

融合外部知识知识图谱问答方法研究

白云天, 郝文宁, 靳大尉, 刘小语

(陆军工程大学 指挥控制工程学院, 江苏 南京 210014)

摘要: 知识图谱问答是自然语言处理领域的热门研究方向之一。现有方法主要存在两大挑战: 一是难以理解复杂的自然语言形式问题, 二是实体表示通常只限于字面含义, 缺乏深入的语义阐释。针对上述问题, 提出一种融合外部知识知识图谱问答方法 DEK-KGQA。首先通过问题知识图谱子图和 QA 上下文构建联合图, 其次利用预训练语言模型计算联合图中节点的相关性评分, 最后引入外部知识, 以增强问答推理过程中的信息交互和推理能力。在 CommonsenseQA 数据集上进行实验验证, 并与现有方法进行比较。实验结果表明, 该方法在常识问答任务中取得了更好的效果, 验证了该方法的有效性。此外, 通过消融实验验证了该方法中各个部分对整体性能的影响。

关键词: 知识图谱问答; QA 上下文; 预训练语言模型; 外部知识

DOI: 10.11907/rjdk.241368

中图分类号: TP391.1

文献标识码: A

开放科学(资源服务)标识码(OSID):

文章编号: 1672-7800(2024)009-0056-07



Study on Knowledge Graph Question Answering Methods Incorporating External Knowledge

BAI Yuntian, HAO Wenning, JIN Dawei, LIU Xiaoyu

(College of Command & Control Engineering, Army Engineering University of PLA, Nanjing 210014, China)

Abstract: Knowledge graph question answering is one of the hot research areas in the field of natural language processing. Existing methods face two main challenges: difficulty in understanding complex natural language questions and limited semantic interpretation of entity representations. To address these challenges, a knowledge graph question answering method called DEK-KGQA is proposed, which integrates external knowledge. First, a joint graph is constructed by combining the question knowledge graph subgraph and the QA context. Then, the relevance scores of nodes in the joint graph are calculated using pre-trained language models. Finally, external knowledge is introduced to enhance information interaction and reasoning ability during the question answering process. Experimental validation is conducted on the CommonsenseQA dataset, comparing the proposed method with existing methods. The results demonstrate that the proposed method achieves better performance in commonsense question answering tasks, validating its effectiveness. In addition, ablation experiments are conducted to evaluate the impact of each component on the overall performance.

Key Words: knowledge graph question answering; QA context; pre-trained language model; external knowledge

0 引言

近年来,知识图谱已成为自然语言处理领域的一个研究热点。知识图谱是一个大规模、高维度的语义网络,其中每个节点代表一个具体的实体或概念,每条边代表两个

实体之间的某种关系。现有的知识图谱包括 Wikidata^[1]、NELL^[2]、Freebase^[3]、Dbpedia^[4]和 YAGO^[5]等。知识图谱问答是指以自然语言形式的问题作为输入,利用知识图谱中的结构化知识给出答案^[6]。

基于知识图谱问答的方法通常分为两种:语义解析和信息检索。语义解析方法的关键在于将非结构化的自然

收稿日期:2024-04-13

扫描二维码阅读全文:

基金项目:国防工业技术发展计划项目(JCKY2020601B018)

作者简介: 白云天(2000-),男,陆军工程大学指挥控制工程学院硕士研究生,研究方向为自然语言处理;郝文宁(1971-),男,博士,陆军工程大学指挥控制工程学院教授、博士生导师,研究方向为大数据技术及应用、数据挖掘和深度学习;靳大尉(1979-),男,硕士,陆军工程大学指挥控制工程学院教授,研究方向为大数据技术及应用;刘小语(2001-),女,陆军工程大学指挥控制工程学院学生,研究方向为军事大数据工程。本文通讯作者:郝文宁。



语言问题转换成计算机可以理解的结构化查询语句形式,然后与知识图谱中的关系、实体进行匹配,得出最终答案。该方法需要先对自然语言问题进行命名实体识别、实体消歧等操作,其次将其转化为结构化查询语句,再与知识图谱中的数据进行匹配。然而,如果实体未能正确识别和对齐,可能导致查询错误,进而影响回答的准确性。信息检索方法的关键在于将自然语言问题与知识图谱中的实体和关系转化为特征向量,并使用相似度匹配的方式检索和排序答案。该方法通过将问题和知识库中的实体与关系映射到一个低维空间中,实现了问题与知识图谱的语义匹配,从而能够有效地找到与问题相关的信息,并生成相应的答案。该方法可以更好地捕捉语义信息。然而,由于实体节点通常由简短的词汇构成,缺乏丰富的语义信息,导致在实体匹配时往往只能进行简单的字符串匹配,难以捕捉实体之间的深层次语义关系,从而影响了答案的准确性。

因此,为了提高知识图谱问答的准确度,本文综合运用语义解析和信息检索方法,提出DEK-KGQA(Descriptive Knowledge for Knowledge Graph Question Answering)方法。首先用预训练语言模型对问题一选项对进行特征提取,以捕获问题与答案之间的语义关系和上下文信息,然后结合知识图谱构建联合图,利用预训练语言模型计算联合图中每个三元组(头实体、关系、尾实体)的相关性分数,并选择得分最高的三元组。此外,引入外部知识库中的知识来增强实体的语义理解和具体解释。最后,将问题、选项、选择的三元组和定义输入到预训练语言模型albert中,计算每个选项作为正确答案的概率。

1 相关工作

传统基于规则的语义解析方法通过预定义的语法和语义规则将自然语言转换成机器可理解的形式。Hakimov等^[7]使用基于依赖关系解析树的推理和建议生成方法,通过迭代局部搜索、结合实体链接和查询构造的方法,生成问题的语义表示和对应的查询语句。Unger等^[8]、Walter等^[9]使用基于词汇树邻接语法的语义解析方法解释用户查询,并生成对应的SPARQL查询,从而从知识图谱中检索相关信息。但上述方法在处理大规模数据集和复杂查询时存在一些限制。为了解决此问题,Yih等^[10]提出使用查询图来解析问题,并将其与知识库的图结构进行匹配。查询图是一种图形化表示,与知识库的图形结构类似,可以更好地表示复杂的多跳和约束问题。在对候选解析结果的评分上,Zhu等^[11]使用树状编码器对候选查询结构进行建模,并编码实体和关系在查询中的顺序,以更好地区分给定问题的候选查询。然而,上述方法未能很好地捕捉到句子中的语义关系,导致生成的查询结果不够准确。

基于信息检索的方法通常使用关键词匹配和相关性

排序来检索与用户输入问题相关的知识图谱实体、属性及关系。例如,使用模块化双向注意力记忆网络(BAMnet)来完成知识图谱问答任务^[12]。该方法将问题和知识图谱编码为一个共同的嵌入空间,通过问题实体与候选答案实体之间的答案路径和子图的表示,获取与问题最相关的信息。由于该方法仅考虑了与问题主题实体直接相关的候选答案实体,对于涉及多跳推理的问题,可能无法捕捉到答案所需的间接关系。为了解决此问题,研究者们从知识库中提取相关信息,构建子图进行推理^[13]。然而,推理路径可能非常复杂,正确路径可能不在子图中,从而导致推理困难和子图的不完整性。为了提高推理效果,Xu等^[14]提出DEKCOR模型,利用完整知识图谱和从Wikipedia等文本语料中提取与问题相关的信息来回答常识问题,通过从知识图谱中检索相关观念的描述,并将文本信息作为上下文信息融合到实体表示中,以获得更精确的常识推理。Yasunaga等^[15]构建异构图,将语言模型和知识图谱相结合,通过基于注意力机制的图神经网络(Graph Neural Networks, GNN)多轮消息传递捕捉节点之间的关系。Sun等^[16]使用双向注意力模块融合来自语言模型和图神经网络的编码表示,并设计了一个剪枝模块从子图中删除无关节点。Wang等^[17]设计了一种交互方案,通过初始化的交互表示,使语言模型和图神经网络的每一层都可以进行双向信息交换。Choi等^[18]提出一种称为掩蔽实体模型分类机制的方法,用于预测知识图谱中三元组的掩蔽实体,以解决知识图谱不完整的问题。

本文综合考虑上述方法,提出融合外部知识的知识图谱问答方法DEK-KGQA。按照Feng等^[19]的方法获取知识图谱子图,加入问题中的相关实体以构建联合图,通过语言模型获取问题实体与知识图谱子图中各个实体的相关程度,以获取得分最高的三元组(头实体,关系,尾实体),并参考文献[14]中引入外部知识的方法,将其作为预训练语言模型的额外输入。通过对外部知识的引入,可以扩展模型的知识范围,丰富模型对知识的理解。

2 方法实现

2.1 相关概念及问题描述

知识图谱 G 是一个有向图 (V, E) ,其中 V 表示实体的集合, E 表示关系的集合。对于每个关系 $r \in E$,存在一对有序实体 $(h, t) \in V \times V$,称为三元组,其中 h 是头实体, t 是尾实体,一个三元组可以表示为 (h, r, t) 。

给定一个自然语言问题 q 、一个知识图谱 G 和一个候选答案集合 $A = \{a_1, a_2, \dots, a_n\}$,知识图谱问答任务是找到一个答案 $a \in A$,使得 a 是 q 在 G 中的正确答案。

表1给出了本文方法用到的符号及其表示。

2.2 方法概述

DEK-KGQA架构如图1所示。对于某个问题 q 及其某

Table 1 Symbols and their descriptions in this paper
表1 本文符号及其描述

符号	描述
G	知识图谱
V	实体集合
E	关系集合
h	某个三元组中头实体
r	某个三元组中关系
t	某个三元组中尾实体
q	自然语言问题
a	答案选项
g	知识图谱子图
c	QA上下文节点
$V_{q,a}$	主题实体
V_q	问题实体
V_a	答案实体
U	联合图
$r_{c,q}$	c与某一 V_q 之间的关系
$r_{c,a}$	c与某一 V_a 之间的关系
V_{sub}	G中实体

个答案选项 a , 可以拼接为一个 QA 上下文节点 c 的初始信息 $[q:a]$ 。基于文献[15]的方法, 首先利用预训练语言模型对 c 进行特征提取, 得到其初始化表示, 并针对 q 从知识图谱中检索出子图 g ; 接着, 将 QA 上下文节点 c 和 g 中的问题实体、选项实体相连, 形成一个联合图 U 。为了动态地捕捉 QA 上下文节点 c 和联合图 U 中所有其他节点的关系, 本文使用预训练语言模型计算每一对节点的相关性分数, 并将该得分作为每个节点的额外特征, 由此从知识图谱中获得得分最高的三元组。在找到所有选项的最优三元组后, 基于信息检索方式, 从外部知识库中得到问题实体、选项实体和三元组实体的描述信息, 将问题选项对、问题描述、选项描述、三元组和三元组描述拼接成一个输入序列, 并在输入序列起始位置添加 [CLS] 标记, 每个输入文本的末尾添加 [SEP] 标记, 最终将完整序列输入到 albert 模型, 利用基于注意力的加权求和以及 softmax 层产生一个表示该选项是正确答案可能性的分数^[20]。

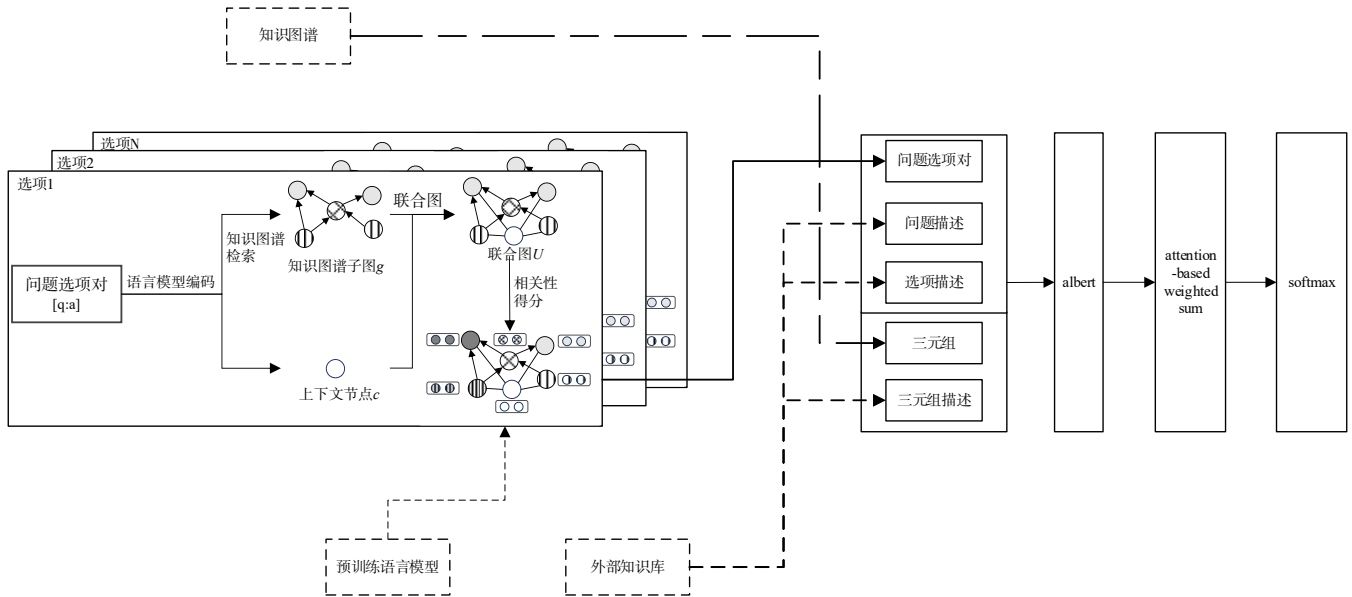


Fig. 1 DEK-KGQA architecture

图1 DEK-KGQA 架构

2.3 联合图表示

首先, 从问题选项对中识别实体提及, 并将其链接到知识图谱中的实体, 以此初始化节点集 V ; 然后, 将出现在任意两个提及实体之间两跳路径中的所有实体添加到 V 中, 保留 V 中节点之间的所有边, 形成知识图谱子图 g 。

为了将知识图谱和 QA 上下文之间的关系显式地表示在同一个图形结构中, 本文引入一个新节点 c 代表 QA 上下文, 并通过两种新的关系类型 $r_{c,q}$ 和 $r_{c,a}$ 将 c 与知识图谱子图 g 中的每个主题实体 $V_{q,a}$ 连接起来形成联合图。具体而言, $r_{c,q}$ 表示 QA 上下文 c 与某一问题实体 V_q 之间的关系, $r_{c,a}$ 表示 QA 上下文 c 与某一选项实体 V_a 之间的关系。联合图 $U=(V_u, E_u)$, 其中有:

$$V_u = V_{q,a} \cup \{c\} \quad (1)$$

U 中每个节点类型为以下 4 种节点类型之一: QA 上下文节点 c 、问题节点 V_q 、选项节点 V_a 和其他节点。本文使用 $text(c)$ 表示 QA 上下文节点的文本, $text(v)$ 表示知识图谱子图节点 $v \in V_{sub}$ 的文本。使用掩码语言模型 RoBERTa 来初始化 QA 上下文节点 c 的嵌入 ($z^{LM} = fenc(text(c))$), 并通过实体嵌入来初始化知识图谱子图 g 中的每个节点^[21]。

数据集 CommonsenseQA 中某个问题的联合图如图 2 所示^[22]。QA 上下文节点 c 为灰色节点, “find”“hamburger”为问题节点 V_q (阴影), “pizza”“fast food restaurant”为选项节点 V_a (黑色), 其他节点为白色节点, $r_{c,q}$ 为稀疏虚线, $r_{c,a}$ 为密集虚线。

2.4 相关性评分

为了提高推理效率和推理的准确性, 本文采用节点相

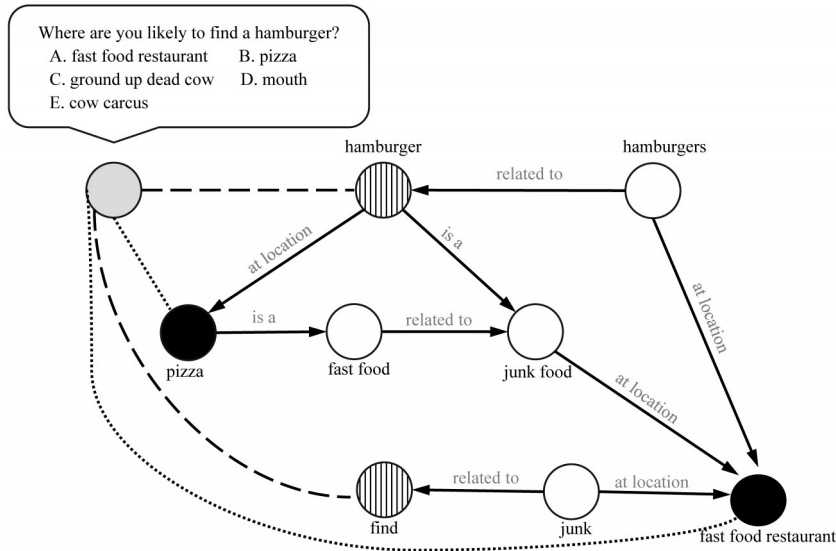


Fig. 2 Joint graph of a question in CommonsenseQA

图2 CommonsenseQA中某个问题的联合图

关性评分的方法,具体如图3所示。

联合图中阴影节点表示问题实体 V_q , 黑色节点表示选项实体 V_a , 白色节点表示其他实体。将QA上下文节点 c 与联合图中任一实体拼接后,通过预训练语言模型计算相关性分数,得到每个实体分数。在获取实体分数后的联合

图中,颜色的深浅表示联合图中实体与QA上下文节点 c 之间相关性的强弱。通过此方法可以排除与当前QA上下文无关的节点,从而减小联合图的规模。该方法使用预训练语言模型评估每个知识图谱子图节点 n 在给定QA上下文时的相关性。

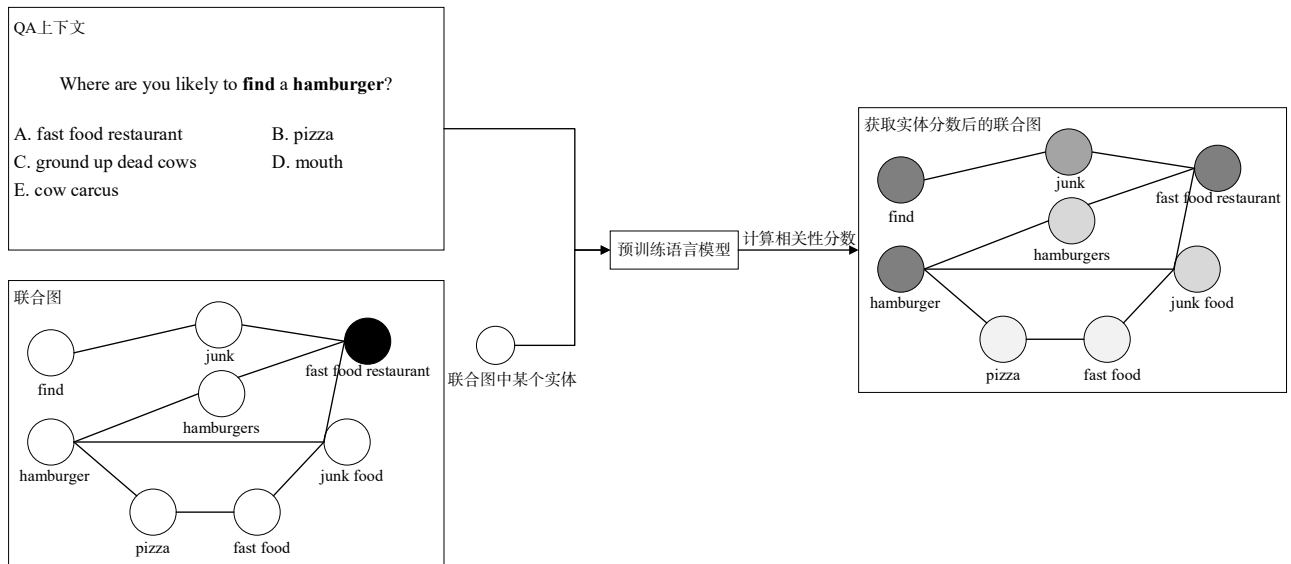


Fig. 3 Process of calculating relevance in the joint graph

图3 联合图中相关性计算过程

对于每个节点 $v \in V_{sub}$, 将其对应的知识图谱子图节点文本 $text(v)$ 和QA上下文节点文本 $text(c)$ 进行拼接,并通过预训练语言模型计算相关性分数 p_c 。具体计算方式如公式(3)所示:

$$p_c = f_{corr}(f_{enc}([text(c); text(v)])) \quad (3)$$

其中,使用基于RoBERTa架构的掩码语言模型作为 f_{corr} ,使用RoBERTa模型作为 f_{enc} 。

对于知识图谱子图中的所有三元组,从知识图谱 ConceptNet^[21]中得到三元组权重 p_t ,通过相关性分数 p_c 、三元

组权重 p_t 和自定义的关系类型权重 p_e 对知识图谱子图中的每个三元组 t 进行评分,得到每个三元组的得分 p 。具体计算方式如公式(4)所示:

$$p = p_{c1} * p_{c2} * p_t * p_e = p_{c1} * p_{c2} * p_t * \frac{N_r}{N} \quad (4)$$

其中: p_{c1} 表示某个三元组中头实体节点 e_h 的相关性分数, p_{c2} 表示同一三元组中尾实体节点 e_t 的相关性分数, r 表示某个三元组 t 的关系类型, N 表示检索到的知识图谱子图中的三元组个数, N_r 表示检索到的三元组中具有关系

类型 r 的三元组个数。最后选择得分最高的三元组 $t_{best} = (e_h, r_{best}, e_t)$, 该三元组反映了在给定的 QA 上下文中与问题和选项最相关的知识。

2.5 上下文信息

在知识图谱中, 实体的表示仅限于其字面含义, 而缺乏深入的语义阐释或上下文相关的具体解释。该现象可能导致信息的不完全性和解读的歧义性, 因为没有足够的信息来揭示实体的复杂性和多维度的含义。为了解决这一问题, 本文采取一种构建性的方法, 即从维基百科这一公认的、全面的知识库中引入相关外部知识。该策略旨在通过维基百科中丰富的信息, 进一步增强现有知识图谱中实体的语义理解和具体解释。

对于问题实体、选项实体以及最优三元组中的头实体和尾实体, 先按照转换为小写和去除标点两种方法进行处

理。然后按照以下形式找到与维基词典中相符合的概念: ①原始形式; ②通过 Spacy 的词形还原形式; ③最后一个词。对于每个实体, 用 d_q 表示问题实体 e_q 的词义, 用 d_c 表示选项实体 e_c 的词义, 用 d_h 表示最优三元组头实体 e_h 的词义, 用 d_t 表示最优三元组尾实体 e_t 的词义。

本文方法将问题选项对、问题描述、选项描述、三元组、三元组描述按照以下格式输入 ALBERT 模型中, $[CLS] q: c [SEP] e_q: d_q [SEP] e_c: d_c [SEP] triple [SEP] e_h: d_h [SEP] e_t: d_t$ 。

图 2 中同一问题的解析如图 4 所示。其中, 粗体表示正确答案, 斜体表示数据集 CommonsenseQA 中增加的额外干扰项。首先, 模型利用知识图谱 ConceptNet 排除干扰项 “cow carcass”, 然后通过三元组知道汉堡是在快餐店中, 从而选择了正确答案。

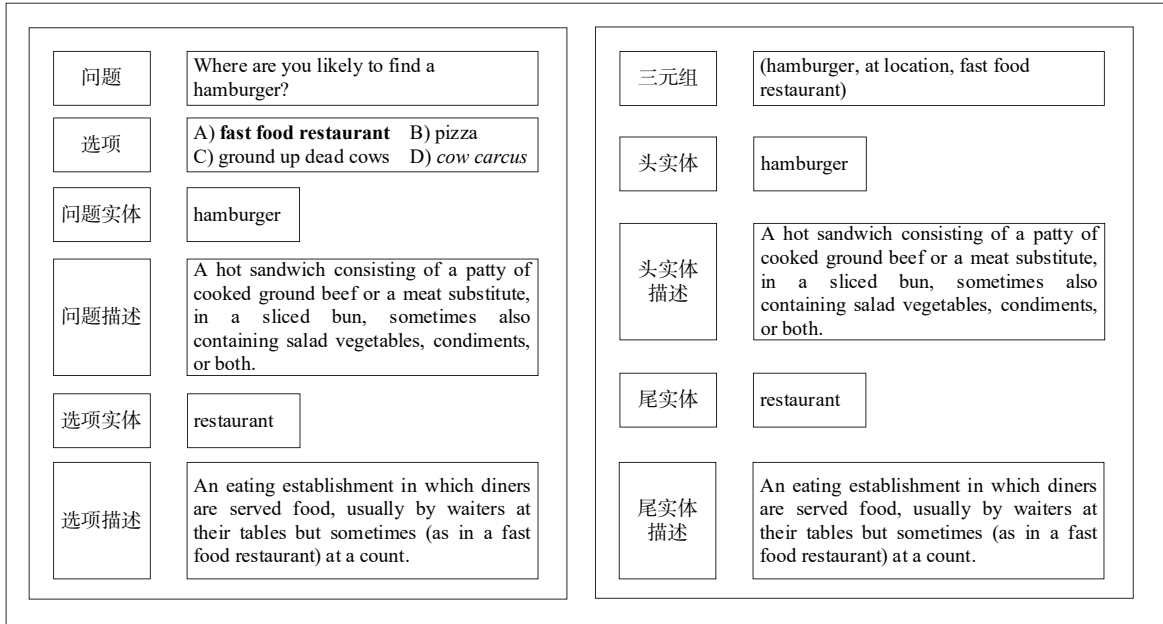


Fig. 4 Example questions and analysis in CommonsenseQA

图 4 CommonsenseQA 示例问题及解析

2.6 推理

在基于 albert 的模型中, 本文使用基于注意力机制的加权求和以及 softmax 层计算问题与选项之间的相关性得分。具体来说, 假设 albert 的输出表示为 (x_0, x_1, \dots, x_m) , 其中每个 x_i 都是一个 d 维向量。首先通过一个线性映射和 tanh 激活函数将这些嵌入转化为 keys, 然后计算这些 keys 与一个查询向量 u 的点积, 得到每个嵌入的注意力权重:

$$q_i = \tanh(u^T * x_i) \quad (5)$$

然后, 通过 softmax 函数对每个嵌入的权重 q_i 进行归一化, 得到最终的注意力权重 α_i :

$$\alpha_i = \text{soft max}(q_i) \quad (6)$$

将注意力权重 α_i 应用于原始的嵌入, 通过加权求和得到最终的上下文向量 v :

$$v = \sum_{i=0}^m \alpha_i x_i \quad (7)$$

其中, α_i 是 x_i 的注意力权重。问题和选项之间的相关

性得分是通过将该上下文向量通过一个线性层并应用 softmax 函数得到的:

$$s = \text{soft max}(v^T b) \quad (8)$$

其中, $b \in R^d$ 是线性层权重, softmax 函数用于计算交叉熵损失函数的所有选项, 从而得到每个选项作为正确答案的概率。

3 实验结果及分析

3.1 数据集

本文使用的常识领域数据集为 CommonsenseQA 数据集。CommonsenseQA 是一个大规模的常识推理问答数据集, 包含 12 102 个问题, 每个问题都有 5 个候选答案。本文遵循文献 [13] 的数据划分进行实验, 其中训练集包含 9 741 个问题和 48 705 个答案, 验证集包含 1 221 个问题和

6 105 个答案,测试集包含 1 140 个问题和 5 700 个答案。

3.2 知识图谱

本文使用的知识图谱为 ConceptNet, ConceptNet 是一个通用的领域常识知识图谱,其数据源主要来自于 Wiktionary、WordNet 和 Open Mind Common Sense。ConceptNet 包含实体总数为 799 273,关系总数为 2 487 810,并包含 42 种实体关系类型,例如 is a、is used for、made of、has a 和 has property 等。

对于每个问题,从该知识图谱中检索一个大小不超过 200 个节点的知识图谱子图,通过该子图获得三元组,计算相关性评分,保留前 5 个得分最高的三元组。

3.3 实验环境设置

实验所用的操作系统为 Ubuntu20.04, GPU 使用 2 张显存为 32GB 的 v100 显卡, Python 版本为 3.8, Pytorch 版本为 2.0.1。实验的超参数具体设置如表 2 所示。

Table 2 Hyperparameter settings

表 2 超参数设置

参数	设置
Batch size	8
Max seq length	192
Lr	1e-5
Weight decay	0.15
Warmup proportion	0.1
Loss Function	CrossEntropy Loss
Optimizer	Adam

3.4 基准方法比较

在 CommonsenseQA 数据集上,将本文方法和不同的基准模型进行比较,实验结果如表 3 所示。从实验结果可以看出,本文方法相较于其他基准方法准确率明显提升,和同组模型 HeadHunter 相比,准确率提高了 3.0%^[24]。组 1 中的方法均为预训练语言模型,不同模型的性能存在差异。这是因为随着模型参数量和训练数据的增加以及模型结构的改进与优化,模型可以学习到更多的语言知识和语言表征。组 2 中均为大型语言模型, BloombergGPT^[25] 专注于金融领域,金融领域训练语料占 51.27%,而 PaLM2^[26] 的训练数据包含来自多种语言和领域的文本数据,更为多样化,且模型参数量远大于 BloombergGPT,因此取得了最佳表现。组 3 方法 STaR 未引入外部知识,而是利用与训练语言模型 GPT-J 的推理能力,用少样本提示的方法迭代引导模型生成高质量的解释^[27]。组 4 中的方法将知识图谱用于问答,通过微调编码器(KagNet^[28])、删除图中不相关节点(MHGRN^[29]、GrapeQA^[30]),取得了不同效果。组 5 中的方法均引入了多种外部知识(包括:文本知识和知识图谱),因此组 5 方法的性能均优于组 4。本文方法相较于组 5 中的其他方法,在计算相关性评分中考虑不同权重对三元组分数的影响,能够有效地捕捉 QA 上下文与知识图谱之间的关系。此外,其考虑了文本知识和知识图谱以增强文本的语义信息,进而提高了回答的准确性。在对比组 2 中的大型语言模型 PaLM2 时,本文方法虽然在知识覆盖范

围上无法与 PaLM2 相媲美,但其在模型规模较小的情况下,提供了模型的可解释性,并在资源利用和成本效益方面展现出明显优势,因而在实验中取得了相对不错的效果。

Table 3 Comparison results with different baseline models on development set and test set

表 3 在开发集和测试集上与不同基线模型比较结果

组别	模型	开发集准确率/%	测试集准确率/%
组 1	BERT	58.3	56.7
	RoBERTa	78.5	72.1
	ALBERT	80.1	73.5
	T5	-	78.1
组 2	BloombergGPT	-	65.5
	PALM2	-	90.4
组 3	STaR	72.5	72.3
	KagNet	73.5	69.0
组 4	MHGRN	74.4	71.1
	GrapeQA	-	73.5
	DRAGON	-	78.2
	Headhunter	83.3	78.4
组 5	DEK-KGQA	81.6	81.4

3.5 消融实验

消融实验结果如表 4 所示。在 DEK-KGQA 中,分别移除外部知识的引入和联合图中的节点相关性评分,考察不同模块对本文方法的影响。若不使用维基词典来增强实体理解,准确率下降到 72.3%,说明外部知识对方法的性能有着显著的正向影响。通过补充外部知识,模型能够更好地理解实体含义。若直接使用所有知识图谱子图中的节点作为输入,准确率下降至 78.7%,说明节点相关性评分在过滤无关的 QA 上下文节点方面存在一定作用,在移除后会引入噪声,降低模型的准确率。

Table 4 Ablation experiments

表 4 消融实验

id	消融实验	准确率/%	增量/%
1	DEK-KGQA	81.6	/
2	No Wiktionary Data	72.3	-11.4
3	No Joint Graph	78.7	-3.5

4 结语

本文提出一种融合外部知识的知识图谱问答方法 DEK-KGQA。首先,本文通过连接问题、选项和知识图谱构建一个联合图,以更好地表示问题与知识之间的关系。然后,使用语言模型计算节点之间的相关性分数,并选取评分最高的三元组获取问题和选项最相关的知识。最后,通过引入外部知识源(维基词典)以增强实体的语义理解。实验结果表明,在 CommonsenseQA 数据集上,本文方法相较于其他基准方法准确率有所提升。

当前方法将输入文本汇集成单一向量,这在一定程度上可能忽略了文本中的细节和上下文信息,影响了 QA 上

下文的推理能力。为了解决此问题,未来工作可以考虑引入更复杂的文本表示方法,例如使用多头注意力机制来捕捉不同词语之间的依赖关系和重要性,从而提升问答的准确性。

为了进一步拓宽模型的知识广度,未来研究还可以考虑引入更多类型的外部知识资源。例如,结合专家知识库、领域特定数据库等,可以提供更精准和细粒度的知识,从而提高模型在专业领域问答的准确性。此外,探索多源知识融合的方法,如知识图谱增强、异构数据整合等,也有助于提升问答系统的性能和泛化能力。

参考文献:

- [1] VRANDEČIĆ D, KRÖTZSCH M. Wikidata: a free collaborative knowledgebase[J]. *Communications of the ACM*, 2014, 57(10): 78–85.
- [2] MITCHELL T, COHEN W, HRUSCHKA E, et al. Never-ending learning [J]. *Communications of the ACM*, 2018, 61(5): 103–115.
- [3] 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.
- [4] LEHMANN J, ISELE R, JAKOB M, et al. DBpedia – a large-scale, multilingual knowledge base extracted from Wikipedia [J]. *Semantic Web*, 2015, 6: 167–195.
- [5] SUCHANEK F M, KASNECI G, WEIKUM G. Yago: a core of semantic knowledge – unifying WordNet and Wikipedia [C]//*16th International World Wide Web Conference*, 2007: 697–706.
- [6] CHEN Z R, WANG X, WANG L, et al. A review of open-domain knowledge graph question answering research [J]. *Journal of Frontiers of Computer Science and Technology*, 2021, 15 (10): 1843–1869.
陈子睿,王鑫,王林,等. 开放领域知识图谱问答研究综述 [J]. *计算机科学与探索*, 2021, 15 (10): 1843–1869.
- [7] HAKIMOV S, UNGER C, WALTER S, et al. Applying semantic parsing to question answering over linked data: addressing the lexical gap [C]//*International Conference on Applications of Natural Language to Data Bases*, 2015:103–109.
- [8] UNGER C, BÜHMANN L, LEHMANN J, et al. Template-based question answering over RDF data [C]//*Proceedings of the 21st International Conference on World Wide Web*, 2012:639–648.
- [9] WALTER S, UNGER C, CIMIANO P, et al. Evaluation of a layered approach to question answering over linked data [C]// *Proceedings of the 11th International Conference on The Semantic Web*, 2012: 362–374.
- [10] YIH W T, CHANG M W, HE X D, et al. Semantic parsing via staged query graph generation: question answering with knowledge base [C]// *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, 2015: 1321–1331.
- [11] ZHU S, CHENG X, SU S. Knowledge-based question answering by tree-to-sequence learning [J]. *Neurocomputing*, 2020, 372: 64–72.
- [12] CHEN Y, WU L, ZAKI M J. Bidirectional attentive memory networks for question answering over knowledge bases [C]//*Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2019: 2913–2923.
- [13] LIN B, CHEN X, CHEN J, et al. KagNet: knowledge-aware graph networks for commonsense reasoning [C]//*2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing*, 2019:2829–2839.
- [14] XU Y, ZHU C, XU R, et al. Fusing context into knowledge graph for commonsense question answering [C]//*Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 2021: 1201–1207.
- [15] YASUNAGA M, REN H, BOSSELU T A, et al. QAGNN: reasoning with language models and knowledge graphs for question answering [C]//*Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2021: 535–546.
- [16] SUN Y, SHI Q, QI L, et al. JointLK: joint reasoning with language models and knowledge graphs for commonsense question answering [DB/OL]. <http://arxiv.org/abs/2112.02732v1>.
- [17] WANG Y, ZHANG H, LIANG J, et al. Dynamic heterogeneous-graph reasoning with language models and knowledge representation learning for commonsense question answering [C]//*Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, 2023: 14048–14063.
- [18] CHOI B, GEUN B, JANG D, et al. MEM-KGC: masked entity model for knowledge graph completion with pre-trained language model [J]. *IEEE Access*, 2021, 9: 132025–132032.
- [19] FENG Y, CHEN X, LIN B, et al. Scalable multi-hop relational reasoning for knowledge-aware question answering [C]// *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 2020: 1295–1309.
- [20] LAN Z Z, MINGDA C, SEBASTIAN G, et al. ALBERT: a lite BERT for self-supervised learning of language representations [DB/OL]. <http://arxiv.org/abs/1909.11942>.
- [21] LIU Z, LIN W, SHI Y, et al. A robustly optimized BERT pre-training approach with post-training [C]// *Proceedings of the 20th Chinese National Conference on Computational Linguistics*, 2021: 471–484.
- [22] TALMOR A, HERZIG J, LOURIE N, et al. CommonsenseQA: a question answering challenge targeting commonsense knowledge [C]// *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2019: 4149–4158.
- [23] SPEER R, CHIN J, HAVASI C. Conceptnet 5.5: an open multilingual graph of general knowledge [C]// *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 2017: 4444–4451.
- [24] LI Y, ZOU B, LI Z, et al. Winnowing knowledge for multi-choice question answering [C]// *Findings of the Association for Computational Linguistics*, 2021: 1157–1165.
- [25] WU S, IRSOY O, LU S, et al. BloombergGPT: a large language model for finance [DB/OL]. <https://arxiv.org/abs/2303.17564>.
- [26] ANIL R, DAI A M, FIRAT O, et al. PaLM 2 technical report [DB/OL]. <https://arxiv.org/abs/2305.10403>.
- [27] ERIC Z, WU Y, MU J, et al. STaR: bootstrapping reasoning with reasoning [DB/OL]. <https://arxiv.org/abs/2203.14465>.
- [28] LIN B Y, CHEN X, CHEN J, et al. KagNet: knowledge-aware graph networks for commonsense reasoning [C]//*Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, 2019: 2829–2839.
- [29] FENG Y, CHEN X, LIN B Y, et al. Scalable multi-hop relational reasoning for knowledge-aware question answering [C]// *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 2020:1295–1309.
- [30] TAUNK D, KHANNA L, KANDRU S, et al. GrapeQA: GRaph augmentation and pruning to enhance question-answering [C]//*Companion Proceedings of the ACM Web Conference 2023*, 2023: 1138–1144.