



# 计算机科学

COMPUTER SCIENCE

## 基于方向感知孪生网络的知识概念先序关系预测方法

杨明, 贺超波, 杨佳琦

引用本文

杨明, 贺超波, 杨佳琦. 基于方向感知孪生网络的知识概念先序关系预测方法[J]. 计算机科学, 2026, 53(2): 39-47.

YANG Ming, HE Chaobo, YANG Jiaqi. [Direction-aware Siamese Network for Knowledge Concept Prerequisite Relation Prediction](#) [J]. Computer Science, 2026, 53(2): 39-47.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

### [基于异构图注意力网络的智能合约漏洞检测方法](#)

Heterogeneous Graph Attention Network-based Approach for Smart Contract Vulnerability Detection

计算机科学, 2026, 53(2): 423-430. <https://doi.org/10.11896/jsjcx.241200144>

### [基于用户静动态兴趣和去噪隐式负反馈的新闻推荐算法](#)

News Recommendation Algorithm Based on User Static and Dynamic Interests and Denoised Implicit Negative Feedback

计算机科学, 2026, 53(2): 152-160. <https://doi.org/10.11896/jsjcx.241200177>

### [融合对比学习的掩码图自编码器](#)

Contrastive Learning-based Masked Graph Autoencoder

计算机科学, 2026, 53(2): 145-151. <https://doi.org/10.11896/jsjcx.250100155>

### [DCL-FKT:融合双重对比学习与遗忘机制的个性化知识追踪模型](#)

DCL-FKT: Personalized Knowledge Tracing via Dual Contrastive Learning and Forgetting Mechanism

计算机科学, 2026, 53(2): 99-106. <https://doi.org/10.11896/jsjcx.250600002>

### [基于双层级对比学习的健壮知识追踪模型](#)

Robust Knowledge Tracing Model Based on Two-level Contrastive Learning

计算机科学, 2026, 53(2): 31-38. <https://doi.org/10.11896/jsjcx.250700196>

# 基于方向感知孪生网络的知识概念先序关系预测方法

杨明 贺超波 杨佳琦

华南师范大学计算机学院 广州 510631

(2023023231@m.scnu.edu.cn)

**摘要** 知识概念先序关系预测旨在通过挖掘知识概念间的语义和拓扑关系信息补全课程知识图谱,从而提升海量教学资源组织和个性化学习路径规划等下游任务的性能。现有的基于特征工程与深度学习的方法在实体语义信息和先序关系方向性建模方面仍存在局限性,知识概念先序关系预测性能仍有提升的空间。针对该问题,设计了一种基于方向感知孪生网络的知识概念先序关系预测方法 DSN-PRL。DSN-PRL 首先采用基于对比学习的预训练语言模型 BERT 学习知识概念的语义表示,然后应用图神经网络聚合多跳拓扑特征以增强层次化结构建模,最后设计方向感知孪生网络,用于学习先序关系的方向性差异以进行预测。在 3 个基准数据集上进行相关实验,DSN-PRL 在多个关键评价指标上均优于现有基线方法,特别是相比表现最优的基线模型 DGPL,其精确率分别提升了 7.3 个百分点,2.7 个百分点和 11.4 个百分点,F1 分别提升了 1.6 个百分点,1.3 个百分点和 4.3 个百分点。

**关键词:** 知识概念先序关系预测;预训练语言模型;对比学习;图神经网络;孪生网络

**中图分类号** TP391

## Direction-aware Siamese Network for Knowledge Concept Prerequisite Relation Prediction

YANG Ming, HE Chaobo and YANG Jiaqi

School of Computer Science, South China Normal University, Guangzhou 510631, China

**Abstract** Prerequisite relation prediction for knowledge concepts seeks to enhance the curriculum knowledge graph by exploring semantic and topological dependencies among concepts, thereby improving downstream applications such as large-scale resource organization and personalized learning path planning. Existing methods, which mainly rely on feature engineering and deep learning, still struggle to effectively model entity-level semantics and the directional nature of prerequisite relations, leaving room for further improvement. To address this problem, this paper proposes a direction-aware siamese network for knowledge concept prerequisite relation learning (DSN-PRL). Firstly, DSN-PRL employs a contrastive learning-based pre-trained language model, BERT, to capture fine-grained semantic representations of knowledge concepts. It then applies a graph neural network to incorporate multi-hop topological features and enhance hierarchical structure modeling. Finally, a direction-aware siamese network is designed to learn the directional distinctions between concept pairs for accurate prerequisite relation prediction. Experiments conducted on three benchmark datasets demonstrate that DSN-PRL outperforms existing baseline methods across multiple key evaluation metrics. In particular, compared with the best baseline model DGPL, DSN-PRL improves precision by 7.3 percentage points, 2.7 percentage points, and 11.4 percentage points, and F1 by 1.6 percentage points, 1.3 percentage points, and 4.3 percentage points, respectively.

**Keywords** Prerequisite relation prediction of knowledge concepts, Pre-trained language model, Contrastive learning, Graph neural network, Siamese network

## 1 引言

随着在线课程服务规模的持续流行与扩大,课程知识图谱逐渐成为整合与管理海量教学资源、支持个性化学习路径规划的重要知识库<sup>[1]</sup>。课程知识图谱以知识概念为实体,通

过构建诸如先序关系、主题关联、知识包含等多种语义关系,形成一个结构化的知识网络。在课程知识图谱的各种语义关系中,知识概念之间的先序关系尤为关键,它是课程知识图谱中认知逻辑的核心骨架,反映了学习内容在认知过程中的先后依赖性<sup>[2]</sup>。例如,在数据结构课程中,“指针”的基本概念是

到稿日期:2025-05-31 返修日期:2025-09-28

基金项目:国家自然科学基金(62477016);广东省基础与应用基础研究基金项目(2024A1515011758,2024A1515140144)

This work was supported by the National Natural Science Foundation of China(62477016) and Guangdong Basic and Applied Basic Research Foundation(2024A1515011758,2024A1515140144).

通信作者:贺超波(hechaobo@foxmail.com)

学习“链表”的基础,其先序关系的明确不仅有助于构建合理的学习路径,也能有效防止学习者因知识衔接不畅而产生学习障碍。

然而,现实中构建的课程知识图谱往往存在知识概念先序关系不完整的问题,不完整的先序关系会影响课程知识图谱的各种下游任务,这一现象在智能课程辅导<sup>[3]</sup>、课程内容规划及教学资源推荐<sup>[4-6]</sup>等应用场景中表现得尤为明显。例如,图1中虚线部分表示“队列”与“层次遍历”之间缺少先序关系,在基于该课程知识图谱进行教学资源推荐时,如果缺失该先序关系,则可能会出现偏差,错误地将与知识概念“层次遍历”相关的教学资源过早推荐给学生,学生会因缺乏必要的前置知识而难以理解,反而降低学习效率。在实际应用中,由于知识概念间先序关系的不完整性,通常需要依赖领域专家标注<sup>[7]</sup>。然而,人工标注方法存在显著的局限性:一方面,专家标注成本高昂且难以规模化;另一方面,面对动态更新的知识体系,静态标注难以适应概念关系的演化。因此,自动化的知识概念先序关系预测成为当前该领域亟待解决的核心问题。该研究旨在通过人工智能技术,自动挖掘知识概念间潜在的先序关系,对现有的知识概念先序关系进行补全,构建完整的课程知识概念体系,减少人工标注成本并提升预测效率<sup>[8]</sup>。

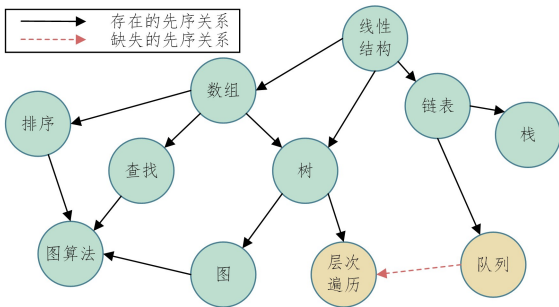


图1 先序关系示例图

Fig. 1 Illustration of prerequisite relationship

在知识概念先序关系预测领域的研究中,已经存在多种有效的方法。早期研究主要依赖特征工程,通过人工构造特征进行预测,然而该方法高度依赖专家知识,泛化能力有限。近年来,深度学习方法被广泛应用于先序关系预测,该方法减少了对传统特征工程的依赖,为先序关系预测提供了更精准的预测能力,在一定程度上提升了模型性能<sup>[9-10]</sup>,但仍存在以下局限性。

1) 现有方法主要依赖知识概念在文本中显式共现的语义信息(如知识概念在文本描述中的显式出现和词频-逆文档频率等),难以捕捉未在文本中直接关联但语义上存在潜在依赖的概念对。例如,在数据结构课程中知识概念“散列函数”与“哈希表”在定义文本中通常未互相提及,但在认知和应用层面,“散列函数”是理解“哈希表”构造和性能的前提概念,二者之间存在隐含的先序依赖关系,传统方法难以发现这类非显式的语义关联。

2) 现有方法未能有效学习先序关系的方向性差异。主流方法通常对概念对(A,B)与(B,A)采用相同或对称的特征处理方式,未能明确区分先序关系与其反向组合之间在语义结构上的差异。这种对称建模策略隐式地假设知识概念之间的

先后顺序可互换,导致模型无法识别非先序关系对所蕴含的弱依赖性甚至无依赖性,限制了对先序关系非对称本质的刻画能力,最终影响预测准确性和概念层级理解的精细度。

针对这些问题,本文提出了一种基于方向感知孪生网络(Direction-Aware Siamese Network, DASNet)的知识概念先序关系预测方法 DSN-PRL。与现有方法不同,DSN-PRL 首先采用 BERT<sup>[11]</sup> 进行对比预训练,并设计了具备先序方向敏感性的对比预训练策略,用于学习知识概念间隐含的语义差异。然后,结合图神经网络学习知识概念间的拓扑结构信息。最后,设计方向感知孪生网络,针对知识概念对采用正反序拼接策略,使模型能够显式建模先序关系与其反向组合的差异。综上所述,本文的主要工作包括:

- 1) 提出一种具有方向感知的对比预训练策略,通过引入正反序样本构造,引导 BERT 学习蕴含先序方向性差异的概念表示,以捕捉概念间隐含的语义关联;
- 2) 设计了方向感知孪生网络(DASNet),采用了知识概念对的正序与反序拼接策略,并引入了可学习方向权重机制,以显式建模并区分先序关系与其反向组合的语义差异;
- 3) 在 3 个真实世界的数据集上进行了广泛的实验研究,结果表明,DSN-PRL 优于现有基线方法。

## 2 相关工作

在知识概念先序关系预测领域,现有方法主要划分为两类:基于特征工程的方法与基于深度学习的学习框架。前者通过结构化特征空间构建实现关系推理,后者则依托神经网络架构实现数据表征的自动学习与端到端预测。这两类方法的演进过程体现了从手工设计特征到数据驱动建模的范式转变。

### 2.1 基于特征工程的先序关系预测方法

基于特征工程的方法是早期兴起的一类方法,依赖手工设计特征来描述知识概念间的关系。例如,Pan 等<sup>[12]</sup>设计了语义特征、上下文特征、结构特征这 3 种主要的特征类别来计算 MOOC 课程知识概念之间的语义相关性;Liang 等<sup>[13]</sup>通过分析两个知识概念之间的引用差异,提出了“参考距离”这一度量方法,能够有效衡量知识概念之间的先序关系。在后续研究中,Liang 等<sup>[14]</sup>进一步结合图结构和文本特征,提出了一种基于主动学习策略的先序关系预测方法,该方法显著提升了模型的性能。然而,人工特征设计难以覆盖复杂多样的知识领域,导致模型的扩展性和通用性不足,难以应对大规模动态变化的知识领域。

### 2.2 基于深度学习的先序关系预测方法

与传统的特征工程方法相比,深度学习方法凭借其强大的表征学习能力,能够更有效地学习知识概念的深层语义表示,并在大规模数据集上表现出更好的稳定性和可扩展性。例如,Roy 等<sup>[15]</sup>受潜在狄利克雷分配方法的启发,提出 PL-DA 模型以获取知识概念的潜在表示。此外,Li 等<sup>[16]</sup>通过构建资源知识概念图并结合变分图自编码技术,提升了模型的预测能力和泛化性能。Jia 等<sup>[17]</sup>使用预训练词嵌入获取知识概念与学习资源的向量表征,并构建异质图建模学习资源与知识概念之间的复杂依赖关系。此外,Zhang 等<sup>[18]</sup>通过 phrase2vec 学习知识概念的初始表征,并引入多头注意力机

制捕捉资源知识概念图中各节点之间的复杂依赖关系。Mazumder等<sup>[19]</sup>通过在异构图上使用节点注意力机制,有效捕获了知识概念间的先序关系。Qu等<sup>[20]</sup>针对现有方法在处理图同构问题时存在的局限性,提出了一种基于置换等变的有向图神经网络模型,用于知识概念先序关系预测。与此同时,Du等<sup>[21]</sup>采用基于提示学习的方法,通过设计提示函数,提升了模型在知识概念先序关系预测中的性能。虽然现有的研究方法已经取得一定进展,但在语义信息建模与先序关系方向性方面,仍存在局限性。例如,文献<sup>[10]</sup>借助知识概念间的提及信息预测先序关系,却忽略了未显式提及知识概念间的深层语义关联。文献<sup>[9,17]</sup>根据知识概念间的相似性特征进行先序关系预测,但由于其对先序关系和反向组合采用相同的特征处理方式,模型无法有效学习先序关系的方向性信息。

### 3 DSN-PRL 方法

#### 3.1 问题定义

知识概念先序关系是课程知识图谱中的一种特定语义关系,其预测本质上可以视为图上的链接预测问题。设知识概念先序关系图为有向图  $G=(C,R)$ , 节点集合  $C=\{C_1, C_2, \dots, C_m\}$  是教育资源中的一组知识概念, 边集合  $R=$

$\{(c_i, c_j)\}_{i,j \in m} \subseteq C \times C$ : 仅当存在边  $(c_i, c_j)$  时, 表示概念  $c_i$  是  $c_j$  的先序知识。

基于以上表示, 先序关系预测的问题可以形式上定义为: 给定知识概念先序关系图, 目标是学习一个函数  $F_\theta: C \times C \rightarrow \{0, 1\}$ , 它可以通过映射知识概念对  $(C_i, C_j)$  到一个二分类, 来预测  $C_i$  是否是  $C_j$  的先序知识概念。具体来说, 如果  $F_\theta$  值为 1, 则认为  $C_i$  是  $C_j$  的先序知识概念, 反之不是。

#### 3.2 模型框架

DSN-PRL 模型的整体架构如图 2 所示, 由 BERT、图卷积神经网络(GCN)和方向感知孪生网络(DAS-Net)3 部分构成。首先, 基于对比学习策略对预训练语言模型 BERT 进行微调, 以学习知识概念间的语义差异。通过引入对比学习任务, 模型不仅能够建模知识概念文本的语义相似性, 还能有效区分先序关系方向差异化的概念对。此过程优化了 BERT 参数, 确保生成的知识概念编码能够在后续预测任务中提升准确性。

完成微调后, 通过 BERT 对知识概念的文本描述进行编码, 并将其作为先序关系图中节点的初始特征。随后, 利用图神经网络对先序关系图进行表示学习, 提取知识概念间的多跳依赖逻辑并优化其表征。最后, 方向感知孪生网络基于增强后的向量表示推断知识概念对间的先序关系。

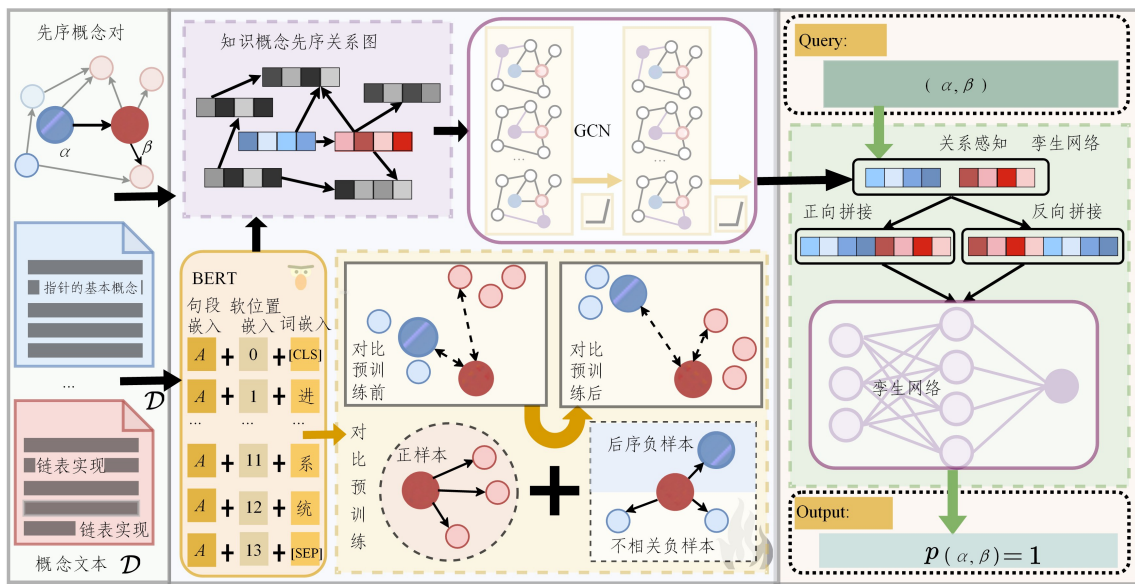


图 2 DSN-PRL 的整体框架  
Fig. 2 Overall framework of DSN-PRL

#### 3.3 基于对比学习的 BERT 预训练

对比学习在多个领域中已被广泛应用, 并通过大量实验验证了其性能和泛化能力<sup>[22-24]</sup>。作为一种提升模型区分能力的学习范式, 它能够有效捕捉知识概念间的语义差别, 特别适用于需要刻画概念隐含关联的任务。针对先序关系预测, DSN-PRL 构建了一种对先序依赖敏感的对比预训练框架: 在原生 BERT 预训练之上, 设计了方向感知的正反向样本构造策略, 以显式刻画概念间的先后依赖并学习更具区分性的概念嵌入。受 Wu 等<sup>[25]</sup>提出的 CPNC 方法的启发, DSN-PRL 保留了其基于节点相似性和关系存在性的正负样本生成方式, 同时额外引入“反向组合非先序”硬负例, 并配合随机无关

负样本, 共同构成多负例排序损失。该策略既拉近了真正先序对的表示, 又推远了反向与无关对的距离, 不仅强化了模型对先序关系的辨别能力, 也使得 BERT 学到的概念表示对语义顺序更为敏感。

为了更直观地说明上述正反向样本构造是如何发挥作用的, 举例描述其在训练中的应用细节。具体而言, 当以 A 为样本中心时, 将  $(A, B)$  作为先序正样本, 反映二者的紧密依赖; 当以 B 为样本中心时, 将  $(B, A)$  作为反向非先序负样本, 以此加强对方向差异的挖掘。此外, 每对正样本还配以多个通过随机选择不相关概念生成的负样本, 对比学习通过“拉近正样本, 拉远反向及无关负样本”策略, 既提升了相似度计算

的准确性,又加强了模型在不同上下文中对概念独立性的理解。训练目标采用多负例排序损失函数作为模型优化目标,具体形式为:

$$L = - \sum_{i=1}^N \log \frac{e^{D(u_i, v_i)}}{\sum_{j=1}^N e^{D(u_i, v_j)} + e^{D(u_i, \bar{v}_j)}} \quad (1)$$

其中,  $\mathbf{u}$  和  $\mathbf{v}$  是正样本对的嵌入向量,  $\bar{\mathbf{v}}$  是负样本的嵌入向量。  $D(u, v)$  代表两知识概念的余弦相似度, 具体的计算式为:

$$D(u, v) = \frac{(\mathbf{X}(u) \cdot \mathbf{X}(v))^T}{(\|\mathbf{X}(u)\| \cdot \|\mathbf{X}(v)\|)} \quad (2)$$

其中,  $\mathbf{X}(u)$  和  $\mathbf{X}(v)$  代表对应知识概念的特征向量表示。

### 3.4 BERT 编码器

针对先序关系预测任务, 本文采用 BERT 基础模型对概念的文本标识进行编码。在该任务中, BERT 通过捕捉知识概念的上下文信息生成高质量的嵌入表示, 能够有效建模知识概念的整体语义信息, 这对于知识概念先序关系预测领域复杂的文本描述具有重要意义。由于每个知识概念的文本描述可能包含多个词元, 为了得到一个完整的知识概念表示, 需要对这些词元的嵌入进行整合, 以形成更具表达力的知识概念向量。因此, 需要对所有词元嵌入进行均值池化, 其计算式如下:

$$\mathbf{e} = \frac{1}{n} \sum_{i=1}^n \mathbf{e}_i \quad (3)$$

其中,  $\mathbf{e}_i$  表示词元的嵌入表示,  $n$  表示词元的总数。均值池化用于合并每个词元的嵌入, 从而生成整个句子的表示。在对比学习微调阶段, BERT 根据各知识概念的文本生成嵌入表示, 并进一步通过与正负样本对的相似性优化来捕捉知识概念间的依赖关系。这种设计使得 BERT 生成的嵌入能够更好地表征知识概念间的语义差异, 为先序关系

预测提供了可靠的基础。

### 3.5 基于 GCN 的知识概念先序图表示学习

为了学习知识概念间的局部依赖和全局拓扑信息, 采用图卷积神经网络(GCN)对知识概念的表示进行建模。DSN-PRL 将 BERT 生成的知识概念嵌入作为图神经网络的初始节点特征, 通过图神经网络进一步编码和强化其语义信息。具体而言, 采用图卷积网络对先序关系图进行建模, 每个节点的表示会随着网络层数的加深而不断更新, 从而融合来自邻居节点的结构信息和语义信息。先序关系图可以形式化表示为  $G = \langle C, R \rangle$ , 其中  $C$  代表每个知识概念节点的集合,  $R$  表示知识概念之间存在的先序关系边。BERT 生成的初始知识概念表示被用作图神经网络的输入, GCN 通过聚合邻居节点的信息来更新节点表示, 以进一步学习知识概念间的结构化关系。GCN 的节点更新计算式如下:

$$\mathbf{h}_v^{(k)} = \sigma(\mathbf{W}^{(k)} \mathbf{h}_v^{(k-1)} + \sum_{u \in N(v)} \frac{1}{C_{uv}} \mathbf{W}^{(k)} \mathbf{h}_u^{(k-1)}) \quad (4)$$

其中,  $\mathbf{h}_v^{(k)}$  表示节点  $v$  在第  $k$  层的表示,  $\mathbf{h}_v^{(k-1)}$  表示节点在第  $k-1$  层的表示;  $N(v)$  表示节点  $v$  的邻居节点集合;  $C_{uv}$  通常是节点  $u$  对节点  $v$  的边的归一化因子, 确保信息传递的平衡(如节点的度数);  $\mathbf{W}^{(k)}$  表示在第  $k$  层的可学习权重矩阵;  $\sigma$  表示激活函数, 这里使用的是 ReLU。

### 3.6 方向感知孪生网络

为进一步区分先序与其反向组合关系, 设计了方向感知孪生网络(Direction-Aware Siamese Network, DASNet), 其网络架构如图 3 所示。DASNet 通过对知识概念对的正序-反序拼接, 捕捉先序和其反向组合关系的区别, 采用可学习的加权机制, 动态调节正序-反序拼接的权重, 进一步提升了对知识概念先序关系的建模能力。

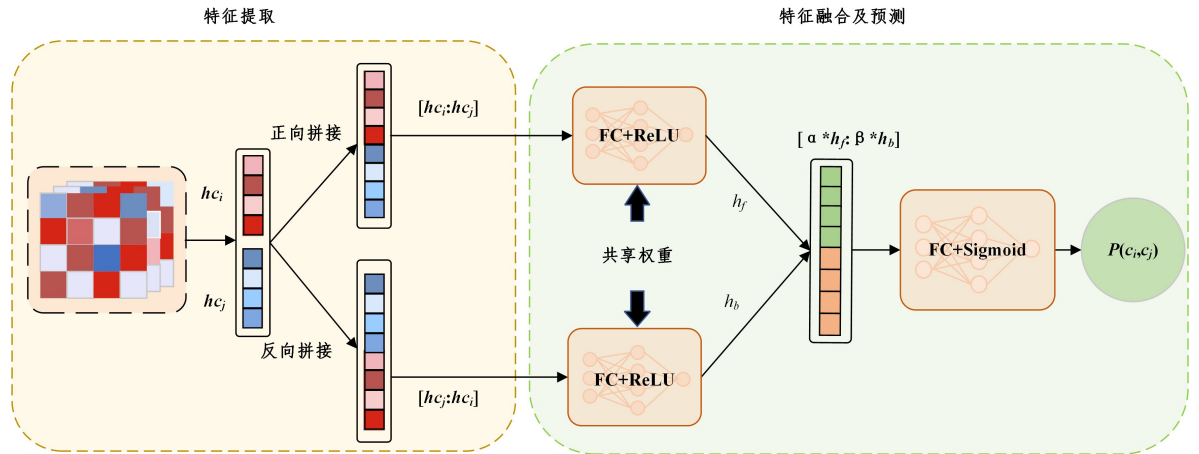


图 3 DASNet 的整体框架

Fig. 3 Overall framework of DASNet

在该框架中, DASNet 将知识概念  $a$  和知识概念  $b$  的嵌入表示分别进行正向拼接和逆向拼接, 以便模型学习到两知识概念间先序关系和其反向组合关系的差异。为学习方向敏感性, 设计双向拼接策略。正向拼接为:

$$\mathbf{h}_{\text{forward}} = \text{Concat}(\mathbf{e}_a; \mathbf{e}_b) \quad (5)$$

其中,  $\mathbf{h}_{\text{forward}}$  表示  $a$  为前驱节点、 $b$  为后继节点的关系模式;  $\text{Concat}$  表示拼接操作。反向拼接为:

$$\mathbf{h}_{\text{backward}} = \text{Concat}(\mathbf{e}_b; \mathbf{e}_a) \quad (6)$$

式(6)表示交换顺序后的逆序关系模式。得到的向量表示将通过共享层提取方向无关的语义共性, 再通过独立变换层解耦方向性依赖差异, 这种层级处理可以避免方向信息污染共性语义。将拼接结果输入共享权重的全连接层, 提取方向无关的共性特征。

$$\mathbf{R}_{\text{forward}} = \text{ReLU}(\mathbf{W}_s \mathbf{h}_{\text{forward}} + \mathbf{b}_s) \quad (7)$$

$$\mathbf{R}_{\text{backward}} = \text{ReLU}(\mathbf{W}_s \mathbf{h}_{\text{backward}} + \mathbf{b}_s)$$

其中,  $\mathbf{W}_s$  为共享权重矩阵,  $\mathbf{b}_s$  为偏置项。该设计通过参数共

享提升模型泛化能力。接着,对共享特征施加独立线性变换以解耦方向语义:

$$\begin{aligned} \mathbf{H}_{\text{forward}} &= \mathbf{W}_f \mathbf{R}_{\text{forward}} + \mathbf{b}_f \\ \mathbf{H}_{\text{backward}} &= \mathbf{W}_r \mathbf{R}_{\text{backward}} + \mathbf{b}_r \end{aligned} \quad (8)$$

其中,  $\mathbf{W}_f, \mathbf{W}_r \in \mathbb{R}^{m \times k}$  为独立权重矩阵;  $\mathbf{b}_f, \mathbf{b}_r$  表示偏置项。该方法可区分正、反向关系的异构特征。最后,引入可学习权重  $\alpha, \beta \in [0, 1]$  来平衡双向信息对预测结果的影响。预测概率的计算式如下:

$$p = \sigma(\mathbf{W}_p \cdot (\alpha \cdot \mathbf{H}_{\text{forward}} + \beta \cdot \mathbf{H}_{\text{backward}}) + \mathbf{b}_p) \quad (9)$$

通过上述拼接策略,模型能够更灵活地适应不同的知识概念关系。对于先序关系的知识概念对,模型会增加正向拼接的权重,减小反向拼接的权重。而对于反向组合关系的知识概念对,则增加反向拼接的权重,减小正向拼接的权重。这样的设计确保模型能够动态调节正负样本之间的关系,从而提升模型在先序学习关系预测任务中的表现。

在预测阶段,通过概率值推断知识概念之间是否存在先序关系,并使用交叉熵损失函数作为优化目标,从而提升预测的准确率,具体计算式如下:

$$L = \frac{1}{|T|} \sum_{(c_i, c_j) \in T} -[y_{ij} \log(p) + (1 - y_{ij}) \log(1 - p)] \quad (10)$$

其中,  $p$  是概率,  $y$  表示是否存在先序关系的标签,  $T$  表示训练集。

## 4 实验结果

### 4.1 数据集

为了评估 DSN-PRL 方法的性能,采用了 3 个基准数据集 MOOC ML<sup>[12]</sup>, LectureBank<sup>[16]</sup> 和 University Course<sup>[26]</sup> 进行实验。MOOC ML 包含 244 个机器学习领域知识概念、7411 个三元组及 1735 个先序关系对;LectureBank 包含 208 个知识概念实体、42750 个三元组及 913 个先序关系对;University Course 是基于美国 654 门计算机科学课程构建的,包含 407 个知识概念、3545 个三元组及 1008 个先序关系对。在此基础上,从维基百科与 MOOCube<sup>[27]</sup> 收集了每个实体的文本语义描述信息,数据集统计信息如表 1 所列。

表 1 数据集统计信息  
Table 1 Statistics of datasets

数据集	概念实体数量	三元组数量	先序关系对数量
MOOC ML	244	7411	1735
LectureBank	208	42750	913
University Course	407	3545	1008

### 4.2 基线模型

1) PREREQ<sup>[15]</sup>: 基于 Siamese 网络架构的监督学习方法,它利用从 Pairwise Latent Dirichlet Allocation (PLDA) 模型获得的知识概念潜在表示来预测在线教育资源中的知识概念先序关系。

2) GAEL<sup>[16]</sup>: 基于图卷积网络框架,通过生成知识概念的嵌入表示并重构邻接矩阵,来实现先序关系的推断与预测。其核心在于利用图结构特征捕捉概念间的潜在依赖模式。

3) VGAE<sup>[16]</sup>: 在基础图自编码器架构中引入变分推断机制,通过随机潜在变量和高斯先验分布建模嵌入表示,同时优化变分下界以提升先序关系预测的鲁棒性。

4) CPRL<sup>[17]</sup>: 结合异构图神经网络与弱监督学习策略,通过融合概念与学习资源的异构关系增强语义表征,并利用双分支网络架构实现先序关系分类。

5) ConLearn<sup>[9]</sup>: 通过微调预训练语言模型获取丰富的上下文语义特征,结合图神经网络的自适应聚合与自注意力机制,无需人工特征即可建模复杂先序依赖。

6) DGPL<sup>[10]</sup>: 基于图神经网络的方法,分别建模概念语义图与先序关系图,通过跨图信息传递机制融合细粒度提及信息与全局拓扑特征,有效解决隐式依赖的识别难题。

### 4.3 实验设置

考虑到数据集中样本数量和知识概念对的差异,为保证实验结果的有效性以及与先前研究对比的公平性,采用不同的数据集划分策略。LectureBank 数据集因样本稀疏、概念对有限,分配 80% 的样本至训练集、10% 的样本至验证集、10% 的样本至测试集;其余数据集因规模较大,采用 60% 的样本为训练集、10% 的样本为验证集、30% 的样本为测试集的划分标准。负样本通过从知识概念集随机抽取无关概念对生成。为缓解类别不平衡,对所有数据集的训练集和测试集分别执行 1.25 倍正样本过采样。最终实验结果基于 5 次独立数据划分的平均值统计得出。

实验在配备 NVIDIA 3090TI GPU 的服务器上进行,优化器采用 Adam 并按数据集设定初始学习率:LectureBank 与 MOOC ML 数据集为 0.001, University Course 数据集为 0.0005。基线模型参数严格遵循原论文设定。选用精确率 (P)、召回率 (R) 和 F1 值作为评价指标:精确率为预测正例中实际正例的比例;召回率为实际正例中被正确预测的比例;F1 值是两者的调和平均数,可以综合反映模型性能。其计算式如下:

1) 精确率

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

2) 召回率

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

3) F1

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (13)$$

其中,  $TP$  代表真正例,即模型准确地将样本预测为正例的数量;  $FP$  表示假正例,即模型错误地将样本预测为正例的数量;  $FN$  为假负例,指模型错误地将实际是正例的样本预测为负例的数量。

### 4.4 实验结果与性能分析

表 2 列出了 DSN-PRL 与多个基线模型在 3 个基准数据集上的性能对比结果,其中加粗字体标注了各项指标中的最优值。可以看出,DSN-PRL 在所有数据集上均取得了最高的精确率和 F1 值。在 MOOC ML, LectureBank 和 University Course 数据集上, F1 值分别达到 0.881, 0.902 和 0.864, 较次优模型 DGPL 分别提高了约 1.6 个百分点, 1.3 个百分点和 4.3 个百分点。这表明 DSN-PRL 在综合平衡精确率与召回率方面具备一定优势,具有良好的整体预测能力。在精确率指标上, DSN-PRL 在 3 个数据集上分别取得了 0.916, 0.925

和 0.912,均高于对比方法,较次优模型 DGPL 分别提高了 7.3 个百分点,2.7 个百分点和 11.4 个百分点。虽然在个别数据集中与 DGPL 或 ConLearn 的差距相对有限,但仍反映出

该方法在识别先序知识概念方面具有更高的准确性。这一结果得益于对比预训练过程中对正负样本语义差异的强化,使模型能更有效地区分先序与非先序关系,从而减小预测误差。

表 2 预测性能的比较结果

Table 2 Comparison results of the prediction performance

数据集	指标	PREREQ	GAE	VGAE	CPRL	ConLearn	DGPL	DSN-PRL
MOOC ML	P	0.448	0.293	0.266	0.800	0.831	0.843	<b>0.916</b>
	R	0.592	0.733	0.647	0.642	0.826	<b>0.889</b>	0.848
	F1	0.510	0.419	0.377	0.712	0.828	0.865	<b>0.881</b>
LectureBank	P	0.590	0.46	0.417	0.861	0.852	0.898	<b>0.925</b>
	R	0.502	0.811	0.575	0.858	0.803	<b>0.881</b>	0.861
	F1	0.543	0.589	0.484	0.860	0.826	0.889	<b>0.902</b>
University Course	P	0.468	0.450	0.470	0.689	0.776	0.798	<b>0.912</b>
	R	<b>0.916</b>	0.886	0.694	0.760	0.782	0.846	0.802
	F1	0.597	0.597	0.560	0.723	0.779	0.821	<b>0.864</b>

另外,在召回率方面,DSN-PRL 在 LectureBank 上达到 0.861,在 MOOC ML 与 University Course 上分别为 0.848 和 0.802,虽然略低于部分基线模型,但在 3 个数据集上的 F1 均取得最优表现。这一现象与模型的训练与预测机制相关。在对比预训练阶段,反向关系被设定为硬负例,并通过多负例排序损失不断拉大正反向表示之间的差异,使得模型逐渐形成了只有在正向证据充分时才输出正例的判别能力。这种偏向在后续 DASNet 中进一步被强化:模型对正反向拼接进行加权建模,并在交叉熵损失下学得一种方向敏感的判定方式,即只有当正向得分显著高于反向得分时,样本才会被预测为先序关系。这样的机制显著降低了假正例的出现,保证了较高的精确率和预测可靠性,但也意味着一部分语义联系较弱、先序依赖关系模糊的真实先序对更难通过判定,从而转化为漏判并反映在召回率的下降上。实验证据也支持这一点:在权重初始化实验中,过强的方向性偏置(如  $\alpha=1.0$ )会显著降低召回率,即使精确率升高,最终 F1 仍不及更均衡的配置;而在 LectureBank 数据集中,先序关系对的数量相较于其他数据集更少,能够提供的有效先序信息并不充分。当方向性约束过强时,模型更容易对先序关系作出更保守的判断,从而遗漏部分真实先序关系,导致召回率下降。而在该数据集上采用对称初始化( $\alpha=0.5$ )时,F1 达到最高值(0.902),说明在先序关系信息相对不足的情境下,均衡的方向性建模能够在保持精确率的同时减轻召回率下降的问题(详见 4.6 节)。整体而言,DSN-PRL 通过硬负例约束与非对称预测在精确率和召回率之间实现了有意义的权衡,以更高的精确率换取有限的召回损失是合理的。

综上,DSN-PRL 通过引入方向敏感的训练机制与方向感知结构,有效提升了模型对知识概念先序关系的推理能力。该方法在多个数据集上性能稳定,尤其在精确率和 F1 值方面表现优越,验证了其在先序关系预测任务中的实际应用潜力。

#### 4.5 消融实验

为了验证方向感知孪生网络和对比预训练在 DSN-PRL 中的有效性,本文设计了两组模型变体,通过逐一替换关键组件构建对照模型,并记录各个变体的性能差异。变体 DSN-

PRL-B 通过移除对比学习策略,仅使用原生 BERT 对知识概念的文本描述进行编码;变体 DSN-PRL-S 则将方向感知孪生网络替换为文献[9,17]中根据知识概念间相似性构建的网络结构,保持其余模块不变。

从图 4 可以看出,DSN-PRL 在 3 个数据集上均取得了最优性能。在 MOOC ML 数据集中,DSN-PRL 的 F1 值为 0.881,分别高于两个变体 DSN-PRL-B(0.837)和 DSN-PRL-S(0.872);在 LectureBank 数据集中,F1 值为 0.902,而两个变体分别为 0.807 和 0.775;在 University Course 数据集中,DSN-PRL 的 F1 值达到 0.864,较 DSN-PRL-B(0.761)和 DSN-PRL-S(0.811)均有所提升。在精确率方面,DSN-PRL 在所有数据集中同样优于两个变体,最高精度差值达到 13.2 个百分点。

从对比结果可以看出,引入对比学习策略能有效提升 BERT 在知识概念先序关系预测中的推理能力,而方向感知孪生网络的结构设计更有助于模型学习知识概念之间的先后依赖差异。两者协同作用使 DSN-PRL 在学习先序关系方面具备更高的建模能力和表达精度。

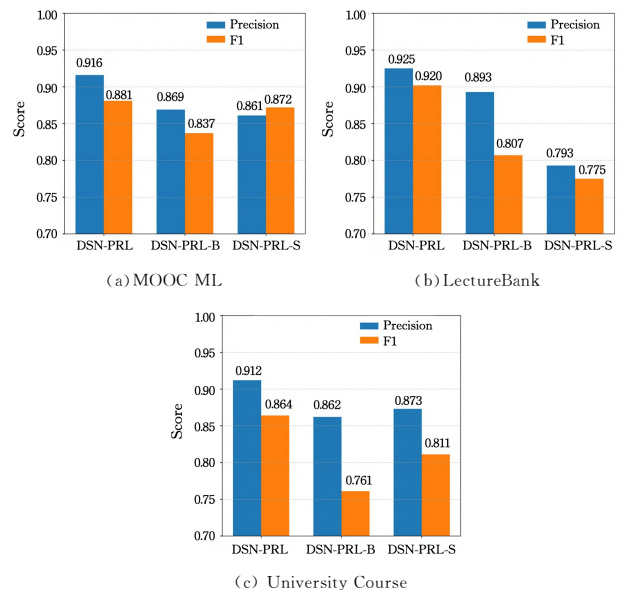


图 4 消融研究

Fig. 4 Ablation study

#### 4.6 先序权重初始值的影响

在方向感知孪生网络(DASNet)中,可学习的正向-反向拼接权重参数 $\alpha$ 和 $\beta$ 用于动态调整正序拼接与逆序拼接的贡献比例。这两个参数的初始值直接影响模型的初始优化方向与收敛效率。为探究不同初始值对模型性能的影响,设计了针对该初始值的对比实验,用于分析初始权重设置对先序关系预测任务的影响,并提出优化建议。基于以上数据集,固定其他超参数,仅调整正向-反向拼接权重 $\alpha$ 和 $\beta$ 的初始值。考虑到实验从均衡到极端的设置,选取了5种具有代表性的初始组合。

1)对称初始化: $\alpha=0.5, \beta=0.5$ ,假设正序与逆序初始贡献均等,旨在通过双向信息平衡验证模型自主学习知识概念先序方向性的能力。此设置作为基线不预设方向偏好,依赖数据驱动模型动态调整权重。

2)正向偏置: $\alpha=0.7, \beta=0.3$ ,假设正序关系对预测更重要,旨在验证预先进行正序方向性引导是否能加速模型对知识概念先序关系模式的学习。

3)反向偏置: $\alpha=0.3, \beta=0.7$ ,探究逆序权重主导时的性能变化,旨在验证预先进行反序方向性引导是否能加速模型对知识概念先序关系模式的学习。

4)极端正向: $\alpha=1.0, \beta=0.0$ ,完全依赖正序拼接,该策略模拟先序关系完全可通过正序语义独立表征的极端假设,旨在验证逆序信息是否对正序依赖的推理具有冗余性,或是否因排除反向干扰而提升预测精度。

5)极端反向: $\alpha=0.0, \beta=1.0$ ,完全依赖逆序拼接,排除正序信息对预测的影响,以验证逆序关系是否具备独立表征先序逻辑的能力。

在此基础上,为精确探索性能最优区间,进一步在潜在峰值附近补充了强正向偏置( $\alpha=0.8, \beta=0.2$ )与弱正向偏置( $\alpha=0.6, \beta=0.4$ )两组实验。该策略旨在以有限实验成本揭示初始权重与性能的内在关系,全面评估权重分配对模型方向性建模能力的影响。每种配置重复实验5次,以平均F1分数及精确率作为性能指标,结果如表3所列。

表3 先序权重初始值的影响研究

Table 3 Research on the influence of the initial value of the prerequisite weight

数据集	指标	对称初始化	正向偏置	反向偏置	强正向偏置	弱正向偏置	极端正向	极端反向
		( $\alpha=0.5, \beta=0.5$ )	( $\alpha=0.7, \beta=0.3$ )	( $\alpha=0.3, \beta=0.7$ )	( $\alpha=0.8, \beta=0.2$ )	( $\alpha=0.6, \beta=0.4$ )	( $\alpha=1.0, \beta=0.0$ )	( $\alpha=0.0, \beta=1.0$ )
MOOC ML	P	0.889	<b>0.916</b>	0.868	0.888	0.893	0.892	0.882
	F1	0.880	<b>0.881</b>	0.879	0.873	0.879	0.877	0.870
LectureBank	P	<b>0.925</b>	0.883	0.880	0.854	0.892	0.939	0.862
	F1	<b>0.902</b>	0.858	0.845	0.844	0.857	0.863	0.842
University Course	P	0.906	<b>0.912</b>	0.911	0.897	0.901	0.904	0.881
	F1	0.860	<b>0.864</b>	0.856	0.861	0.847	0.849	0.842

分析表3所列实验结果可知,权重初始化策略对模型性能的影响与数据特性深度耦合:在标注稀疏的LectureBank数据集中,对称初始化( $\alpha=0.5, \beta=0.5$ )以F1=0.902优于其他配置,其双向均衡特性有效抑制了稀疏标注带来的噪声干扰;而在结构化知识领域(如MOOC ML和University Course),正向偏置( $\alpha=0.7, \beta=0.3$ )较为贴合先序逻辑方向性,在F1上分别取得了0.881和0.864的较优性能。值得注意的是,新增的强正向偏置( $\alpha=0.8$ )与弱正向偏置( $\alpha=0.6$ )的实验结果进一步表明,在MOOC ML和University Course数据集上,性能在 $\alpha=0.7$ 时达到峰值,形成一个相对稳定的高性能区间,其两侧结果( $\alpha=0.6$ 与 $\alpha=0.8$ )略有下降。极端配置中,极端正向( $\alpha=1.0$ )虽在LectureBank的精确率高达0.939,但因召回率显著下降导致F1=0.863低于对称配置,极端反向( $\alpha=0.0$ )则因逆序噪声误导,在所有数据集中表现最差(如University Course-F1=0.842),这表明仅依赖单一方向会破坏模型对双向语义关系的学习能力,且难以平衡精确率与召回率。

#### 4.7 参数设置分析

为分析知识概念嵌入维度对模型性能的影响,本文在3个数据集上选取了128,200,256,512和768这5个典型维度进行实验,结果如图5所示。图5中已分别对3条折线的最优F1分数点做了标注,以便直观对比各数据集的性能峰值。不同维度的选择将直接影响模型语义表示能力与计算

复杂度,是影响性能的重要因素。

从实验结果可知,各数据集在不同维度下的表现各不相同:MOOC ML在中等维度(256)取得最高,表明该任务在保留必要语义的同时,中等维度更有助于模型泛化;University Course数据集则随着维度增大持续提升,768维达到峰值,说明该任务对高维语义表达的依赖更强;LectureBank数据集在200维达到峰值后随维度增加出现回落,表明高维可能引入过拟合或冗余特征。

综上,嵌入维度对不同任务的影响不同:中等维度(200~256)能在多场景下实现性能与效率的平衡;对于需要更丰富语义捕捉的任务(如University Course),高维嵌入(768维)可进一步提升效果。因此,实际应用中应结合数据规模和特性,灵活选择最优嵌入维度。

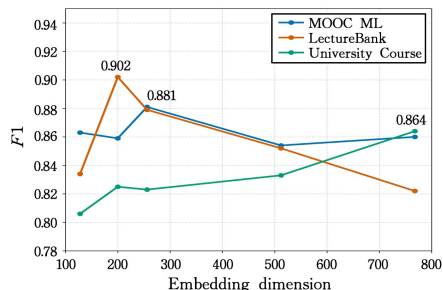


图5 嵌入维度变化分析

Fig. 5 Embedding dimension variation analysis

**结束语** 本文提出了一种基于方向感知孪生网络(DSN-PRL)的知识概念先序关系预测方法,通过对比预训练策略增强语义方向敏感性、图神经网络建模拓扑依赖以及动态权重机制解耦正反向语义差异,显著提升了预测性能。实验结果表明,DSN-PRL在MOOC ML, LectureBank和University Course数据集上的精确率与F1值均优于现有基线方法,尤其在捕捉知识概念文本描述中未显式共现的知识概念对和区分先序关系方向性差异方面表现突出。然而,该方法仍受限于课程知识图谱结构不完整对拓扑建模的影响,且多模块设计导致较高的计算复杂度。未来工作将探索模型轻量化设计以降低计算成本,并研究动态预测机制以适应课程知识概念的变化。

## 参 考 文 献

- [1] ZHOU K Y, MENG Q X, CAO Y Y, et al. Exploration and practice of problem-oriented knowledge graph in college physics curriculum[J]. *College Physics*, 2025, 44(1): 66.
- [2] ZHANG C X, PENG C, LUO M Q, et al. Construction of mathematics course knowledge graph and its reasoning[J]. *Computer Science*, 2020, 47(S2): 573-578.
- [3] WANG S, LIU L. Prerequisite concept maps extraction for automatic assessment[C] // *Proceedings of the 25th International Conference Companion on World Wide Web*. 2016: 519-521.
- [4] LIN Y, ZHANG W, LIN F, et al. Knowledge-aware reasoning with self-supervised reinforcement learning for explainable recommendation in MOOCs[J]. *Neural Computing and Applications*, 2024, 36(8): 4115-4132.
- [5] GONG J B, WANG S, WANG J L, et al. Attentional graph convolutional networks for knowledge concept recommendation in MOOCs in a heterogeneous view[C] // *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2020: 79-88.
- [6] GUAN Q, XIAO F, CHENG X, et al. Kg4ex: An explainable knowledge graph-based approach for exercise recommendation [C] // *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*. 2023: 597-607.
- [7] BERGAN J R, JESKA P. An examination of prerequisite relations, positive transfer among learning tasks, and variations in instruction for a seriation hierarchy[J]. *Contemporary Educational Psychology*, 1980, 5(3): 203-215.
- [8] TANG X S, CHENG L Y, ZHANG C H, et al. Application of large language models in the automated construction of disciplinary knowledge graphs[J]. *Journal of Beijing University of Posts and Telecommunications (Social Sciences Edition)*, 2024, 26(1): 125.
- [9] SUN H, LI Y, ZHANG Y. ConLearn: contextual-knowledge-aware concept prerequisite relation learning with graph neural network[C] // *Proceedings of the 2022 SIAM International Conference on Data Mining*. 2022: 118-126.
- [10] XU G L, BAI R J. Learning with double graph for concept prerequisite discovering [J]. *Data Analysis and Knowledge Discovery*, 2024, 8(5): 38-45.
- [11] DEVLIN J, CHANG M W, LEE K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding [C] // *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 2019: 4171-4186.
- [12] PAN L, LI C, LI J, et al. Prerequisite relation learning for concepts in moocs[C] // *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*. 2017: 1447-1456.
- [13] LIANG C, WU Z, HUANG W, et al. Measuring prerequisite relations among concepts[C] // *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. 2015: 1668-1674.
- [14] LIANG C, YE J, WANG S, et al. Investigating active learning for concept prerequisite learning[C] // *Proceedings of the AAAI Conference on Artificial Intelligence*. 2018: 7913-7919.
- [15] ROY S, MADHYASTHA M, LAWRENCE S, et al. Inferring concept prerequisite relations from online educational resources [C] // *Proceedings of the AAAI Conference on Artificial Intelligence*. 2019: 9589-9594.
- [16] LI I, FABBRIO A R, TUNG R R, et al. What should i learn first: Introducing lecturebank for nlp education and prerequisite chain learning[C] // *Proceedings of the AAAI Conference on Artificial Intelligence*. 2019: 6674-6681.
- [17] JIA C, SHEN Y, TANG Y, et al. Heterogeneous graph neural networks for concept prerequisite relation learning in educational data[C] // *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 2021: 2036-2047.
- [18] ZHANG J, LIN N, ZHANG X, et al. Learning concept prerequisite relations from educational data via multi-head attention variational graph auto-encoders[C] // *Proceedings of the 15th ACM International Conference on Web Search and Data Mining*. 2022: 1377-1385.
- [19] MAZUMDER D, PAIK J H, BASU A. A graph neural network model for concept prerequisite relation extraction[C] // *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*. 2023: 1787-1796.
- [20] QU X, SHANG X, ZHANG Y. Concept prerequisite relation prediction by using permutation-equivariant directed graph neural networks[C] // *Proceedings of Machine Learning Research*. 2024: 39-47.
- [21] DU F, HU W, QIN C J, et al. A prompt-based approach for discovering prerequisite relations among concepts[C] // *Proceedings of the 3rd International Conference on Internet Technology and Educational Informatization*. 2024: 1-13.
- [22] CHEN T, KORNBLITH S, NOROUZI M, et al. A simple framework for contrastive learning of visual representations [C] // *Proceedings of the 37th International Conference on Machine Learning*. 2020: 1597-1607.
- [23] KHOSLA P, TETERWAK P, WANG C, et al. Supervised con-

trastive learning[C]//Proceedings of Advances in Neural Information Processing Systems. 2020:18661-18673.

[24] WANG Z, SHEN Y, ZHANG Z, et al. Relative contrastive learning for sequential recommendation with similarity-based positive sample selection[C]//Proceedings of the 33rd ACM International Conference on Information and Knowledge Management. 2024:2493-2502.

[25] WU S W, SHEN X Q, XIA R. Commonsense knowledge graph completion via contrastive pretraining and node clustering[C]//Findings of the Association for Computational Linguistics. 2023:13977-13989.

[26] LIANG C, YE J B, WU Z H, et al. Recovering concept prerequisite relations from university course dependencies[C]//Proceedings of the 31st AAAI Conference on Artificial Intelligence. 2017:4786-4791.

[27] YU J, LUO G, XIAO T, et al. MOOC Cube: A large-scale data repository for NLP applications in MOOCs[C]//Proceedings of

the 58th Annual Meeting of the Association for Computational Linguistics. 2020:3135-3142.



**YANG Ming**, born in 1999, master candidate. His main research interests include educational knowledge graph and graph neural networks.



**HE Chaobo**, born in 1981, professor, Ph. D supervisor, is a distinguished member of CCF (No. 13911D). His main research interests include graph data mining and intelligent education.

(责任编辑:喻黎)

## 2025年“CCF最高科学技术奖”评选结果公告

“CCF最高科学技术奖”授予在计算机科学、技术和工程领域取得重大突破,成就卓著、贡献巨大的资深中国计算机科技工作者。

CCF奖励委员会决定授予中国科学院计算技术研究所李国杰研究员、中国人民解放军国防科技大学吴泉源教授2025年“CCF最高科学技术奖”,以表彰他们为中国计算机事业的发展做出的卓越贡献。

特此公告。

中国计算机学会  
2026年1月13日

附:



李国杰,中国科学院计算技术研究所研究员,CCF会士。

获奖理由:李国杰研究员是我国计算机科技领域具有战略思维的杰出科学家,在高性能计算机和通用处理器领域的研制和产业化方面做出引领性贡献,探索出一条适合中国国情的跨越式发展道路。他重视战略谋划,为中国计算机事业的创新发展和人才培养做出了重大贡献。



吴泉源,中国人民解放军国防科技大学教授。

获奖理由:吴泉源教授是我国逻辑级数字系统模拟技术的先驱者,通过逻辑级模拟解释技术创建了系统软件模块的正确性调试与效能分析工具,用于保障计算机软硬件系统的高质量研制。他也是我国面向对象分布计算中间件技术的引领者,为中国计算机事业的创新发展做出了重大贡献。

据 CCF 微信公众号