

Research Article

A BERT-Based Approach for Extracting Prerequisite Relations among Wikipedia Concepts

Youheng Bai ^{1,2} **Yan Zhang** ^{1,2} **Kui Xiao** ^{1,2} **Yuanyuan Lou** ^{1,2} and **Kai Sun** ^{1,2}

¹School of Computer Science and Information Engineering, Hubei University, Wuhan 430062, China

²Hubei Province Educational Informatization Engineering Research Center, Hubei University, Wuhan 430062, China

Correspondence should be addressed to Kui Xiao; xiaokui@hubu.edu.cn

Received 15 August 2021; Revised 27 September 2021; Accepted 18 October 2021; Published 29 October 2021

Academic Editor: Chunlai Chai

Copyright © 2021 Youheng Bai et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Concept prerequisite relation prediction is a common task in the field of knowledge discovery. Concept prerequisite relations can be used to rank learning resources and help learners plan their learning paths. As the largest Internet encyclopedia, Wikipedia is composed of many articles edited in multiple languages. Basic knowledge concepts in a variety of subjects can be found on Wikipedia. Although there are many knowledge concepts in each field, the prerequisite relations between them are not clear. When we browse pages in an area on Wikipedia, we do not know which page to start. In this paper, we propose a BERT-based Wikipedia concept prerequisite relation prediction model. First, we created two types of concept pair features, one is based on BERT sentence embedding and the other is based on the attributes of Wikipedia articles. Then, we use these two types of concept pair features to predict the prerequisite relations between two concepts. Experimental results show that our proposed method performs better than state-of-the-art methods for English and Chinese datasets.

1. Introduction

In recent years, the emergence of online learning platforms and e-learning resources has injected new impetus into people's learning. Online learning models have gradually become more popular. Research related to this field has also received considerable attention. As everyone has a different knowledge background, the challenge faced by online learners is usually how to choose learning resources and how to rank them. Typically, each learning resource explains one or more of the leading knowledge concepts. Concepts in a field are usually learned progressively, from simple to complex and from abstract to concrete. Usually, the order of learning resources is determined by the relations between main concepts. This kind of relationship between concepts is generally called a concept prerequisite relation. A prerequisite is usually a concept or requirement before one can proceed to the following one. A prerequisite relation is a natural dependency among concepts when people learn, organize, apply, and generate knowledge [1–3].

The learning order between concepts is determined by their prerequisite relations. As for knowledge in a given field, a directed acyclic graph can illustrate its concept prerequisite relations. The concept appears as a node, and the direction of its arrow represents the prerequisite relations between the concepts. For the concept pair (A, B) in the teaching field, if concept B is the prerequisite relation for concept A, then you first learn concept B before learning concept A. It can be written as $A \leftarrow B$. As shown in Figure 1, neural network (A) relies on concepts such as gradient descent, partial differential, and differential equation. These concepts also rely on differential (B). In other words, before learning neural network, differential equation is needed.

In a classroom course, the instructor will explain each central concept to students according to the inherent order of the concepts. Additionally, the instructor may also spend some time explaining some background knowledge-related concepts to help students understand current knowledge concepts. However, students may not receive assistance

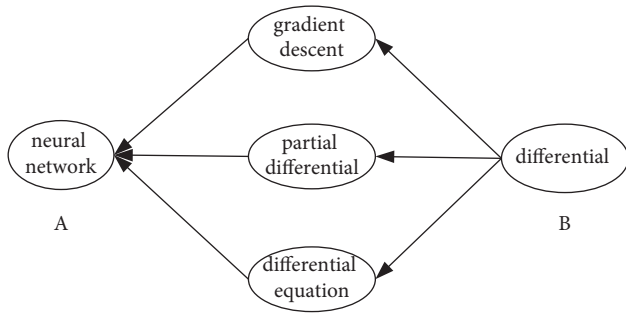


FIGURE 1: Example of concept prerequisite relation (differential equation concept is the prerequisite relation of neural network concept).

from instructors in online courses. For example, when students learn the Vue.js, they usually need to master the HTML and CSS first; when they learn the Java Spring Boot, they usually need to master the Maven first. There is usually a prerequisite relation between two different learning resources. In addition, when people browse a Wikipedia article, they often open the pages of other articles to learn more about the background of the current article. Between Wikipedia articles, there is usually a prerequisite. Due to a lack of understanding of prerequisite relations between different concepts, people may be unable to complete courses or understand the content of Wikipedia articles.

In this article, we propose a method for extracting concept prerequisite relations from Wikipedia using BERT. We used concepts from Wikipedia, and each concept has its own Wikipedia article. Compared with courses on online learning platforms, Wikipedia’s main concepts are easier to extract in an automated way. Furthermore, because Wikipedia has a unique knowledge structure, we can extract the characteristics of concept pairs and analyze the prerequisite relations between concepts easier.

Our main contributions include the following:

- (1) A novel metric to measure the prerequisite relations among Wikipedia concepts superior to the existing methods
- (2) A Chinese dataset annotated with prerequisite relations between pairs of Wikipedia concepts

The structure of this article is as follows. Section 2 reviews past works on the task of concept prerequisite relations extraction with Wikipedia and MOOC. The problem definition of concept prerequisite relations is in Section 3. Section 4 elaborates on the methodology. Section 5 describes datasets and preparation techniques and our experimental results and analysis. Section 6 is the concluding remarks and future work.

2. Related Work

The concept prerequisite relations determine the order in which knowledge is learned and the order in which documents are read. Nowadays, concept prerequisite relations extraction can be used in different kinds of education-related

tasks [4], including curriculum planning [5, 6], learning resources recommendation [7, 8], knowledge tracing [9], and so on. Additionally, there is also a lot of research related to concept prerequisite relation extraction.

The area that researchers pay most attention to is to extract the prerequisite relations between Wikipedia concepts. Talukdar and Cohen [10] utilized three types of features for concept pairs, including WikiHyperlinks, WikiEdits, and WikiPageContent, and then used the MaxEnt classifier to predict prerequisite relations among Wikipedia concepts. Liang et al. [1] studied the problem of measuring prerequisite relations among concepts and proposed the RefD metric to capture the relation. RefD means reference distance, and it uses the page links in Wikipedia to model the prerequisite relation by measuring how differently two concepts refer to each other. Zhou and Xiao [11] employed Wikipedia page links, categories, article content, and time attributes of Wikipedia articles to create features and then predict concept prerequisite relations. Sayyadiharikandeh et al. [12] used the clickstream of human navigation among articles on Wikipedia to infer concept prerequisite relations. In addition, many similar studies have used machine learning methods to predict prerequisite relations between Wikipedia concepts [13–17]. A common problem with these methods is that all of them require experts to manually design the features of the concept pairs.

Besides Wikipedia, some researchers have tried to extract concepts from various learning resources and analyze the prerequisite relationships between concepts. Pan et al. [18] manually extracted the main knowledge concepts of the course from the MOOC video and used the sequence and frequency of appearance of the concepts as features to analyze the prerequisite relations between the concept pairs. Wang et al. [13] extracted the main knowledge concepts from the textbooks, linked these concepts with Wikipedia articles, and then identified the prerequisite relations between the concepts. Liang et al. [14] explored the content of the course introductions on the university website, investigated how to recover concept prerequisite relations on the university website, investigated how concept prerequisite relations are derived from course dependencies, and proposed an optimization-based framework to address the problem. Furthermore, other similar studies use the dependency relationship between learning resources to predict the prerequisite relations between knowledge concepts [1, 2].

As mentioned above, all methods based on machine learning need to use manual design concepts to predict prerequisites. This usually causes other factors that can be used to infer the prerequisite relationship to be ignored. There is a possibility that deep learning will outperform machine learning in this regard since deep learning methods can automatically extract features from raw data. Miaschi et al. [19] used Word2Vec to convert the two concepts into vectors and input the vectors into two LSTM networks, respectively, to obtain the features of the concept pair and predict the prerequisite relations of the concept pair. However, Word2Vec only treats a concept as a normal word. Compared with Word2Vec, BERT [20] can better explore the semantic meaning of a concept, and the contextualized

vectors that BERT generates can also be used to infer concept prerequisite relations.

In this paper, we use the BERT sentence embedding based on contextual embedding to automatically extract the features of concept pairs. Meanwhile, we also designed some features manually for concept pairs. Both classes of features were employed to infer concept prerequisite relations. Furthermore, we created a Chinese concept pair dataset that can be used to identify the prerequisite relations.

3. Problem Definition

The goal of the concept prerequisite relations identification task is to judge whether there is a dependency between two concepts. For a concept pair (A, B), there are four possible relations between them: (1) A is a prerequisite of B; (2) B is a prerequisite of A; (3) the two concepts are related, but they do not have any prerequisite relation between them; and (4) the two concepts are unrelated [10]. In previous studies, researchers usually converted this task into a binary classification problem for processing. They were simply judging whether A is a prerequisite of B. It can be defined as

$$\text{Preq}(B, A) = \begin{cases} 1, & A \text{ is a prerequisite of } B \\ 0, & 0. \end{cases} \quad (1)$$

$\text{Preq}(B, A) = 1$ means that A is a prerequisite of B. In other words, before people can learn about concept B, they must master concept A while $\text{Preq}(B, A) = 0$ means that A is not a prerequisite of B. In this article, we will also turn the concept prerequisite relations identification problem into a binary classification task to deal with.

Moreover, the concepts we use are Wikipedia concepts. Each concept has a corresponding Wikipedia article. The concept is the title of the article.

4. Wikipedia Concept Prerequisite Relations Prediction Method

This section presents our proposed concept prerequisite relations prediction model (AFs + MFs). The structure of the model is illustrated in Figure 1. The input of the model is composed of two types of concept pair features, including features extracted automatically (AFs) and features extracted manually (MFs). Precisely, we extract two BERT sentence embeddings and Wikipedia-based features from concept pairs. First, the model inputs the AFs of the concept pairs into two LSTMs, and the two output vectors of LSTMs are concatenated with MFs. Then, these features are input to a fully connected layer to accomplish concept prerequisite relations recognition.

4.1. Features Extracted Automatically. As a big data pre-training transformation language model of the bidirectional transformer, the application of BERT has significantly improved performance on several NLP tasks. Particularly, sentence-BERT [21] introduces pooling to the token embeddings generated by BERT to generate fixed-size sentence embeddings, obtaining state-of-the-art performance in

many fields, including text similarity and classification problems.

Articles in Wikipedia concepts typically contain a number of sentences, each containing deep semantic information. Hence, we use BERT to generate sentence embeddings as the feature extracted automatically from the concept.

More specifically, first, for the first k words or Chinese characters of the Wikipedia concept article, the BERT tokenizer is used with a maximum sequence length of 500 to obtain the token representation. Then, we generate a concept BERT sentence embedding by inputting tokens as the input of the BERT model (vector size = 768). The two BERT sentence embeddings of the concept pair are used as inputs to the neural network, which is passed to the two 32-unit LSTMs. LSTM can be used to create some feature information not included in automatic feature design and achieve deeper concept feature extraction.

4.2. Features Extracted Manually. As a multilingual open knowledge base, Wikipedia has the characteristics of multi-user collaborative editing, dynamic updating, and complete coverage. Wikipedia's concepts are described through articles with corresponding titles, and the articles contain links, categories, and redirects (synonyms) in the content. Researchers can use this information to extract feature information from concept pairs.

By manually extracting the structural features of concept pairs from Wikipedia articles, we can analyze the prerequisite relations between the two concepts. Therefore, we extract three types of concept pair features from Wikipedia article information: text features, links features, and category features. These features are as follows:

- (i) *Concept Appearance Count (#1, #2).* It is the number of times the concept A/B appears in the Wikipedia article of the concept B/A.
- (ii) *Whether the First Sentence Appears (#3, #4).* The first sentence of the Wikipedia article on concept B/A mentions concept A/B.
- (iii) *Jaccard Similarity (#5).* The Jaccard similarity of articles with two concepts (A, B) is calculated.
- (iv) *LDA (#6, #7).* Shannon entropy of the LDA vector of A/B: In the information world, the higher the Shannon entropy, the more information can be transmitted [19]. Using [lda](https://pypi.org/project/lda/) <https://pypi.org/project/lda/>, different LDA topic models are trained for each dataset.
- (v) *Category (#8).* Whether the concept pair (A, B) belongs to the same Wikipedia category.
- (vi) *Link (#9, #10).* For (A, B) concept pairs, whether the concept B/A article refers to concept A/B, that is, contains a link to concept A/B.
- (vii) *Link in/out of A (#11, #12).* The number of links to concept A in Wikipedia ("link in") and the number of links to other terms in articles of concept A ("link out") are calculated.

(viii) *Link in/out of B (#13, #14)*. Same as above, the number of links in/out of concept B is calculated.

A note should be made that features #1–#7 and #9–#10 are taken from literature [14], feature #5 is taken from literature [19], and features #8 and #11–14 are taken from literature [16]. Previously, only the English dataset had been validated on these features, and this article will make an evaluation of these features on both the Chinese and English datasets simultaneously.

4.3. AFs + MFs: Concept Prerequisite Relations Prediction Model. Based on the above design and analysis, for a concept pair (A, B), the model (Figure 2) separates the concept prerequisite relations prediction into the following steps:

- (1) The first k words or Chinese characters of the concept pair (A, B) Wikipedia articles is first obtained, and the sentences S_A and S_B are generated.
- (2) Then, the sentence is divided into individual words or Chinese characters, and they are labeled separately, and BERT is used to encode them to generate sentence embedding in V_A and V_B .
- (3) V_A and V_B are input to two LSTMs, and hidden features are further extracted. The output of the two LSTMs is vectors O_A and O_B of size 32 each.
- (4) A 14-dimensional vector $V_{(A,B)}$ is created using the Wikipedia features extracted manually, including links, text content, and category features.
- (5) Finally, O_A , O_B , and $V_{(A,B)}$ are concatenated ($O_A \oplus O_B \oplus V_{(A,B)}$) as 78-dimensional features and input to a fully connected layer with a sigmoid activation function to realize concept prerequisite relation prediction.

5. Experiments

5.1. Datasets and Implementation Details. For our research, we used a public dataset, AL-CPL, which is an English dataset designed by Chen et al. [9] in their research. The dataset consists of two-category concept pair sets and prerequisite relation labels from four different fields. The fields are data mining, geometry, physics, and precalculus. Each data item is formalized as a triple (A, B, Label), which is the concept A, B, and the prerequisite relation label, respectively. Each concept in the dataset has a corresponding article in Wikipedia. The left half of Table 1 shows detailed information about the AL-CPL dataset.

In addition, we also want to verify whether the proposed method performs well in other languages. By using the AL-CPL English dataset, this paper creates the CH-AL-CPL Chinese dataset. First, the English Wikipedia article corresponding to each concept in the AL-CPL dataset is found, and then the Chinese article corresponding to each concept based on the cross-language links is found in Wikipedia.

However, Chinese Wikipedia articles are only a small fraction of those on English Wikipedia. Thus, the collection

of Chinese concept pairs obtained by directly using cross-language links is not only small but also has a significant issue of data category imbalance. Due to this, this paper uses the transitivity and asymmetry of the concept prerequisite relations to expand the number of the Chinese dataset.

- (1) *Transitivity*. Concept B is a prerequisite of concept A, concept C is a prerequisite of concept B, and then concept C is a prerequisite of concept A
- (2) *Asymmetry*. If concept B is a prerequisite for concept A, then concept A cannot be a prerequisite for concept B

By combining transitivity and asymmetry, we can increase the number of categories in the dataset and balance the ratio of categories. In Table 1, the right half shows the detail of concept pairs in each domain of the CH-AL-CPL. The CH-AL-CPL dataset has been published on GitHub <https://github.com/lycyhrc/CH-AL-CPL>.

In the experiment, all models were implemented with Keras. Using the bert-as-service (<https://bert-as-service.readthedocs.io/>) sentence encoding service, we generated a 768-dimensional sentence embedding for the first k words or Chinese characters in Wikipedia concept articles. The sentences were tokenized with NLTK [22]. In order to train the model, the following parameters are set: 50 training epochs, 0.01 learning rate, 32 dimensions of the hidden layer, and 0.2 dropout rate. Adam optimization is used to train the model, and L2 regularization is used to prevent overfitting.

In the AFs model, two 768-dimensional sentence embeddings of the Wikipedia concept pair (A, B) are input to two 32-unit LSTMs, and the fully connected layer is used to receive the output of the LSTMs to identify prerequisite relations. Besides, the #1–16 manual features of the Wikipedia concept pair (A, B) are combined, which then sends them to a fully connected layer for prerequisite relations prediction. Further, the AFs + MFs model concatenates the output of the LSTM of the AFs model and the manual features of MFs to complete the prerequisite relations recognition.

5.2. Experimental Result and Analysis. In our experiment results, we compared our method with the following typical concept prerequisite relations prediction baselines:

- (1) *Reference Distance (RefD)* [1]. The basic idea of this method is that each concept can be represented by its collection of related concepts in the concept space if most of the related concepts of concept A refer to concept B. The related concept of concept B rarely refers to concept A, then concept A may depend on concept B. The author constructs related links for each related concept and refers to the EQUAL weight and TF-IDF weight method to identify the prerequisite relations between the two concepts, so we selected the best-performing TF-IDF weight method.
- (2) *Machine Learning (AT)* [14]. This method uses link-based and text-based features extracted from Wikipedia pages and then uses Naive Bayes (NB),

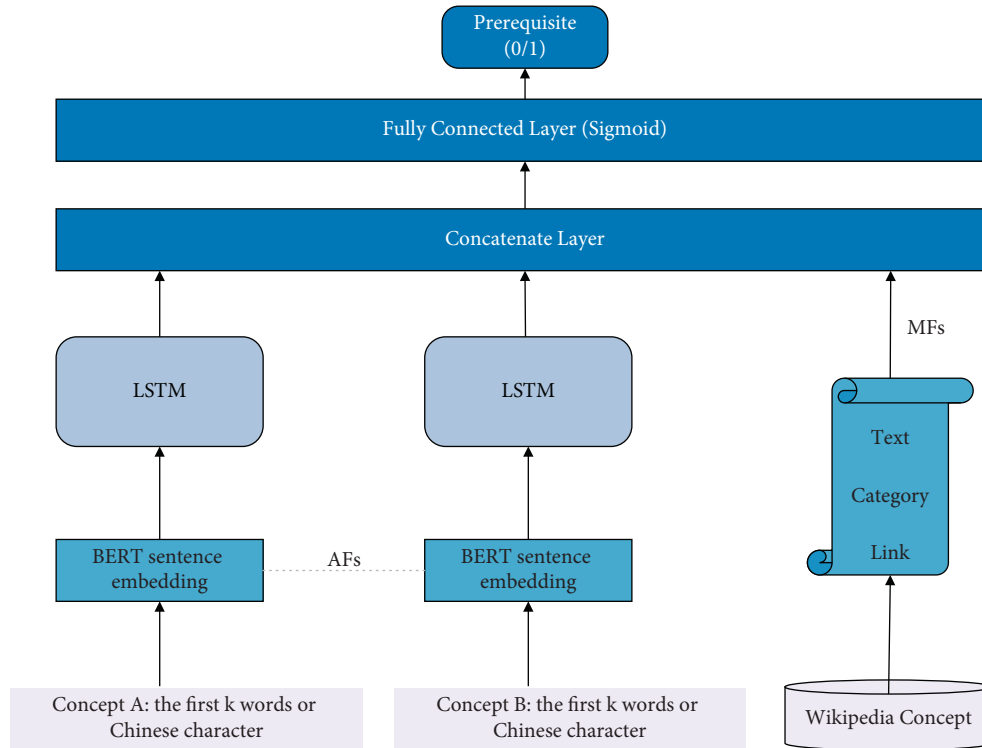


FIGURE 2: Concept prerequisite relations prediction model (AFS + MFS).

TABLE 1: The number of concepts pairs and prerequisite pairs in the dataset AL-CPL.

Domain	AL-CPL			CH-ALCPL		
	#Pairs	#Positive pairs	#Negative pairs	#Pairs	#Positive pairs	#Negative pairs
Data mining	826	292	534	1151	493	658
Geometry	1681	524	1157	3330	1825	1505
Physics	1962	487	1475	2958	1091	1867
Precalculus	2060	699	1361	3200	1431	1769
All	6529	2002	4527	10639	4840	5799

logistic regression (LR), support vector machine (SVM), and random forest (RF) four classifiers to train a concept prerequisite relations prediction model on the AL-CPL dataset, which is used to analyze the prerequisite relations between concept pairs. We use the results of the best-performing random forest classifier (RF) report directly as the basis for comparison.

- (3) *Neural Network (RS)* [16]. This method is based on a neural network, and the UNIGE_SE team is responsible for the sharing task of EVALITA 2020 PRELEARN. The author proposes eight category features based on the content and structural features of Wikipedia, using Italian datasets and deep learning models to analyze the prerequisite relations between concepts. As a comparative algorithm, these feature values are recalculated on Chinese and English datasets, and the author's method is used to

predict the prerequisite relations of English and Chinese concept pairs as the comparative algorithm of this paper.

- (4) *AFs and MFS*. Besides the above baseline, to verify the effectiveness of the proposed method's automatic features and manual features, respectively, this paper also predicts the concept prerequisite relations for the two types of features separately. Specifically, we use a fully connected layer to receive automatic and manual features as input, thereby achieving prerequisite relations prediction individually.

On the AL-CPL and CH-CL-CPL datasets, we conducted a concept prerequisite relations prediction experiment. The performance of the model is evaluated using 5-fold cross-validation. In comparison to a baseline model, the most widely used performance metrics are Precision (P), Recall (R), and *F1* score (*F1*). Tables 2 and 3 show the results of

TABLE 2: Performance comparison of the AL-CPL dataset.

Domain metric	Data mining			Geometry			Physics			Precalculus		
	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>
RefD	52.8	77.5	66.3	42.4	62.3	50.4	49.9	49.6	49.4	75.1	69.4	72.1
RS	62.7	68.2	65.1	70.6	76.0	73.0	53.0	62.3	57.1	69.9	76.9	73.1
AT	80.7	73.3	76.7	95.0	84.7	89.5	85.2	59.3	69.9	90.2	87.1	88.6
AFs	81.7	85.3	83.3	88.4	90.6	89.4	73.8	80.2	76.7	88.3	91.7	89.9
MFs	71.1	78.0	74.3	73.6	83.0	77.9	61.3	80.2	69.3	77.3	82.5	79.8
AFs + MFs	86.5	87.3	86.9	93.7	92.5	93.1	81.4	83.8	82.6	91.6	92.7	92.1

TABLE 3: Performance comparison of the CH-AL-CPL dataset.

Domain metric	Data mining			Geometry			Physics			Precalculus		
	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>	<i>P</i>	<i>R</i>	<i>F</i>
RefD	55.6	56.4	56.0	74.1	62.9	68.1	63.4	69.5	66.3	71.5	65.6	68.4
RS	79.4	79.7	79.6	92.7	92.8	92.7	83.9	83.0	83.3	91.0	88.6	89.8
AFs	89.4	91.0	90.2	96.9	97.5	97.2	89.2	88.3	88.7	93.6	94.4	94.0
MFs	74.9	80.3	77.5	92.0	95.3	93.6	76.8	89.0	82.4	89.1	90.8	89.9
AFs + MFs	90.7	91.9	91.2	97.8	97.9	97.8	92.0	94.1	93.0	96.6	96.8	96.7

evaluating different baselines under different performance metrics for AL-CPL and CH-AL-CPL, respectively.

As shown in Tables 2 and 3, our method significantly outperforms all the baselines in all the metrics on English and Chinese datasets (except AL’s Precision metric).

From Table 2, we can find that our method achieves the best performance against all baselines on all domains, except for the Precision metric of the geometry and physics domain. The *F1* score of AFs + MFs leads AL by about 3.6%, 3.7%, 5.9%, and 3% in each of the four areas. In geometry and physics, Precision is the best probably because these two fields have the rich text and link features.

Based on Table 3, we observe that our method outperforms all baselines in all metrics and achieves the best result. Table 3 reports the evaluation metrics for the four domains. The CH-AL-CPL dataset, which is expanded by transitivity and asymmetry, has the most significant number of prerequisites. The performances obtained are generally better than CH-CAL-CPL. In addition, since the author did not give the code of the AT method from [14], some features in Chinese cannot be calculated, so we did not report the experiment results of the AT method in CH-AL-CPL.

As an excellent NLP language model, BERT replaces the encoder with the decoder by using a two-way transformer encoder. By using this method, the feature encoding effect of the words in the sentence is greatly improved. Compared with previous models, such as Word2Vec, the trained BERT model has a deeper contextual understanding. A context-based semantic feature is suitable for detecting text features of Wikipedia concepts. Moreover, as a concept pair, we also cannot ignore the rich links and category relationships between the two. Overall, through the combination of the two types of features, the performance of the concept prerequisite relations model can be further improved.

5.3. Ablation Study. In order to demonstrate that the length of the Wikipedia article influences the automatic extraction of features, we conducted an ablation experiment by varying the value of *k* (the first *k* words or characters), from 100 to 500. The experiment results are shown in Figure 3.

As shown in Figure 3, increasing *k* will increase the *F1* score of the AFs model. Particularly when *k* = 400, the model is the most effective in predicting the four domains. After the *k* value exceeds 400, however, the textual information of the concept is incorporated into other background knowledge, which affects the performance of the model to a certain extent. According to the CH-AL-CPL dataset, geometry and precalculus have the best *F1* scores when *k* = 500. This may be due to the relatively long average length of articles in this domain.

Additionally, the experiment explored the role of manual features in the MFs model. There are three types of features that are designed, content-based (#1–#7), categorical-based (#8), and link-based (#9–#14). In the experiment, one feature type was removed each time and compared with the MFs model. The result is shown in Figure 4.

Figure 4 illustrates how the prediction performance decreases to varying degrees after reducing a specific type of feature group. After removing the link-based features from the AL-CPL dataset, the decline in the four fields reached 10.5%, 12.7%, 9.1%, and 7.4%, respectively, indicating that the link relations between concepts play a crucial role in the prediction of prerequisite relations.

As a result of the difference in the CH-AL-CPL dataset, content-based features have decreased by 7.0%, 11.3%, 8.8%, and 6.1% in the four domains. Since Chinese Wikipedia has fewer words than English Wikipedia, less text information has a more significant impact on text features. Moreover, removing the features with a category has the most negligible

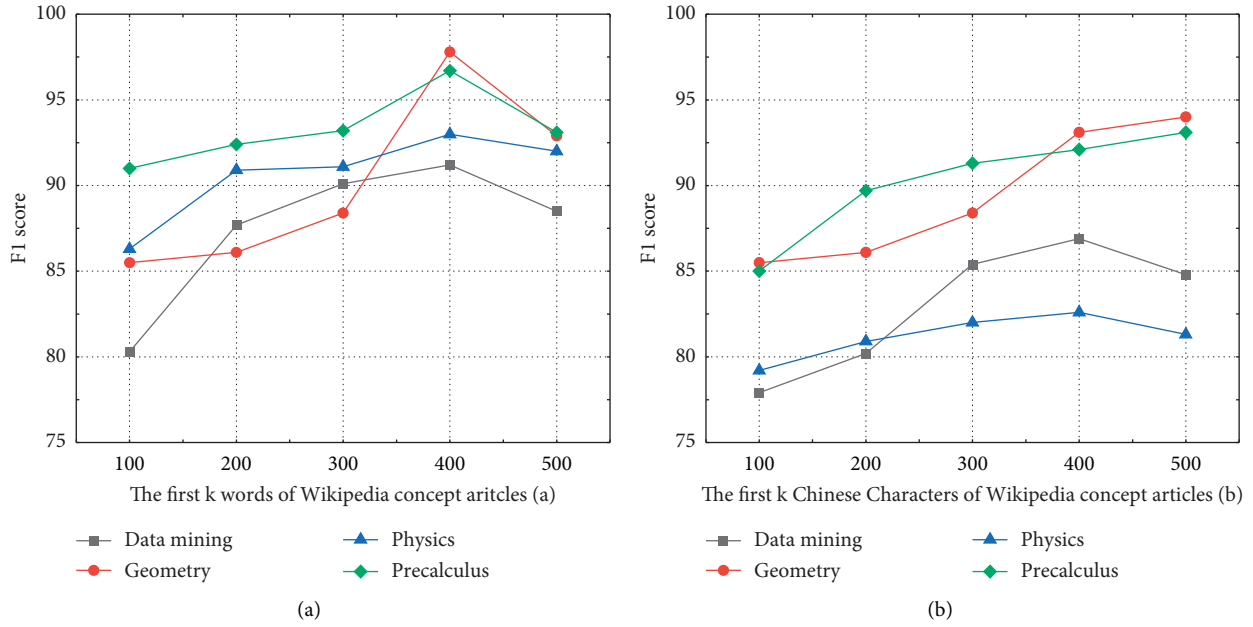


FIGURE 3: The effect of altering the value of k on the $F1$ score of the AFs model: (a) AL-CPL; (b) CH-AL-CPL.

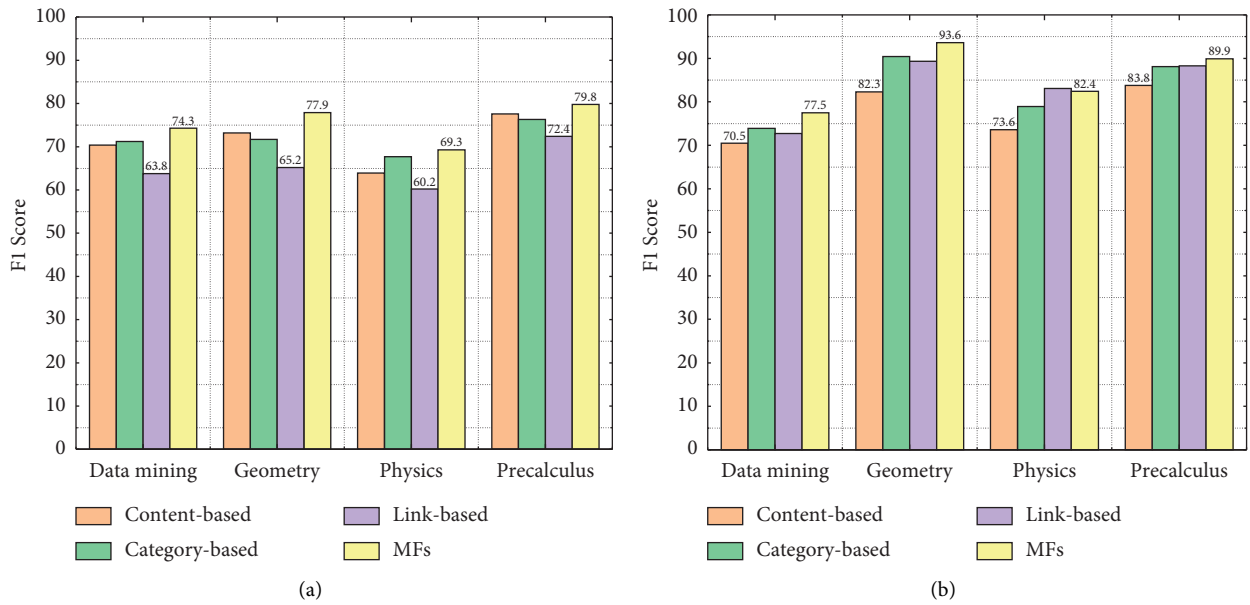


FIGURE 4: Contribution of each feature group to the MFs model in (a) AL-CPL and (b) CH-AL-CPL.

effect on the MFs model mainly because we design too few category-based features.

6. Conclusion and Future Work

In this paper, we propose a novel concept prerequisite relations prediction method called AFs + MFs, which combines the BERT sentence embedding (AFs) of the concept article and Wikipedia-based features (MFs). Furthermore, we designed a Chinese prerequisite relations dataset to verify

the effectiveness of the method. The experiment results show that our method achieves state-of-the-art results on four domains. In addition, we have conducted effectiveness studies on AFs and MFs separately.

In the future, we plan to identify the concept prerequisite relations of non-Wikipedia concepts. Moreover, some learning resources, such as MOOCs and e-lectures, contain multiple concepts. The following research question is how to recommend learning resources by considering concept prerequisite relations.

Data Availability

The CSV data used to support the results of this study have been stored in the GitHub repository (<https://github.com/lycyhrc/CH-AL-CPL>). These data are the concept prerequisite relations dataset of the Chinese version of AL-CPL. There are no restrictions on access to these data.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 61977021), the Technology Innovation Special Program of Hubei Province (Nos. 2018ACA133 and 2019ACA144), and the Teaching Research Project of Hubei University (No. 202008).

References

- [1] C. Liang, Z. Wu, W. Huang, and C. Lee Giles, "Measuring prerequisite relations among concepts," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 1668–1674, Lisbon, Portugal, September 2015.
- [2] S. Laurence and E. Margolis, *Concepts and Cognitive Science. Concepts: Core Readings*, pp. 3–81, MIT Press, Cambridge, Massachusetts, USA, 1999.
- [3] C. Hu, K. Xiao, Z. Wang, S. Wang, and Q. Li, "Extracting prerequisite relations among wikipedia concepts using the clickstream data," in *Knowledge Science, Engineering and Management. KSEM 2021. Lecture Notes in Computer Science*, H. Qiu, C. Zhang, Z. Fei, M. Qiu, and SY. Kung, Eds., vol. 12815Cham, Springer, 2021.
- [4] J. Gordon, L. Zhu, A. Galstyan, P. Natarajan, and G. Burns, "Modeling concept dependencies in a scientific corpus," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pp. 866–875, Berlin, Germany, August 2016.
- [5] B. Golshan, E. Papalexakis, and R. Agrawal, "Toward data-driven design of educational courses: a feasibility study," *Journal of Educational Data Mining (JEDM)*, vol. 8, no. 1, pp. 1–21, 2016.
- [6] H. Li, T. Wang, W. Pan et al., "Mining key classes in java projects by examining a very small number of classes: a complex network-based approach," *IEEE Access*, vol. 9, pp. 28 076–28 088, 2021.
- [7] C. Liang, J. Ye, Z. Wu, B. Pursel, and G. Giles, "Recovering concept prerequisite relations from university course dependencies," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pp. 4786–4791, San Francisco, USA, February 2017.
- [8] R. Manrique, B. Pereira, and O. Marino, "Towards the identification of concept prerequisites via knowledge graphs," in *Proceedings of the 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pp. 332–336, Santiago, Chile, December 2019.
- [9] P. Chen, Y. Lu, V. W. Zheng, and Y. Pian, "Prerequisite-driven deep knowledge tracing," in *Proceedings of the 2018 IEEE International Conference on Data Mining (ICDM)*, pp. 39–48, IEEE, Singapore, November 2018.
- [10] P. Talukdar and W. Cohen, "Crowdsourced comprehension: predicting prerequisite structure in Wikipedia," in *Proceedings of the Seventh Workshop on Building Educational Applications Using NLP*, pp. 307–315, Montréal, Canada, June 2012.
- [11] Y. Zhou and K. Xiao, "Extracting prerequisite relations among concepts in wikipedia," in *Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, Budapest, Hungary, July 2019.
- [12] M. Sayyadiharikandeh, J. Gordon, J. L. Ambite, and K. Lerman, "Finding prerequisite relations using the wikipedia clickstream," in *Proceedings of the Companion Proceedings of The 2019 World Wide Web Conference*, pp. 1240–1247, New York, NY, USA, May 2019.
- [13] S. Wang, O. Alexander, Z. Wu et al., "Using prerequisites to extract concept maps from textbooks," in *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pp. 317–326, New York, NY, USA, October 2016.
- [14] C. Liang, J. Ye, S. Wang, B. Pursel, and C. Lee Giles, "Investigating active learning for concept prerequisite learning," in *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, pp. 7913–7919, New Orleans, USA, February 2018.
- [15] F. Gasparetti, C. De Medio, C. Limongelli, F. Sciarrone, and M. Temperini, "Prerequisites between learning objects: automatic extraction based on a machine learning approach," *Telematics and Informatics*, vol. 35, no. 3, pp. 595–610, 2018.
- [16] A. Moggio and P. Andrea, "UNIGE_SE@PRELEARN: utility for automatic prerequisite learning from Italian Wikipedia," in *Proceedings of the 7th Evaluation Campaign of Natural Language Processing and Speech Tools for Italian*, pp. 1–8, December 2020.
- [17] C. Liang, J. Ye, H. Zhao, B. Pursel, and C. Lee Giles, "Active learning of strict partial orders: a case study on concept prerequisite relations," 2018, <https://arxiv.org/abs/1801.06481>.
- [18] L. Pan, C. Li, J. Li, and J. Tang, "Prerequisite relation learning for concepts in MOOCs," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, pp. 1447–1456, Vancouver, Canada, July 2017.
- [19] A. Miaschi, C. Alzetta, F. A. Cardillo, and F. Dell'Orletta, "Linguistically-driven strategy for concept prerequisites learning on Italian," in *Proceedings of the Fourteenth Workshop on Innovative Use of NLP for Building Educational Applications*, pp. 285–295, Florence, Italy, August 2019.
- [20] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "Bert: pre-training of deep bidirectional transformers for language understanding," 2018, <https://arxiv.org/abs/1810.04805>.
- [21] N. Reimers and I. Gurevych, "Sentence-bert: sentence embeddings using siamese bert-networks," 2019, <https://arxiv.org/abs/1908.10084>.
- [22] S. Bird and E. Loper, "NLTK: the natural language toolkit," in *Proceedings of the ACL 2004 on Interactive Poster and Demonstration Sessions*, Barcelona, 2004.