基于大语言模型的知识图谱补全综述

白寅嵩

（渤海大学 计算机科学与人工智能学院，辽宁 锦州 121013）

摘　要：知识图谱是人工智能领域的重要研究方向，其核心任务之一是知识图谱补全。该任务涵盖三元组分类、链接预测和关系预测，旨在填补缺失的头实体、尾实体或关系。随着大语言模型在自然语言处理领域的快速发展，知识图谱补全与大语言模型相结合的研究日益受到关注，涌现出多种能有效处理图谱补全任务的模型。首先，阐述知识图谱补全的基本概念；其次，探讨与大语言模型融合的知识图谱补全模型；再次，介绍基于大语言模型的知识图谱补全模型的常用数据集与评估方法；最后，总结基于大语言模型的知识图谱补全模型的挑战，并对该领域的未来发展方向进行展望。研究成果可为相关研究者提供该领域现状的整体参考，助力后续研究进一步突破现有方法局限，推动知识图谱补全技术的发展与应用。

关键词：大语言模型；知识图谱补全；自然语言处理

DOI：10. 11907/rjdk. 261028　　扫描二维码阅读全文：

中图分类号：TP182　　**文献标识码：**A　　**文章编号：**1672-7800(2026)007-0214-07

Review of Knowledge Graph Completion Methods Based on Large Language Models

BAI Yinsong

(*School of Computer Science and Artificial Intelligence*, *Bohai University*, *Jinzhou* 121013, *China*)

Abstract: Knowledge graph is an important research direction in the field of artificial intelligence, and one of its core tasks is knowledge graph completion. This task encompasses triple classification, link prediction, and relation prediction, aiming to fill in missing head entities, tail entities, or relations. With the rapid development of large language models in the field of natural language processing, research on the integration of knowledge graph completion with large language models has received increasing attention, and various models that can effectively handle graph completion tasks have emerged. Firstly, the basic concept of knowledge graph completion is elaborated; secondly, knowledge graph completion models integrated with large language models are discussed; thirdly, common datasets and evaluation methods for knowledge graph completion models based on large language models are introduced; finally, the challenges of knowledge graph completion models based on large language models are summarized, and future development directions in this field are prospected. The research results can provide a comprehensive reference for the current status of this field for relevant researchers, help subsequent research to further break through the limitations of existing methods, and promote the development and application of knowledge graph completion technology.

Key Words: large language model; knowledge graph completion; natural language processing

0　引言

知识图谱是一种大规模的图结构拓扑知识库，其基本组成单元以三元组形式表示，具体格式为(头实体，关系，尾实体)。例如，在三元组(特朗普，总统，美国)中，“特朗普”是头实体，“美国”是尾实体，“总统”是关系，这种结构清晰地展示了两个实体之间的关联。知识图谱的概念最早由谷歌于2012年提出，推动了语义搜索引擎的发展[1]。其在知识导向型任务中发挥着关键作用，代表性大规模知识图谱包括YAGO[2]、DBpedia[3]、Freebase[4]和WordNet[5]等。知识持续演进更新，但知识图谱的更新往往存在滞后现象。为确保信息的时效性，知识图谱补全技术应运而生。传统知识图谱补全方法在性能提升方面存在明显局限性，而大语言模型在各种学习任务中展现出卓越能力。因此，部分研究者提出将知识图谱与大语言模型相结合，

收稿日期：2026-01-15

作者简介：白寅嵩(1998-)，男，CCF会员，渤海大学计算机科学与人工智能学院硕士研究生，研究方向为知识图谱补全。

以提高知识图谱补全效能,这一领域已取得诸多进展。本文首先明确定义知识图谱补全的核心概念,介绍传统补全模型分类以及代表性实例;其次介绍基于大语言模型的知识图谱补全模型;再次介绍该类模型常用的数据集与评估方法;最后深入分析该类模型面临的挑战以及未来的发展方向,以期为后续相关研究提供清晰的梳理与有价值的参考。

1　传统知识图谱补全方法

在智能问答系统中,一个完整的知识图谱能够提供更全面精准的回答。反之,若图谱存在缺失,系统可能难以作答或提供错误答案。知识图谱补全旨在填补缺失的实体(包括头实体和尾实体)以及它们之间的关系,能够提升知识图谱的完整性与准确性,大体可分为符号补全方法和深度学习补全方法两种。

符号补全方法结合数学领域的思想,通过数学公式来设计得分函数和损失函数,可分为基于表示学习、基于逻辑规则和基于张量分解的方法。其中,基于表示学习的方法将实体和关系映射到一个低维向量空间中,通过在这些向量上进行运算来预测缺失的实体或关系,代表性方法为TransE[6]、TransH[7]等。基于逻辑规则的方法依据从知识图谱中提取的关系来定义规则,并利用这些规则推断需要补全的知识,代表性方法为AMIE[8]等。基于张量分解的方法主要涉及CP分解与Tucker分解两种技术,其使元组以低维矩阵的形式嵌入欧氏空间,以便于链接预测,代表性方法为SimplE[9]等。

深度学习补全方法结合人工智能技术对知识图谱进行补全,可分为基于强化学习、基于神经网络和基于语言模型的方法。其中,基于强化学习的方法通过设计合理的奖励函数来优化智能体的学习过程,其引入额外的奖励信号或调整奖励函数结构,以提升知识图谱补全性能,代表性方法为DeepPath[10]等。基于神经网络的方法将多领域技术融入知识图谱处理中,包括卷积神经网络(Convolutional Neural Network,CNN)[11]、循环神经网络(Recurrent Neural Network,RNN)[12]和图神经网络(Graph Neural Network,GNN)[13]。CNN通过卷积操作自动从知识图谱中提取局部特征。RNN擅长处理序列数据,能有效捕捉数据中的长期依赖关系。专门为处理图结构数据设计的GNN在知识图谱补全中具有独特优势。由于知识图谱本身是图结构,GNN可直接在其上操作,通过聚合相邻节点的特征信息为每个节点(实体)生成更具表现力的嵌入表示[14-15]。基于语言模型的方法利用在大规模语料上训练的语言模型所获取的丰富语义信息与上下文关联来理解实体与关系的含义,代表性方法包括KG-BERT[16]等。预训练语言模型通过海量文本数据训练,积累了丰富的语言知识与语义表示能力,能够学习实体间的内在关联[17-18]。

2　基于大语言模型的知识图谱补全方法

大语言模型是指使用大量文本数据训练的深度学习领域模型,可以理解文本含义或生成文本内容。其与预训练语言模型的主要区别在于模型参数量,大语言模型一般包含数十亿乃至数万亿个参数,具有比预训练语言模型更为复杂的架构,能更好地理解输入指令,生成更为优质的文本。

基于大语言模型的知识图谱补全方法在多个领域都有应用,主要包括:①智能问答系统。基于大语言模型的知识图谱技术可显著增强智能问答与对话系统在理解和回应复杂查询方面的能力,通过补全知识图谱中的信息缺失,能够为问答与对话提供更丰富、更准确的背景信息,从而使系统进行更全面、更高质量的回应;②推荐系统。知识图谱补全结果可丰富推荐系统中用户与物品的特征信息,进而提升推荐的准确性与个性化水平。系统通过挖掘用户与物品之间的潜在关联,能够识别用户的隐性兴趣及物品间的内在联系,最终推送更贴合用户需求的推荐内容。

根据补全流程中的角色功能定位,即大语言模型是直接生成答案,还是辅助排序、推理或与其他模型组件协同工作,基于大语言模型的知识图谱补全方法可划分为端到端生成、推理引导、排序生成以及协同融合4类。

2.1　端到端生成方法

该类方法的核心思想是将知识图谱补全任务直接构建为大语言模型的文本生成问题,利用大语言模型的推理能力,通过设计特殊的指令使其直接生成缺失的答案。例如,Alqaaidi等[19]提出的RPLLM直接将大语言模型作为关系分类器,将实体文本输入模型,并利用大语言模型的词元化器将实体文本转换为数字ID,通过分析模型在这些ID上的输出概率来预测关系。Yao等[20]提出的KG-LLM将知识图谱嵌入大语言模型,并运用指令使模型采用事实性问答的提示范式,通过多轮交互为目标生成答案,从而直接解决实体和关系缺失的问题。Chepurova等[21]提出的KGCT5-with-neighbors模型显式地引入目标实体的局部邻域结构信息,将目标实体及其邻居实体共同作为输入上下文,使大语言模型在生成答案时不仅能利用语义信息,而且能感知图谱的局部拓扑结构,从而提升对复杂图谱结构的补全性能。

2.2　推理引导方法

该类方法利用大语言模型的复杂推理能力来分解有一定难度的知识图谱补全任务,尤其适用于需要多步逻辑推理的场景。大语言模型并不直接输出最终答案,而是先生成推理链或中间证据,从而提升补全过程的可靠性和可解释性。例如,Luo等[22]提出的Chain of History方法将大语言模型与时序知识图谱相结合,通过构建结构化的历史

信息(如模式匹配历史、实体增强历史等)作为提示,引导大语言模型进行时序推理,有效解决了动态关系中实体状态演变的推理难题。Yuan 等[23]利用大语言模型进行规则挖掘,从大量候选三元组中识别出最可能成立的一组事实,这对于补全知识图谱中的关联簇非常有效。

2.3 排序生成方法

该类方法结合了传统知识图谱嵌入模型的高效排序能力与大语言模型的深层语义理解能力,通常采用“检索—排序—生成”的模式,兼顾了效率与精度。首先,传统模型从大规模候选实体中筛选出 Top-K 个最相关的实体;其次,大语言模型基于语义上下文对这些候选实体进行精排并生成最终答案。例如,Liu 等[24]提出的 DIFT 方法利用在知识图谱上预训练的传统嵌入模型来微调大语言模型,使大语言模型能够吸收图谱的结构化语义知识,从而提升其实体补全和关系推理能力。Li 等[25]提出的 CATS 模型首先利用实体类型等元信息对候选答案进行过滤,其次提取并编码目标实体的局部子图,使大语言模型基于丰富的上下文信息进行最终决策,增强了推理的准确性。Li 等[26]提出的 KGR3 模型包含检索、推理和重排序 3 个模块,检索模块保证效率,推理模块负责复杂推理,重排序模块整合前两步信息进行最终决策,系统性解决了大语言模型在处理结构化数据时的局限性。

2.4 协同融合方法

该类方法旨在深度整合文本语义与图谱结构信息,使大语言模型的通用知识与知识图谱的精确结构化知识相互增强,以获得最优补全性能。例如,Zhang 等[27]提出的知识前缀适配器 Kopa 通过结构预训练提取实体与关系的结构信息以生成嵌入表示,随后将提取的信息作为序列输入大语言模型,实现了结构信息与文本信息的深层耦合。Yang 等[28]提出的 GS-KGC 采用问答格式设计子图提取策略,有效处理了一对多、多对多关系等复杂场景。Madhushani 等[29]利用上下文学习与拓扑信息来提升大语言模型在知识图谱补全任务上的表现。Wei 等[30]提出的 KICGPT 将大语言模型与传统的结构感知模型相结合,缓解了知识图谱中长尾实体和关系的数据稀疏性问题。Zheng 等[31]提出的 KG-CF 将知识图谱中的多步推理路径转化为可被大语言模型理解的文本序列,并利用大语言模型的语义理解能力来筛选和评估这些路径,从而实现对缺失三元组的预测。

基于大语言模型的知识图谱补全方法综合比较见表 1。

Table 1 Knowledge graph completion methods based on large language models

表 1 基于大语言模型的知识图谱补全方法

类型	模型	优点	缺点	应用场景
端到端生成	RPLLM[19] KG-LLM[20] KGCT5-with-neighbors[21]	简单灵活,零样本/少样本能力强	事实幻觉,输出不稳定	面向普通用户信息检索应用
推理引导	Chain of History[22] Yuan 等[23]	可处理复杂逻辑	推理链的可靠性和可控性	金融、医疗、司法等领域
排序生成	DIFT[24] CATS[25] KGR3[26]	判别精准	效率仍有瓶颈	企业级知识平台构建
协同融合	Kopa[27] GS-KGC[28] Sehwag 等[29] KICGPT[30] KG-CF[31]	事实一致性高	计算成本高	专业领域知识库

3 传统方法与基于大语言模型方法的区别

传统知识图谱补全方法与基于大语言模型的知识图谱补全方法的区别有以下 4 点:①核心思想的区别。传统方法遵循符号主义和连接主义路径,其假设知识隐藏在图谱的结构中。例如,TransE 模型认为关系向量应近似于头实体向量到尾实体向量的平移。该类方法高度依赖图谱本身的结构完整性。基于大语言模型的方法则更偏向于经验主义。大语言模型在预训练阶段便从非结构化文本中吸收了海量人类知识,包括实体间的隐含关系和常识。因此,知识图谱补全任务可以转化为使大语言模型理解和生成与这些知识相关的文本问题;②数据来源的区别。传统方法仅以图谱作为来源,无法利用图谱之外的文本信息,如果某个实体或关系在图中连接稀疏(长尾问题),模型便很难学习到其有效的表示,导致性能下降。基于大语言模型的方法则具有多重数据来源,即使图谱中的三元组不完整,其亦能通过从新闻、百科等文本中学到的知识推理出事实,从而补全缺失链接,因而对稀疏图谱和新兴实体有更好的鲁棒性;③优势的区别。传统方法计算高效且模型参数量小,推理速度快,可解释性较强。特别是基于规则的方法推理路径清晰,严格基于已有图谱事实进行推断,不易编造信息。基于大语言模型的方法具有强大的语义理解能力,能处理同义词、复杂关系和隐含常识,且零样

本/少样本学习能力强,无需大量特定训练数据即可执行任务,同一模型可适应多种不同关系的预测;④缺点的区别。传统方法的主要缺点为模型规模小,难以处理大规模数据和复杂关系,有些模型还面临同义词和长尾等问题。基于大语言模型的方法的缺点在于容易产生幻觉,出现编造信息等问题。

4　常用数据集与评估标准

4.1　常用数据集

4.1.1　静态数据集

静态知识图谱数据集包括FB15K、FB15K-237、WN11和WN18RR,具体信息见表2。其中,FB15K数据集源自FreeBase知识图谱,包含14 951个实体、1 345种关系和592 213个三元组,涵盖人物、地点、影视作品等多个领域。该数据集中存在大量可逆三元组,导致划分训练集、测试集和验证集时易产生信息泄漏,从而影响模型性能的准确评估。为应对这些挑战,FB15K-237数据集作为子集被提取出来,其包含14 541个实体、237种关系和310 116个三元组。 WN11是基于WordNet构建的知识图谱数据集,其一个大型英语词汇语义网络,通过同义关系、上下位关系等语义关系组织词汇,包含38 096个实体、11种关系、125 744个三元组。 由于WN18在训练过程中存在与FB15K类似的问题(尤其在处理反向关系时),为提升模型的泛化能力构建WN18RR数据集,其为WN18的子集,包含40 943个实体、11种关系和93 003个三元组。

Table 2　Information of static knowledge graph datasets

表2　静态数据集信息

数据集	实体数	关系数	训练集边数	验证集边数	测试集边数
FB15K	14 951	1 345	483 142	50 000	59 071
FB15K-237	14 541	237	272 115	17 535	20 466
WN11	38 096	11	112 581	2 609	10 544
WN18RR	40 493	11	86 835	3 034	3 134

4.1.2　动态数据集

与静态数据集不同,动态数据集在三元组数据的基础上加入时间戳。该类数据集包括GDELT、ICEWS14和ICEWS05-15,具体信息如表3所示。GDELT数据集涵盖地缘政治与地缘社会事件的时间流变信息,包含事件属性及事件间的关联。ICEWS是基于新闻报道及其他公开信息构建的国际事件数据集,ICEWS14为该数据集的特定版本/子集,包含政治、军事、经济、社会等多领域国际事件的详细信息,提供了事件发生时间、地点以及参与方等具体数据。ICEWS05-15同样源自于ICEWS数据库,涵盖了2005-2015年间的国际事件数据,包含该时段内政治、经济、社会等多领域的大量信息。

在动态数据集中,时间戳指某个事件发生的特定时间点,其为一个三元组打上了精确的时间标记,记录了该关系在哪个时间点成立或有效。时间跨度指某个事实或关系持续有效的整个时间段。模型在动态数据集的评估主要围绕引入时间约束的预测任务展开,给定一个不完整的四元组,要求模型在特定时间点T下预测出缺失的尾实体,从而测试其能否在正确的时间上下文内理解关系;或给定头实体、尾实体和时间戳,使模型预测它们之间的关系。此外,动态数据集还用于评估更复杂的时序推理能力,检验模型是否能理解事件的继承性和互斥性。例如,某一事件在时间T1开始,T2结束,模型应能推理出在T2之后该关系失效,评估重点在于模型是否会生成时间上矛盾的元组。时序模式推断与预测任务还要求模型基于历史事件序列预测未来可能发生的事件。例如,结合某地区在一段时间内紧张关系升级的历史,模型能否预测未来发生“军事冲突”或“和平谈判”的可能性?这评估了模型在时序模式中进行外推的能力。

Table 3　Information of dynamic knowledge graph datasets

表3　动态数据集信息

数据集	实体数	关系数	时间戳数	时间戳跨度	训练集边数	验证集边数	测试集边数
GDELT	500	20	366	15 min	2 735 685	341 961	341 961
ICEWS14	7 128	230	365	24 h	72 826	8 941	8 963
ICEWS05-15	10 488	251	4017	24 h	386 962	46 275	46 092

采用动态数据集对基于大语言模型的知识图谱补全方法进行评估能直接暴露其固有局限:首先,该类数据集可以精准评估大语言模型的静态知识截止日期如何影响其对新近事件的认知。例如,采用ICEWS05-15数据集中2015年的数据测试一个基于2020年数据训练的大语言模型可以直接衡量其动态适应性;其次,该类数据集可以评估大语言模型是真正理解时序逻辑,还是仅基于训练数据中的文本共现进行猜测。例如,模型能否正确推理出事件发生的先后顺序及其因果性,而非简单匹配关键词。

4.2　评估标准

知识图谱补全任务的评估指标主要包括平均倒数排名(Mean Reciprocal Rank MRR)、平均排名(Mean Rank MR)与Hits@k。MRR为核心评估指标,其首先计算每个测试样本中正确答案在预测结果中的倒数排名;其次对所有样本的倒数排名取平均值。计算公式为:

$$MRR = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{rank_i} \tag{1}$$

式中:N为测试集样本总数;$rank_i$表示第i个查询中正确答案的预测排名。

MR指在知识图谱补全任务中,所有测试样本的正确答案在模型预测结果中的平均排名。计算公式为:

$$MR = \frac{1}{N}\sum_{i=1}^{N}rank_i \tag{2}$$

Hits@k用于衡量正确答案出现在预测结果前k位的样本比例。计算公式为:

$$Hits@k = \frac{1}{N}\sum_{i=1}^{N}I(rank_i \leqslant k) \tag{3}$$

式中：$I()$为指示函数，当条件满足时取值为1，否则为0。通常在k取1、3、10时进行多维度评估。

5 面临的挑战以及应对策略

5.1 生成结果的准确性与可靠性问题

在知识图谱补全过程中，大语言模型生成的结果可能包含噪声或错误信息，该类误差会影响模型的判断与推理，导致补全结果不准确。若训练过程中使用了包含错误信息的文本，大语言模型可能会无意中将这类错误信息融入知识图谱补全结果中。大语言模型在生成补全结果时还有可能会虚构事实，这种现象被称为“幻觉”。在缺乏准确支持信息的情况下，模型可能生成看似合理但实际错误的内容。在知识图谱补全任务中，大语言模型可能编造不存在的实体细节或关系，这在需要严格准确知识的应用场景中会带来显著风险。例如，大语言模型在补全“谁提出了万有引力定律”时，可能因其训练数据中的间接关联或错误组合而生成“哥白尼提出了万有引力定律”这一错误答案，这种幻觉对下游应用的危害是直接的。在智能问答系统中，如果基于被污染的知识图谱进行回答，系统可能会向用户提供完全错误的信息，严重损害系统的可信度。在推荐系统中，虚假的三元组可能导致错误的关联推荐，严重影响用户体验和商业收益。对此，可以利用检索增强生成（Retrieval Augmented Generation RAG）等技术先让模型从一个庞大的信息文档库中检索对应信息，再利用这些信息指导结果生成，以提升结果的准确度和可靠性。此外，还可以要求模型说明思考过程以及生成结果的依据。

5.2 知识更新滞后与动态适应性局限

现实世界中的知识持续演变，为保持准确性与时效性，知识图谱必须定期更新。然而，大语言模型的训练与更新过程复杂，需要大量计算资源和时间，这导致基于大语言模型的知识图谱补全任务在知识更新上存在显著滞后。此外，大语言模型在适应动态演化的知识图谱方面存在明显局限。当知识图谱的结构或内容发生变化时，这些模型通常需要进行重新训练以适应新的配置，这无疑增加了大量时间与资源成本。因此，基于大语言模型的知识图谱补全方法在应对该类动态变化时面临挑战，最终可能影响知识图谱的实际应用效果。

时序知识图谱的核心特征为时间戳信息，记录了事实在特定时间点或时间段内的有效性。新生事件层出不穷，但大语言模型的训练数据具有静态的截止日期，对于该日期之后出现的新实体、新关系或已有实体的新状态缺乏认知。即使获得了新知识，如何将其高效地注入大语言模型也是一个巨大挑战。时序知识图谱要求更新机制能够快速、频繁地集成新数据。对于ICEWS、GDELT这类每日甚至每小时都在产生新事件的数据集，如果每次更新都需要对大语言模型进行数天甚至数周的全量微调，那么更新速度将远远落后于事件发生的速度，使图谱失去时效性价值，这种高昂的成本也使持续学习和增量更新变得困难。此外，大语言模型在学习新知识时可能发生“灾难性遗忘”，即记住了新信息，却丢失了旧知识，这对于需要长期历史上下文进行推理的时序任务来说是致命的。时序知识图谱补全不仅是填充新事实，更关键的是要理解事件随时间演变的逻辑和模式。大语言模型在预训练时主要学习语言统计规律，而非显式的时序逻辑，因此在理解和推理严格的时序依赖关系方面存在不足。

在动态适应性方面，持续学习方法可在一定程度上缓解灾难性遗忘问题，未来可将其应用于知识图谱补全中。例如，将外部数据源（新闻API、学术数据库、企业文档流）定义为智能体的交互环境。智能体不再被动等待训练数据，而是主动感知环境变化，通过定期扫描数据源，利用轻量化变化检测模型来识别可能与现有知识图谱相关的新实体、新关系或新事实。将被感知到的变化转换为模型可处理的结构化（头实体，关系，尾实体，时间）候选四元组，并存入一个体验缓冲区，以待后续更新时使用。

5.3 计算资源与时间成本高昂

大语言模型的训练与推理需要庞大的计算资源支持，需依赖大规模GPU集群，并消耗大量电力。面对Freebase、WikiData等包含巨量实体（如Freebase包含超过2.5亿个实体）的大规模知识图谱，即使是进行简单的知识图谱补全链接预测推理任务，也需要对海量候选实体进行生成或排序。多次调用大语言模型会导致惊人的计算开销和极长的响应时间，难以在需要快速响应的生产环境中进行规模化部署。对于计算资源有限的研究机构或中小型企业而言，从头训练或对大语言模型进行有效微调均极为困难。

未来研究应致力于设计更轻量的模型架构，研究基于混合专家（Mixture of Experts，MoE）的大语言模型架构用于知识图谱补全[32]。在MoE模型中，对于每个输入，只有少部分专家（即网络子模块）被激活进行计算。这种稀疏激活特性非常适合知识图谱补全任务，因为不同类型的关系（如“出生于”和“首席执行官”）可以由不同的专家网络进行处理，从而在保持模型总体容量（参数量）的同时显著降低单个推理的计算开销。亦可研究如何通过知识蒸馏技术将大语言模型中蕴含的复杂语义知识和推理能力迁移到更小、更高效的专用学生模型中。例如，训练一个轻量化学生模型，使其在知识图谱补全任务中的输出与大型教师大语言模型的输出尽可能接近，从而实现资源受限环境下的部署。大语言模型自身不用进行全参数微调，而是仅训练一个附着其上的小型适配器来学习新知识，这将极大降低计算和存储成本，使频繁更新成为可能。

5.4 可解释性与透明度缺失

大语言模型在知识图谱补全中的推理与决策过程通常缺乏可解释性。该类模型本质上如同“黑箱”，难以解释

其如何得出特定的补全结果。例如,在进行医学知识图谱补全时,大语言模型生成一个医学判断,人们无从得知这个判断是基于一篇严谨的学术论文,还是源于某个未被证实的网络论坛帖子,这种可解释性的缺失影响了知识图谱补全结果的评估与优化。由于研究者无法完全理解大语言模型的决策过程,改进模型、提升补全结果的准确性与可靠性变得尤为困难。在数据隐私与安全性要求严格的场景中,这种透明度的缺乏会引发更多担忧,因为其难以确保大语言模型在补全过程中符合相关规范与标准。

可通过神经—符号结合的路径提升模型的可解释性。例如,利用大语言模型的语义理解能力从文本中生成或验证符号逻辑规则,然后将这些人类可读的规则作为约束,引导或验证大语言模型的补全结果,这能为模型的决策提供明确的推理路径,显著提升可解释性。亦可发展事后解释技术,为大语言模型的补全决策生成归因图或关键证据,阐明是知识图谱中的哪些已有三元组或文本描述对当前预测贡献最大。

5.5　伦理问题及对策

大语言模型的训练数据源自互联网文本,其中不可避免地包含现实社会中的各种偏见(如性别、种族、地域、职业等)。在知识图谱补全过程中,模型可能无意间学习并放大这些偏见,从而在补全的知识图谱中系统性强化刻板印象和歧视性关联。后续可开发专门的偏见检测工具包,针对知识图谱补全任务构建涵盖不同社会维度的测试基准,定期对模型的补全结果进行系统性偏见审计。在模型训练或微调阶段引入去偏见技术,如对抗性学习,使模型学习消除敏感属性信息,减少生成结果中的偏见内容。研究并实施约束生成技术,将公平性、中立性等伦理规范作为硬约束或软目标融入模型生成过程中,确保输出结果符合伦理要求。

大语言模型的“黑箱”特性使补全结果出现错误或造成损害时难以追溯和界定责任。为解决这一问题,可将伦理准则明确嵌入知识图谱补全系统的设计与开发全流程中,建立从需求分析、设计、实现到部署的伦理检查清单。大力推动模型的可解释性研究,要求关键领域的知识图谱补全系统必须提供其补全决策的推理依据或证据,增强透明度和可信度。鼓励学术界、工业界、政策制定者和社会组织共同参与,建立针对人工智能生成知识的行业标准、认证体系和监管框架,明确各方的责任与义务。

6　结语

大语言模型为知识图谱补全领域提供了新的发展动力,其弥补了传统知识图谱补全技术的一些不足,促使新模型性能更加优异,也将是未来一段时间知识图谱补全领域的主要发展趋势。当然,该领域还面临着诸多挑战。相信随着模型架构的持续演进与训练策略的不断优化,知识图谱补全技术与大语言模型将在未来实现更深层次的有机融合,推动知识图谱朝着更完整、更准确、更动态的方向演进。

参考文献:

[1] MAHDISOLTANI F, BIEGA J, SUCHANEK F M. A knowledge base from multilingual Wikipedias-yago3[R]. France: Telecom ParisTech, 2014.

[2] GERHARD W. A core of semantic knowledge unifying wordnet and wikipedia[C]//Proceedings of the 16th International Conference on World Wide Web, 2007: 697-706.

[3] AUER S, BIZER C, KOBILAROV G, et al. DBpedia: a nucleus for a web of open data[C]//The Semantic Web, 2007: 722-735.

[4] BOLLACKER K, EVANS C, PARITOSH P, et al. Freebase: a collaboratively created graph database for structuring human knowledge[C]//Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data, 2008: 1247-1250.

[5] FELLBAUM C. WordNet: an electronic lexical database[M]. Cambridge: MIT Press, 1998: 678-686.

[6] BORDES A, USUNIER N, GARCIA-DURAN A, et al. Translating embeddings for modeling multi-relational data[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems, 2013: 2787- 2795.

[7] WANG Z, ZHANG J, FENG J, et al. Knowledge graph embedding by translating on hyperplanes[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2014: 1112-1119.

[8] GALÁRRAGA L A, TEFLIOUDI C, HOSE K, et al. AMIE: association rule mining under incomplete evidence in ontological knowledge bases[C]//Proceedings of the 22nd International Conference on World Wide Web, 2013: 413-422.

[9] KAZEMI S M, POOLE D. SimplE embedding for link prediction in knowledge graphs[DB/OL]. https://arxiv.org/abs/1802.04868.

[10] XIONG W, HOANG T, WANG W Y. DeepPath: a reinforcement learning method for knowledge graph reasoning[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 2017: 564-573.

[11] KIM Y. Convolutional neural networks for sentence classification[DB/OL]. https://arxiv.org/abs/1408.5882.

[12] ELMAN J L. Finding structure in time[J]. Cognitive Science, 1990, 14(2): 179-211.

[13] BENJAMÍN S L, REIF E, PEARCE A, et al. A gentle introduction to graph neural networks[EB/OL]. https://distill.pub/2021/gnn-intro/.

[14] DETTMERS T, MINERVINI P, STENETORP P, et al. Convolutional 2D knowledge graph embeddings[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2018: 1811-1818.

[15] SCHLICHTKRULL M, KIPF T N, BLOEM P, et al. Modeling relational data with graph convolutional networks[C]//European Semantic Web Conference, 2018: 593-607.

[16] YAO L, MAO C, LUO Y. KG-BERT: BERT for knowledge graph completion[DB/OL]. https://arxiv.org/abs/1909.03193.

[17] ZHANG Z, HAN X, LIU Z, et al. ERNIE: enhanced language representa-

tion with informative entities [DB/OL]. https://arxiv. org/abs/1905. 07129.

[18] LIU W, ZHOU P, ZHAO Z, et al. K-bert: enabling language representation with knowledge graph[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2020: 2901-2908.

[19] ALQAAIDI S K, KOCHUT K J. Relations prediction in knowledge graph completion using large language models[C]//Proceedings of the 2024 8th International Conference on Information System and Data Mining, 2024: 122-127.

[20] YAO L, PENG J, MAO C, et al. Exploring large language models for knowledge graph completion [DB/OL]. https://arxiv. org/abs/2308. 13916.

[21] CHEPUROVA A, BULATOV A, KURATOV Y, et al. Better together: enhancing generative knowledge graph completion with language models and neighborhood information[C]//Findings of the Association for Computational Linguistics, 2023: 5306-5316.

[22] LUO R, GU T, LI H, et al. Chain of history: learning and forecasting with LLMs for temporal knowledge graph completion[DB/OL]. https://arxiv. org/abs/2401. 06072.

[23] YUAN Y, XU Y, ZHANG W. Is large language model good at triple set prediction? an empirical study[C]//2024 IEEE International Conference on Knowledge Graph, 2024: 470-476.

[24] LIU Y, TIAN X, SUN Z, et al. Finetuning generative large language models with discrimination instructions for knowledge graph completion [C]//International Semantic Web Conference, 2024: 199-217.

[25] LI M, YANG C, XU C, et al. Context-aware inductive knowledge graph completion with latent type constraints and subgraph reasoning[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2025: 12102-12111.

[26] LI M, YANG C, XU C, et al. Retrieval, reasoning, re-ranking: a context-enriched framework for knowledge graph completion[C]//Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies, 2025: 4349-4363.

[27] ZHANG Y, CHEN Z, GUO L, et al. Making large language models perform better in knowledge graph completion[C]//Proceedings of the 32nd ACM International Conference on Multimedia, 2024: 233-242.

[28] YANG R, ZHU J, MAN J, et al. Exploiting large language models capabilities for question answer-driven knowledge graph completion across static and temporal domains[DB/OL]. https://arxiv. org/html/2408. 10819v1.

[29] MADHUSHANI S U, PAPASOTIRIOU K, VANN J, et al. In-context learning with topological information for knowledge graph completion[DB/OL]. https://arxiv. org/abs/2412. 08742.

[30] WEI Y, HUANG Q, ZHANG Y, et al. Kicgpt: large language model with knowledge in context for knowledge graph completion[C]//Findings of the Association for Computational Linguistics: EMNLP 2023, 2023: 8667-8683.

[31] ZHENG Z, DONG Y, WANG S, et al. KG-CF: knowledge graph completion with context filtering under the guidance of large language models [DB/OL]. https://arxiv. org/abs/2501. 02711.

[32] SHAZEER N, MIRHOSEINI A, MAZIARZ K, et al. Outrageously large neural networks: the sparsely-gated mixture-of-experts layer[DB/OL]. https://arxiv. org/abs/1701. 06538.

（责任编辑：尹晨茹）