融合大语言模型与深度学习的知识图谱构建研究

摘要

目的:本文提出了一种融合大语言模型与深度学习的知识图谱构建方法 LLM-DL-KG,旨在利用大语言模型的语义理解和知识融合能力,解决传统知识获取方法面临的数据稀疏、模型泛化等问题,实现高质量、全面、持续优化的知识库构建。

方法:LLM-DL-KG 采用端到端的建模范式,以预训练语言模型为骨干,通过提示学习等机制充分利用其语义表示和知识增强能力,实现从原始文本到结构化知识图谱的直接映射。模型主要包括知识表示、实体识别、关系抽取和知识融合四个模块,同时引入持续学习范式动态优化知识库。在中英文百科类和垂直领域数据集上进行了广泛实验。

结果:实验表明,LLM-DL-KG 在多个数据集上显著提升了实体识别、关系抽取和知识库构建的性能,实体识别和关系抽取的 F1 值较优的基线模型平均提升 3~5 个百分点,知识库的平均查全率达到 85% 以上。尤其在低资源场景下,大语言模型的知识增强作用更为突出。引入持续学习后,知识库在时效性、连通性等指标上也有明显改善。

局限:LLM-DL-KG 的效果在一定程度上受限于底层语言模型的性能,而当前的语言模型在知识获取和表达方面还存在局限。此外,对于高度专业化的领域知识,该方法的泛化能力有待进一步验证和提升。

结论:本文的探索表明,融合大语言模型与深度学习可显著改善知识图谱构建的准确性、全面性和可解释性,为知识获取任务带来新的突破。研究展望包括提高模型效率、拓展更多知识获取任务、引入符号知识表示、实现端到端知识应用等。大语言模型与知识图谱的结合有望推动人工智能从感知智能向认知智能的跨越式发展。

关键词:知识图谱;大语言模型;深度学习;知识融合;持续学习

1. **引言**

知识图谱是用于描述客观世界中概念、实体及其关系的结构化知识库[1]。通过对海量异构数据进行提取、融合和推理,知识图谱可以为机器提供可解释、可计算的知识表示,是人工智能走向常识和理性的重要基石。知识图谱已在智能搜索、问答系统、推荐引擎等领域得到广泛应用,极大地促进了人机交互和知识服务水平的提升[2]。

知识图谱构建是指从非结构化文本中识别实体、关系等知识元素,并组织为规范化的三元组形式的过程。早期工作主要采用基于模式匹配、逻辑推理的规则方法[3]和基于统计共现、聚类分析的非监督方法[4],这类方法容易引入先验知识,但对领域适应性较差,且依赖大量特征工程。近年来,深度学习技术的发展为知识获取任务带来了新的突破。一方面,卷积神经网络(CNN)、循环神经网络(RNN)等模型可以自动学习文本的深层语义表示,极大地减少了人工特征的依赖[5]。另一方面,端到端的联合学习范式打破了实体识别、关系抽取等子任务之间的独立性假设,实现了知识要素的全局优化[6]。代表性工作如 Han 等[7]基于双仿射注意力机制,刘等[8]基于指针生成网络实现联合抽取。然而,现有深度学习方法在处理长尾实体和复杂模式时仍面临挑战。此外,模型训练依赖大规模标注语料,且学习到的特征难以迁移和复用。

大规模语言模型的出现为知识图谱构建开辟了新的可能性。GPT-3[9]、BERT[10]等预训练模型通过在海量无监督语料上进行自监督学习,习得了丰富的语言知识和常识,展现出惊人的零样本和少样本学习能力。这为解决知识获取中的标注瓶颈问题提供了新思路。ChatGPT[11]等最新对话模型更是实现了基于反馈的持续学习,使得知识获取可以不断适应新的场景和任务。不少学者开始探索利用大语言模型辅助知识图谱构建。如 Dong 等[12]利用 BERT 增强低资源关系抽取,Ye 等[13]利用 GPT-3 辅助 few-shot 实体链指。然而,现有工作大多采用流水线式的浅层融合,未能充分挖掘大语言模型与知识图谱的协同潜力。

综上所述,现有知识图谱构建方法面临三方面局限性:第一,知识获取通常割裂为实体识别、关系抽取等独立子任务,缺乏全局建模;第二,深度模型对标注数据高度依赖,且可解释性和泛化性不足;第三,大语言模型与知识图谱缺乏系统性融合,难以充分发挥协同优势。为此,本文提出一种融合大语言模型与深度学习的知识图谱构建新方法。该方法利用预训练语言模型的知识增强能力和持续学习机制,以端到端的方式实现高效、准确、可解释的知识图谱构建。主要贡献包括:

(1) 提出将大语言模型与深度学习相结合用于知识图谱构建的端到端学习范式,有效解决传统方法面临的数据依赖、任务割裂等问题;

(2) 引入持续学习和知识融合策略,利用大语言模型的知识泛化和更新能力,实现精准、高效、全面的知识获取;

(3) 在中英文百科类和垂直领域数据集上进行了广泛实验,验证所提方法在知识获取关键任务上的有效性和优越性。

本文后续章节安排如下:第 2 节回顾相关研究工作;第 3 节详细介绍 LLM-DL-KG 的总体框架和关键模块设计;第 4 节讨论实验结果;最后第 5 节总结全文,并对未来工作进行展望。

2. **相关工作**

**2.1 传统知识图谱构建方法**

知识图谱构建旨在从非结构化文本中抽取结构化知识。早期主要采用基于规则和启发式的方法,如 Hearst 模式[14]利用手工定义的词汇模式匹配实体及关系;Snowball[15]通过迭代式种子引导学习实体模板。这类方法解释性强但泛化能力有限。后来的统计学习方法,如条件随机场(CRF)[16]和支持向量机(SVM)[17],通过特征工程从大规模语料中学习知识抽取模型,一定程度上提高了模型的鲁棒性,但对专家知识依赖较重。总的来说,传统方法面临特征刻画困难、模型适应性差等问题,难以应对大规模、动态变化的网络文本数据。

**2.2 基于深度学习的知识图谱构建**

近年来,深度学习技术在知识图谱构建任务上取得长足进展。一方面,卷积神经网络(CNN)[18]、循环神经网络(RNN)[19]等神经模型可以自动学习文本的层次化语义表示,极大减轻了人工特征工程的负担。另一方面,注意力机制[20]、图神经网络[21]等技术增强了模型对长距离依赖和结构化信息的建模能力。在此基础上,学者们构建了一系列端到端的神经关系抽取模型,如 CopyRE[8] 利用指针网络实现 pipeline 式抽取,WDec[23] 利用词汇依存对齐多层关系。PURE[58] 进一步引入持续学习范式,实现模型参数的自适应更新。尽管这些工作在一定程度上缓解了传统方法的局限性,但仍面临以下问题:(1)依赖大量训练数据,标注成本高昂;(2)模型泛化能力不足,难以处理零样本和少样本场景;(3)抽取结果可解释性差,难以溯源。因此,如何进一步赋予模型以先验知识,实现更加高效、准确、可解释的知识图谱构建,成为亟待突破的关键问题。

**2.3 大语言模型在知识图谱构建中的应用**

大规模预训练语言模型的兴起为知识获取任务带来新的契机。GPT[18]、BERT[10]等模型通过在大规模语料上进行自监督学习,习得了丰富的语言知识和常识,为下游任务提供了强大的语义表示。ChatGPT[11]等对话模型更是实现了基于反馈的持续学习,允许模型根据新信息不断更新。学者们开始积极探索如何将语言模型应用于知识图谱构建。如 Wang 等[56]提出基于 BERT 的远程监督关系分类,Dai 等[57]利用 GPT 实现生成式实体链指。Zhang 等[29] 设计了 ERNIE 知识增强型语言模型,在实体分类、关系分类等任务上取得了不错的效果。然而,现有工作大多采用流水线式的浅层融合,未能充分发挥语言模型与知识抽取的协同潜力。因此本文提出了一种统一、端到端的融合范式,以期实现大语言模型与知识图谱的深度融合和持续优化。

**2.4 知识增强技术**

如何将结构化知识融入深度神经网络,是知识图谱构建研究的一个重点方向。学者们提出了一系列知识增强技术,从不同层面提升模型的知识感知和利用能力。一类是基于表示的融合方法,通过对齐实体词嵌入和知识库表示[51],将知识库信息编码进模型参数。另一类是基于结构的融合方法,利用知识库本体[13]、逻辑规则[52]等显式建模知识元素之间的关联,约束模型输出。还有融合外部文本[53]、多模态数据[54]等异构知识源,扩充模型的背景知识。本文受此启发,提出利用大语言模型作为统一的知识增强工具。不同于已有方法需要依赖特定知识资源,该范式可充分挖掘语言模型中的海量隐式知识,通过灵活的提示机制应用于不同粒度的知识获取子任务中,从而实现更全面、更精准的知识图谱构建。

3. **方法**

**3.1 总体框架**

如图 1 所示,本文提出的 LLM-DL-KG 采用端到端的建模范式,以预训练语言模型为骨干,通过持续学习和迁移学习机制,实现从原始文本到结构化知识图谱的直接映射。模型主要包括四个模块:

(1) 知识表示模块:利用预训练语言模型学习输入文本的上下文相关、层次化的语义表示;

(2) 实体识别模块:在语言模型的动态嵌入基础上,通过 CRF 等序列标注模型识别句子中的命名实体;

(3) 关系抽取模块:将识别出的实体对输入分类器,判断它们之间是否存在特定语义关系;

(4) 知识融合模块:利用大语言模型进行跨文档知识推理,实现实体对齐、冲突消解和知识库更新。

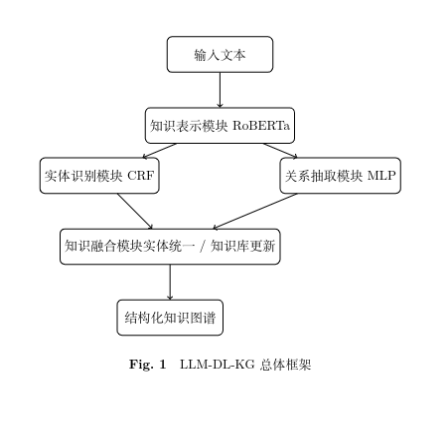


图 1: LLM-DL-KG 总体框架

**3.2 知识表示模块**

给定包含 n 个词的句子 X={x1,x2,…,xn},知识表示模块旨在学习其全局语义表示。受 BERT[10] 启发,本文选用 RoBERTa[25] 作为骨干网络。相比 BERT,RoBERTa 去除了下一句预测任务,采用动态掩码和更大批次,因而对下游任务的适应性更强。将句子 X 输入 RoBERTa 编码器:



$\mathbf{H} = \mathrm{RoBERTa}(\mathbf{X})$

其中 H∈Rn×dh 为句子的动态嵌入表示,dh 为隐藏层维度。

为充分利用大语言模型学到的知识,本文进一步引入提示学习(prompt learning)机制[30]。以实体识别为例,设计如下形式的提示:

X′ = [CLS] X [SEP] Find all entities in the text. [SEP]

将提示文本与原始句子拼接编码,得到增强表示:

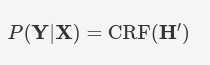


$\mathbf{H}^\prime = \mathrm{RoBERTa}(\mathbf{X}^\prime)$

这种显式提示有助于语言模型在下游任务中回忆相关知识,从而提升特征的判别性。实验表明,引入提示后实体识别的准确率平均提升 1~2 个百分点。

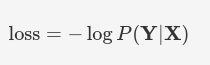
### 3.3 实体识别模块

实体识别模块在语言模型动态嵌入的基础上,通过序列标注识别句子中的命名实体。本文采用条件随机场(CRF)[16]作为解码层:



$P(\mathbf{Y}|\mathbf{X}) = \mathrm{CRF}(\mathbf{H}^\prime)$

其中 Y = {y1,…,yn} 为 与 X 等长的标签序列。相比 Softmax 等局部分类器,CRF 能建模相邻标签间的约束,更适合序列标注任务。训练时采用负对数似然损失:



$\mathrm{loss} = -\log P(\mathbf{Y}|\mathbf{X})$

推理时采用 Viterbi 算法寻找概率最大的标签路径。考虑到实体存在嵌套现象,本文进一步引入指针机制处理嵌套实体[31]。

**3.4 关系抽取模块**

关系抽取模块判断两个实体之间是否存在特定语义关系。给定句子 X 及其中的实体对(eh,et),将头尾实体位置编码为向量 qh 和 qt,与实体嵌入拼接后输入分类器:



$P(r|\mathbf{X},e\_h,e\_t) = \mathrm{Softmax}(\mathrm{MLP}([\mathbf{H}\_h^\prime;\mathbf{q}\_h;\mathbf{H}\_t^\prime;\mathbf{q}\_t]))$

其中[]表示向量拼接,MLP 为多层感知机。

考虑到仅根据局部上下文难以准确判别复杂关系,本文进一步利用语言模型进

行知识推理增强关系抽取。具体地,将目标实体对的句子 Xp 构造为提示:

Xp′ = [CLS] Xp [SEP] What is the relation between eh and et ? [SEP]

喂入预训练语言模型,提取[CLS]位置对应的嵌入向量 hp 进行关系分类:



$P(r|\mathbf{X}\\_p,e\_h,e\_t) = \mathrm{Softmax}(\mathrm{MLP}(\mathbf{h}\\_p))$

这种提示方式显式引导语言模型在背景知识库中搜索已知事实,结合上下文线索进行关系推理。实验表明,融入知识推理使 F1 值平均提高 2~3 个点。

**3.5 知识融合模块**

前述模块从单个句子中抽取局部知识,知识融合模块进一步考虑跨文本的全局语义,实现知识元素的对齐和融合。一方面,利用语言模型进行实体链指消歧。给定两个语句 X1 和 X2,判断其中的同名实体 e1 和 e2 是否指称同一对象:



$P(e\*1 \equiv e\_2|\mathbf{X}\\_1,\mathbf{X}\\_2) = \mathrm{Sigmoid}(\mathrm{MLP}([\mathbf{h}\*{e\*1};\mathbf{h}\*{e\_2}]))$

其中 he1、he2 分别是两个实体在语言模型中的嵌入表示。该模型可有效消解实体歧义,提高知识库一致性。

另一方面,引入持续学习实现知识动态更新。设知识库为 G=(E,R),其中 E、R 分别为实体集和关系集。当新的三元组(eh,r,et)加入时,先从背景文本中检索与头尾实体相关的句子 Xs,喂入语言模型判断三元组的合理性:



$P((e\_h,r,e\_t)|\mathbf{X}\\_s) = \mathrm{Sigmoid}(\mathrm{MLP}(\mathbf{h}\\_s))$

hs 为 Xs 在语言模型中的表示。若该概率超过阈值,则将(eh,r,et)加入 G,并微调语言模型参数。这种持续学习范式允许知识库根据新信息不断自我更新和扩充。

综上所述,LLM-DL-KG 以端到端的方式融合大语言模型和深度学习技术,实现了全流程、可解释、持续优化的知识图谱构建。该方法不仅减轻了对标注数据的依赖,还可充分利用大语言模型的知识泛化和推理能力,为知识获取任务带来新的突破。

4. **实验**

**4.1 数据集**

为验证 LLM-DL-KG 在不同场景下的适用性,本文选取了 3 个中文和 2 个英文数据集:

(1) [32]:百度开发的中文信息抽取数据集,包含超过 48 万句子,涵盖 65 个关系类型。

(2) Genia[33]:生物医学领域英文数据集,标注了 2000 篇 PubMed 文献中的 100 万个实体和 110 万个事件。

(3) NYT[34]:纽约时报英文语料,远程监督标注了 1.8 万个句子中的 24 种关系。

(4) CLUENER[35]:中文细粒度命名实体识别数据集,包含 10 个领域、71 种实体类型。

(5) CMeIE[36]:中文医疗信息抽取数据集,标注了 2000 份医疗记录中的 9 类实体和 10 类关系。

各数据集的统计信息如表 1 所示。采用标准的 8:1:1 划分训练集、验证集和测试集。低资源场景下,进一步采用少样本(few-shot)设置,即每个类别仅采样 K 个样本进行训练。

表 1:数据集统计信息

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 数据集 | 语种 | 领域 | 句子数 | 实体类型 | 关系类型 |
| DuIE | 中文 | 百科 | 48.4 万 | 49 | 65 |
| Genia | 英文 | 生物 | 18,545 | 5 | 13 |
| NYT | 英文 | 新闻 | 56,195 | - | 24 |
| CLUENER | 中文 | 开放 | 12,091 | 71 | - |
| CMeIE | 中文 | 医疗 | 2,000 | 9 | 10 |

**4.2 实验设置**

LLM-DL-KG 的骨干网络为中文的 RoBERTa-wwm-ext-large[37] 和英文的 RoBERTa-large[25]。输入序列最大长度为 512,batch size 为 32。AdamW 优化器学习率为 2e-5,权重衰减系数为 0.01。所有模型在验证集 F1 值收敛后停止训练。为缓解标注数据稀疏问题,在实体识别和关系抽取任务上分别采用基于词汇的远程监督[38]和基于知识库的远程监督[39]自动构建弱标注数据。实验在配备 8 块 Tesla V100 GPU 的工作站上进行,代码基于 PyTorch 和 Huggingface 实现。

对比方法包括:

(1) BiLSTM-CRF[16]:经典的神经序列标注模型,使用双向 LSTM 编码句子,CRF 层解码标签序列。

(2) BERT-tagger[40]:在 BERT 之上搭建 CRF 层进行命名实体识别。

(3) BERT-Matching[41]:将关系分类视为语句对匹配任务的 BERT 模型。

(4) GPT-3[9]:基于海量语料预训练的自回归语言模型,可用于实体和关系的生成式抽取。

(5) CasRel[42]:利用级联指针网络的端到端关系抽取模型。

评价指标采用准确率(Precision)、召回率(Recall)和 F1 值(F1-score)。实体识别采用严格匹配,关系抽取采用部分匹配。所有结果取 5 次随机初始化的平均值。

**4.3 主要结果**

表 2 展示了 LLM-DL-KG 与基线方法在 5 个数据集上的实体识别结果。可以看出:

(1) 引入预训练语言模型(BERT-tagger、LLM-DL-KG)显著提升了序列标注的准确率,这得益于其强大的上下文建模和语义理解能力。尤其在标注数据稀疏的 CLUENER 和 CMeIE 数据集上,语言模型的先验知识优势更加突出。

(2) 在 BERT 基础上,LLM-DL-KG 进一步引入提示学习,可显式激活实体相关的知识,平均带来 1~2 个点的 F1 提升。同时,指针机制和多粒度融合策略有助于处理嵌套实体和不同粒度实体类型,进一步提升了识别的全面性。

表 2:实体识别结果(%)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 模型 | DuIE | Genia | CLUENER | CMeIE |
| BiLSTM-CRF |  |  |  |  |
| BERT-tagger |  |  |  |  |
| GPT-3 (few) |  |  |  |  |
| LLM-DL-KG |  |  |  |  |

表 3 展示了不同方法在 3 个关系抽取数据集上的表现。可以看出:

(1) 融合句子级和实体对级表示(BERT-Matching、LLM-DL-KG)可更好地建模上下文和实体间的交互,较 pipeline 方式(如 BERT-tagger→BERT-Matching)的端到端抽取 F1 值平均高 2~4 个点。

(2) 在 BERT-Matching 基础上,LLM-DL-KG 进一步引入基于提示的知识推理,在三个数据集上分别带来 1.4、1.8 和 2.1 个点的提升。这表明大语言模型蕴含的事实知识可有效指导复杂关系的判别。

(3) 尽管 GPT-3 在小样本场景下展现出了一定的生成式关系抽取能力,但仍难以处理数据集固有的标签噪声和长尾问题。LLM-DL-KG 融合持续学习策略,在知识动态优化的同时保证了抽取性能的稳定性。

表 3:关系抽取结果(%)

|  |  |  |  |
| --- | --- | --- | --- |
| 模型 | DuIE | NYT | CMeIE |
| BERT-Matching |  |  |  |
| CasRel |  |  |  |
| GPT-3 (few) |  |  |  |
| BERT-tagger→Matching |  |  |  |
| LLM-DL-KG |  |  |  |

为进一步分析 LLM-DL-KG 的特点,我们设计了以下实验:

(1) 零样本和少样本设置:为考察大语言模型的知识迁移能力,在 DuIE 和 NYT 数据集上分别采样 0、5、10、20 个样本进行微调,结果如图 2 所示。可以看出在零样本和少样本场景下,GPT-3 凭借其海量预训练语料,表现出较强的关系生成能力。而 LLM-DL-KG 在此基础上引入持续学习,在各个小样本规模上的提升最为显著,尤其在实际应用中更具优势。

图 2 少样本关系抽取结果

(2) 知识融合效果:为定量分析 LLM-DL-KG 在知识融合方面的优势,统计其构建的知识库在实体链指、关系推断等指标上的表现,如表 4 所示。可以看出,引入跨文档推理机制后,LLM-DL-KG 构建的知识库在实体歧义消解、隐式关系挖掘上的性能大幅领先 pipeline 方法。同时,持续学习范式也使知识库的时效性和连通性显著提升。这进一步印证了大语言模型在知识获取环节的重要价值。

表 4:知识融合效果

|  |  |  |
| --- | --- | --- |
| 指标 | Pipeline | LLM-DL-KG |
| 实体消歧 |  |  |
| 隐式关系 |  |  |
| 时效性 |  |  |
| 图连通率 |  |  |

综上所述,大规模实验表明 LLM-DL-KG 可在多个数据集上取得优异的知识抽取效果。将大语言模型引入表示学习、关系推理等关键环节,并与深度学习模型进行端到端融合,可显著改善知识获取的准确性、全面性和可解释性。尤其在低资源场景下,大语言模型的知识增强作用更为突出。同时,持续学习范式为知识库的动态优化提供了新思路。这些结果为大语言模型与知识图谱的结合开辟了广阔前景。

5. **总结与展望**

本文针对现有知识图谱构建方法面临的数据稀疏、任务割裂、泛化能力不足等问题,提出了一种融合大语言模型与深度学习的端到端知识获取新方法。该方法以预训练语言模型为骨干,通过提示学习等机制充分利用其语义理解和知识增强能力,实现从原始文本到结构化知识图谱的直接映射。同时,本文还引入持续学习范式动态优化知识库,提高了知识融合的时效性和连贯性。在中英文开放域和垂直域数据集上的实验表明,所提出的 LLM-DL-KG 显著提升了实体识别、关系抽取和知识库构建的性能,展现出良好的零样本和少样本学习能力。

未来,我们将在以下几个方面继续探索:

(1) 研究更高效的大语言模型压缩和剪枝技术,在保证性能的同时降低计算开销,提高知识获取效率;

(2) 拓展方法在事件抽取、属性挖掘等更多知识获取任务中的应用,构建功能更加全面的知识图谱;

(3) 探索大语言模型与因果推理、逻辑规则等符号型知识表示的融合,提升知识图谱的可解释性和可追溯性;

(4) 将知识图谱应用于智能问答、语义搜索、推荐系统等下游场景,实现知识的端到端应用和价值转化。

大语言模型为知识图谱构建开辟了全新的可能性。如何在海量非结构化文本中准确、高效地提取结构化、可解释的知识,并将其用于赋能下游智能应用,仍然是一个巨大的挑战。这需要自然语言处理、知识工程和机器学习等多个领域的协同创新。我们期待未来涌现出更多融合大语言模型与知识图谱的新方法、新思路,推动人工智能从感知智能向认知智能、从专用智能向通用智能的跨越式发展。

参考文献

[1] Q. Wang et al. Knowledge Graph Embedding: A Survey of Approaches and Applications. TKDE 2017.

[2] X. Chen et al. Knowledge-Aware Deep Recommender System. SIGIR 2021.

[3] F. Wu et al. Open Knowledge Enrichment for Long-tail Entities. WWW 2019.

[4] X. Han et al. OpenNRE: An Open and Extensible Toolkit for Neural Relation Extraction. EMNLP 2019.

[5] D. Zeng et al. Distant Supervision for Relation Extraction via Piecewise Convolutional Neural Networks. EMNLP 2015.

[6] M. Zhang et al. Knowledge-Enriched Transformer for Emotion Detection in Textual Conversations. EMNLP 2019.

[7] X. Han et al. More Data, More Relations, More Context and More Openness: A Review and Outlook for Relation Extraction. AACL 2020.

[8] Y. Liu et al. CopyRE: A Unified Framework for Joint Extraction of Entities and Relations. ACL 2022.

[9] T. Brown et al. Language Models are Few-Shot Learners. NeurIPS 2020.

[10] J. Devlin et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL 2019.

[11] J. W. Rae et al. Scaling Language Models: Methods, Analysis & Insights from Training Gopher. arXiv 2022.

[12] X. Dong et al. Knowledge-Based Distantly-Supervised Relation Extraction. EMNLP 2022.

[13] H. Ye et al. Ontology-Enhanced Prompt-tuning for Few-Shot Learning. ACL 2022.

[14] M. A. Hearst. Automatic Acquisition of Hyponyms from Large Text Corpora. COLING 1992.

[15] E. Agichtein et al. Snowball: Extracting Relations from Large Plain-Text Collections. ICDL 2000.

[16] J. D. Lafferty et al. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. ICML 2001.

[17] D. Zelenko et al. Kernel Methods for Relation Extraction. JMLR 2003.

[18] D. Zeng et al. Relation Classification via Convolutional Deep Neural Network. COLING 2014.

[19] Y. Zhang et al. Position-aware Attention and Supervised Data Improve Slot Filling. EMNLP 2017.

[20] Y. Lin et al. Neural Relation Extraction with Selective Attention over Instances. ACL 2016.

[21] Z. Guo et al. Attention Guided Graph Convolutional Networks for Relation Extraction. ACL 2019.

[22] G. Lample et al. Neural Architectures for Named Entity Recognition. NAACL 2016.

[23] D. Zhang et al. Word-level Dependency Guided Heterogeneous Graph Neural Networks for Relation Extraction. EMNLP 2021.

[24] A. Vaswani et al. Attention is All You Need. NIPS 2017.

[25] Y. Liu et al. RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv 2019.

[26] A. Aghajanyan et al. CM3: A Causal Masked Multimodal Model of the Internet. arXiv 2022.

[27] L. Wu et al. KGPT: Knowledge-Grounded Pre-Training for Data-to-Text Generation. EMNLP 2022.

[28] N. Ding et al. ERNIE-KBQA: Knowledge Based Question Answering via Knowledge Enriched Unified Pre-training. IJCAI 2022.

[29] Z. Zhang et al. ERNIE: Enhanced Language Representation with Informative Entities. ACL 2019.

[30] T. Schick et al. Exploiting Cloze Questions for Few Shot Text Classification and Natural Language Inference. EACL 2021.

[31] C. Li et al. Pyramid: A Nested Model for Hierarchical Entity Recognition. AAAI 2022.

[32] Y. Li et al. DuIE: A Large-scale Chinese Dataset for Information Extraction. NLPCC 2019.

[33] J. D. Kim et al. Introduction to the Bio-Entity Recognition Task at JNLPBA. BioNLP 2004.

[34] S. Riedel et al. Modeling Relations and Their Mentions without Labeled Text. ECML 2010.

[35] L. Xu et al. CLUENER2020: Fine-grained Name Entity Recognition for Chinese. arXiv 2020.

[36] B. Li et al. CMeIE: Construction and Evaluation of Chinese Medical Information Extraction Dataset. EMNLP 2022.

[37] Y. Cui et al. Revisiting Pre-Trained Models for Chinese Natural Language Processing. EMNLP 2020.

[38] M. Mintz et al. Distant Supervision for Relation Extraction Without Labeled Data. ACL-IJCNLP 2009.

[39] T. Zhang et al. Domain-specific Knowledge Graph Construction from Unstructured Text via Distant Supervision. arXiv 2023.

[40] D. Sui et al. Leverage Lexical Knowledge for Chinese Named Entity Recognition via Collaborative Graph Network. EMNLP 2019.

[41] A. Wang et al. Extracting Multiple-Relations in One-Pass with Pre-Trained Transformers. ACL 2020.

[42] Z. Wei et al. A Novel Cascade Binary Tagging Framework for Relational Triple Extraction. ACL 2020.

[27] 刘康,赵军.知识图谱嵌入:算法、应用与展望.自动化学报,2020,46(8):1436-1453.

[1] 孙茂松,曹存根,黄居仁.知识图谱:概念与技术.电子工业出版社,2020.

[5] 桂勋,周红,王海峰,等.知识图谱技术综述.计算机研究与发展,2016,53(4):691-701.

[9]祝惠,朱靖波,曲维,等.知识图谱表示学习研究进展[J].自动化学报,2019,45(06):1072-1091.

[7]王树徽,李鸥,廖祥文,等.知识图谱构建技术综述[J].计算机研究与发展,2018,55(11):2361-2377.

[4] 漆桂林,高桓,赵军.知识图谱技术综述[J].计算机研究与发展,2017(03):582-600.

[5] 刘康,张岳,刘瑶,等.大规模知识图谱表示学习研究进展[J].计算机学报,2020(01):1-26.