network data to enrich the user's representation vectors, thus improving behavioral prediction. Zhang et al. [19] used GCN to analyze user social network data to enrich the semantic representation of users and to build a knowledge graph, which leads to a vector representation of the knowledge, and finally fuses the user and knowledge representation vectors for sharing behavior prediction.

2. **Theoretical analysis.**

2.1. **Knowledge graphs represent learning techniques.** A KG has the structural characteristics of a graph and consists of nodes and edges [20], where each node represents an entity and each edge corresponds to a link between entities. Each node represents an entity, each edge corresponds to a link between entities, and every two nodes and one edge constitute a piece of knowledge, i.e., a triple (h, r, t) . KG models the relationship between knowledge in the form of a triple, and the semantic information it contains can be of great help in predicting sharing behavior. To overcome the data sparsity issue of KG, traditional data processing approaches, such as the unique heat code representation, are susceptible to data sparsity, and the computational results may be inaccurate and complicated. In contrast, representation learning methods project entities and relations into vector spaces for computation, effectively alleviating the data sparsity problem and significantly improving computational efficiency. The KG representation model is mainly a translational distance model [21], which is centered on the distance metric and inscribes the distance between entities through a scoring function, based on which the reasonableness of the triad is judged [22]. Commonly used models for translational distance representation include the TransE, TransH, TransR and TransD models [23], while the TransD model is an improvement of the other three models and consists of two parts, the entity semantic space and the relation semantic space, as shown in Figure 1. However, considering different entity types, the TransD model sets two different projection

matrices A and B for the head and tail entities to perform the corresponding calculations as

shown in Equation (1) and Equation (2), where wh ,wt , wr

is the projection vector of head and

tail entities and relations, respectively, and I is the unit matrix.

$$M = I + w_h w_h^T \tag{1}$$

$$M = I + w_t w_t^T \tag{2}$$

Similar to other representation models, the score function for TransD is as follows.

$$f_r(h,t) = \|M_r hh + r - M_r tt\|_{1/2} \tag{3}$$

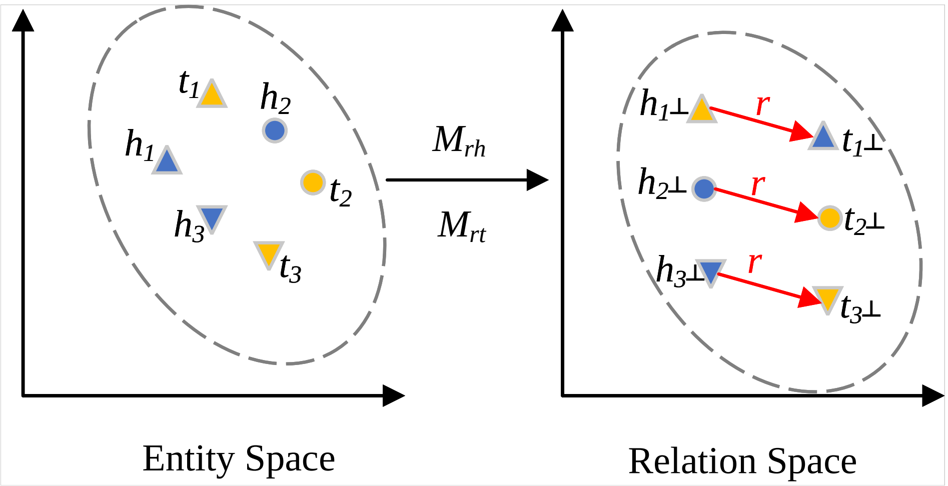

Figure 1. TransD model structure

2.2. **Graph convolutional network.** In deep learning, convolution is usually operated on Euclidean data, however, in practical applications many are non-Euclidean data, such as graph- structured data [24]. Social networks and KGs can then be represented as graph structures. There are many difficulties in applying convolution on graph structure, as shown in Figure 2, the Euclidean data is a regular matrix structure, but the graph structure does not have such a regular structure, in the graph, each node has a different number of neighbors, which makes it impossible to fix the convolution's function to achieve the weight sharing in the graph structure.

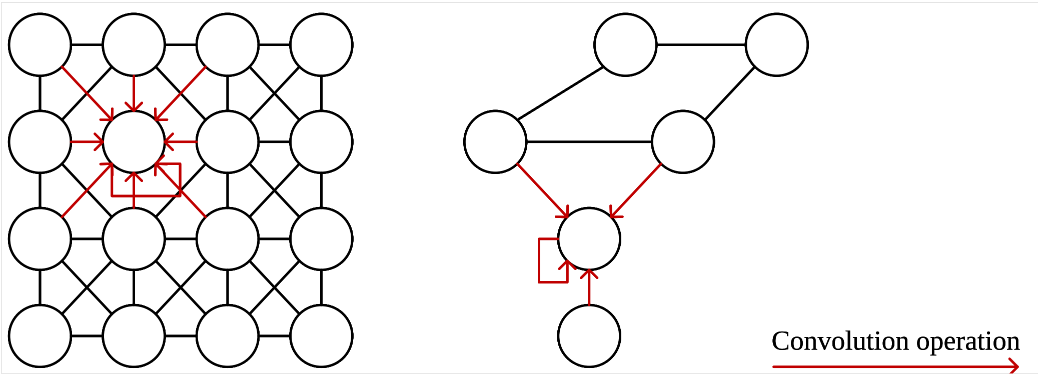

Figure 2. The difference between Euclidean spatial convolution and graph convolution

To implement the convolution operation on graphs, GCN [25] tries to define convolution in spectral space using the convolution theorem with the following equation.

$$f * g = \mathcal{F}^{-1}\big(\mathcal{F}(f) \cdot \mathcal{F}(g)\big) \tag{4}$$

where $*$ is the convolution operator, $\cdot$ is the dot product operation, $\mathcal{F}(f)$ denotes the Fourier transform of f, and $\mathcal{F}^{-1}(f)$ denotes the Fourier inverse transform. The convolution operation is represented on the graph as follows.

$$x *_G y = U\big((U^T x)\,\varepsilon\,(U^T y)\big) \tag{5}$$

where $*_G$ denotes the convolution operator on the graph, x is the signal in the nodal domain, $U^T x$ is the representation of x in spectral space, y is the convolution kernel signal on the nodal domain, $U^T y$ is the representation of the convolution kernel signal in the spectral domain, and ε represents the elemental product of the signal.

Let $U^T y = [\theta_1, \dots, \theta_n]$, $g = \mathrm{diag}([\theta_1, \dots, \theta_n])$, the final graph convolution can be expressed as follows.

$$x *_G y = U\, g_\theta\, U^T x \tag{6}$$

where g_θ denotes the convolution kernel in the spectral domain.

3. English learning knowledge graph construction and representation learning.

3.1. English learning knowledge graph construction. To construct the KG of English learning knowledge, this paper models the knowledge points used in English learning and establishes the mapping relationship between different entity structures in the knowledge base, the technical architecture of KG is shown in Figure 3. Taking the entity knowledge point of English learning as a spatial structure, it is represented as $G = (E, R, T)$, $E = (e_1, e_2, \dots, e_n)$ is the entity set, n represents n entities in the knowledge graph, is the set of relations, and $R = (r_1, r_2, \dots, r_m)$ represents the number of relations has m. Each triplet in the set is represented by (h, r, t), with h, r, t representing the head, relation, and tail entities, respectively. Then the transformation function of the entity in space is shown as follows.

$$g(h, r, t) = \|h + r - t\| \tag{7}$$

where g represents the transformation function of the entity in space. In the process of conversion, the greater the value of $g(h,r,t)$, the higher the degree of matching between English learning knowledge points and facts. To concretized the KG and establish the semantic relationship between different knowledge points, the transformed entity information is used, and the final English learning KG is shown in Equation (8).

$$K = g(h,r,t)ht^{\alpha} \tag{8}$$

where K represents the constructed English learning KG; α represents the physical dimension corresponding to the knowledge point.

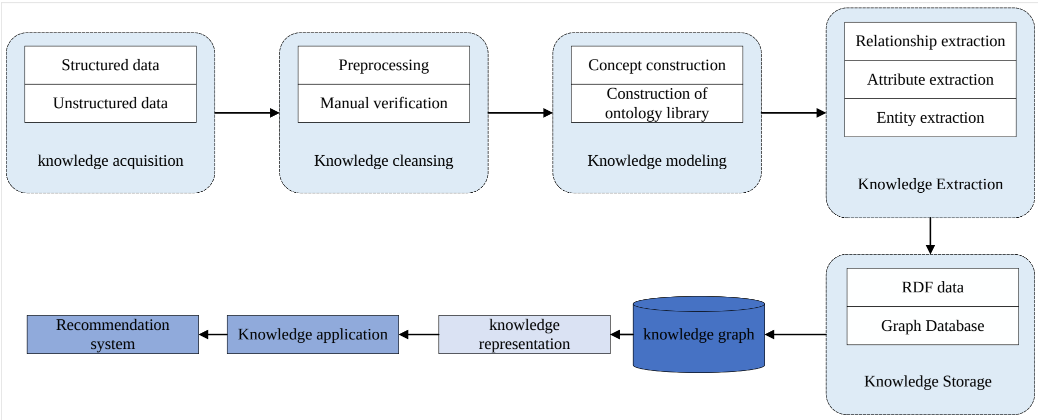

Figure 3. The technical architecture of KG

3.2. **Learning English knowledge graph representation based on enhanced TransD Model.** The current TransD ignores the underlying knowledge hidden by the wealth of additional information, resulting in an inaccurate representation of the entity and relational vectors and an inability to train the model more effectively. Thus, relied on relational path information [26] and hierarchical relationship structure [27], this paper optimizes the TransD model (ETransD), adds the representation of entity information into the path, introduces the hierarchical relationship structure information, and combines the TransD model that can handle complex relationships. Integrate into one model to improve entity and relationship representation in KG.

There is a relational path $p = \{r_1, e_1, r_2, \ldots, e_{n-1}, r_n\}$ between the entity h and t, where m indicates that this path has m relationships and $m-1$ entities. The vector representation of the path is obtained by adding the relationship on the path and the vector of the entity by the addition operation.

$$p = r_1 + e_1 + r_2 + \ldots + e_{n-1} + r_n \tag{9}$$

For the two relationship information between (h,r,t), the score function is defined as follows.

$$E(h,p,t) = E(h,r,t) + E(h,P,t) \tag{10}$$

where $E(h,r,t)$ is the score function for (h,r,t) in KG, $E(h,P,t)$ is the score function for P with multiple relational paths between entities, as shown below

$$E(h,P,t) = \frac{1}{Z} \sum_{p \in P(h,t)} R(p \mid h,t)\, E(h,p,t) \tag{11}$$

where Z is the normalization factor, $R(p \mid h,t)$ represents the possibility of the existence of a relational path P between (h,t), and $E(h,p,t)$ is the score function of a single path between entities.

For the goal of addressing the issue of entity sparsity and enhance the representation ability of KG, the three-layer hierarchical relationship structure is used to effectively integrate rich hierarchical information into the relational vector. By forming relationships into relational clusters, similar relationships are very close in semantic connection, and learning them together is conducive to knowledge sharing and obtaining more valuable information. For (h, r, t), the vector representation of r is the sum of the vector representation r_c of the relation cluster, the relation-specific vector representation r', and the vector representation r_s of the subrelationships $r_{\mathrm{HRS}} = r_c + r' + r_s$. Thus, the vector representation of each relation can make use of information from the three levels of structure, which ultimately leads to a score function that adds hierarchical information to the structured base.

$$E_{\mathrm{HRS}}(h, r, t) = \|h + r_c + r' + r_s - t\|_{1/2} = \|h + r_{\mathrm{HRS}} - t\|_{1/2} \tag{12}$$

Then, the relational path information and the hierarchical relationship structure information are unified to obtain the comprehensive score function based on the relational path information and the hierarchical relationship structure information, as shown below.

$$E = E_{\mathrm{HRS}}(h, r, t) + E(h, p, t) \tag{13}$$

4. Research on knowledge sharing behavior of English learning based on knowledge graph in social network environment.

4.1. GCN-based user feature extraction for social networks. Intending to the existing knowledge sharing behavior prediction methods ignoring the user social relationship and the problem of insufficient knowledge feature extraction, this paper carries out in-depth extraction of the user features in the social network through GCN, and combines with ETransD to mine the feature information of the shared knowledge to supplement the knowledge feature vectors in the prediction task, and at the same time, uses the attention mechanism to aggregate the user and shared knowledge vectors to get the potential feature vectors, and puts the potential feature vector and user vector to calculate the inner product to get the vector similarity, and the final prediction probability of knowledge sharing behavior is obtained in the form of inner product through the prediction function. The overall model structure of the designed knowledge sharing behavior prediction method is indicated in Figure 4.

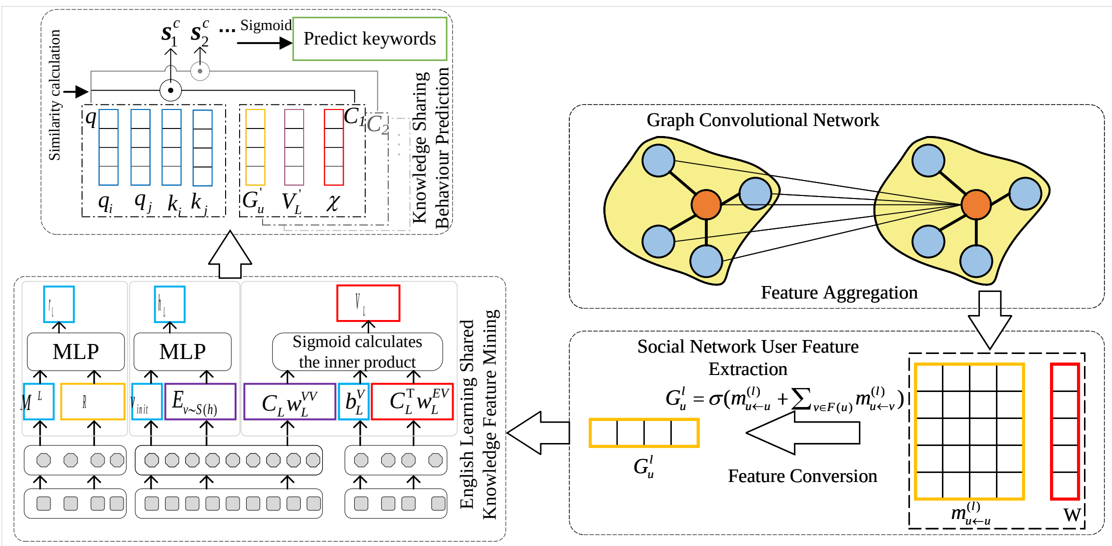

Figure 4. The entire model structure of the designed prediction method

Knowledge sharing behaviors of English learning among users are influenced by their social relationships [28], and social information contains rich information about user preferences, which can be used as a complement to knowledge sharing data. To better use social relationships to model the feature representation of users, this paper uses GCN for the embedding propagation of user nodes in the social network graph, which is divided into message creation and message aggregation, where each user can indirectly aggregate information about its K-order neighbors through K message passes.

For two users u and v who are directly connected on a social network and have a social relationship between them, using GCN to learn the social network graph, the L-th message passing process from v to u is defined as follows.

$$m_{u\leftarrow v}^{(l)} = f\big(g_u^{l-1},\, g_v^{l-1},\, p_{uv}\big) \tag{14}$$

where $m_{u\leftarrow v}^{(l)}$ is the embedding of the message, $f(\cdot)$ is the message function which receives as input the $L-1$-order embeddings g_u^{l-1} and g_v^{l-1} on the social network graph, $g^0 = q$ is the initialized user embedding representation, and p_{uv} is the decay coefficient.

To learn the weight coefficients between users flexibly and accurately, the messages passed between users are defined through GCN and attention mechanisms as follows.

$$m_{u\leftarrow v}^{(l)} = \beta_{uv}^{l-1}\, g_v^{l-1} W \tag{15}$$

where β_{uv}^{l-1} is the weight of the effect of v on u at the L-th message propagation, calculated as follows.

$$\beta_{uv}^{l-1} = \frac{\exp\Big(\mathrm{Relu}\big(a^\top\, [\, W g_u^{l-1} \;\; W g_v^{l-1}\,]\big)\Big)}{\sum\limits_{w\in\mathcal{F}(u)} \exp\Big(\mathrm{Relu}\big(a^\top\, [\, W g_u^{l-1} \;\; W g_w^{l-1}\,]\big)\Big)} \tag{16}$$

where $w \in \mathcal{F}(u)$ is the first order neighbor of user u. $g_u^{l-1}, g_v^{l-1}, g_w^{l-1}$ is the $L-1$-order embedding of users u, v, w respectively, the W-matrix performs feature transformation on the embedding, and $\|\|$ denotes the splicing operation.

After the message creation, in the message aggregation phase, the user u aggregates the received messages to update his embedding as follows.

$$G_u^l = \sigma\left(m_{u\leftarrow u}^{(l)} + \sum_{v\in\mathcal{F}(u)} m_{u\leftarrow v}^{(l)} \right) \tag{17}$$

where G_u^l is a feature representation where u aggregates information about its neighbors and σ is an activation function. The self-feature can be preserved by passing a message $m_{u\leftarrow u}^{(l)} = h^{s,l-1}W$ to itself.

4.2. **ETransD-based feature mining of shared knowledge for English language learning.** After embedding the representation of social network users, this paper firstly employs ETransD, the KG representation learning model combined with optimization in Section 3.2, to mine the feature information of shared knowledge in order to complement the knowledge feature vectors in the task of predicting knowledge behaviors in English learning. Firstly, (h, r, t) is taken as input and features are extracted from the head entity h and relation r using multilayer perceptron respectively, and tail entity features are predicted and taken as output, which is calculated as follows.

$$\hat{t} = M_L\left(\begin{bmatrix} h \\ r_L \end{bmatrix} \right) \tag{18}$$

where r represents the initial features of the relationship, $r_L = M_L(R)$, $\hat{t}$ is the vector of the predicted tail entity. $h^L = \mathbb{E}_{v\sim\mathcal{S}(h)}\big[\mathcal{C}^L(v_{\mathrm{init}}, h)[[e]]\big]$, h^L are the head entity feature representations corresponding to the knowledge, $\mathcal{S}(h)$ denotes the set of knowledge vectors associated with the head entity h, v is the head entity vector, h denotes the shared knowledge ID, and $\mathcal{C}^L$ stands for the corresponding crossover matrix, as shown in Equation (19). The actual feature vectors of the target entity and the tail entity are combined, and the inner product between them is calculated by the Sigmoid function to obtain the score similarity function, as shown in Equation (20).

$$C^L = \nu^L (e^L)^\top = \begin{bmatrix} \nu_L^{(1)} e_L^{(1)} & \cdots & \nu_L^{(1)} e_L^{(d)} \\ \vdots & \ddots & \vdots \\ \nu_L^{(d)} e_L^{(1)} & \cdots & \nu_L^{(d)} e_L^{(d)} \end{bmatrix} \tag{19}$$

$$\text{score}(h, r, t) = f(t, \hat{t}) = \sigma\big(t^\top \hat{t}\big) \tag{20}$$

where t is the feature vector of the tail entity. The resulting partition function is normalized to obtain the final prediction function, and the value domain is mapped between 0 and 1, d is the dimension of the corresponding vectors s and h for knowledge and entities, and for latent features ν_l and entities e_l, the feature vectors of the shared knowledge in the next layer are output by projecting the cross-feature matrices into the latent layer space representation.

$$V^L = C^L w_{VV}^L + (C^L)^\top w_{EV}^L + b_V^L = \nu^L (e^L)^\top w_{VV}^L + e^L (\nu^L)^\top w_{EV}^L + b_V^L \tag{21}$$

where $w_\cdot^L$ is the training weights and $b_\cdot^L$ is the bias parameter.

4.3. **Feature aggregation and prediction of knowledge sharing behavior.** To achieve efficient aggregation of social network user features and shared knowledge features, this paper uses the attention mechanism to process them to obtain the potential feature vector χ after the fusion of user and shared knowledge vectors, which is firstly computed by the query vector Q and the key vector K. The user vector is transformed to be represented by the intermediate nodes G_u and V^L of the knowledge vectors, as shown in Equation (22).

$$q_i = W_Q G_u, \quad q_j = W_Q V^L, \quad k_i = W_K G_u, \quad k_j = W_K V^L \tag{22}$$

The fused potential eigenvector χ is obtained by combining the W_Q query vector transformation matrix, the W_K key vector transformation matrix, and then the intermediate nodes G_u and V^L with the following calculation.

$$G_u' = \text{att}_{1,1}\, G_u + \text{att}_{1,2}\, V^L \tag{23}$$

$$V_L' = \text{att}_{2,1}\, G_u + \text{att}_{2,2}\, V^L \tag{24}$$

$$\chi = \lambda G_u' + \mu V_L' \tag{25}$$

where $\text{att}_{i,j}$ denotes the relative weights of the intermediate features and

$$\text{att}_{i,j} = \frac{\exp(q_i k_j^\top)}{\sum \exp(q_i k_j^\top)}, \quad \lambda \text{ and } \mu \text{ denote the balancing factors.}$$

The fused potential feature vectors are used as inputs to the lower layer, and then the inner product of the fused vectors and user vectors is computed to obtain the vector similarity, and this is used to predict the user knowledge sharing behavior.

$$\hat{y}_{uv} = \sigma\big((G_u \cdot V^L) \parallel \chi\big) \tag{26}$$

where σ denotes the Sigmoid activation function.

After establishing the knowledge sharing behavior prediction model, this paper alternates the feature aggregation and shared knowledge feature mining modules for independent training, and the loss function is designed as three parts, and the loss function of the feature aggregation module is as follows.

$$L_{uv} = \sum_{u \in U,\, v \in V} f(\hat{y}_{uv}, y_{uv}) \tag{27}$$

where $f(\cdot)$ is the cross-entropy loss function and $\hat{y}_{uv}$ is the prediction function, the result of the inner product is calculated by the activation function. The loss function of shared knowledge feature mining module is as follows.

$$L_{\mathrm{KG}} = \sum_{(h,r,t) \in \mathcal{G}} \mathrm{score}(h,r,t) \;-\; \sum_{(h',r',t') \notin \mathcal{G}} \mathrm{score}(h',r',t') \;+\; \sum_{n} r \log(r') \tag{28}$$

The goal of this module is to increase the scores of all positive triples while decreasing the scores of all negative sample triples. The final loss function can be expressed as follows.

$$L = L_{uv} + \rho_1 L_{\mathrm{KG}} + \rho_2 \frac{\|W\|^2}{2} \tag{29}$$

The last term of the loss function is a regularization term to prevent overfitting, which consists of two balancing parameters ρ_1 and ρ_2. ρ_1 is the ratio of the learning rates of the two tasks and ρ_2 is the hyperparameter. The gradient is decreased by the optimizer function and the values of the parameters in the model are constantly updated.

5. Experiments and analysis of results.

5.1. Comparison of the accuracy of knowledge sharing behavior prediction.

This paper uses 3694 user data, 53978 English learning knowledge, and 9805 knowledge sharing records crawled from online learning platforms in literature [29] as the experimental dataset, and divide this dataset into training set, testing set, and validation set according to the ratio of 7:2:1. The experiments were conducted with Windows 10 operating system, 64-bit processor, 1,755 MHz 24GD6 GeForce RTX 2080 Ti GPU, 256 GB of physical memory, PyCharm2020 compilation environment, Python version Python3, and Pytorch framework. To improve the robustness and convergence speed of the model, the experiments use Dropout and batch regularization for entity embedding, and choose L2 regularization to prevent overfitting. The learning rate is set to 0.01, the embedding dimension of entities and relations is set to 100, and the batch size is 128.

For ease of analysis, the method in the literature [14] is denoted as MT1, the method in the literature [18] as MT2, the method in the literature [19] as MT3, and the proposed method as MT4, and the prediction is performed using the mean accuracy rate (MAP), the overall accuracy rate (OA), the overall precision rate (OP), the overall recall (OR), the overall F1 value (OF1), and the ROC curves accuracy assessment. This paper has statistically counted the MAP of knowledge sharing behaviors of the four models on each Batch size of the test set using box- and-line diagrams as shown in Figure 5. The MAP values of MT1, MT2, MT3 and MT4 are 0.825, 0.863, 0.881 and 0.918 respectively, MT4 has the highest MAP value and the best prediction performance, and the performance of MT2 and MT3 is almost comparable, MT1's prediction accuracy is significantly lower than that of the other three models, and although it constructs an English language knowledge graph for the online learning platform, does not consider the platform's the influence of users' social relationships on shared knowledge in the platform, so the prediction accuracy is not high.

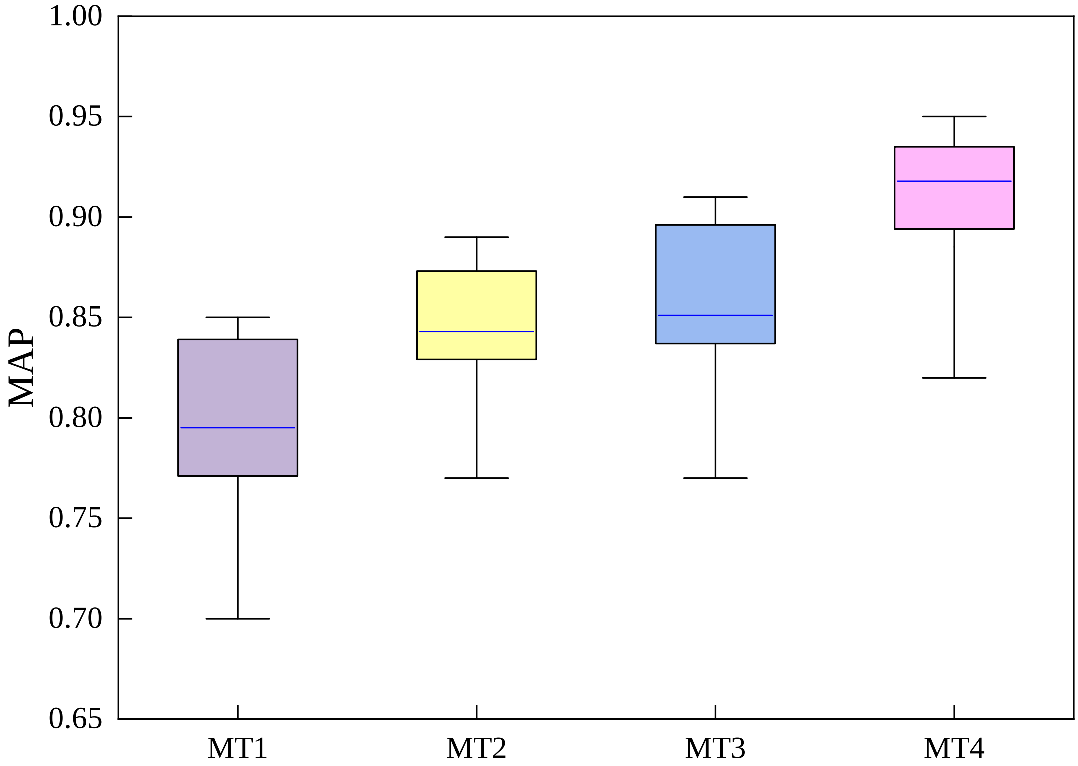

Figure 5. MAP of different knowledge sharing behavior prediction methods

The prediction accuracy indexes of different methods are shown in Table 1. All the prediction accuracy indexes of MT4 are higher than the other methods. The OF1 of MT4 is 0.911, which is improved by 0.06, 0.022, and 0.018 compared with MT1, MT2, and MT3, respectively. Both MT1 and MT2 are based on GNN for feature extraction of English KGs, the difference is that MT2 considers social network data and enriches the representation of user feature vectors, so the accuracy of MT2 is higher than that of MT1. MT3 uses GCN for user and knowledge feature mining of social network graphs and English KGs respectively, but it does not take into account the user's neighborhood information and does not optimize the representation learning model for KGs, and the prediction accuracy is not as good as that of MT4, which in general performs the best in terms of prediction accuracy.

Table 1. Prediction accuracy indexes of different knowledge-sharing behaviors

Indicator	MT1	MT2	MT3	MT4
OA	0.824	0.861	0.894	0.933
OP	0.865	0.896	0.883	0.916
OR	0.838	0.882	0.904	0.907
OF1	0.851	0.889	0.893	0.911

Further, this paper compares the ROC curves and AUC values of the four prediction methods in order to comprehensively analyze the prediction ability of the various methods on sharing behaviors, and obtains the comparison graph shown in Figure 6. The horizontal coordinate is specificity (False Positive Rate, FPR) and the vertical coordinate is sensitivity (True Positive Rate, TPR). The AUC values of all methods are greater than

0.9, indicating that all four methods are effective, and the larger the AUC value, the higher the prediction accuracy of the model, and the AUC value of MT4 is improved by 3.3%, 2%, and 2.1% compared with that of MT1, MT2, and MT3, respectively, which indicates that the prediction accuracy of MT4 is better than that of the other three models.

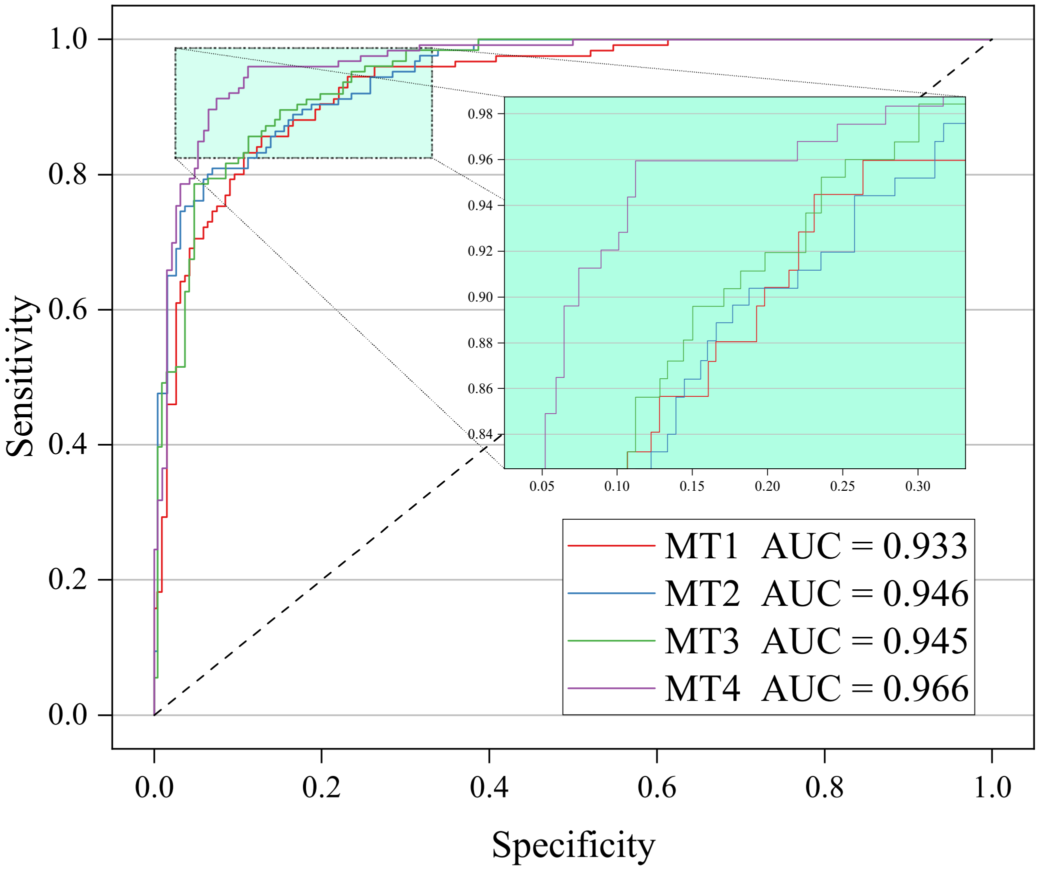

Figure 6. ROC curves for different knowledge sharing behavior prediction methods

5.2. **Comparison of errors in forecasting.** In addition to visualizing the prediction capability of the four methods through the above prediction accuracy metrics, further in-depth analyses of the prediction error metrics of the coefficient of determination R2, the mean absolute error MAE, the mean square error MSE, and the root mean square error RMSE are also needed to assess the prediction performance in a more comprehensive way. The prediction error indicators for each method are shown in Table 2.

Table 2. Prediction error comparison

Indicator	MT1	MT2	MT3	MT4
R2	0.875	0.913	0.937	0.962

As can be seen from Table 2, all the accuracy evaluation indexes of MT4 are better than the other three methods, in which the MAE and MSE of MT4 are 0.084 and 0.135, respectively, which are reduced by 0.112 and 0.124, respectively, compared with MT1, 0.059 and 0.079, respectively, compared with MT2, and 0.025 and 0.031, respectively, compared with MT3. This suggests that a knowledge representation that incorporates users' social relationships and KG is more suitable for solving the problem of predicting knowledge-sharing behaviors in English language learning. Comparing the coefficient of determination R2, the closer R2 is to 1, the higher the prediction accuracy of the model is, and the R2 of MT4 is 0.962, which is the closest to 1, and it is improved by 9.9%, 5.3%, and 2.7% compared to MT1, MT2, and MT3, respectively. MT4 not only uses the ETransD for knowledge feature mining, but also utilizes the attention mechanism to merge the user features with the knowledge features and to highlight the critical features, which greatly improves the prediction accuracy.

6. **Conclusion.** English is a subject covering a large amount of knowledge, and knowledge sharing among students can significantly improve English learning ability. Aiming at the problems that the existing knowledge sharing behavior prediction methods ignore the social factors of users and the knowledge features are not sufficiently extracted, this paper investigates a KG-based knowledge sharing behavior prediction method for English learning in the social network environment. Firstly, the English learning KG is modelled to establish the mapping relationships of different entities. On this basis, the TransD model is enhanced by uniting the relationship path information and the hierarchical relationship structure (ETransD), and the vector representation of the relationship is converted into the sum of relationship vectors at three levels by obtaining the potential semantic information of the multi-step relationship paths in order to enhance the ability of knowledge representation learning. Then GCN is used to capture social network user features and combined with ETransD to mine the feature information of shared knowledge. Finally, the attention mechanism is used to aggregate user and shared knowledge features, and the final prediction probability of knowledge sharing behavior is obtained in the form of inner product through the prediction function. The experimental outcome implies that the suggested method has high prediction accuracy and low prediction error, and can be better applied to the field of knowledge sharing behavior prediction.

REFERENCES

[1] S. Wang and R. A. Noe, "Knowledge sharing: A review and directions for future research," *Human Resource Management Review*, vol. 20, no. 2, pp. 115–131, 2010.
[2] D. Li and V. Edwards, "English language teaching and educational reform in Western China: A knowledge management perspective," *System*, vol. 47, pp. 88–101, 2014.
[3] W. J. Ali, "Factors Affecting on Knowledge Sharing Among Undergraduate Students," *Cihan University-Erbil Journal of Humanities and Social Sciences*, vol. 5, no. 1, pp. 75–81, 2021.
[4] M. Ramzan, Z. K. Javaid, and A. A. Ali, "Perception of Students about Collaborative Strategies Employed by Teachers for Enhancing English Vocabulary and Learning Motivation," *Pakistan JL Analysis & Wisdom*, vol. 2, p. 146, 2023.
[5] R.-C. Shih and S.-J. Lou, "The development and application of a knowledge sharing behavior model for Taiwanese junior high school English teachers," *African Journal of Business Management*, vol. 5, no. 30, p. 12066, 2011.
[6] Z.-Y. Chen, Z.-P. Fan, and M. Sun, "A hierarchical multiple kernel support vector machine for customer churn prediction using longitudinal behavioral data," *European Journal of Operational Research*, vol. 223, no. 2, pp. 461–472, 2012.
[7] Y. Cai, Y. Yang, and W. Shi, "A predictive model of the knowledge-sharing intentions of social Q&A community members: a regression tree approach," *International Journal of Human–Computer Interaction*, vol. 38, no. 4, pp. 324–338, 2022.
[8] S. Chen, N. Andrienko, G. Andrienko, L. Adilova, J. Barlet, J. Kindermann, P. H. Nguyen, O. Thonnard, and C. Turkay, "LDA ensembles for interactive exploration and categorization of behaviors," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 9, pp. 2775–2792, 2019.
[9] C.-W. Chang, C.-Y. Chang, and Y.-Y. Lin, "A hybrid CNN and LSTM-based deep learning model for abnormal behavior detection," *Multimedia Tools and Applications*, vol. 81, no. 9, pp. 11825–11843, 2022.

[10] Y. Tong and Z. Zhan, "An evaluation model based on procedural behaviors for predicting MOOC learning performance: students' online learning behavior analytics and algorithms construction," *Interactive Technology and Smart Education*, vol. 20, no. 3, pp. 291–312, 2023.

[11] C. Zhong, F. Xiong, S. Pan, L. Wang, and X. Xiong, "Hierarchical attention neural network for information cascade prediction," *Information Sciences*, vol. 622, pp. 1109–1127, 2023.

[12] A. Chennamaneni, J. T. Teng, and M. Raja, "A unified model of knowledge sharing behaviours: theoretical development and empirical test," *Behaviour & Information Technology*, vol. 31, no. 11, pp. 1097–1115, 2012.

[13] A. Sheth, S. Padhee, and A. Gyrard, "Knowledge graphs and knowledge networks: the story in brief," *IEEE Internet Computing*, vol. 23, no. 4, pp. 67–75, 2019.

[14] A. Rossi, D. Barbosa, D. Firmani, A. Matinata, and P. Merialdo, "Knowledge graph embedding for link prediction: A comparative analysis," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 15, no. 2, pp. 1–49, 2021.

[15] B. Liang, H. Su, L. Gui, E. Cambria, and R. Xu, "Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks," *Knowledge-Based Systems*, vol. 235, 107643, 2022.

[16] M. Ramezani, M.-R. Feizi-Derakhshi, and M.-A. Balafar, "Text-based automatic personality prediction using KGrAt-Net: a knowledge graph attention network classifier," *Scientific Reports*, vol. 12, no. 1, 21453, 2022.

[17] L. Cao, H. Zhang, and L. Feng, "Building and using personal knowledge graph to improve suicidal ideation detection on social media," *IEEE Transactions on Multimedia*, vol. 24, pp. 87–102, 2020.

[18] Q. M. Tran, H. D. Nguyen, T. Huynh, K. V. Nguyen, S. N. Hoang, and V. T. Pham, "Measuring the influence and amplification of users on social network with unsupervised behaviors learning and efficient interaction-based knowledge graph," *Journal of Combinatorial Optimization*, vol. 44, no. 4, pp. 2919–2945, 2022.

[19] L. Zhang, W. Zhang, L. Wu, M. He, and H. Zhao, "SHGCN: Socially enhanced heterogeneous graph convolutional network for multi-behavior prediction," *ACM Transactions on the Web*, vol. 18, no. 1, pp. 1–27, 2023.

[20] T.-Y. Wu, H. Li, S. Kumari, and C.-M. Chen, "A Spectral Convolutional Neural Network Model Based on Adaptive Fick's Law for Hyperspectral Image Classification," *Computers, Materials & Continua*, vol. 79, no. 1, pp. 19–46, 2024.

[21] T.-Y. Wu, A. Shao, and J.-S. Pan, "CTOA: Toward a Chaotic-Based Tumbleweed Optimization Algorithm," *Mathematics*, vol. 11, no. 10, 2339, 2023.

[22] T.-Y. Wu, H. Li, and S.-C. Chu, "CPPE: An Improved Phasmatodea Population Evolution Algorithm with Chaotic Maps," *Mathematics*, vol. 11, no. 9, 1977, 2023.

[23] Z. Tan, X. Zhao, Y. Fang, and W. Xiao, "GTrans: generic knowledge graph embedding via multi-state entities and dynamic relation spaces," *IEEE Access*, vol. 6, pp. 8232–8244, 2018.

[24] S. Zafeiriou, M. Bronstein, T. Cohen, O. Vinyals, L. Song, J. Leskovec, P. Liò, J. Bruna, and M. Gori, "Guest editorial: Non-euclidean machine learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 2, pp. 723–726, 2022.

[25] S. Zhang, H. Tong, J. Xu, and R. Maciejewski, "Graph convolutional networks: a comprehensive review," *Computational Social Networks*, vol. 6, no. 1, pp. 1–23, 2019.

[26] Y. Shen, N. Ding, H.-T. Zheng, Y. Li, and M. Yang, "Modeling relation paths for knowledge graph completion," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 11, pp. 3607–3617, 2020.

[27] H. A. Santoso, S.-C. Haw, and Z. T. Abdul-Mehdi, "Ontology extraction from relational database: Concept hierarchy as background knowledge," *Knowledge-Based Systems*, vol. 24, no. 3, pp. 457–464, 2011.

[28] S. Panahi, J. Watson, and H. Partridge, "Towards tacit knowledge sharing over social web tools," *Journal of Knowledge Management*, vol. 17, no. 3, pp. 379–397, 2013.

[29] A. Alsayat and H. Ahmadi, "A hybrid method using ensembles of neural network and text mining for learner satisfaction analysis from big datasets in online learning platform," *Neural Processing Letters*, vol. 55, no. 3, pp. 3267–3303, 2023.