A Prompt-based Approach for Discovering Prerequisite Relations Among Concepts

Fangrui Du[1], Wenxin Hu[*2], Changjiu Qin[3], Zheng Xiao[4]

{[1]623725909@qq.com, [2]wxhu@cc.ecnu.edu.cn, [3]51215903107@stu.encu.edu.cn, [4]51255903091@stu.ecnu.edu.cn}

East China Normal University Shanghai, Shanghai, China

Abstract. The concept prerequisite learning task is a crucial study in smart education and plays an important role in online course design, course guidance, and learning material recommendation systems. In general, concept prerequisite relations are the determination of whether two concepts have a prerequisite relationship, usually as a binary task. Nowadays, learning concept representations from pre-trained models has become a new trend in concept prerequisite learning (CPL) tasks. However, too many handicraft features are still required to discover the concept features in previous studies. In this paper, we propose a concept prerequisite relationship discovery method based on prompt learning, in which we design four prompt functions, mapping the predicted labels to the existing labels through answer engineering after the model training. Conducting thorough experiments across three publicly available benchmarks reveals enhancements of up to 13% in F1 score. It shows that the prompt learning approach can effectively improve the prediction of prerequisite relations and provide a new idea for the study of concept prerequisite relations tasks.

Keywords: Concept prerequisite learning, prompt learning, smart education.

1 Introduction

Today, with the proliferation of open-source knowledge learning platforms such as MOOC, Coursera, and NetEase Open Classes, people are increasingly learning online. Faced with plenty of knowledge materials, it can massively increase the learner's learning efficiency if the best study path could be shown. However, the order of learning resources is determined by the relations between their central concepts [1]. Therefore, one of the key steps toward the study recommendation system is the concept prerequisite learning (CPL) task which learns the prerequisite relation between two concepts. In other words, if concept A has a prerequisite relation to concept B, learners should learn concept A before concept B.

CPL was first introduced by Talukdar and Cohen[2], who learned the concept from Wikipedia, represented it as a probabilistic planning problem, and found prerequisite relations by machine learning. Previous works mainly focus on various materials, including Wikipedia articles[3], university course dependency[4], or MOOC[5], to learn the relationship between the two concepts. Nonetheless, these resources either necessitate extra pre-processing and refinement steps or encompass an excessive amount of unstructured text, introducing further complexities to prerequisite education. In a recent study by K. Xiao[6], the Bert model was implemented to acquire an understanding of the prerequisite relations between concepts, and the study validated

ITEI 2023, November 24-26, Zhengzhou, People's Republic of China

DOI 10.4108/eai.24-11-2023.2343577

the efficacy of employing pre-training models for concept acquisition. Therefore, this paper aims to determine whether two given concepts contain prerequisite relations with the help of pre-train learning models.

In recent studies, Prompt-learning has become a new paradigm in modern NLP, which provides a pre-trained learning model with "task description". The main idea of prompt learning is to unify the training objectives of the downstream task into a prediction problem for the mask language model. It entails devising a prompt template, merging it with the original input sentences, and employing a pre-trained language model to forecast the masked words within the prompt template. As a simple and effective method, prompt learning has worked well for many natural language processing tasks, especially low-resource situations.

We employ prompt models and propose a model named PCPL for prerequisite relationship discovery. The main contributions of this paper can be summarized as follows:

(1) We are the first to propose the discovery of conceptual prerequisite relations using the prompt learning method. Considering the prerequisite relationship citation scenarios, we designed four prompt template functions to explore the effect of different types of template functions on prerequisite relation learning. We have proposed PCPL that necessitates no additional external knowledge and relies solely on the input of two concepts to determine prerequisite relationships.

(2) Due to the advantage of prompt learning under few-shot learning, it is the first attempt of prerequisite relationship learning with low-resource situation. In order to simulate low-resource scenarios, we evaluate the method in a few-shot learning context. Remarkably, our prompt-based model performs admirably even in situations with limited training samples.

(3) A comprehensive series of experiments has been conducted across a variety of different pre-trained models. We explore the performance of diverse pre-trained models on varying datasets by employing distinct prompt templates. The results demonstrate that prompt-based models consistently achieve state-of-the-art performance on educational datasets.

2 Related Work

2.1 The concept prerequisite learning works

Previous works mainly focus on three paradigms to learn prerequisite relations. The first paradigm models the concept meaning through textual content[7,8]. It calculates the correlations with the means of features engineering, which entirely relies on manual construction and needs several materials to discover the relationship. The second paradigm models the concept and relationship into neural networks such as GCN and VGAE[9,10,20], which transfers the CPL task into the link prediction task. In this paradigm, automatic feature acquisition for end-to-end classification is implemented. However, manual work is still required to design a reasonable network structure, and the model's performance depends on a large number of datasets. Due to the popularity of transfer learning, the third paradigm models learn the concepts through pre-trained models such as Bert, Roberta, etc. Some latest studies[11] employ manual and semantic features together and construct various neural networks to predict prerequisite relationships. To some extent, they combine these three paradigms in all. In this paper, we employ prompt engineering in which pre-training method is used to learn the concept features automatically without constructing additional manual features.

2.2 Prompt Works

In order to measure the knowledge learned by pre-trained models such as Bert and Roberta, researchers have proposed prompt learning, where the model is asked to answer mask questions by reconstructing the input. prompt-based learning approaches have yielded good results for many natural language processing tasks[12,13,14]. Schick et al. [12] propose manual construction of templates and knowledge distillation using a large unannotated external corpus for document-level sentiment classification tasks. Hu et al. [15] propose a tag-word mapper that uses an external knowledge base to extend the classification task. Shin et al. [16] propose a gradient-guided approach to automatically construct prompt templates for downstream tasks. In most cases, pre-trained models adapt to various downstream tasks, while the prompt is the various downstream tasks that accommodate the pre-trained language model. The idea of the prompt is to reconfigure the downstream task to fit the pre-trained language model without the need to fine-tune it. Few or zero samples are the most significant advantages for prompt engineering. In contrast to the traditional paradigm of pre-training fine-tuning, prompt-based languages are a boon for low-resource languages as they do not require task-defined parameters and do not require much annotation information. In this paper, we use prompting to reconstruct the model's input, and let the model give reasonable answers through cloze questions.

3 Methodology

In this section, we introduce the details of our method: prompt-based CPL framework. The prompt-based approach to prerequisite relationship discovery uses the pretrained model to transform prerequisite relationship discovery into a masked prediction task, where categories are judged by the output of [mask] positions. Prompt learning includes the construction of a prompt, the construction of a verbalizer and training stage.

The prompt-based CPL framework proposed in this paper is illustrated in Figure 1. For concept prerequisite learning task, the input comprises two distinct knowledge points. A complete input sentence is obtained by adding a template of our design, and the new sentence is fed into the pre-trained language model to obtain the probability distribution of the [mask] position about the whole word list. We define the prerequisite relationship through the verbalizer, and let the model predict the probability of the words filled by the verbalizer to obtain the final predicted category.

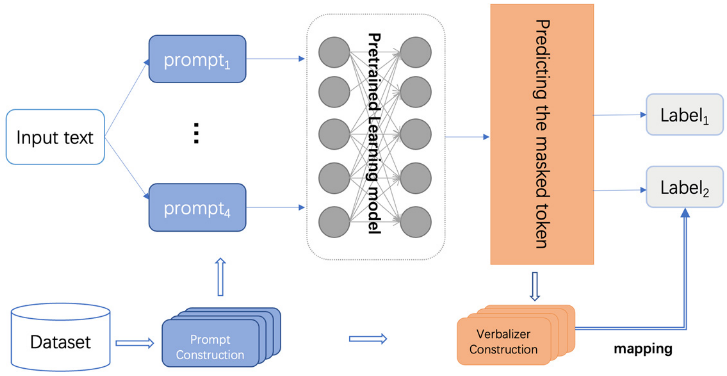

Fig.1. prompt-based CPL framework.

3.1 Problem definition

The objective of the concept prerequisite learning task is to ascertain if there exists a correlation between two concepts. In the previous study, researchers usually convert it into a binary classification problem, that can be defined as:

$$Preq(A, B) = \begin{cases} 1, & A \text{ is a prerequisite of } B \\ 0, & A \text{ is not prerequisite of } b \end{cases} \tag{1}$$

$Preq(A, B) = 1$ means A is a prerequisite of B. In other words, if people want to master concept B, they must master concept A beforehand.

3.2 Construction of Prompt Function

Since the idea of prompt learning is used in our framework, we first need to construct a prompt project. There are various ways to set up prompt templates, such as handwritten templates, automatic discrete templates, automatic continuous templates, etc. This paper is designed in the form of handwritten templates. Therefore, we build a prompting function P to modify the input text x, the input after constructing prompt is:

$$X' = P(X) \tag{2}$$

The prompt function works as follows: Firstly, we design a template, which is a textual string that has two slots: input slots [X] and answer slots [Z]. Such as in the case of CPL tasks where x1 is “tree” and x2 is “binary tree”, the template may take a form as “[X1] and [X2] have [Z] relationship”. Then the input sentence X’ would become “tree and binary tree have [Z] relationship”. In prompt learning, if input texts come entirely before the answer slots[Z], we name it the prefix prompt. Otherwise, if answer slots are in the text, we call it a cloze prompt. In the CPL task, we define two prefix prompts and two cloze prompts in Table1.

Specially, in order to distinguish whether there is a prerequisite relationship, there is such a teaching scenario, the teacher will teach the knowledge of "binary tree", frequently mention the knowledge of "tree", because "tree" is the prerequisite knowledge of binary tree, so that the "tree" and "binary tree" appear in the same semantic scenario. By applying the concept of prompt-based learning, we can recreate educational scenarios through the construction of templates. This enables pre-trained language models to gain a comprehensive understanding of the contextual settings in which concepts appear, ultimately allowing them to learn the prerequisite relationship between two concepts. We categorize the citation scenarios between two concepts into two types: explicit citation scenarios and implicit citation scenarios. Explicit citation scenarios involve a teacher explicitly indicating the prerequisite relationship between two concepts in the educational context. Implicit citation scenarios require inferring the prerequisite relationship between two concepts based on the teacher's discourse. According to the prompt type and citation scenarios, we define four prompt templates.

Table 1. CPL prompt templates.

Number	Prompt type	citation Scenarios	Template
Template 1	cloze prompt	explicit	[X1] and [X2] have [mask] relationship
Template 2	prefix prompt	explicit	Do [X1] and [X2] have prerequisite relationship? [mask]
Template 3	prefix prompt	Implicit	We should learn/study [x1] first, then learn [x2]. Right？ [mask]
Template 4	cloze prompt	Implicit	Learning [X1] is [mask] for learning [X2].

Through formulating the prompt, we acquire the reconstructed text X′, which comprises the masked tokens. Take "tree and binary tree have [Z] relationship" for example, after the model accepts and reads the input sequence, it first replaces it with [MASK] at the position where the prerequisite relationship is judged. Then, the model inserts the [CLS] and [SEP]. Where [CLS] is inserted at the head of the sentence and also as a marker for the start of the sentence, and [SEP] is checked at the end of the sentence.

3.3 Construction of Verbalizers

In this section, we design some possible answers to fill the answer slots. In prompt learning, Z could be a set that ranges from the entirety of the language. Function $f_{fill}(x',z)$ is the process by which we fill in the position [Z] with potential answers. For CPL tasks, the answersets vary with different templates, we define Verbalizers as table 2.

Table 2. CPL verbalizers.

Prompt templates	Verbalizers(1)	Verbalizers(0)
Template1	Prerequisite	No-prerequisite
Template2	Yes	No
Template3	Yes	No
Template4	Important/crucial	irrelevant

After that, a verbalizer is constructed that maps each label to a word from the masked language model. In the end, We explore the set of potential responses, z, by assessing the likelihood of their associated completed prompts using a pre-trained learning model, the calculation process is shown as follows:

$$z\text{=search}z\epsilon ZPffillx',z;\theta \tag{3}$$

Where θ is the parameters of the language model, $f_{fill}(x',z)$ indicates the insertation of answer z into the input text x', $P(f_{fill}(x',z);\theta)$ denotes the probability of each answer z being inserted into the input text x' under the parameter θ, $search$ indicates the search function of the search probability. We find the highest scoring text $\hat{z}$ to maximise the learning model score.

3.4 Prediction process and Training loss

During the training phase, the pre-trained model predicts the probability distribution $p(v \mid X')$ for all predicted words. Subsequently, We calculate the average probability of forecasted words belonging to identical categories, resulting in the corresponding label probability distribution $p(l \mid X)$. The loss function is defined as follows:

$$L(y, p) = -\sum_{i=0}^{k} y_i \log(p_i) \tag{4}$$

where k is the number of labels and y represents the ground truth. The process of prediction is shown as Figure 2.

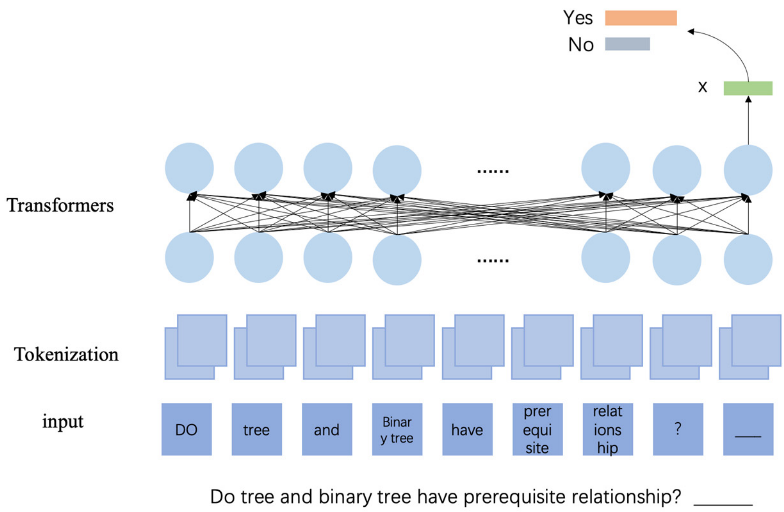

Fig.2. The process of prediction.

4 Experiments

In this section, we perform a series of experiments on publicly available datasets to assess the effectiveness of our suggested approach and compare it with the current state-of-the-art methods for identifying prerequisite relationships.

4.1 Datasets

We evaluate our approach on the following three public datasets:

MOOC ML: an English dataset that includes concepts extracted from video subtitles of Coursera's courses. The concepts are mostly from Machine Learning and Data Structure & Algorithms.

Lecture Bank: an English dataset collected from online lecture files that cover 60 courses in Natural Language Processing.

University Course: an English dataset that comes from university course dependency and its concepts are extracted using Wikipedia Miner. The concepts are mostly within the field of computer science.

4.2 Experimental Setup and Evaluation Metrics

In order to assess the ability of the model to predict prerequisite relationships, we randomly sliced the dataset at a rate of 6:4, constructing the training set and test set, shown as table3. For the purpose of addressing the imbalance problem, we oversample the positive examples and negative samples in two strategies as follows.

$$Strategy\ 1:\ \ if\ A \rightarrow B\ and\ B \rightarrow C, then\ A \rightarrow C$$

$$Strategy\ 2:\ \ if\ A \rightarrow B, then\ B\ ! \rightarrow\ A$$

The first strategy is based on the prerequisite relationship chaining rule: If concept A is a prerequisite knowledge point of concept B, concept B is a prerequisite knowledge point of concept C, then concept A is a prerequisite knowledge point of concept C. The second strategy indicates that if concept A is the prerequisite knowledge point of concept B, then concept B must not be the prerequisite knowledge point of concept A.

Then, we insert the training set and test set into each position of the template.

Table 3 datasets.

Dataset	concepts	Training set	Test set
MOOC ML	244	1882	1256
Lecture Bank	205	2570	1714
University Course	407	16011	10675

The PCPL model presented is constructed using the OpenPrompt toolkit, a resource developed by the Natural Language Processing Laboratory at Tsinghua University. We set the learning rate as 1e-4, epoch as 10, weight decay as 1e-2, max sequence Length as 256. We also use Adam as an optimizer.

Accuracy, Precision, Recall, and macro F1 Score are used as evaluation metrics. The detailed formulas are as follows.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{5}$$

$$Precision = \frac{TP}{TP+FP} \tag{6}$$

$$Recall = \frac{TP}{TP+FP} \tag{7}$$

$$F1 = \frac{2*Precision*Recall}{Precision+Recall} \tag{8}$$

In this context, TP denotes the count of correctly identified positive samples, FP signifies the count of negative samples erroneously categorized as positive, TN represents the count of accurately identified negative samples, and FN indicates the count of positive samples inaccurately classified as negative.

4.3 Performance of PCPL on Three datasets

For the pre-trained learning model, we employ the Bert-base, t5-base, gpt2 in three datasets.

Bert-base[19] is an advanced language model created by Google, excels in comprehending word contexts within sentences through extensive training on an extensive text corpus. Its bidirectional nature, which takes into account the contextual information to the left and right of a word during prediction, empowers it to capture intricate linguistic connections effectively.

T5[17] is a text-to-text model developed by Google that frames all NLP tasks as a text-to-text problem. It uses a unified architecture where both input and output are treated as text strings. This model has a single pre-trained architecture that can be fine-tuned for various NLP tasks such as translation, summarization, text classification, and more.

GPT2[20] is a generative language model developed by OpenAI. It's known for its ability to generate coherent and contextually relevant text based on a given input or prompt. It has been used for tasks like text generation, chatbots, and text completion.

Table 4-6 shows the experimental results with different templates and different pretrained models for classification on MOOC, University Course and Lecture Bank.

We can conclude from the three tables that (1) No matter which pretrained models is employed, all three datasets have the best precision value on the template 4, which indicates that implicit referenced scenarios are more consistent with the pre-trained model training contexts. it is explainable because there are seldom scenarios that the relationship between two concepts is pointed out directly. (2) On the MOOC ML and Lecture Bank, the Bert model achieves the best training results on all four templates, which is due to the fact that Bert, as a bi-directional encoder, was initially designed for understanding tasks, and prerequisite relationship prediction usually requires that the model be able to understand contextual information in the text, which is compatible with Bert's task. In contrast, gpt2, as a generative model, is a unidirectional encoder and is not applicable to the prerequisite relation understanding task. (3) It’s worth noting that on University Course dataset, T5 achieves the best performance, which indicates that the best model for different datasets is not necessarily consistent and there are only subtle differences in the effects of the different models.

Table 4. The performance of PCPL on MOOC ML.

Model	Template	Accuracy	Precision	Recall	F1_recall
	T1	0.8416	0.8478	0.84	0.8403
	T2	0.8439	0.8492	0.8425	0.8429
Bert	T3	0.8415	0.8443	0.8405	0.8409
	T4	**0.859**	**0.8612**	**0.8582**	**0.8586**
	AVG	0.8465	0.8506	0.8453	0.8457
	T1	0.8288	0.8324	0.8276	0.8279
T5	T2	0.8288	0.829	0.8284	0.8286
	T3	0.8511	0.8515	0.8507	0.8509

	T4	0.8431	0.848	0.8418	0.8422
	AVG	0.8379	0.84	0.8371	0.8374
GPT2	T1	0.8383	0.8391	0.8389	0.8384
	T2	0.828	0.828	0.8282	0.828
	T3	0.836	0.837	0.8353	0.8356
	T4	0.8376	0.8387	0.8369	0.8372
	AVG	0.835	0.8357	0.8348	0.8348

Table 5.The performance of PCPL on Lecture Bank.

Model	Template	Accuracy	Precision	Recall	F1_recall
Bert	T1	0.9154	0.9155	0.9155	0.9154
	T2	0.9119	0.9123	0.9121	0.9119
	T3	0.9014	0.9022	0.9011	0.9013
	T4	**0.9212**	**0.9228**	**0.9216**	**0.9212**
	AVG	0.9125	0.9132	0.9126	0.9125
T5	T1	0.9037	0.9037	0.9038	0.9037
	T2	0.9084	0.9085	0.9085	0.9084
	T3	0.9177	0.918	0.918	0.9177
	T4	0.9049	0.9051	0.905	0.9049
	AVG	0.9087	0.9088	0.9088	0.9087
GPT2	T1	0.9142	0.9146	0.9141	0.9142
	T2	0.91	0.91	0.909	0.91
	T3	0.9078	0.908	0.908	0.9078
	T4	0.9078	0.9078	0.9077	0.9078
	AVG	0.9099	0.9101	0.9097	0.9099

Table 6. The performance of PCPL on University Course.

Model	Template	Accuracy	Precision	Recall	F1_recall
Bert	T1	0.9614	0.9629	0.9614	0.9613
	T2	0.9347	0.939	0.9348	0.9345
	T3	0.9632	0.9641	0.9633	0.9632
	T4	0.962	0.9632	0.962	0.9619
	AVG	0.9553	0.9573	0.9554	0.9552
T5	T1	0.9647	0.965	0.9647	0.9647
	T2	**0.9649**	0.9653	**0.965**	**0.965**
	T3	0.9644	0.9651	0.9644	0.9644
	T4	0.9639	**0.9654**	0.964	0.9639
	AVG	0.9645	0.9652	0.9645	0.9645
GPT2	T1	0.9609	0.9621	0.961	0.9609
	T2	0.962	0.962	0.962	0.962
	T3	0.9617	0.9625	0.9617	0.9617
	T4	0.8878	0.8904	0.8878	0.8876
	AVG	0.9431	0.9443	0.9431	0.943

4.4 Comparisions Against State-of-the-Arts methods

We compared our framework with several state-of-the-art approaches, including PREQEQ, Bert, VGA, CPRL.

PRERE[18]is a method for predicting unknown concept antecedents from labelled concept prerequisite and course prerequisite data. PREREQ uses pairwise linked LDA models to obtain vector representations of concepts and Siamese networks to predict unknown concept antecedents

Bert[19] uses a pair of concept names and fine-tuned them to classify the prerequisite relations using a binary cross-entropy loss.

VGA[10] is a graph-based, unsupervised educational relationship extraction model that treats concepts as nodes in a graph and relationships as edges in a graph, and discovers prerequisite relationships through graph autoencoders.

CPRL[21] is the latest CPL method that achieves state-of-the-art results by creating a heterogeneous graph to represent concepts beyond the learning object. It uses a R-GCN to encode nodes and a Siamese network to identify preconditions.

Table 7. A Comparative Analysis of Performance: Prompt-Based Model vs. Baseline Approaches.

	MOOC ML			Lecture Bank			University Course		
	precision	recall	F1 score	precision	recall	F1 score	precision	recall	F1 score
PREREQ	0.448	0.592	0.510	0.590	0.502	0.543	0.468	0.916	0.597
Bert	0.746	0.753	0.761	0.838	0.837	0.838	0.838	0.835	0.835
VGAE	0.496	0.496	0.495	0.706	0.673	0.669	0.570	0.562	0.539
CPRL	0.800	0.642	0.712	0.861	0.858	0.860	0.689	0.760	0.723
Our best model	**0.861**	**0.858**	**0.859**	**0.923**	**0.922**	**0.9212**	**0.965**	**0.965**	**0.965**

For MOOC ML and Lecture Bank, our best model is under the Bert pretrained model and T4 prompt template, for University Course, our best model is under the T5 pretrained model and T2 prompt templates. The experimental results for all assessed models across these three datasets, with Precision, Recall, and F1 score referring to macro averages, are summarized in Table 7. Notably, as shown in Table 7, our proposed model consistently attains state-of-the-art performance in three datasets, irrespective of the prompt template used. From the results we can conclude that (1) When compared to non-pretrained models like PREREQ, VGAE, and CPRL, Bert models outperform them due to their comprehensive language representation acquired through extensive pre-training on a vast corpus. (2) In comparison to the Bert model, our prompt-based model exhibits superior performance. It effectively leverages the knowledge acquired during the pre-training stage, allowing it to not only accurately decipher implicit intent but also demonstrate applicability across various domains.

4.5 Few-shot performance

In order to replicate low-resource scenarios found in real-world applications, we employed a random sampling approach, selecting k (50,100,200,300,500) instances for the training set, with the remaining examples designated for the test set. Recognizing that various selections of pretrained models and prompt templates can exert an influence on test performance, we maintained consistency throughout this experimental phase by utilizing the Bert pretrained model in combination with the T4 template.

Table 8. Performance on few-shot learning

Datasets	Training set	Proportions	Accuracy	Precision	recall	F1 score
MOOC ML	50	1.6%	0.4983	0.2491	0.5	0.332
	100	3.2%	0.5001	0.2504	0.5	0.3337
	200	6.4%	0.5044	0.2522	0.5	0.3352
	300	9.6%	0.7174	0.7392	0.7188	0.7116
	500	15.9%	**0.7843**	**0.7905**	**0.7849**	**0.7833**
Lecture Bank	50	1.2%	0.6925	0.7161	0.6926	0.6839
	100	2.3%	0.786	0.7887	0.7863	0.7857
	200	4.7%	0.7906	0.7947	0.7901	0.7897
	300	7.0%	0.8187	0.8373	0.8199	0.8166
	500	11.7%	**0.8547**	**0.8548**	**0.8546**	**0.8546**
University Course	50	0.2%	0.4996	0.2498	0.5	0.3331
	100	0.4%	0.5003	0.2502	0.5	0.3334
	200	0.7%	0.6799	0.7305	0.6801	0.6615
	300	1.1%	0.8719	0.8749	0.8719	0.8717
	500	1.9%	**0.8875**	**0.8939**	**0.8874**	**0.887**

The performance of few-shot learning is shown as Table8. Since the size of each dataset is different, the proportion of few shots samples is different too. For University Course dataset, when the training set is reduced to about 1.9% of the original, the accuracy and precision of our model can still reach 88.75% and 89.39%. For the Lecture Bank dataset, the accuracy can reach 81.87% when the training size is only 300. For MOOC ML dataset, four evaluation indicators can achieve more than 70% when the training size is reduced to 300. The experimental results proved that our proposed model presents strong robustness even when the training size is few-shot.

5 Conclusion

In this paper, we introduce an innovative approach to learn prerequisite relations, termed PCPL (Prerequisite Concept Precept Learning), built upon the foundations of Prompt learning. Firstly, according to the scenarios in which the concepts appear in the contextual context. we design explicit citing templates and implicit citing templates, as well as the corresponding verbalizer. Secondly, in response to the four proposed templates, this study conducted experiments using three representative pre-trained language models, namely Bert, GPT-2, and T5. The experimental results demonstrate that the influence of different templates and various pre-trained models on the results is relatively subtle. Thirdly, We compared this paper's method with other benchmark models and improved the f1-recall by 13%, which means that the present method is more effective in discovering the prerequisite relationship between two concepts. Ultimately, our assessment extends to evaluating the models under few-shot conditions. Surprisingly, the prompt-based model demonstrates strong performance, even when the available training samples are limited. In this study, the prompt templates in the experiments were designed manually and were not unique. Different prompt templates will have different prompting effects, and automatically designed templates can be tried in the future.

Acknowledgment. We extend our gratitude to all readers who contributed valuable and insightful feedback on this manuscript. This research has received financial support from the 2023 Teaching Reform Project of East China Normal University "Data-Driven Teaching Evaluation on Computer Foundation Courses". The experimentation phase was made possible through the resources provided by the ECNU Multi-functional Platform for Innovation (Project ID: 001).

References

[1] Kui Xiao, Youheng Bai, and Yan Zhang. 2022. Extracting Precedence Relations between Video Lectures in MOOCs. In Proceedings of the 2022 International Conference on Multimedia Retrieval (ICMR '22). Association for Computing Machinery, New York, NY, USA, 608–614. https://doi.org/10.1145/3512527.3531414

[2] Partha Talukdar and William Cohen. 2012. Crowdsourced comprehension: predicting prerequisite structure in wikipedia. In Proceedings of the Seventh Workshop on Building Educational Applications Using NLP, pages 307–315.

[3] Mohsen Sayyadiharikandeh, Jonathan Gordon, Jose-Luis Ambite, and Kristina Lerman. 2019. Finding Prerequisite Relations using the Wikipedia Clickstream. In Companion Proceedings of The 2019 World Wide Web Conference (WWW '19). Association for Computing Machinery, New York, NY, USA, 1240–1247.

[4] Liang, C., Ye, J., Wu, Z., Pursel, B. and Giles, C. 2017. Recovering Concept Prerequisite Relations from University Course Dependencies. Proceedings of the AAAI Conference on Artificial Intelligence. 31, 1 (Feb. 2017). DOI:https://doi.org/10.1609/aaai.v31i1.10550.

[5] Liangming Pan, Xiaochen Wang, Chengjiang Li, Juanzi Li, and Jie Tang. 2017. Course Concept Extraction in MOOCs via Embedding-Based Graph Propagation. In Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pages 875–884, Taipei, Taiwan. Asian Federation of Natural Language Processing.

[6] Bai, Youheng, Zhang, Yan, Xiao, Kui, Lou, Yuanyuan and Sun, Kai, (2021), A BERT-Based Approach for Extracting Prerequisite Relations among Wikipedia Concepts, Mathematical Problems in Engineering, 2021, issue , p. 1-8.

[7] Liangming Pan, Chengjiang Li, Juanzi Li, and Jie Tang. 2017. Prerequisite Relation Learning for Concepts in MOOCs. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 1447–1456, Vancouver, Canada. Association for Computational Linguistics.

[8] Liang, C., Wu, Z., Huang, W., & Giles, C.L. (2015). Measuring Prerequisite Relations Among Concepts. *Conference on Empirical Methods in Natural Language Processing*.

[9] Irene Li, Alexander R. Fabbri, Swapnil Hingmire, and Dragomir R. Radev. 2020.R-VGAE: Relational-variational Graph Autoencoder for Unsupervised Prerequisite Chain Learning. In Proceedings of the 28th International Conference on Computational Linguistics. 1147–1157.

[10] Irene Li, Alexander R. Fabbri, Robert R. Tung, and Dragomir R. Radev. 2019. What Should I Learn First: Introducing LectureBank for NLP Education and Prerequisite Chain Learning. In Proceedings of the Thirty-Third AAAI Conference on ArtificialIntelligence. 6674–6681

[11] Kui Xiao, Youheng Bai, and Yan Zhang. 2022. Extracting Precedence Relations between Video Lectures in MOOCs. In Proceedings of the 2022 International Conference on Multimedia Retrieval (ICMR '22). Association for Computing Machinery, New York, NY, USA, 608–614. https://doi.org/10.1145/3512527.3531414

[12] HAN X, ZHAO W L, DING N, et al. PTR: Prompt Tun-ing with Rules for Text Classification[EB/OL]. arXiv Preprint: 2105.11259, 2021
[13] SCHICK T, SCHÜTZE H. Exploiting Cloze-questions for Few-shot Text Classification and Natural Language Inference[C]//Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume. Stroudsburg, PA, USA: Asso- ciation for Computational Linguistics, 2021: 255-269. DOI: 10.18653/v1/2021.eacl-main.20.
[14] SHIN T, RAZEGHI Y, LOGAN IV R L, et al. Auto- Prompt: Eliciting Knowledge from Language Models with Automatically Generated Prompts[C]//Proceedings of the 2020 Conference on Empirical Methods in Natu- ral Language Processing (EMNLP). 2020: 4222-4235. DOI: 10.18653/v1/2020.emnlp-main.346.
[15] HU S, DING N, WANG H, et al. Knowledgeable Prompt-tuning: Incorporating Knowledge into Prompt Verbalizer for Text Classification[C]//Proceedings of the 60th Annual Meeting of the Association for Compu- tational Linguistics (Volume 1: Long Papers). 2022: 2225-2240. DOI: 10.18653/v1/2022.acl-long.158
[16] LI C, GAO F, BU J, et al. SentiPrompt: Sentiment Knowledge Enhanced Prompt-tuning for Aspect-based Sentiment Analysis [J]. arXiv Preprint: 2109.08306, 2021. DOI: 10.48550/arXiv.2109.08306
[17] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. J. Mach. Learn. Res. 21, 1, Article 140 (January 2020), 67 pages.
[18] Sudeshna Roy, Meghana Madhyastha, Sheril Lawrence, and Vaibhav Rajan. 2019. Inferring concept prerequisite relations from online educational resources. In Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence (AAAI'19/IAAI'19/EAAI'19). AAAI Press, Article 1183, 9589–9594. https://doi.org/10.1609/aaai.v33i01.33019589
[19] Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. ArXiv, abs/1810.04805.
[20] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (n.d.). Language Models are Unsupervised Multitask Learners.
[21] Chenghao Jia, Yongliang Shen, Yechun Tang, Lu Sun, and Weiming Lu. 2021. Heterogeneous Graph Neural Networks for Concept Prerequisite Relation Learning in Educational Data. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 2036–2047, Online. Association for Computational Linguistics.