

一种基于异构图神经网络和文本语义增强的实体关系抽取方法

彭勃, 李耀东, 龚贤夫, 李浩

引用本文

彭勃, 李耀东, 龚贤夫, 李浩. 一种基于异构图神经网络和文本语义增强的实体关系抽取方法[J]. 计算机科学, 2024, 51(6A): 230700071-5.

PENG Bo, LI Yaodong, GONG Xianfu, LI Hao. [Method for Entity Relation Extraction Based on Heterogeneous Graph Neural Networks and TextSemantic Enhancement](#) [J]. Computer Science, 2024, 51(6A): 230700071-5.

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)

[面向产线AI质检的少样本评测方法研究和验证](#)

Study and Verification on Few-shot Evaluation Methods for AI-based Quality Inspection in Production Lines

计算机科学, 2024, 51(6A): 230700086-8. <https://doi.org/10.11896/jsjcx.230700086>

[基于CRF的中文语法错误诊断系统的实现与应用](#)

Implementation and Application of Chinese Grammatical Error Diagnosis System Based on CRF

计算机科学, 2024, 51(6A): 230900073-6. <https://doi.org/10.11896/jsjcx.230900073>

[基于BERT和CNN的药物不良反应个案报道文献分类方法](#)

Literature Classification of Individual Reports of Adverse Drug Reactions Based on BERT and CNN

计算机科学, 2024, 51(6A): 230400049-6. <https://doi.org/10.11896/jsjcx.230400049>

[DUWe:动态未知词嵌入方法在Web异常检测中的应用](#)

DUWe:Dynamic Unknown Word Embedding Approach for Web Anomaly Detection

计算机科学, 2024, 51(6A): 230300191-5. <https://doi.org/10.11896/jsjcx.230300191>

[WiCare:一种非接触式的老人如厕跌倒监测模型](#)

WiCare:Non-contact Fall Monitoring Model for Elderly in Toilet

计算机科学, 2024, 51(6A): 230700044-8. <https://doi.org/10.11896/jsjcx.230700044>

一种基于异构图神经网络和文本语义增强的实体关系抽取方法

彭勃¹ 李耀东¹ 龚贤夫¹ 李浩²

1 广东电网有限责任公司电网规划研究中心 广州 510080

2 四川大学计算机学院 成都 610065

摘要 信息化时代,如何从海量自然语言文本中提取结构化信息已经成为研究热点。电力系统中繁杂的知识信息需要通过构建知识图谱来解决,而实体关系抽取是其上游的信息抽取任务,其完成度直接关系到知识图谱的有效性。而随着深度学习的不断发展,利用深度学习技术来完成实体关系抽取任务的研究逐渐展开并取得了良好的效果。然而目前依然存在文本语义应用不完全等问题。针对这些问题本文尝试提出了一种基于异构图神经网络和文本语义增强的实体关系抽取方法,该方法使用词节点与关系节点学习语义特征,并通过 BRET 与预训练任务分别获得两种节点的初始特征,使用多层图网络结构迭代更新,并在每一层中使用基于多头注意力机制的信息传递实现两种节点的交互。通过该模型与其他实体关系抽取在两个公开数据集上实验对比,所提模型取得了预期效果,在多种情境下普遍优于对比模型。

关键词: 深度学习;自然语言处理;知识图谱;实体关系抽取;异构图神经网络;文本语义增强

中图分类号 TP391

Method for Entity Relation Extraction Based on Heterogeneous Graph Neural Networks and Text Semantic Enhancement

PENG Bo¹, LI Yaodong¹, GONG Xianfu¹ and LI Hao²

1 Grid Planning and Research Center of Guangdong Power Grid Co.,Guangzhou 510080,China

2 College of Computer Science,Sichuan University,Chengdu 610065,China

Abstract In the era of information technology,extracting structured information from massive natural language texts has become a research hotspot. The complex knowledge information in the power system needs to be solved by constructing a knowledge graph,and entity relation extraction is the upstream information extraction task,whose completeness directly affects the effectiveness of the knowledge graph. With the continuous development of deep learning,research on using deep learning techniques to solve entity relation extraction tasks has gradually been carried out and achieved good results. However,there are still problems such as incomplete application of text semantics. This paper attempts to propose an entity relation extraction method based on heterogeneous graph neural network and text semantic enhancement to address these issues. This method uses word nodes and relationship nodes to learn semantic features and obtains initial features of the two types of nodes through BRET and pre-training tasks respectively. It uses a multi-layer graph network structure for iteration and implements the interaction between the two types of nodes by using multi-head attention mechanism for information transmission in each layer. Through experimental comparison with other models on two public datasets,this model achieves the expected effect and generally outperforms other entity relationship extraction models in various scenarios.

Keywords Deep learning,Natural language processing,Knowledge graph,Entity relation extraction,Heterogeneous graph neural networks,Text semantic enhancement

1 引言

大数据信息化时代,随着电网规模愈发庞大,电力系统所产生的知识量也呈现出量级增长。电力系统所产生的知识信息来源繁多且结构复杂,原因主要是调度方法日益复杂,大量的测控通信以及新型装置也在升级与增加,同时电力新业务也在不断拓展。传统的组织形式已经无法很好地满足当前的需求,如何准确且快速地从其中获取有用的信息已经成为了目前的重点研究方向。在自然语言处理(Natural Language

Processing,NLP)中,信息抽取是具有代表性的重要任务,它从给定的语料库中通过命名实体识别(Named Entity Recognition)、实体关系抽取(Entity Relation Extraction)和事件抽取(Event Extraction)等任务^[1]提取命名实体、关系和事件等信息类型。

本文主要关注实体抽取任务,其是许多自然语言处理下游任务的基础。其主体目标是通过模型从无结构的文本中抽取实体以及关系构成三元组。三元组可以用于文本理解、知识库构建、知识图谱构建等研究,本文希望将三元组抽取后使

基金项目:中国南方电网有限责任公司科技项 037700KK52220042(GDKJXM20220906)

This work was supported by the science and Technology Project of China Southern Power Grid Co. Ltd 037700KK52220042(GDKJXM20220906).

通信作者:彭勃(1339512151@qq.com)

用知识图谱技术将其应用在电力系统之中,以便处理日益繁杂的知识信息。

传统的关系抽取方法主要围绕人工设计的特征和核函数展开。基于特征的方法利用人工定义的特征模式设计分类器来完成实体关系的分类。Kambhatla^[2]使用最大熵模型构建了实体关系分类器,结合了来自文本的各种词汇、句法和语义特征。Culotta等^[3]则提出了从数据库规则中以概率方式学习先验知识,再通过条件随机场来抽取关系。基于特征的方法往往依赖于模型设计者的经验,并且对上下文信息的利用有限。为了减少对人工构造特征的依赖,研究者们开始尝试更多基于机器学习的实体关系抽取方法。Zhang等^[4]提出了在解析树中嵌入句法结构特征,在关系提取中取得了较好的效果,这些特征可以通过卷积核来捕捉。

随着人工智能的不断发展,深度学习在实体抽取任务领域的应用表现也十分出色。早期的实体关系抽取常常采用 pipeline 的方式,用两个子任务先后完成命名实体识别和关系分类^[5-7]。但在实体抽取任务不断进步的同时,仍然存在一些难点与不足:在含有三元组重叠的复杂语境下模型的表现不佳;对全局的文本信息利用并不充分;实体识别与关系抽取任务在联合模型中交互不足。因此,研究效果更好的实体关系抽取方法对多个领域的发展具有重要意义。

本文基于深度学习的方法,提出了一种基于异构图神经网络和文本语义增强的实体关系抽取模型。该异构图神经网络包含词节点以及关系节点。词节点通过预训练语言模型获得初始表征,关系节点通过关系嵌入预训练任务获得初始表征,并且使用基于多头注意力机制的信息传递在两者之间进行信息交互,通过多层图网络结构迭代更新来解决字词与关系间交互不足的问题。此外,本文也详细介绍了特征增强策略以及使用二进制标注器对三元组标注的过程。最终在多个公开数据集上进行实验与对比,结果表明该模型在面对包含三元组的复杂语句时具有优势。

2 相关工作

2.1 知识图谱

知识图谱是由 Google 提出并应用的,在提出后立刻在金融、医药等领域得到了广泛应用^[8-9]。知识图谱实现了对知识的智能化处理,包括对信息的抽取、推理、检索等。

知识图谱本身是以图的形式组织,图谱中基本组成单位为“实体-关系-实体”三元组。将实体表示为节点,关系表示为边,通过相互连接,就可以构成一张语义网^[10-11]。

2.2 预训练模型

预训练语言模型指在大规模语料库上训练而成的语言模型。它的目的是学习自然语言中的统计规律和语义信息,对于多种自然语言处理任务(如句子分类、命名实体识别、机器翻译等)可以作为底层模型进行迁移学习。一般来说,预训练语言模型的运作方式是将大规模语料库输入模型中进行训练,使得模型能够从这些文本数据中学习到多种语言层次的特征。在训练完成后,模型可以被微调到适应某个特定任务(如认知机器人对话,文本分类、生成对话等),从而缩短模型的训练时间、提升模型的性能。

以 OpenAI 提出的 GPT 模型为代表的基于参数迁移的迁移学习仍然是预训练语言模型的主流,其主要是基于 Transformer 的 Decoder 训练,但是训练过程中没有考虑语境

中的下文信息。Devlin 等提出的 BERT^[12]使用最为广泛,它使用 Transformer Encoder 结构并借助 Masked Language Model, Next Sentence Prediction 两种任务实现双向语言模型构建,完成预训练。

2.3 实体抽取关系方法

Yu 等^[13]使用 span-based 标注策略将头实体识别和尾实体关系识别两个子任务转换为序列标注问题,其只考虑了从头实体到尾实体和关系的单向预测,导致缺乏一定的交互性。Zeng 等^[14]提出的 OrderCopyRE 将三元组视为 token 序列,则原来的实体关系抽取任务就变为按某种顺序输出三元组的文本生成任务,具体使用强化学习方法学习抽取的顺序。强化学习的奖惩机制最终使得预测的三元组与标签三元组尽可能保持一致。然而该方法存在暴露偏差问题,具体来看,训练阶段与预测阶段的前序三元组来源不同,这会导致偏差在预测阶段三元组的生成过程中不断累积,影响最终表现。Gupta 等^[15]将文本以 token 为单位展开至二维表格中,其中的元素即对应的 token pair。模型最终也以表格形式编码或解码,在此基础上提出了一种基于表格填充方法的多任务循环神经网络建模实体识别和关系分类任务之间的依赖。该方法借助 token pair 学习上下文,循环神经网络框架帮助上下文感知,实现了 token 级别的关系学习,这意味着一个句子可以建立多个关系实例。

GCN(Graph Convolutional Network)是一种基于图结构数据的神经网络模型。图神经网络在文本分类情感分析、问答系统等领域已经取得了不错的效果,同样在实体关系抽取任务中也取得了不错的效果。Zhang 等^[16]借助图神经网络的思路,在经过剪枝的依存句法树上进行 GCN 运算以抽取实体和关系。剪枝树可以去除长文本中的冗余信息,但也会面临相关信息丢失的问题。Hong 等^[17]构建了一个基于 GCN 的端到端模型,通过序列标注获得句子实体后,提出了一种关系感知注意机制表示实体间的关系,并最终将实体和关系分别作为节点与边建立图。值得注意的是,该方法还考虑了关系的方向,在多个公开数据集都取得了显著的效果。

本章主要介绍了本文所涉及与采用的相关技术和算法。相比 Zhang 等^[4]的研究,本文结合注意力机制,将实体关系转化为图结构,并设计了一种异构图神经网络与注意力机制相结合的方法。

3 基于异构图神经网络和文本语义增强的实体关系抽取方法

本章使用的模型结构基于 Zhao 等的工作^[18]进行了一定的修改,本文模型 RIFRE+ 的整体结构如图 1 所示。模型首先将文本转换为异构图结构,图中包括词节点和关系节点两种类型的节点,并使用 BERT 与预训练的关系嵌入分别获得词节点和关系节点的初始向量表示。模型主要包括一个多层异构图网络,在每一层都用多头注意力机制在关系节点以及词节点之间进行信息传递,在层间迭代更新节点信息。之后将增强处理后的节点特征用于二进制标注器,识别输入语句中的主体与客体,最后输出满足筛选条件的三元组。

3.1 异构图构建

将输入的语句定义为 $W = \{\omega_1, \omega_2, \dots, \omega_n\}$, ω_i 表示输入文本中的一个词,即一个 token,其中 n 表示文本的长度。将其放入预训练模型 BERT 获得输出特征向量 $H = \{h_1, h_2, \dots, h_n\}$, $h_i \in R^{(d \times 1)}$ 。其中, d 表示特征向量的

维度, h_i 表示词节点的初始向量。

关系节点首先需要通过一个预训练模块,在三元组(subject, relation, object)中,主体与客体可以结合关系去预测另一个。

本模块设计了一个位置预测任务,整合主体的特征表示与关系嵌入预测客体在句子中的位置。在这里,关系嵌入通过随机初始化定义初始值,在任务中不断迭代更新后得到优化的嵌入向量。其具体公式表示为:

$$y_{\text{sub}}^r = y_{[\text{CLS}]} \text{Maxpool}(y_{\omega_1}, y_{\omega_2}, \dots, y_{\omega_{n+k}}) \oplus y_r \quad (1)$$

$$P_{\text{obj}}^r = \text{Sigmoid}(W_p y_{\text{sub}}^r + b_p) \in R^{(n \times 1)} \quad (2)$$

其中, [CLS] 表示整个句子上下文的信息,位置预测任务中也融合了“[CLS]”的特征表示来更好地建模句信息; \oplus 表示向量拼接运算; k 表示构成主题的词数; y_r 表示关系 r 的嵌入; W_p 表示参数矩阵; b_p 表示偏执向量; y_{sub}^r 表示主体的特征; P_{obj}^r 表示每个 token 对应客体的预估分数。

该模块使用交叉熵(Cross Entropy)作为损失函数,损失函数的具体定义为:

$$\text{Loss}_{\text{pre}} = -\frac{1}{n} \sum_{i=1}^n [h_{\text{obj}}^r(i) \log p_{\text{obj}}^r(i) + F_{\text{pre}}] \quad (3)$$

$$F_{\text{pre}} = (1 - h_{\text{obj}}^r(i)) \log(1 - p_{\text{obj}}^r(i)) \quad (4)$$

在输入语句中, n 表示句子长度,每个位置都获得预测概率 $p_{\text{obj}}^r(i)$,表示其是否属于该客体, $h_{\text{obj}}^r(i)$ 表示句中索引位置 i 的 token 与客体关联的标签。

该预训练模块将三元组与语句信息纳入考虑,能够获得表示较好的关系向量。用 E 表示预训练后的关系嵌入矩阵,则关系节点的向量表示为:

$$[r_1, r_2, \dots, r_M] = W_r E([r_1', r_2', \dots, r_M']) + b_r \quad (5)$$

其中, M 代表数据集中的关系类别数量, r_i 表示初始关系节点向量, r_i' 表示第 i 种关系的独热向量, W_r 与 b_r 皆是可训练的参数。

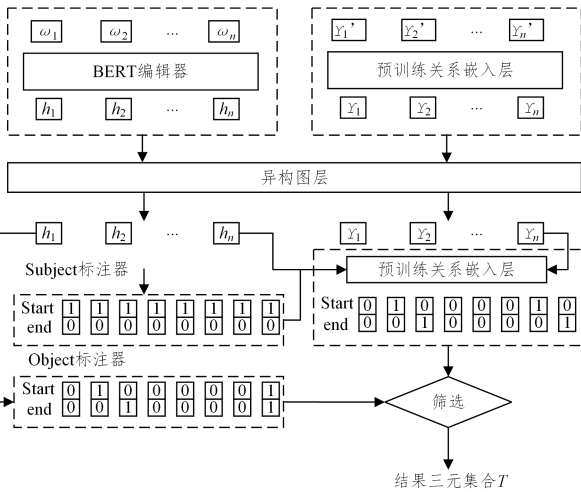


图1 RIFRE+结构图

Fig.1 Overall structure of RIFRE+

3.2 多层异构图网络

本文计划使用多层图网络结构,主要目的是让词节点与关系节点得到充分的交互,使它们之间的语义信息进行迭代更新。在异构图中,词节点 u_i 和关系节点 v_j 中的每个节点分别与另一类的所有节点相连,用以更新自身,节点间的信息交互采用图注意力运算,具体过程如式(6)和式(7)所示:

$$a_{ij} = W_a [W_p u_i \parallel W_q v_j] \quad (6)$$

$$a_{ij} = \frac{\exp(a_{ij})}{\sum_{k \in N(u_i)} \exp(a_{ik})} \quad (7)$$

其中, \parallel 表示拼接, $N(u_i)$ 即 u_i 和与 u_i 相邻的所有节点, W_a , W_p , W_q 皆为可训练的参数。为了关注不同子空间中的编码表示信息,使用注意力系数对节点所有邻居加权求和得到节点的增强特征,具体操作如下:

$$u_i' = \sum_{j \in N(u_i)} \frac{a_{ij}^k}{\sum_{j \in N(u_i)} a_{ij}^k} W_k v_j \quad (8)$$

其中, a_{ij}^k 表示第 k 个注意力头归一化的注意力系数, W_k 是对应的权重矩阵。之后,通过能保留非线性和重要信息的门控机制控制原始与增强特征融合,具体过程如式(9)和式(10)所示:

$$g_i = \text{Sigmoid}(W_g [u_i \parallel u_i']) \quad (9)$$

$$\tilde{u}_i = g_i \odot u_i' + (1 - g_i) \odot u_i \quad (10)$$

其中, \tilde{u}_i 表示最终输出特征, \odot 表示 Element-Wise Production。最终用 NIP(Nodes Information Passing)描述上述过程。接下来,关系节点通过经过 NIP 的相邻词节点获取信息更新,在不同层之间添加了残差连接来缓解梯度消失。具体过程如式(11)和式(12)所示:

$$\tilde{r}_j^l = \text{NIP}(r_j^l, \{h_i^l\}, j \in N(r_j^l)) \quad (11)$$

$$r_j^{l+1} = \tilde{r}_j^l + r_j^l \quad (12)$$

其中, r_j^l 表示初始关系节点, \tilde{r}_j^l 表示初始节点下一层图结构的更新向量, r_j^{l+1} 表示该层获得的最终特征向量。

获取了新的关系节点向量,这意味着其可以用来指导词节点获得新的节点信息。正如图二所示,每层都包含了关系节点和词节点的更新。

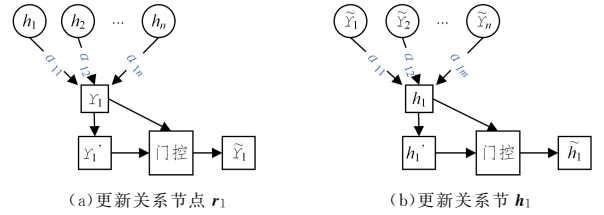


图2 关系节点词节点更新示意图

Fig.2 Diagram of relationship node and word node update

第 l 层的更新过程如式(13)–式(16)所示:

$$\tilde{r}_j^{l+1} = \text{NIP}(r_j^l, \{h_i^l\}, i \in N(r_j^l)) \quad (13)$$

$$r_j^{l+1} = \tilde{r}_j^{l+1} + r_j^l \quad (14)$$

$$\tilde{h}_j^{(l+1)} = \text{NIP}(h_j^l, \{r_j^l\}, j \in N(u_i)) \quad (15)$$

$$h_i^{(l+1)} = \tilde{h}_j^{(l+1)} + h_i^l \quad (16)$$

3.3 二进制标注器

本文使用二进制标注器识别实体并组成三元组。其主题为:分别识别可能的主体与客体,使用每个可能的主体结合关系找到相应的客体,如果其结果在客体集合中,则生成三元组,反之则舍弃。具体地,在识别可能的主体与客体过程中,关键的操作为识别其开始与结束位置,具体操作如式(17)和式(18)所示:

$$p_i^{\text{sta}} = \text{Sigmoid}(W_{\text{sta}} \tanh(h_i^l) + b_{\text{sta}}) \quad (17)$$

$$p_i^{\text{end}} = \text{Sigmoid}(W_{\text{end}} \tanh(h_i^l) + b_{\text{end}}) \quad (18)$$

其中, h_i^l 表示多层异构图输出的特征, p_i^{sta} 与 p_i^{end} 分别表示该词节点作为主体开头与结尾的概率, W_{sta} , b_{sta} , W_{end} , b_{end} 为训练权重参数。此外,设置阈值 t ,只有大于阈值才会被认为可能成为主体或客体。然而在复杂语境下往往会出现三元组重

叠的问题,在三元组中,理论上得知其二便可推出最后一个。标注器预测与构成有效三元组关系的客体 o_i 。判断预测结果客体 o_k 是否存在于客体集合,最终公式为:

$$\mathbf{h}'_{ijk} = \text{Tanh}(W_h [\mathbf{h}'_i \parallel \mathbf{r}'_j \parallel \mathbf{s}_k]) + b_n \quad (19)$$

$$p_i^{\text{sta}} = \text{Sigmoid}(W_{\text{sta}} \tanh(\mathbf{h}'_{ijk}) + b_{\text{sta}}) \quad (20)$$

$$p_i^{\text{end}} = \text{Sigmoid}(W_{\text{end}} \tanh(\mathbf{h}'_{ijk}) + b_{\text{end}}) \quad (21)$$

$$T\{(\mathbf{s}_k, \mathbf{r}_j, \mathbf{o}_k)\} = f(\mathbf{O}, \mathbf{o}_k) \quad (22)$$

其中, \mathbf{s}_k 由组成该实体的所有词节点向量平均池化而来, \mathbf{h}'_{ijk} 拼接了词节点 \mathbf{h}'_i 和关系节点 \mathbf{r}'_j , 第 k 个主体候选 \mathbf{s}_k 表示特征, $W_{\text{sta}}, b_{\text{sta}}, W_{\text{end}}, b_{\text{end}}$ 为训练权重参数, p_i^{sta} 与 p_i^{end} 分别表示该词节点作为主体开头与结尾的概率, $f(\mathbf{O}, \mathbf{o}_k)$ 表示最终的结果筛选函数。

4 实验与分析

实验基于深度学习框架并且使用了在 NYT*、WebNLG* 和 ACE 2003 公开数据集。优化器使用 SGD, 其学习率为 1×10^{-3} 。

表 1 RIFRE+ 的主要实验结果

Table 1 Main experimental results of RIFRE+

Methods	NYT*			WebNLG*			ACE 2003		
	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1
Convolution kernel over parse trees							76.32	62.99	69.02
CopyRE	61.0	56.6	58.7	37.7	36.4	37.1			
GraphRel	63.9	69.0	61.9	44.7	41.1	42.9			
OrderCopyRE	77.9	67.2	72.1	63.3	59.9	61.6			
CasRel	89.7	89.5	89.6	93.4	90.1	91.8			
SETFF	91.7	90.8	91.2	92.9	91.9	92.4			
RIFRE	93.6	90.5	92.0	93.3	92.0	92.6			
Our Model	94.4	90.7	92.5	94.5	92.4	93.5			

但是,从整体上看模型的结果较为优秀不等于模型中所有的改进皆对最终的结果具有正向影响。

为了对比并且验证模型中每一部分对结果影响的有效性,本文还在其基础上做了多个消融实验,主要验证了本文提出的词节点关系节点的更新、预训练关系嵌入、多头注意力机制以及结果筛选几个模块的训练结果在两个数据集中各项指标的具体表现,并与本文所设计模型进行对比,可以较为直观地反映各个模块改进对最终结果的影响。

最终对比结果如表 2 所列。通过对比消除相应模块所得到的结果可以看到,移除词节点更新与关系节点更新后 F1-Score 都有所下降,这说明词节点与关系节点需要互相交换信息。而预训练关系嵌入比随机关系嵌入优势更加明显,原因是其可以提供更多关系语义相关信息。类似地,多头注意力更有助于获取全面的文本语义表达,结果筛选操作有助于任务整体表现的提升。

表 2 RIFRE+ 消融实验结果

Table 2 Ablation study results of RIFRE+

Methods	NYT*			WebNLG*		
	Prec.	Rec.	F1	Prec.	Rec.	F1
本文模型	94.4	90.7	92.5	94.5	92.4	93.5
词节点更新	93.2	90.1	91.6	93.4	91.7	92.6
关系节点更新	93.1	90.0	91.5	93.2	91.6	92.3
预训练关系嵌入	93.5	89.9	91.7	93.4	91.3	92.4
多头注意力	93.8	90.4	92.1	93.8	92.1	93.0
结果筛选	93.3	90.7	92.0	93.6	92.6	93.1

此外,为了探究并对比本文所提出的模型在复杂语境中的表现,将数据集中按照不同的三元组个数 (N) 以及不同的三元组重叠模式 (Normal, SEO, EPO) 进行划分,再进行对比实验,结果如表 3 所列。通过对比可知,本文模型在多种重叠模式下都取得了较好的结果。在按照不同三元组个数划分中,本文模型在绝大部分情况下都优于 SETFF。但在句子中三元组个数增加后,SETFF 表现更优。可能的原因是其使用的表格填充方法在识别时对互相彼此的影响更小。

表 3 RIFRE+ 在不同三元组个数和重叠模式语句下的 F1-Score

Table 3 F1-Score of RIFRE+ on sentences with different overlapping patterns and different numbers of triples

Method	NYT*						WebNLG*					
	Normal	SEO	EPO	N=1	N=2	N \geq 5	Normal	SEO	EPO	N=1	N=2	N \geq 5
CopyRE	66.0	48.6	55.0	67.1	58.6	30.0	59.2	33.0	36.6	59.2	42.5	30.0
GraphRel	69.6	51.2	58.2	71.0	61.5	41.1	65.8	38.3	40.6	66.0	48.3	32.1
OrderCopyRE	71.2	69.4	72.8	71.7	72.6	45.9	65.4	60.1	67.4	63.4	62.2	55.7
CasRel	87.3	91.4	92.0	88.2	90.3	83.7	89.4	92.5	94.7	89.3	90.8	90.9
SETFF	89.1	93.2	93.3	89.0	91.6	91.2	89.2	93.2	94.9	89.7	91.8	92.7
Our Model	91.1	94.2	93.9	90.9	93.4	90.9	90.1	94.7	95.1	90.6	93.1	92.5

结束语 本文以深度学习相关技术为基础,针对实体关系抽取任务中存在的三元组重叠、关系识别不准确、文本语义差异表达以及全局关联信息利用不足等问题,提出了基于异构图神经网络的模型,用于增强模型对文本语义特征的学习能力。该方法首先需要预训练语言模型 BRET 获取词节点初始特征,再通过预训练关系嵌入初始化关系节点,用多头注意力机制传递信息更新词节点与关系节点,多层异构图网络进行节点的迭代更新,其输出结果用于二进制标注器识别输入句中的三元组,最后通过筛选提高准确度。本文通过多个对比实验证明了所提模型的优势,以及各个阶段操作的必要性,能够为知识图谱的构建提供坚实的基础。但是该模型的计算所需时间较长,后续阶段,需要针对模型参数等方面进行缩小操作,以加快训练与推理速度。

参 考 文 献

- [1] LIU K. A survey on neural relation extraction[J]. Science China Technological Sciences, 2020, 63(10): 1971-1989.
- [2] KAMBHATLA N. Combining lexical, syntactic, and semantic features with maximum entropy models for information extraction[C]// Proceedings of the ACL Interactive Poster and Demonstration Sessions, 2004: 178-181.
- [3] CULOTTA A, MCCALLUM A, BETZ J. Integrating probabilistic extraction models and data mining to discover relations and patterns in text [C] // Proceedings of the Human Language Technology Conference of the NAACL, 2006: 296-303.
- [4] ZHANG M, ZHANG J, SU J. Exploring syntactic features for relation extraction using a convolution tree kernel [C] // Proceedings of the Human Language Technology Conference of the NAACL, 2006: 288-295.
- [5] CHAN Y S, ROTH D. Exploiting syntactico-semantic structures for relation extraction [C] // Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics; Human Language Technologies, 2011: 551-560.
- [6] RINK B, HARABAGIU S. Utd: Classifying semantic relations by combining lexical and semantic resources [C] // Proceedings of the 5th International Workshop on Semantic Evaluation, 2010: 256-259.
- [7] ZENG D, LIU K, LAI S, et al. Relation classification via convolutional deep neural network [C] // 25th International Conference on Computational Linguistics; Technical papers (COLING 2014). 2014: 2335-2344.
- [8] SHEN Y L, ZHANG P, LI X Y, et al. Grid robust state estimation based on multi-source data fusion [J]. Guangdong Electric Power, 2019, 32(9): 146-153.
- [9] YANG S L, FU W Z, DONG J. Research on the incremental knowledge base system in intelligent decision support system [J]. Journal of Hefei University of technology, 2004, 27(4): 339-343.
- [10] XU B, ZHANG Y M. System condition-based maintenance decision-making model considering severe weather conditions [J]. Guangdong Electric Power, 2019, 32(6): 61-69.
- [11] HUANG C L, WANG F, WANG J J. Intelligent online hydro-power dispatching based on knowledge reasoning [J]. Automation of Power Systems, 2010, 34(21): 50-54.
- [12] DEVLIN J, CHANG M W, LEE K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding [C] // Proceedings of NAACL-HLT, 2019.
- [13] YU B, ZHANG Z, SHU X, et al. Joint extraction of entities and relations based on a novel decomposition strategy [C] // European Conference on Artificial Intelligence, 2020: 2282-2289.
- [14] ZENG X, HE S, ZENG D, et al. Learning the extraction order of multiple relational facts in a sentence with reinforcement learning [C] // Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), 2019: 367-377.
- [15] GUPTA P, SCHÜTZE H, ANDRASSY B. Table filling multi-task recurrent neural network for joint entity and relation extraction [C] // 26th International Conference on Computational Linguistics; Technical Papers (COLING 2016). 2016: 2537-2547.
- [16] ZHANG Y, QI P, MANNING C D. Graph convolution over pruned dependency trees improves relation extraction [C] // Conference on Empirical Methods in Natural Language Processing, 2018.
- [17] HONG Y, LIU Y, YANG S, et al. Improving graph convolutional networks based on relation-aware attention for end-to-end relation extraction [J]. IEEE Access, 2020, 8: 51315-51323.
- [18] ZENG X, ZENG D, HE S, et al. Extracting relational facts by an end-to-end neural model with copy mechanism [C] // Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2018: 506-514.
- [19] FU T J, LI P H, MA W Y. GraphRel: Modeling text as relational graphs for joint entity and relation extraction [C] // Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy: Association for Computational Linguistics, 2019: 1409-1418.
- [20] WEI Z, SU J, WANG Y, et al. A novel cascade binary tagging framework for relational triple extraction [C] // Annual Meeting of the Association for Computational Linguistics, 2019.
- [21] ZHAO K, XU H, CHENG Y, et al. Representation iterative fusion based on heterogeneous graph neural network for joint entity and relation extraction [J]. Knowledge-Based Systems, 2021, 219: 106888.



PENG Bo, born in 1991, postgraduate, engineer. His main research interest is power grid planning.