

廖凤露, 孙婷婷. 大模型赋能图书关系网络建设新范式[J]. 智能计算机与应用, 2026, 16(2): 95-102. DOI: 10.20169/j.issn.2095-2163.25071605

# 大模型赋能图书关系网络建设新范式

廖凤露, 孙婷婷

(电子科技大学 图书馆, 成都 611731)

**摘要:** 针对传统元数据驱动的关系网络建设方法在图书语义深度挖掘和隐藏关系识别能力上的局限, 以及大模型在垂直应用中的幻觉现象, 提出大模型赋能的图书关系网络建设新范式, 构建更具语义密度和知识发现能力的关系网络, 为学科交叉研究和知识创新服务提供新思路。首先, 在语义特征增强层, 综合提示词、模型开发和生成策略三种方式来提高语义特征生成的可靠性。其次, 在关系网络生成层, 基于上下文语义向量构建关系模型, 突破元数据限制以挖掘深层语义关联。实验表明, 该范式具备高度可靠的图书语义信息增强能力, 并且能够发掘图书之间的隐藏关系, ROUGE 平均得分取得了 118% 的性能提升,  $R_{BERT}$  突破 71%, 提高了 3.47 个百分点; 发现跨类别隐藏关系超过所有相关关系的 40%。

**关键词:** 关系网络; 语义特征; 大语言模型; 幻觉检测与缓解; 微调; LoRA

**中图分类号:** TP18 **文献标志码:** A **文章编号:** 2095-2163(2026)02-0095-08

## LLMs empower a new paradigm for constructing book relationship networks

LIAO Fenglu, SUN Tingting

(Library, University of Electronic Science and Technology of China, Chengdu 611731, China)

**Abstract:** In view of the limitations of traditional metadata-driven relational network construction methods in the ability to deeply mine book semantics and identify hidden relationships, as well as the illusion phenomenon of large language models in vertical applications, this study proposes a new paradigm for the construction of book relational networks enabled by large language models, constructing a relational network with higher semantic density and knowledge discovery capabilities, and providing new ideas for interdisciplinary research and knowledge innovation services. In the semantic feature enhancement layer, comprehensive prompt words, model development and generation strategies are used to improve the reliability of semantic feature generation. In the relationship network generation layer, a relationship model is constructed based on contextual semantic vectors to break through metadata limitations to mine deep semantic associations. Experiments show that this paradigm has a highly reliable ability to enhance book semantic information and can discover hidden relationships between books. The average ROUGE score has achieved a performance improvement of 118%, and  $R_{BERT}$  has exceeded 71%, an increase of 3.47 percentage points; and cross-category hidden relationships exceeded 40% of all related relationships.

**Key words:** relationship network; semantic features; LLMs; hallucinations detection and mitigation; fine-tuning; LoRA

## 0 引言

以 ChatGPT、DeepSeek 为代表的大语言模型 (Large Language Models, LLMs), 通过海量数据预训练与自监督学习, 展现出强大的上下文理解、逻辑推理和泛化能力<sup>[1]</sup>, 彻底革新了自然语言处理 (Natural Language Processing, NLP) 的建模方式。2024 年和 2025 年《政府工作报告》中明确提出开展“人工智能+”行动, 支持大模型广泛应用, 鼓励人工

智能技术和经济社会各领域的深度融合, 推动各行业的智能化转型升级。图书馆作为信息资源中心和知识服务核心场所, 充分利用大模型能力重新定义知识关系网络建设的新范式具有重要意义, 是下游知识服务、资源推荐等任务的基础。

传统知识网络建设数据大多来源于元数据且通过  $n$ -gram (连续  $n$  个词的组合) 的相似度来建立关系。但是首先, 这种方法的数据字段类别较少, 无法刻画对象的全貌; 其次相似度的计算缺乏语义和

**基金项目:** 2023 年度四川省教育信息技术研究课题 (DSJZXKT194)。

**作者简介:** 孙婷婷 (1987—), 女, 硕士, 工程师, 主要研究方向: 教育信息化, 智慧图书馆。

**通信作者:** 廖凤露 (1993—), 女, 硕士, 助理工程师, 主要研究方向: 信息化建设, 智慧图书馆。Email: fliao@uestc.edu.cn。

**收稿日期:** 2025-07-16

上下文感知,无法发掘对象之间的隐藏关系。大模型在处理语言任务时泛化能力显著,能够通过小样本、甚至是零样本学习实现知识迁移<sup>[2]</sup>,是语义特征扩充的基座模型。但是受技术架构、训练数据质量等影响,模型幻觉问题难以避免<sup>[3]</sup>,提示工程和模型开发等技术分别从指令优化和算法优化的角度来缓解模型幻觉。Transformer 编码器通过多头自注意力机制(Multi-Head Self-Attention)捕捉文本的上下文依赖关系,将文本转换为蕴含语义与上下文信息的嵌入(Embedding)向量<sup>[4]</sup>,通过计算嵌入向量之间的相似性来构建图书关系网络。

## 1 相关研究

### 1.1 大语言模型简介及其发展历程

大语言模型在海量文本数据的基础上训练,能够高精度地生成类人文本、回答问题并完成其他语言相关任务<sup>[5]</sup>,是架构更复杂、训练数据规模更大、能力更强的语言模型(Language Model, LM)。2017年 Vaswani 等学者<sup>[6]</sup>在论文《Attention is All You Need》中提出 Transformer 架构,通过引入注意力机制(Attention Mechanism)来捕捉序列中各个元素之间的依赖关系,解决了传统序列模型(如 RNN/LSTM)在长程依赖和计算效率上的瓶颈,成为大语言模型发展的核心基础。而后分别基于 Transformer 的编码器(Encoder)和解码器 Google 推出了 BERT 模型<sup>[7]</sup>,OpenAI 推出了 GPT 模型<sup>[8]</sup>。ChatGPT<sup>[9]</sup>基于 GPT-3.5 和 InstructGPT 专门为多轮对话进行微调,开创了对话式 AI 的人机交互模式,被称为“ChatGPT 时刻”。

2025年1月发布的 DeepSeek-R1 模型<sup>[10]</sup>通过专家混合(MoE)架构减轻训练负担和强化学习(RL)增强推理能力。与 GPT-4o 和 OpenAI-o1 等最先进的大语言模型相比,DeepSeek-R1 不但推理能力强大,同时训练成本极低,即使没有巨额计算费用也可以实现高性能推理。DeepSeek-R1 蒸馏模型用“满血版”模型生成的中间数据来微调参数较小的模型,例如 Qwen-32B、Llama-70B 等,确保小参数模型在推理任务中也表现出色,同时足够轻量化以便本地部署。DeepSeek-R1 打破了先前尖端 AI 的闭源性质,使得 LLMs 得以普及化,推动了各行各业 AI 应用的蓬勃发展。

### 1.2 大语言模型在语义表征中的应用

将大语言模型应用于语义表征既可以挖掘深层次的语义信息,又可以提高特征构建、提取和选择的

效率。

(1)语义特征的增强与扩充:根据已有元数据生成语义特征,丰富特征维度,为后续分析和模型训练提供更全面的数据。

Liu 等学者<sup>[11]</sup>利用大模型文本生成能力,分别从 item 和 user 角度进行特征增强,将元数据作为输入要求大模型生成相关摘要以及用户感兴趣的 topic,进行特征扩充。Peng 等学者<sup>[12]</sup>对 LLM 进行了微调和部署,用于电子商务场景中的查询重写,以弥补语义匹配过程中的语义空白,尤其是 long-tail 查询。

(2)特征编码:利用大模型将文本转换为高维语义向量,实现文本特征编码,以便于多特征的融合与关联关系的计算等。

U-BERT<sup>[13]</sup>用 BERT 模型将用户写过的评论、用户元数据、当前推荐的 domain 数据向量化后编码一起形成用户的个性化表征,然后用于下游推荐任务。LEARN<sup>[14]</sup>提出了基于 Transformer Decoder 实现的 User-Tower 和 Item-Tower 架构,用于生成 user 特征编码和 item 特征编码。

### 1.3 大语言模型幻觉检测及缓解技术

大语言模型幻觉(hallucinations)是指大模型在生成文本时,编造出看似合理或可信但与事实相悖或用户输入不一致的信息的现象<sup>[15]</sup>,通常被称为“一本正经地胡说八道”。幻觉产生的原因可归纳为内部因素和外部因素。

大模型本质上是概率模型,是依据用户输入以及上下文来预测下一个输出,模型倾向于输出统计概率高但未必符合事实的内容,这一特性是幻觉产生的内在根源。训练数据若存在缺失、偏差或错误,会致使模型学习到错误信息,进而成为引发幻觉的外部诱因。大模型幻觉检测技术旨在识别模型生成内容中的虚假信息,为缓解幻觉提供依据。可以通过分类器、不确定性的度量等方式来检测。牛津大学研究团队近期提出的语义熵驱动大模型幻觉检测框架<sup>[16]</sup>借鉴信息熵的思想,认为幻觉内容因缺乏真实知识支撑,会呈现较高的语义不确定性,对同一问题的多个回答聚类后计算语义熵,如果熵值较高,说明答案各不相同,可能是幻觉。

为缓解模型幻觉问题,研究者们提出了多种技术,Tonmoy 等学者<sup>[17]</sup>梳理了大量大语言模型幻觉消除技术后将其分为提示工程和模型开发两类。其中,提示工程方法通过调整指令以便从 LLM 中获得最佳输出,包括基于检索增强的方法、基于反馈的自

我完善策略、提示微调方法,该过程可以提供特定的上下文和预期结果,以缓解模型幻觉。模型开发方法通过改进算法和提升数据质量的方式减轻幻觉,包括新的解码策略、基于知识图谱的优化、增加新的损失函数成分和监督微调等。

## 2 大模型赋能图书关系网络建设新范式

### 2.1 整体架构

大模型赋能的图书关系网络建设新范式依托大模型强大的自然语言理解、上下文推理与知识生成能力,是构建图书资源间多维语义关系网络的新型

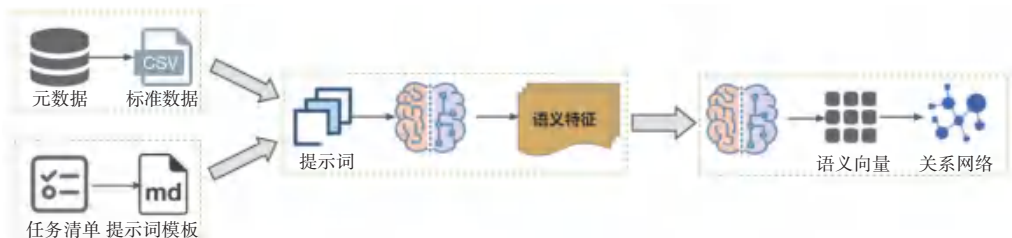


图 1 整体架构

Fig. 1 General framework

### 2.2 大模型赋能图书语义特征增强

在大量自然语言数据和概括任务中训练得到的大语言模型,拥有强大的语言理解和生成能力,即使在没有原文的情况下依然可以根据用户的输入生成摘要,但是这种方式生成的内容可能与事实不符。本研究提出的大模型增强图书语义特征方法(Large Language Model for Enhanced Book Semantic Features, LLM-EBSF)从提示词、模型和生成策略三个途径来提高大模型生成内容的准确性。

#### 2.2.1 提示词模板构建

提示词是用户输入给模型的指令或问题,用来引导模型生成结果。好的提示词不仅可以提升大模型性能,使其更好地满足用户需求;还能提高大模型的可靠性,避免模型幻觉问题的产生。结构化提示词结构层级分明、可读性强,能够帮助用户更轻松地写出性能更好的提示词。基于 LangGPT<sup>[18]</sup> 等结构化提示词框架设计用于图书摘要生成的提示词模板如图 2 所示。在图 2 中,使用 markdown 文件标识符“###”和“[ ]”分别标识标题和变量,标题用来标识层次结构,包括 5 个部分:角色、图书基本信息、任务目标、要求、参考摘要,变量包括书名、作者、中图法分类类别、参考摘要,需要从元数据中获取后替换。

#### 2.2.2 LoRA 参数微调大模型

LoRA (Low-Rank Adaptation)<sup>[19]</sup> 冻结预训练模

知识组织范式,能够提升图书关系网络的语义粒度,发现图书间的隐含关联关系。整体架构如图 1 所示,核心部分包括语义特征增强和关系网络生成。

(1) 语义特征增强:首先对元数据处理得到标准化的数据以及根据任务生成提示词模板,然后用标准化数据替换提示词里的变量得到每个书目的语义特征增强提示词,最后分批次将提示词输入大语言模型生成结果。

(2) 关系网络生成:用大语言模型将增强后的语义特征转化为高维语义向量,基于语义向量计算图书间的相似度,构建关系网络。

型的原始参数,仅在模型关键层旁引入低秩 (Low-Rank) 可训练矩阵,训练效率高、资源消耗低,新增 0.01%~0.50% 参数就能表现出与全参数微调相当的性能。在小样本上进行 LoRA 微调既保证了预训练模型的通用能力,又提高了模型对特定领域任务的适配度。

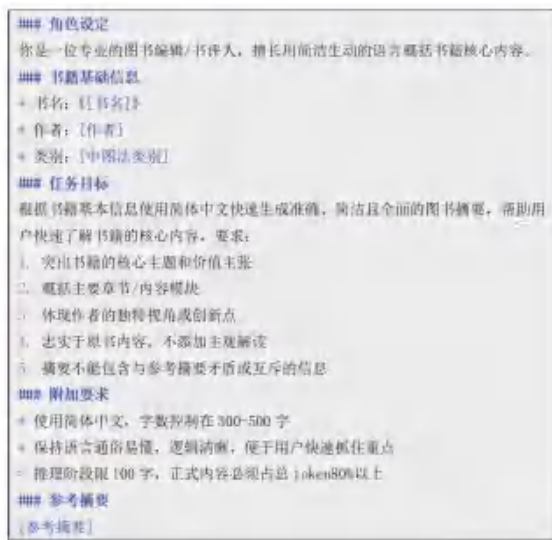


图 2 图书摘要生成的提示词模板

Fig. 2 Prompt template for book summaries generation

研究推得的公式为:

$$y = W^T x + \Delta W^T x = W^T x + (AB)^T x \quad (1)$$

其中,  $\mathbf{W} \in R^{(d \times k)}$  表示固定预训练模型的参数矩阵, 这里  $d$  表示输入维度,  $k$  表示输出维度;  $\Delta\mathbf{W}$  表示训练微调参数, 被分解为2个低秩矩阵  $\mathbf{A} \in R^{(d \times r)}$  和  $\mathbf{B} \in R^{(r \times k)}$ , 这里  $r \ll \min(d, k)$ 。所以显著减少了可训练参数的数量。

### 2.2.3 “自监督”摘要增强生成策略

“自监督”摘要增强生成策略见图3。首先将提示文本经分词器转换为模型可以处理的 token 序列, 然后 encode 编码后输入到大语言模型中生成摘要词向量, 经 decode 解码后转换回人类可读的文本。“自监督”生成策略是在摘要生成后增加备选摘要之间的语义相似性校验, 如果相似性得分  $s$  小于阈值(threshold), 认为内容存在潜在幻觉; 如果此时重试次数  $t$  小于最大重试次数(max\_retries), 大模型按照相同的提示词重新生成备选摘要, 直到相似性得分  $s$  大于等于阈值(threshold) 或者重试次数  $t$ , 等于最大重试次数(max\_retries), 最后挑选相似性得分最大的摘要作为最终摘要。

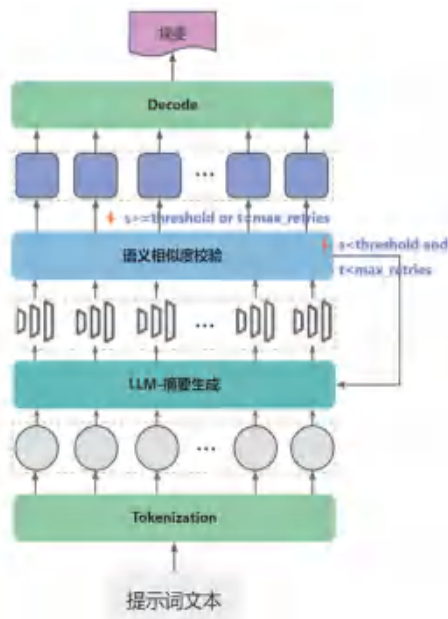


图3 “自监督”摘要增强生成策略

Fig. 3 Strategy for “Self-supervised” summary enhancement generation

在此基础上, 研究推得, 第  $i$  个摘要和第  $j$  个摘要之间的相似度  $s_{ij}$  计算方法可以表示为:

$$s_{ij} = \frac{\mathbf{i} \cdot \mathbf{j}}{\|\mathbf{i}\| \cdot \|\mathbf{j}\|} \quad (2)$$

其中,  $\mathbf{i} = (i_1, i_2, i_n)$  和  $\mathbf{j} = (j_1, j_2, j_n)$  表示语义向量。进一步推得每个摘要的平均相似度  $\bar{s}_i$  计算方法公式为:

$$\bar{s}_i = \frac{1}{n-1} \sum_{j=1, j \neq i}^n s_{ij} \quad (3)$$

## 2.3 大模型赋能图书关系网络建设

自然语言处理(NLP)任务执行的前提是对输入文本编码, 将其转化为模型能够识别的张量, 这些张量包含了文本的语义信息和上下文信息, 能够反映文本的整体特征、内部结构、主题、情感等信息。所以将 Transformer 编码器的最后一层输出或者多层输出进行组合, 得到文本的语义特征向量来计算相似性, 方法如图4所示, 主要包括2个阶段:

(1) 特征提取阶段。利用 BERT/GPT 等预训练语言模型对图书摘要文本进行深度语义编码, 生成具有上下文感知能力的特征向量。

(2) 相似性计算阶段: 通过计算特征向量间的余弦相似度(详见式(1)), 建立图书间的关联关系。

与传统基于词汇相似性计算方法相比, 该方法能够充分考虑词汇的语义关系和上下文语境, 更精准地反映图书之间的内在联系, 为图书关系网络的建设提供更可靠的依据, 有助于挖掘出更有价值的图书关联信息。

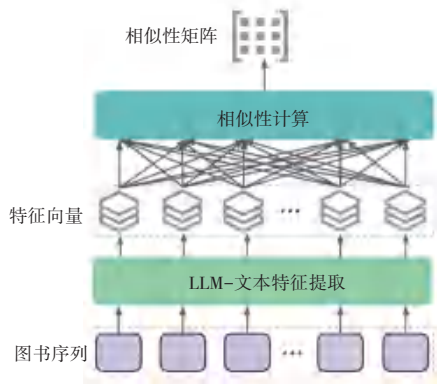


图4 基于大模型的图书关系网络构建方法

Fig. 4 Method for constructing book relationship network based on LLMs

## 3 实验

### 3.1 实验准备

#### 3.1.1 数据说明

原始数据来源于电子科技大学图书馆, 考虑计算成本和时间等因素, 从 60 万条有参考摘要的中文图书中随机抽取 4 万条数据作为样例, 覆盖 22 个中图法分类类别, 字段包括: 书名、作者、参考摘要、中图法分类类别。

#### 3.1.2 评价指标

(1) ROUGE: 通过计算生成文本与参考文本之间

的重叠情况来评估文本生成质量的指标<sup>[20]</sup>, 偏向于 Recall。以 ROUGE - N 为例, 计算方式具体如下:

$$\text{ROUGE} - N = \frac{\sum_{S \in |\text{ReferenceSummaries}|} \sum_{\text{gram}_n \in S} \text{Count}_{\text{match}}(\text{gram}_n)}{\sum_{S \in |\text{ReferenceSummaries}|} \sum_{\text{gram}_n \in S} \text{Count}(\text{gram}_n)} \quad (4)$$

其中, ReferenceSummaries 为样本的一段参考摘要; S 为其中的一个句子; Count<sub>match</sub>(gram<sub>n</sub>) 是生成文本和一组参考文本中共同出现的 n-gram 的最大值。

(2) BERTscore: 是一种基于模型的评测方法<sup>[21]</sup>。通过 BERT 模型提取生成文本和参考文本的词嵌入, 这里分别记为  $C = (c_1, c_2 \dots, c_m)$  和  $R = (r_1, r_2 \dots, r_n)$ , 对于生成文本中的每个词, 找到参考文本中相似度最高的词, 计算平均相似度作为精准度  $P_{\text{BERT}}$ , 推得的数学公式如下:

$$P_{\text{BERT}} = \frac{1}{|C|} \sum_{c_m \in C} \max_{r_n \in R} c_m^T r_n \quad (5)$$

对于参考文本中的每个词, 找到生成文本中相似度最高的词, 计算平均相似度作为召回率  $R_{\text{BERT}}$ , 推得的数学公式如下:

$$R_{\text{BERT}} = \frac{1}{|R|} \sum_{r_n \in R} \max_{c_m \in C} c_m^T r_n \quad (6)$$

BERTScore 中, 还会用到  $P_{\text{BERT}}$  和  $R_{\text{BERT}}$  的调和平均数  $F_{\text{BERT}}$ , 推得的数学公式如下:

$$F_{\text{BERT}} = 2 \frac{P_{\text{BERT}} R_{\text{BERT}}}{P_{\text{BERT}} + R_{\text{BERT}}} \quad (7)$$

### 3.1.3 参数设置

本次实验选择 DeepSeek-R1-Distill-Qwen-32B 模型作为摘要生成预训练模型, bge-large-zh-v1.5 作为文本向量模型。LoRA 微调和摘要生成的部分重要参数设置见表 1, 包括 LoRA 微调预训练模型过程中使用的 Lora 配置参数、训练参数等, 摘要生成 input 参数和 generation 参数的设置, 主要用于控制模型生成结果的可靠性和随机性等。

表 1 实验参数设置

Table 1 Experimental parameter settings

类别	参数	含义	值
LoRA 微调	$r$	低秩矩阵的秩	16
	lora_alpha	缩放因子, 对原始模型权重影响程度	32
	num_train_epochs	模型训练轮数	3
摘要生成	max_length	最大输入词元长度	500
	max_new_tokens	最多额外生成的词元数量	500
	num_generation	同一个 prompt 生成摘要数量	5
	max_retries	最大重试次数	10
	threshold	相似度校验阈值	0.75

## 3.2 实验结果及分析

### 3.2.1 语义特征增强结果

“自监督”摘要增强生成的重试次数和相似性结果分布如图 5 所示。由图 5 可知, 图书摘要任务的首次生成结果展现出良好可信度, 即 80% 以上的任务首次产出的 5 个备选摘要之间的平均相似度大

于 0.75; 若首次生成结果未达标, 最多经过 4 次重试, 99% 以上的任务均能产出可靠结果。最终相似性的结果位于 [0.77, 0.98] 区间内, 以 0.9 为中心向两侧平滑延展, 呈现出典型的正态分布形态, 75% 的样本相似性高于 0.86, 说明大部分样本的相似性处于优良水平, 且离散程度较低, 数据稳定性强。

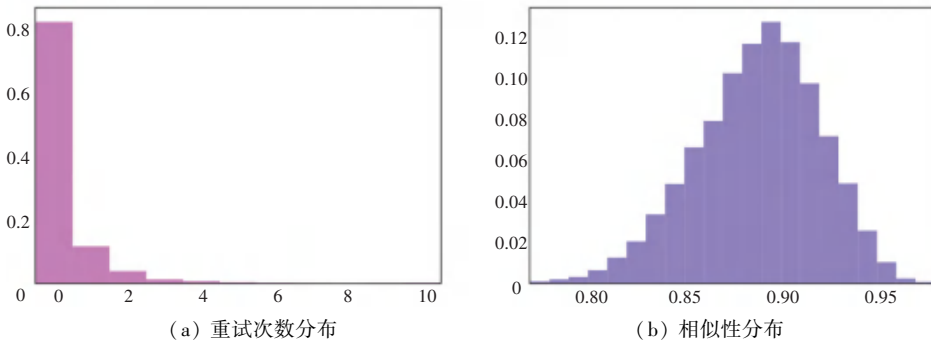


图 5 “自监督”摘要增强生成结果分布

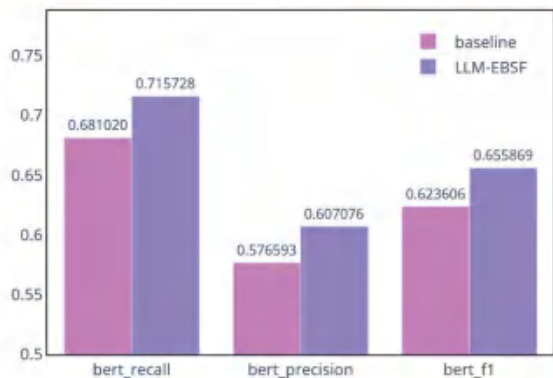
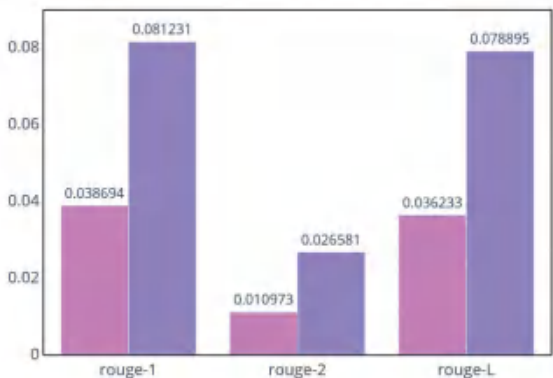
Fig. 5 Distribution of “Self-supervised” summary enhancement generation

图书语义特征增强的准确性结果如图 6 所示。图 6(a)是 ROUGE 得分结果,图 6(b)是 BERT 得分结果,将未优化的预训练模型的输出结果作为 baseline。“LLM-EBSF”方法在 ROUGE 和 BERT 指标上的表现均优于 baseline 模型。

(1) ROUGE 指标: ROUGE-1 (8.12%)、ROUGE-2(2.66%)和 ROUGE-L(7.89%)分别实现

110%、142%和 118%的相对性能提升,表明其在语义覆盖度和文本连贯性上的显著优化。

(2)BERT 指标:  $R_{BERT}$ 、 $P_{BERT}$  和  $F_{BERT}$  得分分别提升至 71.57%、60.71%和 65.59%,较基准方法提升 3.47、3.05 和 3.23 个百分点,验证了其深层语义理解能力的增强。



(a) ROUGE 得分

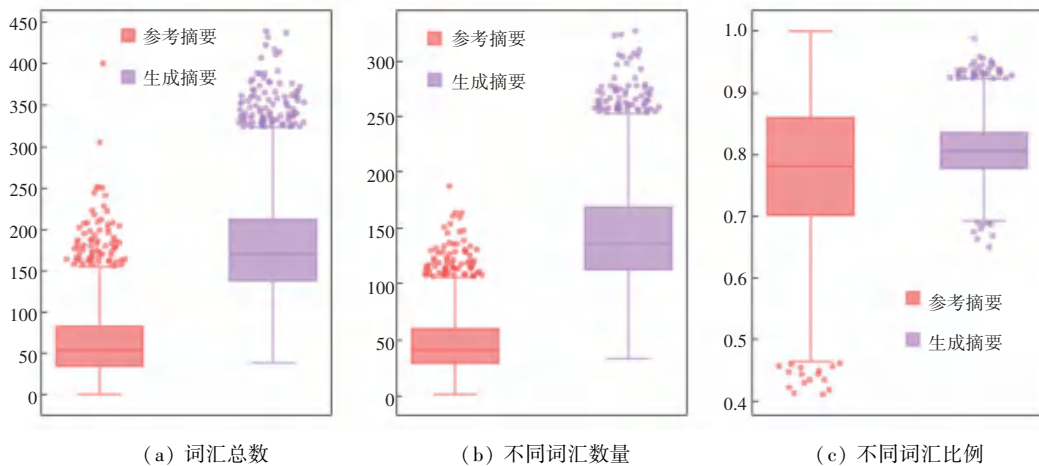
(b) BERT 得分

图 6 语义特征增强的准确性结果

Fig. 6 Accuracy of semantic feature enhancement

图书语义特征增强的多样性结果如图 7 所示。图 7(a)是词汇总数分布结果,图 7(b)是不同词汇数量分布结果,图 7(c)是不同词汇数量占总词汇量的比例分布结果。原始文本经过 jieba 分词器处理后得到词汇量的统计结果。生成摘要在词汇总数和不同词汇数量上均显著高于参考摘要,且分布更为广泛,存在更多的高值离群点。其中,参考摘要的词汇总数集中在 [35, 83] 之间,中位数为 54,不同词汇

数量集中在 [29, 60] 之间,中位数为 41;生成摘要的词汇总数集中在 [138, 212] 之间,中位数 170,不同词汇数量集中在 [113, 169] 之间,中位数为 137。分析可知,生成摘要的不同词汇比例分布更集中,50% 以上的数据汇聚在 0.77~0.84 之间,且中位数值更大(约 0.81)。综上,表明生成摘要在内容丰富性和词汇多样性方面表现更优。



(a) 词汇总数

(b) 不同词汇数量

(c) 不同词汇比例

图 7 语义特征增强的多样性结果

Fig. 7 Diversity of semantic feature enhancement

### 3.2.2 图书语义关系结果

随机选取 4 000 本书来构建关系网络,共产生

7 998 000 个关系对,875 073 个非同类别关系对,  $\langle i, j, \text{similarity} \rangle$  表示一个关系,其中 similarity

表示图书*i*和*j*的语义相关度,如果大于0.75、图书*i*和图书*j*是相关的,大于0.9即为强相关。图书语义关系构建结果见表2,共发现相关关系对93 082个,其中强相关关系对422个;发现同类别相关关系对55 446个,非同类别相关关系对37 636个,后者约为相关关系对的40.4%。与之相比,使用IF-IDF模型能发现的最大相关度仅为0.46,且相关度超过0.2的关系对仅有17组。说明基于大语言模型的方法在关联发现数量和强度上均展现明显优势,能够挖掘出更为丰富的语义关联。

表2 图书语义关系构建结果

Table 2 Results of semantic relationships between books

参数	数值
样本数量	4 000
关系对总数	7 998 000
非同类别关系对数量	875 073
相关关系对数量	93 082
强相关关系对数量	442
同类别相关关系对数量	55 446
非同类别相关关系对数量	37 636

## 4 结束语

本研究提出了一种基于大语言模型的图书关系网络建设新范式,通过语义特征增强与关系网络生成的双路径架构,突破了传统元数据驱动方法在数据维度和语义挖掘深度上的瓶颈,并通过实验证实了该方法的可行性。其核心贡献体现为:

(1)理论价值。本研究为大模型赋能知识网络建设提供了新思路。首先,大模型增强语义特征方法在扩充图书语义特征的同时融合了提示微调、模型微调和自我完善的生成策略等技术,进一步加强了模型的可靠性。其次,基于上下文语义向量构建的图书关系网络可挖掘传统元数据(如分类号、作者、题目等)无法捕获的深层关联。

(2)实践价值。构建的关系网络可应用于图书馆知识服务创新,例如通过可视化图谱揭示学科交叉热点,以及为个性化资源推荐系统提供语义级别的支持。

尽管本研究取得了一定的成果,但仍存在一些局限性。首先,本研究主要基于文本数据构建图书关系网络,对于图表、公式等多模态数据的融合考虑不足。其次,对于一些专业性较强、领域知识复杂的图书,其增强效果可能有限。最后,大模型的训练和生成需要

大量的计算资源和数据支持,这在一定程度上限制了本研究的数据和模型规模。未来将持续从多模态特征扩充、领域知识增强等方面优化研究,同时搭建完整的电子科技大学图书馆图书关系网络图谱,揭示图书之间的关系以及推进个性化资源推荐系统的研发。

## 参考文献

- [1] TOUVRON H, LAVRIL T, IZACARD G, et al. Llama: Open and efficient foundation language models [J]. arXiv preprint arXiv, 2302.13971, 2023.
- [2] BROWN T, MANN B, RYDER N, et al. Language models are few-shot learners [J]. Advances in Neural Information Processing Systems, 2020, 33: 1877-1901.
- [3] 张铮, 刘晨旭. 大模型幻觉: 人机传播中的认知风险与共治可能 [J]. 苏州大学学报(哲学社会科学版), 2024, 45(5): 171-180.
- [4] 周成彬, 刘忠宝. 基于语义信息共享 Transformer 的古文机器翻译方法 [J]. 情报工程, 2022, 8(6): 114-127.
- [5] ENKELEJDA K, KATHRIN S, STEFAN K, et al. ChatGPT for good? On opportunities and challenges of large language models for education [J]. Learning and Individual Differences, 2023, 103: 102274.
- [6] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]// Advances in Neural Information Processing Systems. Long Beach, USA: NIPS Foundation, 2017: 5998-6008.
- [7] 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 (Long and Short Papers). ACL, 2019: 4171-4186.
- [8] RADFORD A, NARASIMHAN K, SALIMANS T, et al. Improving language understanding by generative pre-training [EB/OL]. (2018-01-01). <https://api.semanticscholar.org/CorpusID:49313245>.
- [9] OpenAI. Introducing ChatGPT [EB/OL]. (2022-11-30). <https://openai.com/index/chatgpt/>.
- [10] GUO D, YANG D, ZHANG H, et al. Deepseek - r1: Incentivizing reasoning capability in llms via reinforcement learning [J]. arXiv preprint arXiv, 2501.12948, 2025.
- [11] LIU Qijiong, CHEN Nuo, SAKAI T, et al. A first look at LLM-powered generative news recommendation [J]. arXiv preprint arXiv, 2305.06566, 2023.
- [12] PENG Wenjun, LI Guyang, JIANG Yue, et al. Large language model based long-tail query rewriting in taobao search [C]// Companion Proceedings of the ACM Web Conference. New York: ACM, 2024: 20-28.
- [13] QIU Z, WU X, GAO J, et al. U-BERT: Pre-training user representations for improved recommendation [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(5): 4320-4327.
- [14] JIA Jian, WANG Yipei, LI Yan, et al. Knowledge adaptation from large language model to recommendation for practical industrial application [J]. arXiv preprint arXiv, 2405.03988, 2024.
- [15] 赵月, 何锦雯, 朱申辰, 等. 大语言模型安全现状与挑战 [J]. 计算机科学, 2024, 51(1): 68-71.
- [16] FARQUHAR S, KOSSSEN J, KUHN L, et al. Detecting hallucinations in large language models using semantic entropy

- [J]. *Nature*, 2024, 630: 625–630.
- [17] TONMOY S M, ZAMAN S M, JAIN V, et al. A comprehensive survey of hallucination mitigation techniques in large language models[J]. *arXiv preprint arXiv*, 2401.01313, 2024.
- [18] WANG Ming, LIU Yuanzhong, LIANG Xiaoyu, et al. LangGPT: Rethinking structured reusable prompt design framework for LLMs from the programming language[J]. *arXiv preprint arXiv*, 2402.16929, 2024.
- [19] HU E J, SHEN Yelong, WALLIS P, et al. LORA: Low-rank adaptation of large language models[J]. *arXiv preprint arXiv*, 2106.09685, 2021.
- [20] LIN C Y. Rouge: A package for automatic evaluation of summaries [C]// *Workshop on Text Summarization Branches Out, Post-Conference Workshop of ACL 2004. ACL*, 2004: 74–81.
- [21] ZHANG Tianyi, KISHORE V, WU F, et al. BERTScore: Evaluating text generation with BERT[J]. *arXiv preprint arXiv*, 1904.09765, 2019.