Sensors and Materials, Vol. 35, No. 12 (2023) 1–22
MYU Tokyo

S & M ####

Application of Graph Convolutional Network in the Construction of Knowledge Graph for Higher Mathematics Teaching

Ruixue Yuan,[1] Hongming Li,[2] Zhenyu Sun,[3*] and Huiwen Zhang[4]

[1]School of Economics and Management, Beijing University of Chemical Technology, Beijing 100029 China
[2]College of Education, University of Florida, Gainesville 32601, United States
[3]School of Humanities, University of Chinese Academy of Sciences, Beijing 100049, China

(Received February 2, 2023; accepted December 8, 2023)

Keywords: graph convolutional network, higher mathematics, knowledge graph construction, triples, knowledge classification

In this study, we first perform data preprocessing using higher mathematics textbooks and network resources to complete the construction of the dataset extracted from the knowledge of this mathematical discipline in order to study the application of a graph neural network in the knowledge graph (KG) of higher mathematics. Then, the graph convolutional network (GCN) and attention mechanism are introduced, and the relationship extraction model based on attention-guided GCN and the text classification model based on GCN are established, whereby the automatic extraction of higher mathematical knowledge can be realized. Finally, sensors are used to collect data from students' classroom responses, the data layer construction of a subject KG is realized, and the database is used to realize the storage visualization of triples. The results show that on the basis of the text classification model of GCN, higher mathematical knowledge classification is set, providing a reference example for the application of GCN in higher education. The constructed higher mathematics atlas has a total of 2580 triples, which can be used on the higher mathematics visual query platform. The self-built recognizer recognition test set obtains a P-value of 89.16% in the higher mathematics theorem law entity and a P-value of 97.73% in the test question of higher mathematics. Therefore, the GCN model established here can be effective in the construction of higher mathematical KGs.

1. Introduction

In today's digital age, the rapid development of sensor technology has facilitated the collection and acquisition of a wide range of data. Sensors are devices that can sense and measure various physical quantities in the environment, including, for example, temperature, pressure, light, and sound. In education, the application of sensors also presents great potential. The application of sensors in teaching higher mathematics education is of great significance and potential value. Much information can be obtained about students' learning behaviors, cognitive processes, and reaction times by using sensors to collect data from students during the learning

*Corresponding author: e-mail: sunzhenyu@ucas.ac.cn

ISSN 0914-4935 © MYU K.K.
https://myukk.org/

process. Such data can help to understand students' learning needs, troublesome points, and learning strategies, which will enable appropriate support and guidance for personalized instruction. In higher mathematics teaching, sensors can be used to collect students' dynamic information during problem solving, such as handwriting, mouse trajectories, and eye tracking data. Such data can help us analyze factors, such as students' problem-solving strategies, attention span, and cognitive load, to gain insight into students' mathematical thinking processes and pain points. On the basis of the data, an accurate and comprehensive higher mathematical knowledge graph (KG) can be established to reveal the connections and learning paths among mathematical concepts and provide teachers with targeted teaching strategies and personalized feedback.

The application and popularization of China's Internet technology and computer platform have provided great convenience for developing various fields and industries. A stable and comprehensive development environment has been created to lay a solid foundation for the creation of future development prospects.[1] Graph convolutional network (GCN) is a high-precision data processing method. It is more flexible and changeable than traditional data processing modes. In practical application, errors can be better avoided. GCN can solve existing problems quickly in a complex context.[2] In the context of the new era, online education technology has gradually matured. Compared with traditional teaching methods, such technology has significant advantages in terms of mode, scope, resources, and teachers. On the basis of the above findings, it is also possible to construct a KG in combination with the actual needs of higher mathematics teaching and organize and display the architecture of disciplines in education.[3]

The research objective is to develop teaching methodologies of higher mathematics. From higher mathematics textbooks and network resources, the content information of books is obtained for data cleaning, word segmentation, and other related types of data preprocessing to complete the construction of datasets extracted from higher mathematical knowledge. Then, an entity recognition model is proposed. The GCN and attention mechanism (AM) are introduced. A relational extraction model based on attention-guided graph convolutional network (AGGCN) and a GCN-based text classification model are proposed to realize the automatic extraction of higher mathematical knowledge. Finally, the data layer construction of the subject KG is realized using the extraction model of higher mathematical knowledge. The storage visualization of triples is realized using the database. Our purpose is to construct a subject KG to help users realize the query of related knowledge. A well-structured mathematical KG can help users efficiently memorize, learn, and research mathematical knowledge. It also significantly improves the user's ability to identify and resolve problems. The construction of the KG of higher mathematics can reveal the subject structure, which will assist teaching and promote the development of the education industry. Therefore, our technology has crucial value and significance at the user application and theoretical support levels.

2. Literature Review

The earliest practical application of the KG was in the 1960s by Glanzel *et al.*, who combined scientometrics with the KG to show the structural relationship of scientometrics in the form of graphs.(4) The most widely accepted definition is that "a KG is essentially a knowledge base for a semantic web." Nechaev *et al.* built a library of common sense. They built a multilingual database based on DBpedia based on Wikipedia entries. Many foreign applications are based on DBpedia.(5) The application of a domestic KG is also developing rapidly, such as Zhishi.me. It is the first Chinese KG developed by Shanghai Jiao Tong University in China. Its raw data are based on encyclopedic entries.(6) In addition, there is a bilingual KG developed by Allibaih. It is characterized by a wealth of semantic relationships. It also provides corresponding user interfaces to third parties.(7)

Presently, foreign scholars tend to adopt semisupervised and supervised methods to construct a corpus of the KG by collecting network information on the Internet. Song *et al.* summarized various categories of entities and used the conditional random field model to complete entity recognition.(8) For entity recognition in open domains, Wei *et al.* proposed an unsupervised learning method that identifies entities based on semantic features, using a clustering algorithm to group the identified entities. The algorithm was applied to the keyword autocompletion effect to achieve good results.(9) Duanmu and Xing used a remotely supervised method to complete the acquisition and preprocessing of data to reduce errors caused by manual operations. Also, a semi-automatic method of collecting financial KG data was proposed.(10) Using the existing KG, Tian *et al.* investigated how to screen different data sources to ensure the quality of the created graphs. In addition, the Chinese supervised relationship extraction method was proposed.(11)

In recent years, methods based on graph neural networks (GNNs) have emerged. Coley *et al.* applied GCN to natural language for the first time. GCN does not require pretrained word embeddings; embedding is accomplished by constructing graphs of word-to-word and word-to-sentence relationships among the data. Hence, in this study, GCN is used to extract features to look for relationships between the data.(12) In the process of constructing a KG in education, for small sample data in professional fields, such as mathematics and physics, GNNs can better extract the data characteristics.

Therefore, GCN has been widely used in the construction of KGs for higher mathematics teaching and has achieved some important results. However, some limitations remain. For example, GCN requires a large amount of labeled data for training in the construction of KGs for higher mathematics teaching, which is time-consuming and labor-intensive. In addition, representing the knowledge of higher mathematics as an accurate graph structure is a challenge. Referring to the above research theory, we propose an innovative construction of a KG for higher mathematics using GCN. Its innovations are as follows. Here, manual methods and web crawler technology are used for data collection, and entity extraction is carried out by bidirectional long short-term memory–conditional random field (BiLSTM-CRF). Then, GCN relationship classification is carried out to complete knowledge extraction. In addition to studying the query function, the functions of entity recognition and automatic classification are also proposed. They enable us to save time and visualize hard-to-interpret math problems, which makes comprehension easier for students. In this way, the current research gap in this area is filled.

3. Materials and Methods

3.1 KG

A KG is mainly used to express the properties of a concept and its relationship to other concepts. It represents the thing to be described in the form of symbols. It usually consists of nodes and edges. Nodes represent entities, and edges generally refer to relationships between entities. The main components of a KG are shown in Fig. 1.

The construction of the KG is divided into three stages: data collection, knowledge modeling, and knowledge application. First, data are collected. Automated, semi-automated, and manual data collection methods are adopted for unstructured, semistructured, and structured data, respectively. The data are integrated for preliminary representation through knowledge extraction. Then, an entity-build model is formed by entity alignment. The data specification of the corresponding layer is developed. Finally, the quality of this knowledge is reasoned and evaluated, and the KG is applied.[13]

The subject KG refers to applying knowledge from a subject area to the KG. It mainly contains the relationship among knowledge within a specific subject area. During the construction process, the internal relationships between subjects and subject textbooks, the syllabus and the logical order of knowledge points, and students' cognitive levels must be considered.

3.2 GCN

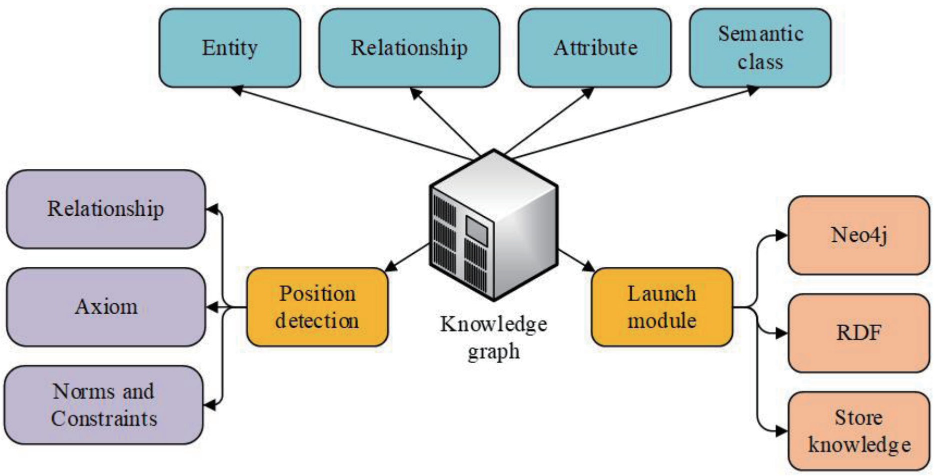

Fig. 1. (Color online) Main components of KG.

GCN is a deep learning model based on graph structure data. It is a network for information propagation and feature extraction on a graph structure through convolution operations. In general, the basic structure of GCN consists of the following key components: input presentation layer, graph convolutional layer, activation function, pooling or aggregation layer, and output layer. The specific model structure is shown in Fig. 2.

As seen in Fig. 2, the entire GCN model can be composed of multiple graph convolutional layers and other additional layers, which are trained and optimized by backpropagation algorithms. The graph coding layer is used to encode the graph structure by integrating the information of nodes and edges into the coded representation of the nodes to capture the global information in the graph structure. The attention guide layer uses the AM to guide the model to focus on important nodes and edges in the graph, and improves the model's attention and weight on key information. The ternary discriminant layer is used to identify the triplet relationship between nodes, such as the relationship between entities or concepts, to further enrich the understanding and expression ability of the model on the KG.

3.2.1 Attention guide layer

The AM is a way to efficiently calculate the impact of important parts of the data. The multiheaded AM can unearth important information that affects the outcome from multiple levels of the data.[14] The adjacency matrix is transformed into an edge-weight connection diagram by attention matrix $\overline{\boldsymbol{A}}$, which can be calculated using

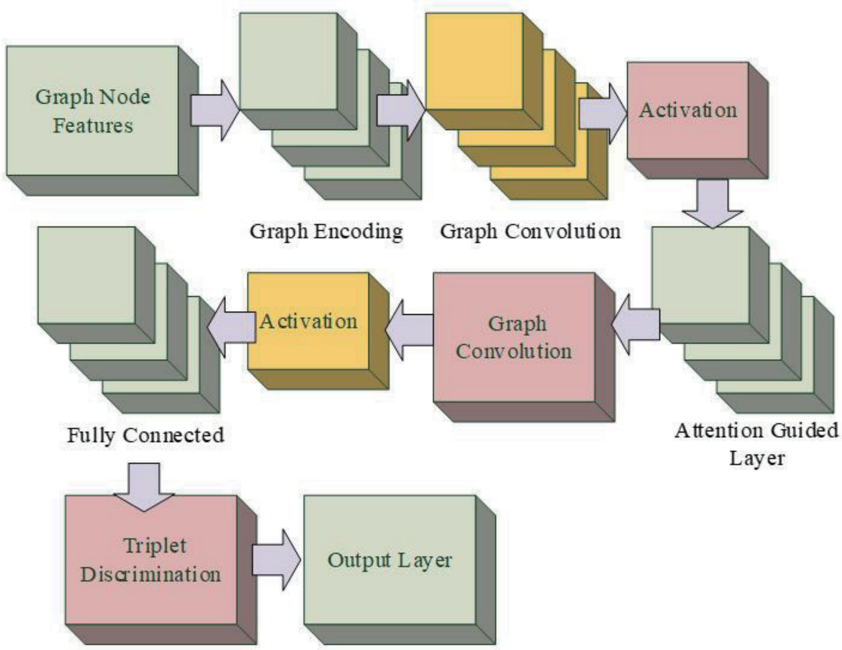

Fig. 2. (Color online) GCN structure.

$$\overline{\boldsymbol{A}} = Softmax\left(\frac{\boldsymbol{Q}\boldsymbol{W}_i^Q \times \left(\boldsymbol{K}\boldsymbol{W}_i^K\right)^T}{\sqrt{d}}\right), \tag{1}$$

where Q and K are graphical feature representations of the l layer. $\boldsymbol{W}_i^Q \in \mathbb{R}^{d\times d}$ and $\boldsymbol{W}_i^K \in \mathbb{R}^{d\times d}$ are all parameter matrices. d is the dimension of the input feature. Accordingly, the original graph of the initial input is converted to a fully connected edge-weight graph. The output vector $\boldsymbol{h}_i^l$ of node i in layer l is obtained as

$$\boldsymbol{h}_i^l = \sigma\left(\sum_{j=1}^{n} \overline{A}_{ij}\boldsymbol{W}^{(l)}\boldsymbol{h}_j^{(i-1)}\right) + \boldsymbol{b}^{(l)}, \tag{2}$$

where $\overline{A}_{ij}$ represents the weight of node i to the edge of node j. $\boldsymbol{b}^{(l)}$is the deviation vector. $\boldsymbol{h}_j^{(i-1)}$ is the initial input word vector. $\boldsymbol{W}^{(l)}$ is the weight matrix for linear transformation. σ is the node coefficient.

3.2.2 Graph coding layer

The graph coding layer is used to encode the text dependence information output by the data preprocessing module into graph structure data to support the model in effectively capturing entity association information from graph structure data.(15) The traditional models used for entity encoding are usually the hidden Markov model (HMM) and bidirectional encoder representation from transformers (BERT) model, used primarily for large-scale unlabeled corpus training and encoding after obtaining rich semantic information of the text. Here, the GCN model is used to encode for comparative analysis. GCN characterizes nodes as neighboring nodes. In the GCN of layer l, the input set $\{\boldsymbol{h}_1^0, \boldsymbol{h}_2^0, ..., \boldsymbol{h}_n^0\}$ is given. Then, $\{\boldsymbol{h}_1^l, \boldsymbol{h}_2^l, ..., \boldsymbol{h}_n^l\}$ is the output set. The output vector $\boldsymbol{h}_i^l$ of node i in layer l is represented by node i and its neighbors.

$$\boldsymbol{h}_i^l = \sigma\left(\sum_{j\ 1}^{n} \overline{A}_{ij}\boldsymbol{W}^{(l)}\boldsymbol{h}_j^{(i\ 1)}\right) + \boldsymbol{b}^{(l)} \tag{3}$$

After the L-layer GCN processing of each node vector, the hidden representation of the node is obtained. A characteristic representation of a sentence can be obtained using these word representations.

$$\boldsymbol{h}_{sent} = f\left(\boldsymbol{h}^{(L)}\right) = f\left(\mathrm{GCN}\left(\boldsymbol{h}^{(0)}\right)\right) \tag{4}$$

$\boldsymbol{h}^{(L)} \in \mathbb{R}^{n\times d}$ represents all hidden representations of the L layer (L is a hyperparameter). The function $f : \mathbb{R}^{n\times d} \to \mathbb{R}^d$ transforms n vectors into a sentence vector. Similarly, the ith entity $\boldsymbol{h}_{ei}$ is calculated using Eq. (5).

$$\boldsymbol{h}_{ei} = f\left(\boldsymbol{h}_{ei}^{(L)}\right) \tag{5}$$

3.2.3 Triad discriminant layer

The triples extraction depends on determining entity pairs and the relationship between entities. The discriminative relationship between entity pairs that remove alias entities constitutes a triple.(16) The triad discriminant layer calculates the overall expression of the GCN-encoded sentence through a dimensional transformation function, which is expressed as

$$\boldsymbol{h}_{sent} = f\left(\boldsymbol{h}^{(L)}\right) = f\left(\mathrm{GCN}\left(\boldsymbol{h}^{(0)}\right)\right). \tag{6}$$

The relationship extraction task can be considered as classifying the relationships of pairs of entities in the description text. The relationship r_{ij} between entity $\boldsymbol{h}_{ei}$ and entity $\boldsymbol{h}_{ej}$ can be calculated by a feedforward neural network $FFNN(\cdot)$.

$$r_{ij} = FFNN\left(\boldsymbol{h}_{ei}, \boldsymbol{h}_{ej}, \boldsymbol{h}_{sent}\right) \tag{7}$$

Here, the Softmax function is used to predict the relationship category of r_{ij}.

$$\overline{r} = P\left(r_{ij} \mid \boldsymbol{h}_{ei}, \boldsymbol{h}_{ej}, \boldsymbol{h}_{sent}\right) = Softmax\left(MLP\left(\boldsymbol{h}_{ei}, \boldsymbol{h}_{ej}, \boldsymbol{h}_{sent}\right)\right) \tag{8}$$

For triples extraction, the predictive model can determine which entities are the same for a given set of entities $\{\boldsymbol{h}_{e1}, \boldsymbol{h}_{e2}, ..., \boldsymbol{h}_{en}\}$. The relationship r_{ij} between $\boldsymbol{h}_{ei}$ and $\boldsymbol{h}_{ej}$ of different entities is output. Then, the forms of the triple $\boldsymbol{h}_{ei}$, r_{ij}, and $\boldsymbol{h}_{ej}$ are generated. Therefore, the objective function is defined as

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m} \lg\left(\overline{r_i}\right) + E_{\langle \boldsymbol{h}_e', \boldsymbol{h}_e \rangle}\left(\arg\max \mid \left\|\boldsymbol{h}_e\right\|_2, \left\|\boldsymbol{h}_i\right\|_2\right), \tag{9}$$

where m is the number of relationship labels. The gradient descent method can be used to find the parameters of the objective function. In addition, the hyperparameter learning rate α is adopted to update the parameters as

$$\theta \leftarrow \theta + \alpha \nabla_{\theta} J(\theta). \tag{10}$$

4. Model Design

4.1 Higher mathematical entity extraction method based on BiLSTM-CRF

Higher mathematical knowledge is divided into two types of entities: objectively existing entities and theorem law entities. The BiLSTM-CRF model is applied to the extraction of higher mathematical knowledge. Figure 3 shows the process flow.

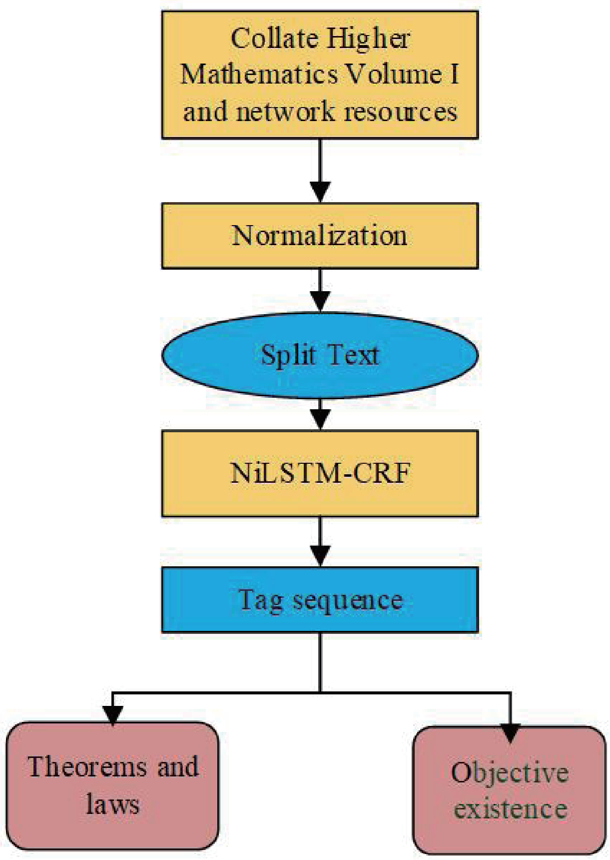

Fig. 3. (Color online) Higher mathematical knowledge extraction process.

There are generally three methods of entity extraction: rule-based, machine-learning-based, and deep learning (DL)-based methods.(17) In this study, a DL-based method is used for complete entity extraction. The BiLSTM-CRF model is used to realize entity extraction in higher mathematics. The CRF model is designed to solve problems such as sequence labeling. This model can fully learn the corpus context information when solving problems, such as sequence labeling.

LSTM introduces three gates: input, output, and forget. The input gate determines which information should be updated. The forget gate selects which information is missing. The output gate determines what information to output to solve the above problem.(18) The LSTM structure diagram is shown in Fig. 4.

LSTM can realize the storage of the required information by controlling the information through the memory unit. Although LSTM solves the vanishing of gradients, it does not consider the impact of the following text on the previous text. Therefore, the prediction results will still not be accurate.(19) To accomplish this, the researchers train LSTM from front to back as well as from back to front. This is called BiLSTM. The information in the context of the sentence can be fully used.

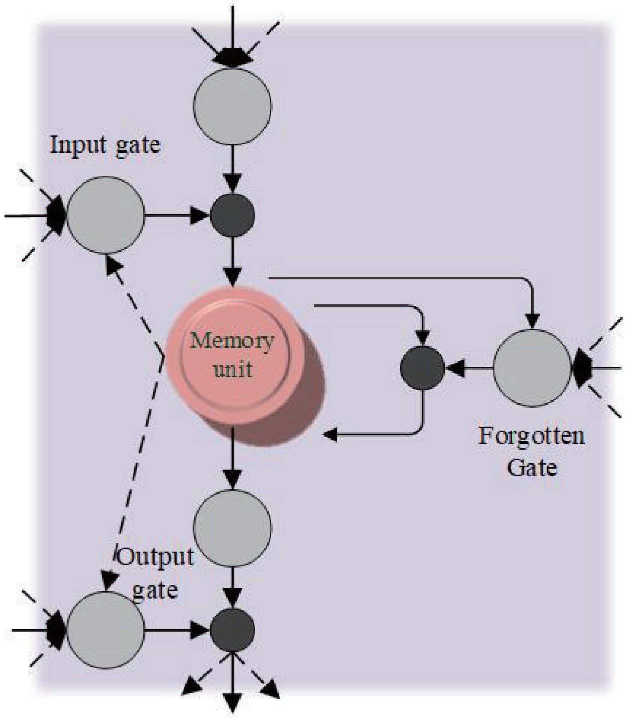

Fig. 4. (Color online) LSTM structure.

Figure 5 shows that BiLSTM is trained from front to back as well as from back to front. Then, the result is input to the output layer. In mathematics, descriptions of logical concepts, such as methods and theorems, are usually clearly indicative, such as the "XX method" and "XX theorem." Therefore, the initial word is crucial in determining whether a text sequence contains a labeled entity and can indicate strong dependences between preceding and following texts. Combining BiLSTM with CRF can improve the accuracy of labels. The structure of the BiLSTM-CRF model is shown in Fig. 6.

As seen in Fig. 6, a text sequence is input to the model. The final output is the label corresponding to the relevant character. Characters are first put into BiLSTM as a word vector. Then, the output of the bidirectional LSTM is input into the conditional random field by implicit projection. Then, the log-likelihood of the output sequence is output.

4.2 Higher mathematics KG construction

The construction of a KG is divided into ontology construction, knowledge extraction, and knowledge storage. There is no exact method for each subprocess. Therefore, it is necessary to choose the method appropriate to the specific field. The detailed construction process is shown in Fig. 7.

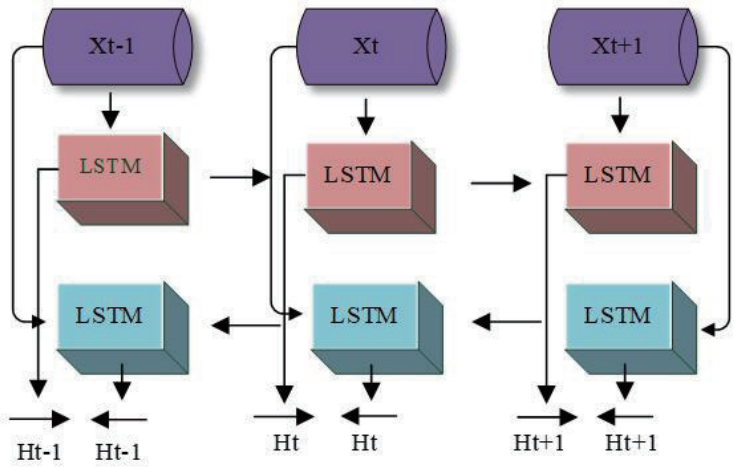

Fig. 5. (Color online) BiLSTM structure.

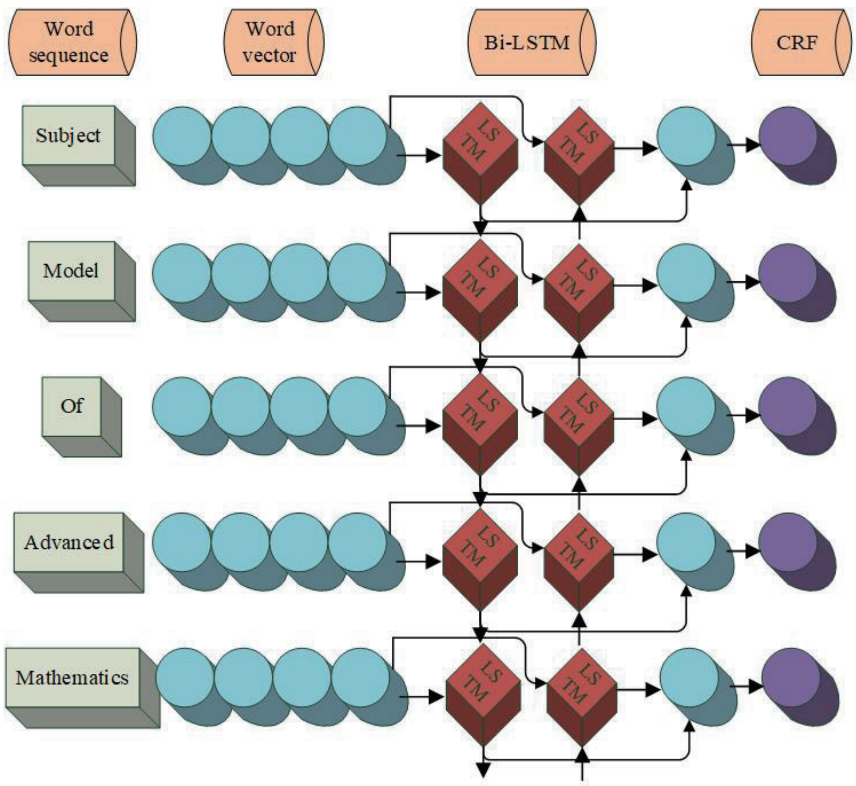

Fig. 6. (Color online) BiLSTM-CRF model structure.

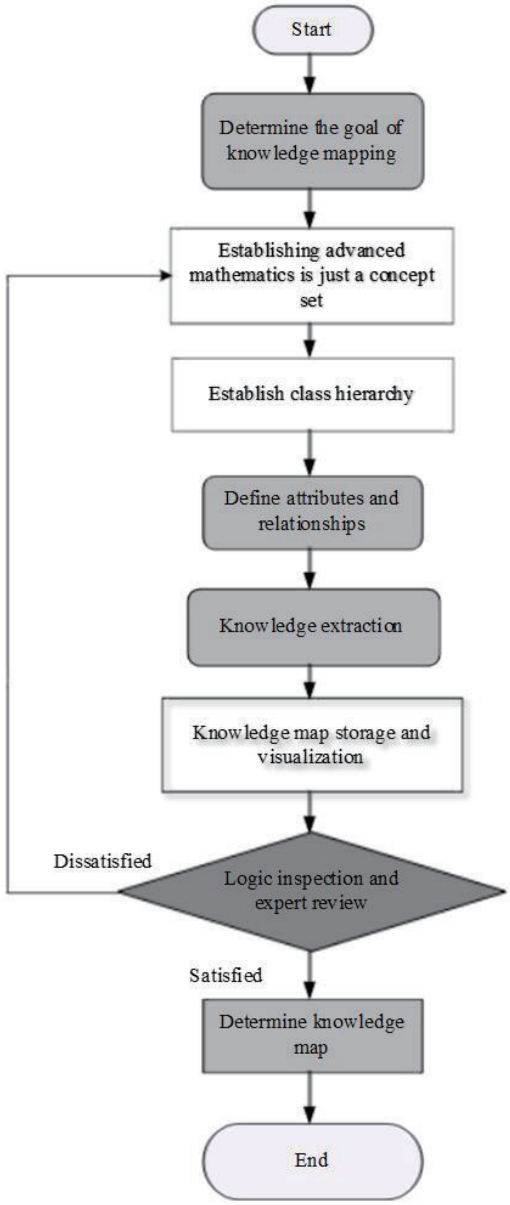

Fig. 7. Process of constructing KG for higher mathematics.

The work in this paper focuses on the subject area of higher mathematics. Therefore, the top-down method is adopted on the basis of an in-depth study of the current mainstream KG construction methods. The KG of higher mathematics is gradually constructed by the process in Fig. 7 under the principle of following the life cycle of the KG and the characteristics of higher mathematics disciplines. The specific construction content mainly comprises the following four tasks.

4.2.1 Corpus construction

When constructing a KG, many corpora in related fields are required for underlying support. A corpus plays a fundamental role in constructing the KG.[20] There are three main sources for obtaining information about higher mathematics: (1) books on higher mathematics, (2) websites related to higher mathematics, and (3) Baidupedia.

4.2.2 Ontology construction

Ontology construction is also known as schema layer construction. Here, the higher mathematical pattern layer is constructed semi-automatically with the help of Protégé ontology construction software for the seven-step process. Figure 8 shows the construction principles and methods.

Here, the ontology construction principle is followed, and the construction idea of the seven-step method is used for reference. Refer to the currently proposed ontology construction method

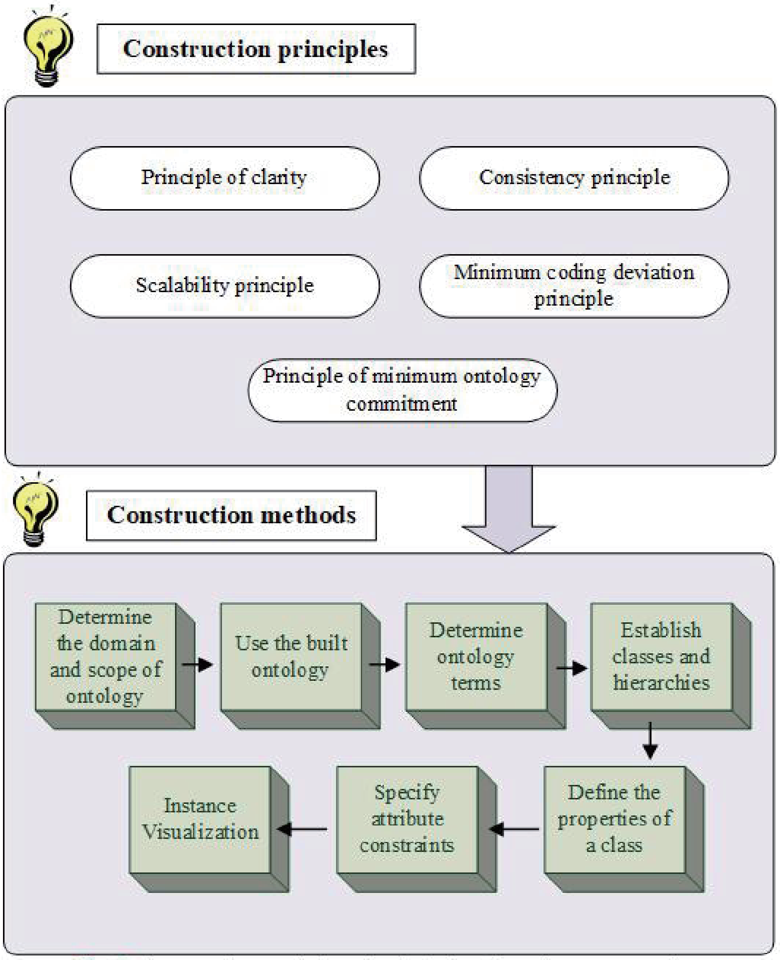

Fig. 8. (Color online) Construction methods and principles of ontology construction.

of knowledge in higher mathematics to determine the ontology construction method applicable to this discipline. Modifications and optimizations will be continued in later practical studies.

4.2.3 Pattern construction

The collected data are collated. Higher mathematical disciplines are studied and analyzed. Then, the higher mathematical knowledge with mathematical properties, concepts, theorems, and definitions is divided into knowledge points, trees, chains, and blocks to represent the knowledge structure of higher mathematics disciplines. This will facilitate follow-up studies. On the basis of the six knowledge modules of higher mathematics subjects, the two categories of knowledge are obtained, as shown in Table 1.

In constructing a KG, the hierarchical division of the classification structure and the definition of attribute relationships of knowledge points are important. Data attributes are generally used as values to represent attributes. The knowledge point attributes of mathematics are mainly the degrees of difficulty and importance. Different analyses are conducted in accordance with different situations to determine their results. The importance of knowledge points is divided into five levels: clarity, understanding, familiarity, mastery, and application. The difficulty of knowledge points is divided into four categories: basic, medium, relatively difficult, and difficult. The data attribute structure of the knowledge points in higher mathematics is illustrated in Fig. 9.

4.2.4 Data construction

The collected datasets are not only from Baidu Encyclopedia. A heart rate sensor is also used to capture students' emotional responses during the learning of higher mathematics. Eye trackers are adopted to record students' gazes during math problem solving. By using these sensors to collect data on students' learning behavior, reaction time, and learning environment, we can identify students' learning difficulties or adjust teaching strategies in a timely manner to provide personalized learning support. Such data collection methods can provide comprehensive and rich information that helps to deeply understand students' performance and characteristics in mathematics learning. In KG construction, after preprocessing the collected sensor data, the behavior patterns of students are mapped to nodes in the KG. The ability level of students is mapped to the attribute values of the nodes, and the interaction between students is mapped to edges in the KG.

Table 1
Classification of knowledge in higher mathematics.

Category	Objectively existential class	Theorem law class
Function, limit, continuous	Concepts and propositions	Calculation
Unary calculus	Concepts and propositions	Solution methods
Ordinary differential equation	Concepts and propositions	Calculation and solution methods
Spatially resolved geometry	Concepts and propositions	Solution methods
Multivariate calculus	Concepts and propositions	Calculation and solution methods
Infinite series	Concepts and propositions	Calculation and solution methods

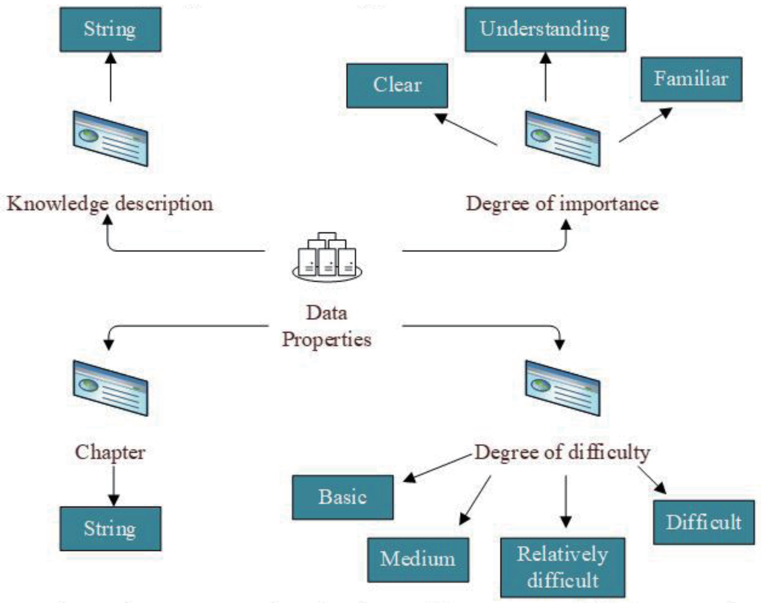

Fig. 9. (Color online) Ontology data properties in subject area of higher mathematics.

The datasets collected here are mainly from Baidupedia. The amount of data is large, and the structure is dispersed owing to its unstructured nature. This leads to inefficient screening when using traditional knowledge extraction methods. Therefore, extracting higher mathematical triples from an unstructured text is highlighted. Here, the relationship extraction model of BiLSTM-CRF entity recognition and AGGCN relationship recognition is performed to complete the automatic extraction of triples. For the field of higher mathematics, the specific process mainly includes entity extraction, relationship extraction, and KG storage. There is no clear word segmentation boundary in Chinese for named entity recognition tasks, so the text should first be segmented. Feature vectors are constructed, and models are trained and tested. The identified entity is saved. All the identified entities saved are used to complete the identification of relationships between entities. Finally, the relationship recognition results are saved to the graph database together with the entities obtained from previous entity recognition. The construction of a higher mathematical KG is preliminarily complete.

4.2.5 Hardware and software environments

The software and hardware architectures of embedded systems suitable for the operation of GNNs require a certain amount of computing power and storage capacity. Therefore, the operating system used here is a Linux-based embedded operating system, Ubuntu Core, to provide a stable operating environment and support the operation of the GNN framework. For the GNN framework, we select PyTorch Geometric. In addition, model compression and optimization techniques, such as quantization and pruning, are used to reduce the calculation and storage requirements of the model, and improve the operation efficiency on the embedded system. The hardware processor is the ARM Cortex-A series with 8G memory. Interfaces that support high-speed data transmission and communication, such as PCIe, Ethernet, and USB, are also used for data exchange and remote access with other systems.

5. Results

5.1 Data collection results

The higher mathematical relation dataset is obtained by manually filtering out the text using the entity relationships determined from the collected data. The dataset has a total of 2580 valid data with Tab as the separator. The relationship is entity 1 to entity 2. The dataset has nine types of relationships: sibling, antecedent, inclusion, opposition, synonymy, successor, transformation, correlation, and causation. The higher mathematical classification dataset is obtained by manually screening the collected data, with a total of 4586 sentence-level knowledge points with Tab as the separator. The data labeling criteria are the same as in the data labeling process described above. Moreover, it is ensured that the knowledge points of this dataset occupy more than 80% of the knowledge points mentioned in the syllabus after repeatedly comparing the syllabus of higher mathematics subjects and consulting in-service university mathematics teachers. The categorical datasets constructed are the limits and continuity of functions, derivatives and differentiation, the differential median theorem, the integration of functions, applications of definite integrals, and infinite series.

5.2 Analysis of entity extraction results

In the experimental training, Adam is selected as the optimizer of the model, with a learning rate of 1e-3, a batch size of 32, 50 epochs, and a dropout of 0.2. We stop the model from training by setting the operation for the validation set to not drop for ten consecutive epochs to prevent the model from overfitting.

Firstly, the running times of different models in the construction of the KG of higher mathematics teaching, shown in Table 2, are compared. As shown in Table 2, the running time of the BiLSTM-CRF model in the construction of the KG of higher mathematics teaching is 120 s, which has obvious advantages over other models. The HMM and CRF models require running times of 180 and 150 s, respectively, and the BiLSTM model requires a running time of 130 s.

Table 2
Running times of different models.

Model	Running time (s)
BiLSTM-CRF	120
HMM	180
CRF	150
BiLSTM	130
BERT-CRF	200
BERT-softmax	220

Moreover, the BERT-CRF and BERT-softmax models require running times of 200 and 220 s, respectively. The advantage of the BiLSTM-CRF model is that it combines BiLSTM and CRF, which can effectively capture the context information and the dependences among labels of sequence data. Compared with other models, the BiLSTM-CRF model can complete the operation more rapidly when dealing with the task of building the KG of higher mathematics teaching, which leads to the improved efficiency and real-time performance of the model.

The identification results of objectively existing entities and theorem law entities are analyzed in comparison with the traditional HMM and BERT models. The results are shown in Figs. 10 and 11.

The comparison of the results in Figs. 10 and 11 shows that the BiLSTM-CRF model has high accuracy in the recognition of mathematical objectively existing entities and theorem law entities. For the identification of mathematical objectively existing entities, the BiLSTM-CRF model achieves an accuracy of 92.5%, which is significantly higher than those of the other models. For the recognition of mathematical logical concept entities, the accuracy of the BiLSTM-CRF model is 89.7%, which again is the highest. The advantage of the BiLSTM-CRF model is that it can capture the context information and label dependences in the sequence data, thereby improving the accuracy of entity recognition. Compared with the other models, the BiLSTM-CRF model has higher accuracy and reliability in identifying mathematical objectively existing entities and theorem law entities.

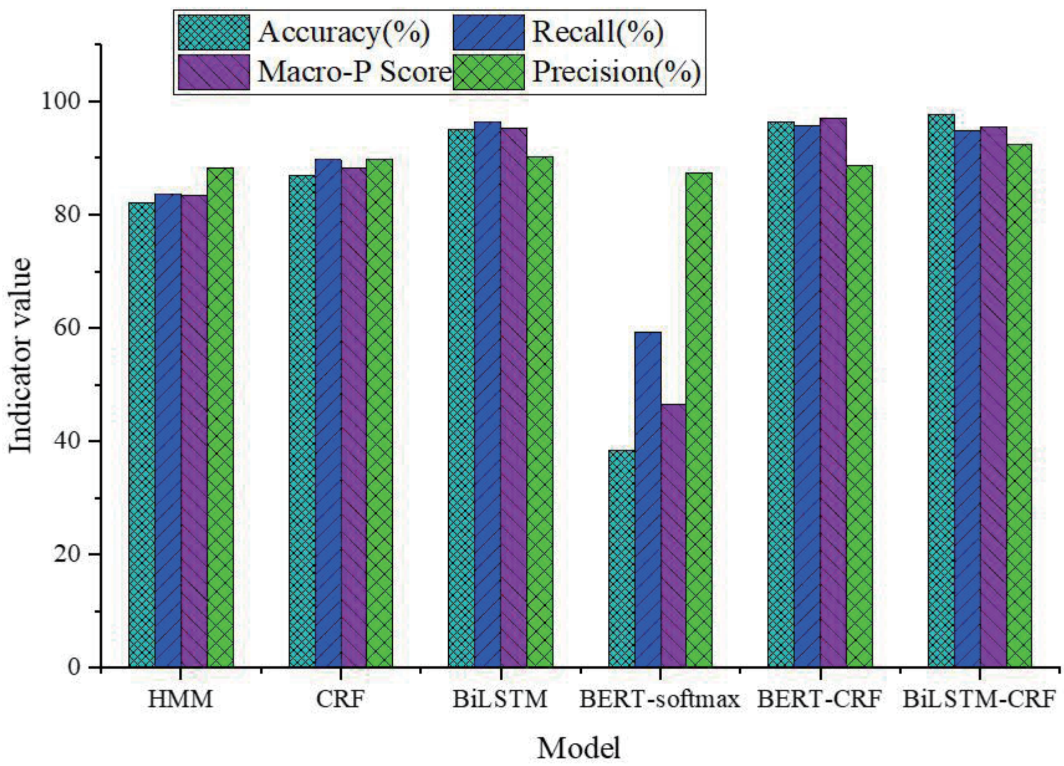

Fig. 10. (Color online) Identification results of mathematically objective existing entities for each method.

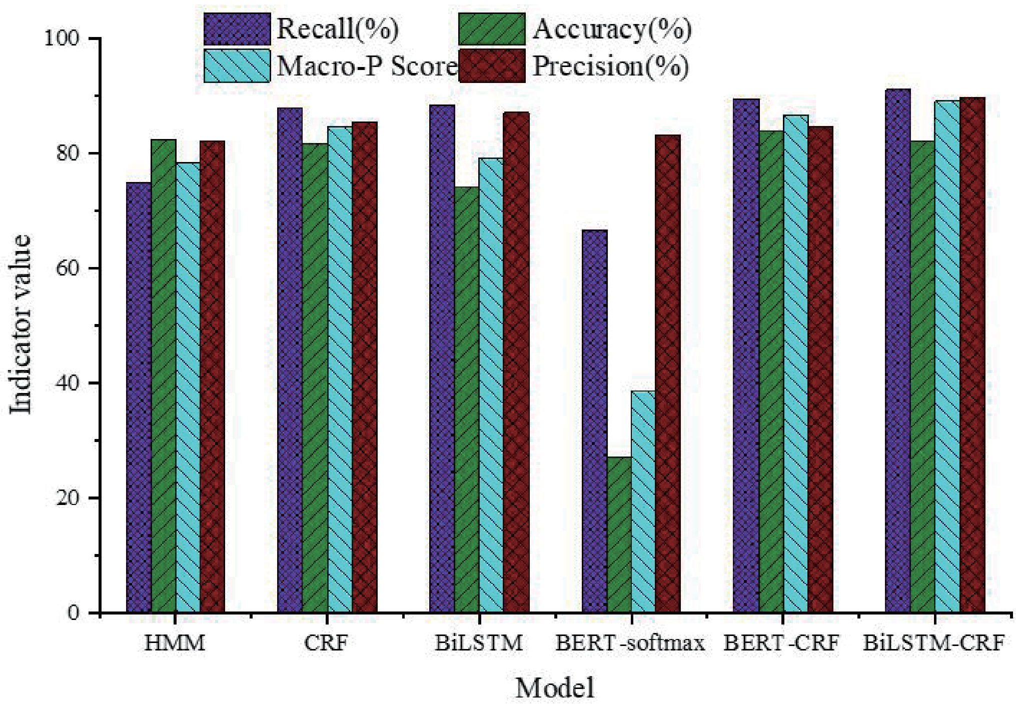

Fig. 11. (Color online) Identification results of mathematical, logical concept entities for each method.

Therefore, it is necessary to use the BiLSTM-CRF model for mathematical information recognition. After vectorizing the Chinese text by word, BiLSTM is used to obtain the bidirectional semantic features of the context. Compared with the traditional named entity recognition model, the BiLSTM-CRF model has a better recognition effect. Regarding the recognition effect, the BiLSTM-CRF model occupies less memory than models that occupy a large portion of the general processing unit memory. Moreover, its effectiveness is not significantly different from that of the traditional model with the highest accuracy. Therefore, we choose to use the BiLSTM-CRF model as the named entity recognition model in the consideration of the experimental equipment and other factors.

5.3 GCN-based results of higher mathematical text classification

For the GCN method proposed here, Adam is used as the optimizer. The learning rate is set to 1e-2. The number of epochs is set to 200. We stop the model from training by setting the operation such that the loss of the model for the validation set will not drop for ten consecutive epochs to prevent the model from overfitting. The results, along with those of the traditional methods, are shown in Fig. 12.

In Fig. 12, the GCN model shows the highest accuracy in the classification of higher mathematical texts. The GCN model achieves an accuracy of 91.6%, which is significantly higher than those of the other models. The accuracies are 85.2% for the BiLSTM model, 83.5% for the CNN model, 87.1% for the BiLSTM Attention model, and 89.3% for the BERT-based

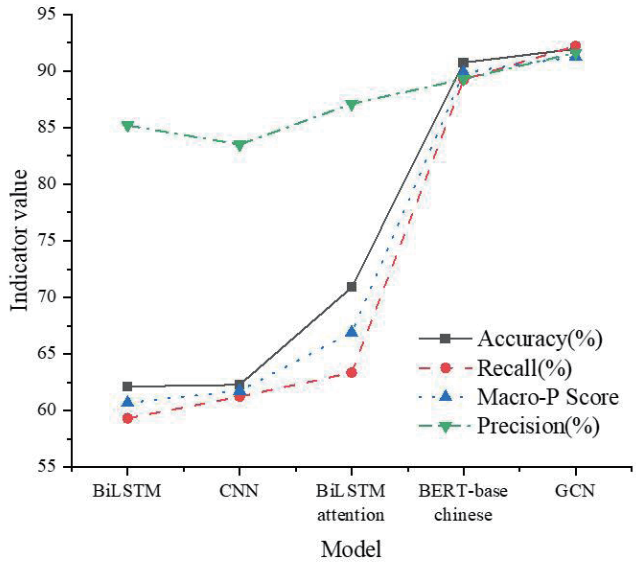

Fig. 12. (Color online) Results for various methods (100%).

Chinese model. The advantage of the GCN model is that it can make full use of the node and edge information in the graph structure data to perform a comprehensive and accurate modeling of higher mathematical texts. GCN can capture the interactions and relationships between nodes and update nodes through the graph convolutional layer to extract richer feature representations. Compared with the other models, GCN shows higher accuracy and performance in higher mathematical text classification tasks.

GCN has achieved good results. Compared with recent mainstream models, it has achieved relatively high results on p-values. The reason is as follows. First, the number of datasets constructed here is very small because of limited resources. Consequently, the model cannot be fully trained. Second, the quality of word embedding greatly affects the model's performance. Existing word embeddings are not specifically aimed at mathematical training. Neither Word2Vec nor BERT can handle word-to-word semantic information well. Therefore, the model cannot extract features. Finally, the model's performance is greatly affected by the imbalance of dataset categories. However, GCN-based models build graph nodes between words and between words and sentences. Such a model learns the connections between the data and classifies these graph nodes independently, which significantly alleviates the above problems.

To further prove the advantages of the model proposed here, current similar excellent research on the construction of a KG is examined here. The works reported in Refs. (21) and (22) are further compared. Similarly, its performance is evaluated on the basis of the above four indicators, and the specific comparison results are shown in Fig. 13.

As shown in Fig. 13, the accuracy of the GCN model is 94.62%, which is significantly higher than 86.34% in Ref. 21 and 89.21% in Ref. 22. This means that the GCN model can predict labels more accurately, improving its overall prediction performance. In addition, the recall rate of the GCN model is 95.27%, which is higher than 88.12% in Ref. 21 and 91.06% in Ref. 22. This

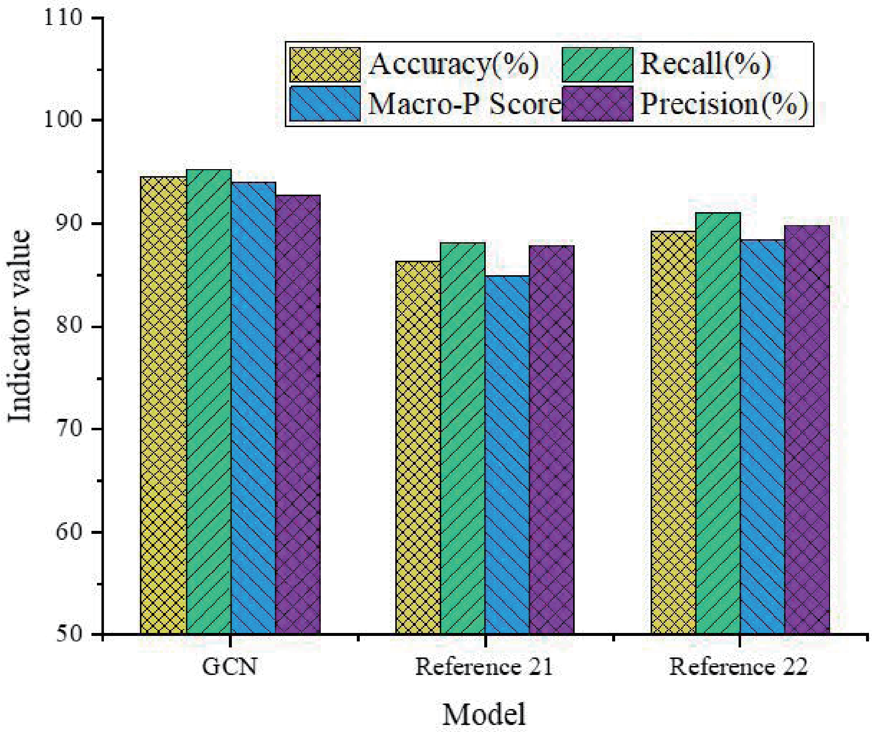

Fig. 13. (Color online) Ratio of training to test sets for six knowledge classifications.

suggests that the GCN model can better capture positive samples and reduce the number of missed relevant information. The macro-P score of the GCN model is 93.98%, which is higher than 84.96% in Ref. 21 and 88.45% in Ref. 22. The macro-P score indicates the average accuracy of the model over all categories, and the GCN model performs better on multiple categories. In terms of accuracy, the accuracy of the GCN model is 92.75%, exceeding 87.89% in Ref. 21 and 89.74% in Ref. 22. This result shows that the GCN model predicts positive samples more accurately and reduces misclassification.

In summary, the GCN model has significant advantages over those in Refs. 21 and 22 in terms of accuracy, recall, macro-P score, and precision. The advantages of the GCN model include higher accuracy, better positive sample capture capabilities, and higher precision, which enable the GCN model to better process and analyze the graph structure data in the construction of the higher mathematical KG, thus improving the overall performance and prediction ability of the model.

In addition, we determine the sentence length and label distributions for the divided training and test sets to analyze the dataset, as shown in Fig. 14.

Sentences that are very long will affect the training effect, so the sentence length of the dataset is determined. The sentence length of the higher mathematical knowledge classification dataset constructed here is concentrated in the normal range of 30–60 words. Except for individual data, the sentence length distributions in the training and test sets are approximately the same. From Fig. 14, the ratio of six categories of the higher mathematical knowledge

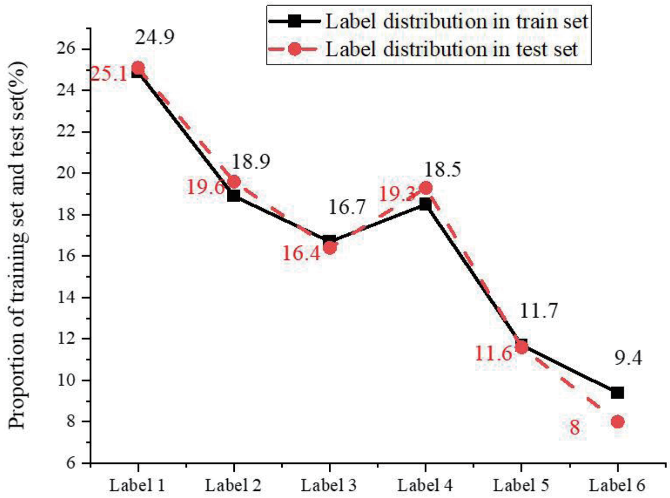

Fig. 14. (Color online) Ratio of labels in training to test sets for six knowledge classifications.

classification dataset constructed here is roughly 2.5:2:1.5:2:1:1. The dataset has the problem of unbalanced categories, which increases the difficulty of model training.

6. Conclusions

The use of sensors to collect and apply data in the teaching of higher mathematics has important background and significance. The effect and quality of higher mathematics teaching can be improved by making full use of sensor technology. The development of personalized teaching can be promoted, students' mathematical thinking and innovation abilities can be cultivated, and more intelligent and personalized solutions can be provided for future education.(23–25) Here, combined with the needs of GCN and mathematics disciplines, the construction and application of KGs of higher mathematics disciplines were carried out. Firstly, the relevant knowledge of higher mathematics in universities was collected and processed. Secondly, the KG of the higher mathematics discipline was constructed in accordance with the knowledge structure level. Then, the application system based on the KG of the mathematics discipline was designed and realized. The application system based on the KG of the mathematics discipline was designed. The system will have considerable application value for gaining the subject knowledge of higher mathematics. The experimental results show the following. (1) The self-built recognition body recognition test set has a P-value of 89.16% for the higher mathematical theorem law entity, and a P-value of 97.73% for the objectively existing test questions of higher mathematics. (2) The higher mathematical KG constructed here has a total of 2580 triples. The higher mathematical visual query platform was realized and can be used. (3) The classification of higher mathematical knowledge is set on the basis of the text classification model of GCN. The P-value is 91.28%, which is far better than that of the traditional text classification model. The disadvantages of the proposed GCN model are as follows. (1) The content of higher mathematics contains many knowledge points. There are no clear standards for the relationships among knowledge points, the division of knowledge points, and the visualization method. Therefore, the content of the extracted knowledge may not be comprehensive. The integration of multiple data sources to build a disciplinary knowledge network will be considered. (2) We have not yet implemented the user interaction module of the higher mathematical knowledge query platform. Therefore, the function that allows users to optimize knowledge ontology and supplement knowledge triples will be designed in the upcoming work. In the future, the KG of higher mathematics subjects will be embedded in the online teaching platform.

Acknowledgments

This work was supported by Inner Mongolia Education Science Planning (No. NGJGH2020277), the Natural Science Foundation of Inner Mongolia (Nos. 2022MS08023 and 2023QN08049), the Mongolian Medicine Synergy Innovation Center Scientific Research Foundation of Inner Mongolia (No. MYYXTYB202108), the Key Technology Project Foundation of Inner Mongolia (No. 2021GG0176), the "Asking, Learning, Practice" Research Projects of Baotou Medical College (No. 2021BYWWJ-YB-22), the "Hualei Plan" Projects of Baotou Medical College (No. 110-2023300737), the Scientific Research Foundation of Baotou Medical College (No. BYJJ-ZRQM 202304), and the Baotou Youth Innovative Talents Project (202229).

References

1 S. Ren, Z. Liu, R. Zhanbayev, and M. Du: J. Econ. Anal. **1** (2022) 50. https://doi.org/10.58567/jea01010004
2 M. Sakai, K. Nagayasu, and N. Shibui: Sci. Rep. **11** (2021) 1. https://www.nature.com/articles/s41598-020-80113-7
3 C. Y. C. Yeh, H. N. H. Cheng, and Z. H. Chen: Res. Pract. Tech. Enhanced Learn. **14** (2019) 1. https://telrp.springeropen.com/articles/10.1186/s41039-019-0100-9
4 W. Glänzel, L. Zhang, and Y. Huang: Scientometrics **126** (2021) 6201. https://link.springer.com/article/10.1007/s11192-020-03821-x
5 Y. Nechaev, F. Corcoglioniti, and C. Giuliano: Prog. Artif. Intell. **7** (2018) 251. https://link.springer.com/article/10.1007/s13748-018-0160-x
6 W. Wang and C. Lu: Soft Comput. **24** (2020) 8173. https://link.springer.com/article/10.1007/s00500-019-04384-7
7 M. Allibaih: J. Arts Humanit. **8** (2019) 17. https://doi.org/10.18533/journal.v8i8.1642
8 S. Song, N. Zhang, and H. Huang: Cluster Comput. **22** (2019) 5195. https://link.springer.com/article/10.1007/s10586-017-1146-3
9 Q. Wei, Y. Zhang, and M. Amith: Health Informatics J. **26** (2020) 21. https://doi.org/10.1177/1460458219869490
10 J. Duanmu and Y. Xing: Remote Sens. **12** (2020) 808. https://doi.org/10.3390/rs12050808
11 F. Tian, D. C. Yang, and Y. Q. Meng: Nucleic Acids Res. **48** (2020) D1104. https://doi.org/10.1093/nar/gkz1020
12 C. W. Coley, W. Jin, and L. Rogers: Chem. Sci. **10** (2019) 370. https://pubs.rsc.org/en/content/articlehtml/2019/sc/c8sc04228d
13 M. M. Ghazal and A. Hammad: Int. J. Constr. Manage. **22** (2022) 1632. https://doi.org/10.1080/15623599.2020.1738205
14 H. Gao, K. Xu, and M. Cao: IEEE Trans. Comput. Social Syst. **9** (2021) 336. https://doi.org/10.1109/TCSS.2021.3102591
15 X. Ling, L. Wu, and S. Wang: ACM Trans. Knowl. Discov. Data (TKDD) **15** (2021) 1. https://doi.org/10.1145/3447571
16 Y. Hong, Y. Liu, and S. Yang: IEEE Access. **8** (2020) 51315. https://doi.org/10.1109/ACCESS.2020.2980859

17 W. Khan, A. Daud, and F. Alotaibi: ETRI J. **42** (2020) 90. https://doi.org/10.4218/etrij.2018-0553
18 X. Yuan, L. Li, and Y. Wang: IEEE Trans. Ind. Inform. **16** (2019) 3168. https://doi.org/10.1109/TII.2019.2902129
19 J. Gu, L. Liang, and H. Song: Sci. Rep. **9** (2019) 1. https://www.nature.com/articles/s41598-019-54495-2
20 B. B. Gupta and M. Quamara: Concurr. Comput. Pract. Exp. **32** (2020) e4946. https://doi.org/10.1002/cpe.4946
21 L. Guo, F. Yan, T. Li, T. Yang, and Y. Lu: Robot. Comput.-Integr. Manuf. **73** (2022) 102222. https://doi.org/10.1016/j.rcim.2021.102222
22 N. Díaz-Rodríguez, A. Lamas, J. Sanchez, G. Franchi, I. Donadello, S. Tabik, D. Filliat, P. Cruz, R. Montes, and F. Herrera: Inf. Fusion **79** (2022) 58. https://doi.org/10.1016/j.inffus.2021.09.022
23 Y. C. Luo, H. Y. Huang, B. W. Chen, and Y. P. Kuo: Sens. Mater. **35** (2023) 1587. https://doi.org/10.18494/SAM4067
24 P. H. Huang, Y. Q. Zhu, L. W. Chao, W. C. Lin, R. X. Zheng, K. W. Lee, and C. J. Huang: Sens. Mater. **35** (2023) 1569. https://doi.org/10.18494/SAM4062
25 C. H. Lin, C. C. Lin, Y. C. Zhan, C. Y. Yeh, and S. H. Wang: Sens. Mater. **35** (2023) 1613. https://doi.org/10.18494/SAM4077