

◎热点与综述◎

重叠实体关系抽取综述

冯 钧, 张 涛, 杭婷婷

河海大学 计算机与信息学院 水利部水利大数据重点实验室, 南京 211100

摘 要: 实体关系抽取能够从文本中提取事实知识, 是自然语言处理领域中重要的任务。传统关系抽取更加关注于单实体对的关系, 但是句子内包含不止一对实体且实体间存在重叠现象, 因此重叠实体关系抽取任务具有重大研究价值。任务发展至今, 总体可以分为基于序列到序列、基于图和基于预训练语言模型三种方式。基于序列到序列的方式主要以标注策略和复制机制的方法为主, 基于图的方式主要以静态图和动态图的方法为主, 基于预训练语言模型的方式主要以BERT挖掘潜在语义特征的方法为主。回顾该任务的发展历程, 讨论分析每种模型的优势及不足点; 结合目前研究的最近动态, 对未来的研究方向进行展望。

关键词: 重叠实体关系抽取; 深度学习; 图神经网络; 预训练语言模型

文献标志码: A **中图分类号:** TP391 **doi:** 10.3778/j.issn.1002-8331.2107-0359

Survey of Overlapping Entities and Relations Extraction

FENG Jun, ZHANG Tao, HANG Tingting

Key Laboratory of Water Big Data Technology of Ministry of Water Resources, School of Computer and Information, Hohai University, Nanjing 211100, China

Abstract: Entity relation extraction can extract knowledge of facts from text, which is an important task in the field of natural language processing. Traditional relation extraction pays more attention to the relation of a single entity pair. However, there are more than one entity pair as well as overlaps between entities. Therefore, it is important to research overlapping entity relation extraction. With the development of the task, there are three classes method: Seq2Seq, graph-based and pre-trained language model. Seq2Seq-based methods make use of labeling strategies and copy mechanisms, graph-based methods are mainly based on static graphs and dynamic graphs, and pre-training language models are methods mining potential semantic features on BERT. This paper reviews the development of the task, discusses and analyzes the pros and cons of each model. It combines with the recent developments of research to prospects for future research directions.

Key words: overlapping entities and relations; deep learning; graph neural network; pre-training language model

近年来大数据得到快速发展, 使得互联网上留存海量的无结构化或半结构化数据。如何通过深层次的挖掘这些数据得到更有价值的信息, 给用户带来更好的体验, 成为了这一领域学者的热点问题。在这种环境下, 信息抽取技术也成为了焦点, 信息抽取任务主要包括两大部分: 实体关系抽取(entity relation extraction)、事件抽取(event extraction)^[1]。而实体关系抽取作为信息抽取技术关键任务之一, 自然引起了广大学者的关注。

实体关系抽取能够从文本中提取事实知识, 是自然语言处理(natural language processing, NLP)领域中重要的任务之一, 对知识图谱构建及其下游应用起着支撑作用, 例如决策支持、智能问答等。现阶段, 实体关系抽取可以分为两大类: 基于传统机器学习和基于深度学习^[2]。表1中整理了目前实体关系抽取方法的优缺点。

(1) 基于传统机器学习

基于传统机器学习的方法主要依赖于自然语言处理工具来提取特征, 对提取到的特征采用相对简单的方

基金项目: 国家重点研发计划(2018YFC0407901); 国家自然科学基金面上项目(61602151)。

作者简介: 冯钧(1969—), 女, 博士, 教授, CCF 会员, 研究方向为时空数据管理、智能数据处理与数据挖掘、水利信息化; 张涛(1997—), 男, 硕士研究生, CCF 会员, 研究方向为实体关系抽取, E-mail: zthohai@hhu.edu.cn; 杭婷婷(1986—), 女, 博士, CCF 会员, 研究方向为知识图谱构建、信息抽取。

收稿日期: 2021-07-21 **修回日期:** 2021-10-28 **文章编号:** 1002-8331(2022)01-0001-11

表1 实体关系抽取方法比较

Table 1 Comparison of entity relation extraction methods

类别	方法	优点	不足
基于机器学习的 实体关系抽取	有监督	应用特定领域,具有较高精准率	依赖人工标注语料,需要大量预处理工作,耗费大量人力
	半监督	一定程度上降低对语料的依赖,更好地适应大规模语料训练任务	存在语义漂移问题,易受初始种子影响
	无监督	不依赖预定义关系、标注数据,具备很好的领域移植性	存在难以描述关系和低频实体召回率低问题
基于深度学习的 实体关系抽取	有监督	依赖人工标注数据,具有较高精准率	训练出来的模型对数据依赖程度高,模型的跨领域泛化能力低
	远程监督	利用远程知识库,极大减少人力成本,领域迁移性强	错误标签和误差传播问题

法就可以获得较好的效果。基于机器学习的方法以数据集标注量又可以划分为有监督、半监督和无监督的方法。①有监督的方法:利用人工标注语料构建语料库,将关系抽取任务视作分类任务,对识别到的候选实体预测关系。常用方法主要分为基于特征向量的方法^[3]和基于核函数的方法^[4]。②半监督的方法:利用少量人工标注的数据构建种子,通过训练大量未标记的语料库,生成新的实例进而增加新的种子,不断迭代,扩大种子规模,从而得到预期的效果^[5]。常用的算法主要有自举(boost strapping)^[6]、协调训练(co-training)^[7]和标注传播(label propagation)的方法^[8]。③无监督的方法:虽然半监督的方法仅使用少量的语料,但是种子的好坏会极大影响后续的结果,在此问题基础上,研究者采用基于聚类的思想,通过自底向上的方法从语料库中抽取出具具有相似性关系的实体对,在从句中选取合适的词语标记关系^[9]。

(2) 基于深度学习

由于传统的机器学习方法依赖自然语言处理工具抽取特征,可能会引起特征提取误差传播问题,极大影响关系抽取问题。为了能够自动化地抽取特征,深度学习的方法受到研究者的关注^[10]。深度学习关系抽取目前分为两大类有监督和远程监督。①有监督的方法:利用人工标注的数据集,通过深度学习的神经网络模型,自动获取句子特征,进而得到较好的训练效果。按照实体关系抽取的子任务完成顺序不同可以分为流水线(pipeline)方法^[11]和联合(joint)学习方法^[12]。②远程监督的方法:利用远程知识库自动化抽取大量的实体对,迅速的扩大知识库的规模。自动化抽取假设句子中实体对的关系与知识库中该实体的关系,但是这种数据标注的过程会带来大量的噪声数据,在抽取过程会产生误差传播^[13]。

基于深度学习的联合抽取方法很好的解决了机器学习方法中存在的误差传播问题,并且考虑到实体识别和关系抽取任务之间的依赖关系。但是大部分的模型将任务限定在单一句子中,并假设句子中只包含一个实体对。然而,这一假设是不切合实际的,句子中包

含了更为复杂的事实关系,不同的关系三元组在一个句子中可能有重叠。事实上,在WebNLG数据集68.2%的句子存在重叠,这会导致模型提取三元组不是很精准。现阶段可以按照实体重叠程度划分为三种类型,包括正常类(normal)、单实体重叠(single entity overlap, SEO)和实体对重叠(entity pair overlap, EPO),如图1所示。S1是属于正常类,其中只包含了一个三元组;S2属于SEO,句中的实体Tennessee同时与实体Martha和Monteagle存在关系,两个三元组发生了重叠;S3属于EPO,句子中的实体对<Jackson Heights, Queens>发生了重叠。

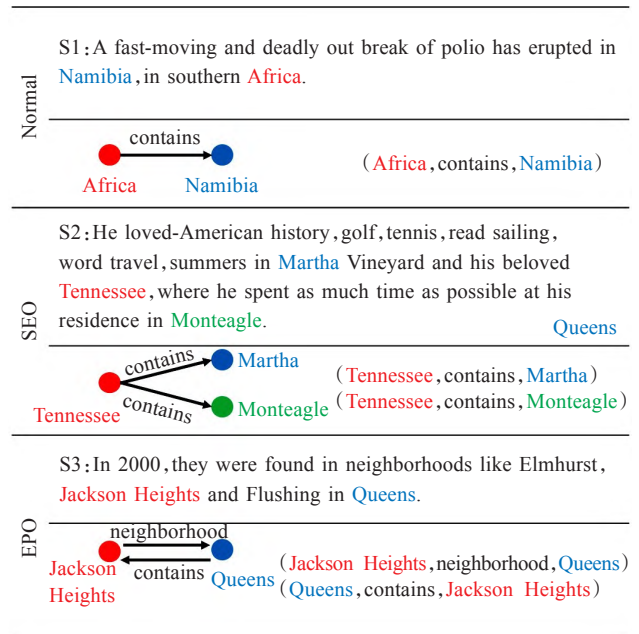


图1 重叠类型分类

Fig.1 Classification of overlapping types

从图1中的实例可以看出,重叠实体关系抽取相比经典关系抽取更加具备挑战性,主要包含以下5个挑战:

(1)暴露偏差,模型采用Seq2Seq的方法,但是实际上任务间仍旧存在级联关系。在实际训练阶段,级联任务均采用真实数据标签输入进行训练,会导致预测阶段和训练阶段的数据出现分布偏差,并产生误差积累。

(2)交互缺失,模型在建模过程中忽略的内在联系可以分为:头实体和尾实体、实体和关系、关系和关系以及三元组间的内在联系。有效建模三元组内在联系,可以减少模型对数据量的要求,减少模型计算负担。

(3)实体完全提取,大多数模型采用部分匹配的方法评估模型性能,当头实体最后一个单词、尾实体和关系被识别正确时,则视作三元组正确。这种评估方式不符合事实且无法应用到实际场景,如何在提取完整实体的同时不影响模型性能,仍然是要面临挑战之一。

(4)嵌套实体识别,句子中存在嵌套实体以及复杂的关系,使得句子十分复杂且不易解析。如何有效识别实体边界以及复杂的关系是现阶段仍面临的挑战之一。

(5)图模型构建,图模型很适用于实体关系提取任务,图的结构在模型整体性能中起着重要作用,而构建高质量和特定任务的图需要具备良好的领域知识。如何动态构建自适应图是极具挑战性的任务。

本文将针对重叠实体关系抽取任务,梳理从经典关系抽取任务到重叠实体关系抽取任务的脉络,对其中作者的核心思想、使用的方法以及方法优势和不足点进行总结,进而为接下来的研究方向提出可能性的建议以及未来可能改进的研究方向。据目前来看,这是针对重叠实体关系抽取的第一个全面综述。

1 重叠实体关系抽取方法

2011年,Hoffmann等人^[8]首次在模型构建过程中考虑到重叠实体关系问题,并引入基于多实例学习的概率图模型解决此问题。经过多年后,重叠实体关系抽取这一问题重要性逐渐被研究者意识到并采用基于序列到序列(sequence to sequence, Seq2Seq)、基于图(graph-based)和基于预训练语言模型(pre-trained language model, PTM)的方法进行深入研究。

1.1 基于序列到序列的方法

基于Seq2Seq的方法最初由Sutskever等人^[14]提出,采用编码-解码架构解决机器翻译问题,是自然语言处理领域中运用最广泛的框架之一,如图2所示。现在应用到关系抽取领域,该架构将非结构化文本转化为词向量作为输入序列,通过编码获取上下文表示,再经过解码,将关系三元组顺序输出。2017年,Zheng等人^[15]提出了一种新颖的标注策略(novel tagging),标注实体信息同时标注关系类型,以此将命名实体识别和关系抽取任务转变为序列标注任务。采用简单的端到端神经网络模型就可以有效的抽取三元组,提高了实体关系抽取任务的 $F1$ 值,达到了42.0%。但该模型忽略了SEO和EPO的情况,在此之后,学者对重叠实体关系抽取任务展开了进一步研究,现按照任务解决方案可以划分为:基于标注策略和基于复制机制。

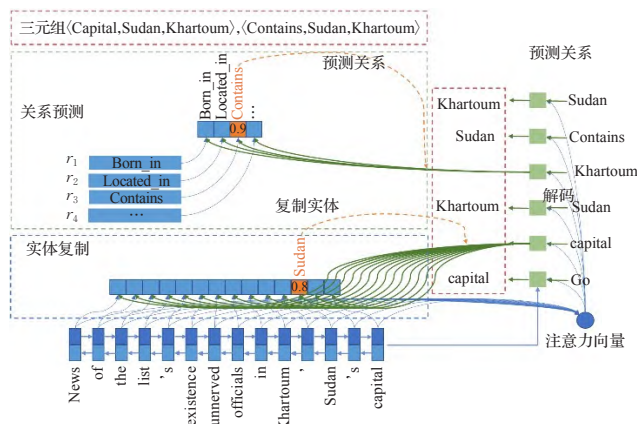


图2 基于序列到序列的经典模型框架

Fig.2 Traditional model framework based on Seq2Seq

1.1.1 基于标注策略

基于标注策略的解决方案主要是受到 novel tagging 模型的启发,将实体关系抽取两阶段任务转变为对实体和关系开始位置和结束位置的共同标注,从而实现联合解码。后续学者经过巧妙调整标注策略使得模型可以解决重叠实体关系抽取任务。现阶段标注多采用 BIESO 标注实体信息,分别表示{实体开始,实体内部,实体结束,单个实体,无关键词},关系类型采用预定义关系词进行标记。根据过往的工作总结,将标注策略按照抽取顺序分为头实体优先抽取、关系优先抽取、三元组建模抽取。

(1)头实体优先抽取

Yu等人^[16]提出了一种新颖的分解式策略,将抽取任务分解为头实体(head entity, HE)提取、尾实体关系(tail entity relation, TER)提取,两个任务共享编码层。通过引入基于跨度距离的标记方案,进一步将两个子任务转变为多序列标注问题,对于句子中的 HE 标记其位置和类型,再根据 HE 去预测 TER,TER 包括两个子序列标注任务,标记尾实体(tail entity, TE)开始位置和关系类型,标记 TE 结束位置和关系类型,最终在 WebNLG 数据集中 $F1$ 值达到了 83.1%。田佳来等^[17]人采用分解策略思想的同时引入了 GLU dilated CNN 对句子进行编码,采用自注意力机制提高模型抽取能力,并成功将模型运用到军事领域,其 $F1$ 值达到了 80.84%。分解策略,头实体、尾实体经过两次解码获得,并未考虑到头尾实体间的相互作用,基于此赵等人^[18]使用多标签策略处理重叠实体关系问题,并在编码层加入了词性信息和句法依存关系,以及引入多头注意力机制,在 NYT 数据集 $F1$ 值达到了 66.7%。头实体优先抽取,再判定实体对之间的关系,会产生大量的冗余实体,并进行了许多不必要的操作。为了进一步探索更好的抽取方式,部分学者采用了关系优先抽取。

(2)关系优先抽取

Lin等人^[19]提出了一种基于关系矩阵约束下的关系

提取模型 RERLC, 该模型认为在同句子内, 相关性强的关系对出现的概率更高。将关系转变为向量, 经过计算关系向量间的夹角计算相关度, 从而建立关系矩阵。在序列标注的过程中获取头实体, 再经关系矩阵计算尾实体。但是, 该方法缺乏建模关系和实体间的相互作用, 并且模型泛化能力不强; 基于此, Zhuang 等人^[20]在关系优先的基础上, 将预测的关系与文本结合进行实体识别, 结合实体对和关系获得关系。并分析数据集, 获得数据中特定的规则, 将规则加入到三元组解码层, 进一步提升原有模型的性能, 该模型结合文本, 引入数据增强, 提高了模型泛化能力, 但是会因为关系标注错误引起误差传播; 在此之上, Yuan 等人^[21]构建了关系注意力网络提取三元组, 首先采用 Bi-LSTM 对句子进行编码, 再通过关系注意力网络获取不同关系下句子特征表示, 经过关系门过滤负样本关系后, 提取到对应的实体。当指定关系识别出多个实体时, 采用启发式规则, 将距离相近的实体自动结合成三元组。该方法通过关系门避免了误差传播, 但该模型在 EPO 问题效果不佳; Liu 等人^[22]相较于文献[21], 设计了一种多头自注意力网络, 充分获取不同关系标签下的关联强度。将检测到的关系类型与实体提取模块的结果融合, 经过预先定义的阈值判断三元组。Luo 等人^[23]将 SEO 进一步划分为 ELS (exclude loop sentences) 和 ILS (include loop sentences), 从而引入了双向树标记的方案。首先, 将句子中具有相同关系的三元组分组在一起, 按照三元组中实体和关系出现的顺序建模成二叉树结构, 最后, 建立二叉树结构和序列标签之间的映射, 从而取得了良好的效果。双向二叉树虽然结构新颖, 但是该结构难以处理 EPO 问题, 模型性能受结构限制。Ma 等人^[24]设计了一种级联双解码联合提取模型, 该模型首先检测文本中存在的关系, 并将它们视为额外的映射, 提取每个关系相应的 HE 和 TE, 结果有着明显的提升。不管是头实体优先抽取或者是关系优先抽取, 都是将关系视作离散标签分配给实体。当数据量不充足情况下, 分类器很难准确分类, 这样会导致模型训练效果不佳。在此问题之上, 出现了三元组建模抽取的方式。

(3) 三元组建模抽取

在 2020 年 ACL 会议上, Wei 等人^[25]设计了级联二进制标记框架 CaseRel, 使模型学习给定关系下 HE 到 TE 之间的映射函数, 从而达到对三元组整体建模的效果。主要流程如下: 首先, 经过两个二进制分类器 (开始位置分类器、结束位置分类器) 标识句子中所有可能的实体。然后, 检查所有可能的关系, 在指定关系下用分类器标识尾实体, 每次标记前会将之前标记的数据清零。该模型在 WebNLG 数据集上 F1 值达到了 91.8%, 对比当时最好的模型有着 30.2% 的提升。该模型设计的通用算法框架, 为重叠实体关系抽取任务引入新的解决方

法, 但是该模型需要对实体进行剪枝, 避免复杂度过高; Wang 等人^[26]在 CaseRel 基础上使用 IDCNN (iterated dilated CNN) 和 Bi-LSTM 对句子进行编码, 并采用注意力机制, 以获得句子中更丰富的语义特征, 来提升模型的效果; Wang 等人^[27]引入了一种新颖的握手标记策略, 针对句子中的任意一个单词以及除它以外其他单词做以下判断: 是否同为一个实体的开始和结束, 给定关系下是否分别为两实体的头部, 给定关系下是否为实体的尾部, 根据以上判断的结果进行标记。该策略通过回答问题的方式很好的解决了复杂实体在重叠实体关系抽取任务中的影响, 在 NYT 数据集 F1 值达到了 92.0%。

1.1.2 基于复制机制

复制机制模仿人类在背诵课文时, 对于难以理解的语句进行死记硬背的方法。在实际应用过程中, 复制机制结合 RNN 或 LSTM 等模型获取到想要复制片段语义特征, 在输入序列中定位位置, 将其直接复制到输出序列中, 以应对特定任务。基于复制机制的方法按照复制内容分为单实体识别复制和多实体识别复制。单实体识别复制指的是在复制的过程中遇到拥有多个单词的实体时, 仅复制实体的最后一个单词。而多实体识别在复制过程中可以完成复制整个实体。

(1) 单实体识别复制

Zeng 等人^[28]提出了一种基于复制机制的端到端模型 CopyRe, 该模型先提取关系再提取实体, 通过复制实体让实体参与到不同的三元组, 并针对不同情况采用了不同的解码策略: 使用单解码器生成所有的三元组序列; 使用多解码器, 其中的每一个解码器对应一个三元组, F1 值提高了 4.82%; 2019 年, Zeng 等人^[29]在自己模型的基础上, 认为提取句子中三元组顺序对抽取结果有着至关重要的影响。在实际训练过程中, 已提取出的三元组作为先验知识会极大影响剩下三元组提取。考虑到以上问题, 在任务中引入了强化学习, 对解码过程中产生的三元组与已有标记三元组进行比较, 以此设定奖励机制, 对模型进行多次迭代, 模型效果得到了提升。

不管是 CopyRe 还是 CopyRe-RL 模型, 在识别和复制实体时仅涉及到实体的最后一个单词, 会极大影响关系类型提取。在 NYT 数据集中 77.44% 的实体不仅仅包含一个词汇, 仅复制单实体是不切实际的, 有必要将任务拓展到多实体识别复制。

(2) 多实体识别复制

Bai 等人^[30]通过多层神经网络进行编码并结合自注意力机制, 设计了双指针网络结构分别识别实体的开始位置和实体的结束位置, 从而让完整的实体参与到复制过程中, 在 NYT 数据集上 F1 值提升了 18.9%, 但是, 该模型并未考虑实体对和关系之间的相互作用; 此外, Zeng 等人^[31]从理论层面剖析了 CopyRE 出现的问题, 采用了非线性激活函数解决首尾实体顺序混乱问题, 引入

多任务学习框架解决实体识别不全的问题,在当时的多个数据集上达到了SOTA,但是该模型预测三元组数量受限制。为了进一步探索,Wang等人^[32]采用RCNN对句子进行编码,通过设定变量阈值避免复制实体过程中HE和TE相同,并将模型成功的应用在了生物领域, $F1$ 值提升了14.4%。Nayak等人^[33]针对于复制机制改进了两种方法:①提出一种新的实体关系表示方案,三元组内采用“;”分隔元组内容,三元组间采用“|”分隔。在解码过程中每步产生一个单词。但是解码器在解码过程中会预测到不存在于当前句子或关系集的标记,为了保证仅仅复制源句子中的实体,引用了一种排除机制,将没有特殊标记的单词相应的SoftMax分数置为0;②考虑到编码器每一次只提取一个单词运行速度较慢且浪费资源,提出了一种基于指针网络的解码框架,该框架用五元组标记实体,分别是<HE开始位置,HE结束位置,TE开始位置,TE结束位置,关系类别>,并且在解码过程中每次提取整个元组。

1.2 基于图的方法

虽然Seq2Seq已经成为NLP中处理各种任务的主要方法,但是有各种各样的NLP问题可以用图形结构来更好地表达。因此,人们对NLP领域中开发新的图卷积神经网络(graph convolutional network, GCN)技术兴趣激增,其中采用的经典模型框架如图3所示。基于图的方法按照图结构又可以划分为静态图方法和动态图方法。静态图指预定义的图结构,动态图指模型在训练过程中通过注意力机制动态的增加节点来完善图。

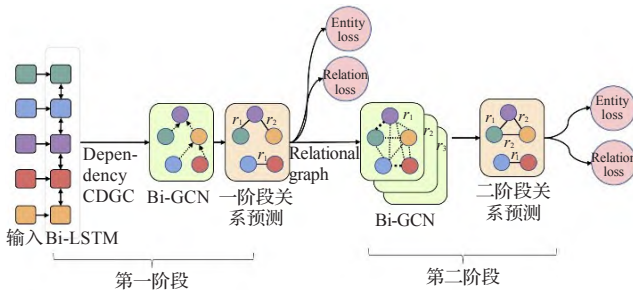


图3 基于图的经典模型框架

Fig.3 Traditional model framework based on graph

(1) 静态图

2018年,Wang等人^[34]为了解决重叠实体关系抽取任务,提出了基于图结构的联合抽取模型,该结构包含一个新颖的转换系统识别实体和关系,并通过权重损失函数建立实体与关系和关系与关系之间的依赖,最终 $F1$ 值达到了50.9%,该模型最早引入图结构解决问题,但是提出的转换系统泛化能力不强;Hong等人^[35]利用BiLSTM获取实体和关系的表示,并设计了一种新的关系感知注意力机制和关系感知GCN,用来获得两个实体跨度之间的关系特征,并利用实体跨度构建图结构,模型的 $F1$ 值达到了72.8%,但是该模型存在实体跨度识

别不全的问题;在此基础上,Wang等人^[36]结合边界预测实体的开始和结束位置,以加强实体跨度表示,使用注意机制来获得两个实体跨度之间关系的特征,构成跨度图,模型最终的 $F1$ 值达到了74.9%;静态图采用预定义图结构,使得模型性能受限,因此开展了动态图的研究。

(2) 动态图

Fu等人^[37]利用依存句法分析将输入句子转变为依存树,将树的邻接矩阵输入到BI-GCN中获取局部特征,从而分别提取到实体和关系,为了形式化计算实体和关系间的相互作用程度,引入了加权GCN,对每个已提取到的关系,计算任意实体对之间边(关系)的权重,从而解决了重叠实体关系问题, $F1$ 值提高了11.0%,依存树直接影响模型的性能,并且模型计算负担大。Fei等人^[38]通过跨度注意层(span attention)获取所有候选实体表示并为每个关系形成跨度图。另外,将这些跨度图输入到图注意力模型(graph attention model, GAT)以动态的学习实体跨度与其关系边的相互作用;此外,Duan等人^[39]提出了一种基于多头自注意力和紧密连接的图卷积网络MA-DCGCN。在该模型中,多头自注意力机制专门用于将权重分配给实体之间的多个关系类型,以确保多个关系的概率空间不相互排斥,并自适应的提取重叠实体间的多种关系。上述模型在将句子转换为图结构时,将实体跨度视为节点,关系视为边,通过GCN对每个实体对间的关系进行了量化,也导致模型在EPO问题上表现不佳。

1.3 基于预训练语言模型

在2018年10月,谷歌提出预训练语言模型^[40],它经过在大规模未标记的语料库上运行无监督学习获得好的特征表示。根据特定的NLP任务,对PTM进行微调就能提取到很好的特征表示,并且已经在NLP领域中获得了广泛的应用,基于预训练语言模型的框架如图4所示。现阶段使用最广泛的预训练语言模型就是BERT(bidirectional encoder representations from transformers),在微调过程中BERT使用预训练的参数进行初始化,然后使用来自下游任务的标记数据对所有参数进行微调。每个下游任务都有单独的微调模型,即使它们使用相同的预训练参数进行初始化。

Li等人^[41]利用BERT模型最后两层输出的隐藏层信息,构建二维矩阵表示特征,在矩阵中掩盖无关实体,获取实体位置信息。对于输入语句中存在的每种关系类型,每个实体对,独立计算实体对是否存在特定关系的概率,以预测句子中存在的多种关系。虽然模型最终的效果良好,但是,该模型时间复杂度高,应考虑引入规则减少计算负担;Sui等人^[42]将联合实体关系抽取任务转变为集合预测问题,从而使减少模型预测三元组顺序带来的负担,并率先引入了非自回归编码器(non-autoregressive decoder)结合二部图匹配损失函数(bipartite

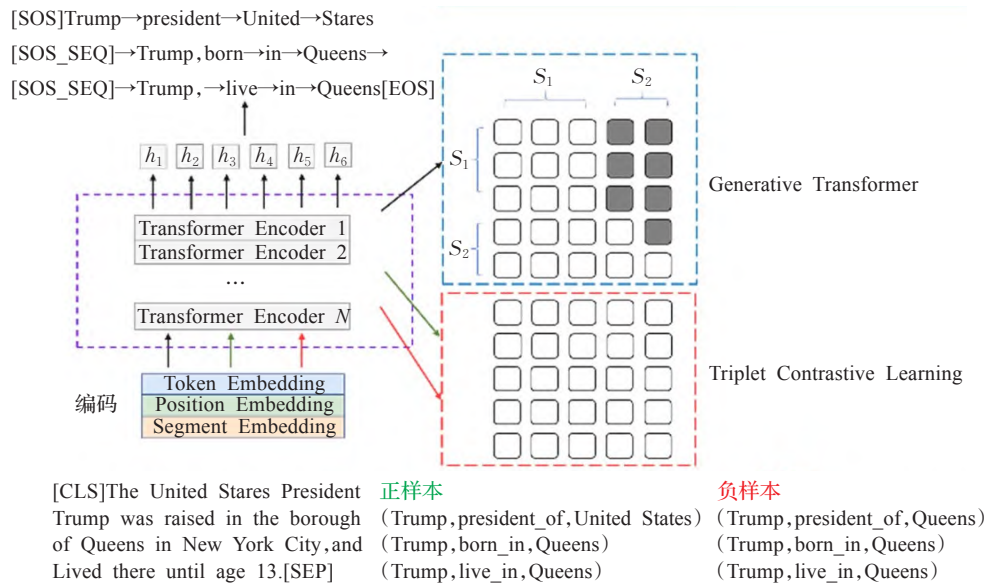


图4 基于预训练语言模型的框架

Fig.4 Framework based on pre-training language models

match loss),使得模型直接输出最终的三元组。同样的,为了直接生成三元组,Liu等人^[43]探索了一种新颖的模型架构,该模型将transformer和指针网络组合以提取实体和关系。此外,引入了语法引导的网络(syntax-guided network)以明确地将句子的语法信息纳入编码器,帮助模型更加关注句子的重要单词。直接生成三元组的模型,虽然减少了模型预测三元组顺序带来的负担,但是模型对数据需求更高,并且缺少对抽取过程的思考,使结果显得并不可靠;在此基础上,Ye等人^[44]用BERT对句子编码解码,使用部分因果掩盖区分开编码解码的表示,使用波束搜索算法生成更多的三元组。在此基础上,提出了一种三元组对抗训练方法,在训练阶段捕获正样本和负样本的特征,学习到的校准算法在验证阶段会过滤掉不符合源句子事实的三元组,提高了结果的可靠性;Hang等人^[45]设计了三个步骤提取重叠实体关系,由BERT获取句子上下文信息共享给下游两个子任务,命名实体识别任务将获取到的参数转换为对应实体标签的概率分布,实体关系抽取任务将获取到的参数转换为相应关系类型的概率分布,以上步骤使得模型在性能上有了提升。但是,模型通过共享参数的方法对下游任务进行训练,更容易产生误差传播。

BERT通过获取双向上下文特征信息,并可以将训练好的模型通过微调应用到下游任务,在原有的基础上获得了很大的提升,但是BERT仍存在许多不足之处。

(1)无法处理长文本,BERT是由12层transformer搭建而成,对于长度超过512字符的句子BERT处理不了,而在实际应用过程中经常出现超过字符限制的长句子,会对抽取任务性能产生影响。

(2)BERT在预训练阶段会随机掩盖句子中15%的词,利用上下文对掩盖掉的词进行预测,加深对上下文

信息记忆的程度。当采用BERT下游任务进行微调时,文本中存在噪音数据,会导致预训练阶段和微调过程存在偏差,影响模型的性能。

2 数据集和评价指标

近年来,用于重叠关系抽取任务实验评估的数据集主要包括NYT纽约时报数据集、WebNLG数据集。

2.1 数据集

NYT数据集:NYT数据集最初是由远程监督的方法生成的新闻语料库,广泛用于重叠实体关系抽取任务。总共有24中预定义的关系,其中长度超过100和不存在三元组的句子被过滤掉,最终训练集、测试集和验证集分别包含56 195、5 000、5 000。

WebNLG数据集:最初为自然语言生成任务而创建,但后来也应用于重叠实体关系抽取任务。这个数据集包含246种预定义的关系,数据集中的每一个句子都包含多个三元组,其训练集、测试集和验证集分别包含5019、500、703。

现阶段研究者使用的数据集是由文献[28]预处理过后发布出来的。如表2所示,本文统计了NYT和WebNLG数据集中分别属于Normal、SEO和EPO句子数量。特别注意的是,同一个句子可以即属于SEO又属EPO。从表

表2 数据集统计

Table 2 Statistics of datasets

类别	NYT		WebNLG	
	训练集	测试集	训练集	测试集
Normal	37 013	3 266	1 596	246
EPO	9 782	978	227	26
SEO	14 735	1 297	3 406	457
Sentence	56 195	5 000	5 019	703
Relation	24		246	

中可以看出, NYT 数据集中 34.13% 的句子存在重叠现象, 其中 51% 的句子属于 EPO, 76.8% 的句子属于 SEO。WebNLG 数据集中 68.2% 的句子存在重叠现象, 其中 0.06% 的句子属于 EPO, 99.5% 句子属于 SEO。

2.2 评价指标

重叠实体关系抽取任务的评价指标包括: 精准率 (Precision)、召回率 (Recall) 和 F 值 (F Measure)。Precision 是从查准率对模型效果进行评估, 其公式如下:

$$Precision_R = \frac{\text{被正确抽取属于关系 } R \text{ 的三元组个数}}{\text{所有被抽取属于关系 } R \text{ 的三元组个数}} \quad (1)$$

Recall 是从查全率对模型效果进行评估, 其公式如下:

$$Recall_R = \frac{\text{被正确抽取属于关系 } R \text{ 的三元组个数}}{\text{实际应被抽取属于关系 } R \text{ 的三元组个数}} \quad (2)$$

其中 Precision 和 Recall 是互补关系, 因此通过 F 值综合考虑 Precision 和 Recall 对模型总体进行评估, 其公式

如下:

$$F_\beta = \frac{(\beta^2 + 1) \cdot Precision \cdot Recall}{\beta^2 \cdot Precision + Recall} \quad (3)$$

β 是调节 Precision 和 Recall 比重的重要参数, 在实际应用过程中, 认为 Precision 和 Recall 同样重要, 所以 $\beta = 1$ 。所以上式表示为:

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (4)$$

表 3 中收集了不同数据集下模型表现, 以及模型优势以及不足点。表中大部分模型在 WebNLG 数据集都能取得良好的效果, 而在 NYT 数据集则相对不佳。WebNLG 数据集特点就是 SEO, 则说明现有的大部分模型都能很好地处理 SEO, 而在 EPO 表现不是理想。从另一方面说明, 模型需要更加庞大的数据去处理多标签实体。从表中可以看出, 对三元组整体建模的模型效果一般优于缺少三元组内在联系考虑的模型, 采用异质图模

表 3 不同数据集及模型性能

Table 3 Different performance of datasets and models

类别	模型	优势	不足	NYT/WebNLG F1 值/%	发表 年份	发表会议/期刊
基于标注 策略	ETL-Span ^[16]	引入分解策略和跨度策略, 捕捉语义依赖关系, 减少噪声	头尾实体交互缺失	78.0/83.1	2020	ECAI
	RERLC ^[19]	关系矩阵, 以向量来表示关系	关系和三元组的交互缺失, 模型泛化能力不强	82.1/86.7	2020	ICCC
	RFTE ^[20]	引入数据增强和规则	关系标注引起误差传播	89.7/91.9	2020	ISPA
	RSAN ^[21]	关系注意力网络, 关系门过滤掉不相关关系	EPO 效果不佳	84.6/82.1	2020	IJCAI
	AttentionRE ^[22]	多头关系注意力网络	实体和三元组间的交互缺失	83.0/87.7	2020	IJCAI
	CASREL ^[25]	二进制框架对三元组整体建模	需要进行实体剪枝	89.6/91.8	2020	ACL
	TP-Linker ^[27]	握手框架, 解决嵌套实体	计算复杂度高	91.9/86.7	2020	ACL
复制机制	CopyRe ^[28]	复制实体让实体参与到不同的三元组, 两种解码策略	模型依赖分词工具	58.7/37.1	2018	ACL
	CopyRe-RL ^[29]	考虑三元组抽取顺序, 引入强化学习	训练时间过长, 实体识别不全	72.1/61.6	2019	ACL
	DPointer ^[30]	双指针网络标记实体开始和结束位置	实体和关系之间的交互缺失	70.9/38.1	2019	Neurocomputing
	CopyMTL ^[31]	多任务学习框架	三元组数量受限	72.0/56.4	2020	AAAI
	PNDcc ^[32]	三元组新的表示方式, 减少内存消耗和运行时间	存在实体和关系不匹配的情况	78.9/70.5	2020	AAAI
图模型	GraphRel ^[37]	加权 GCN	模型性能受依存树影响	61.9/42.9	2019	ACL
	MA-DCGCN ^[39]	多头注意力结合 GCN, 对每个关系建模	关系之间的交互缺失	79.4/66.3	2020	IJCAI
	RIFRE ^[50]	异质图网络不断迭代融合单词节点和关系节点	考虑单词和关系的交互而不是实体与关系	92.0/92.6	2021	Knowledge-Based Systems
预训练语言模型	BSTP ^[43]	Transformer 语法引导网络	模型对数据需求高	87.2/74.9	2021	Journal of Intelligent & Fuzzy Systems
	CGT ^[44]	对比训练方法训练三元组可靠性	模型结构较为复杂	89.1/83.4	2021	AAAI
	BERT-JEORE ^[45]	BERT 共享编码联合提取	共享参数的方法, 更容易产生误差传播	86.5/84.8	2021	Expert Systems with Applications

型的模型效果要大幅度优于同质图。为了进一步展现不同实体重叠程度下模型性能,收集了其中一部分模型进行对比,如图5所示。从图5中可以看出,在NYT数据集中SEO、EPO表现最佳的模型分别为TP-Linker^[27]、RFTE^[20],在WebNLG数据集中SEO、EPO表现最佳的模型分别为RFTE^[20]、TP-Linker^[27]。

3 未来研究方向

目前,重叠实体关系抽取任务虽然已经取得了巨大的成功,但依旧存在着提升空间,值得学者们进一步探索。通过对现有工作进行总结,未来的研究方向包括以下几个方面。

3.1 基于序列到序列的模型

基于序列到序列的模型,从标注策略到复制机制,本质上都是调整抽取策略带来性能的提升,并且现有模型的解决思路已经达到了一个较高的标准,很难获得性能上的大幅度提升。而强化学习不同于现有的有监督模型,它通过反馈机制更新模型。但是,目前只是在重叠实体关系抽取领域进行了简单的尝试,拥有着无限可能。

最初于2018年,Feng等人^[46]使用强化学习有效过滤远程监督数据中的噪声,后经文献[29]转换奖励机制使得强化学习有效解决重叠实体关系抽取问题。最为

典型的是Takanobu等人^[47]于2019年AAAI会议上提出的模型,与其他人不同的是该模型引入强化学习建模实体和关系间的交互,利用层次结构解决重叠问题。高层次强化学习用于基于句子中的特定关系令牌来标识关系。在识别了关系之后,使用低级RL来利用序列标记方法提取与关系相关联的两个实体。多次重复该过程以提取句子中存在的所有关系三元组。

总之,强化学习在重叠实体关系抽取领域已经进行了初步的尝试,并且提供了具有可行性的解决方案,但仍然有着很大的提升空间值得学者探索。

3.2 图神经网络模型

研究表明,重叠实体关系抽取任务在引入图神经网络模型后,模型性能有了很不错的提升。相比于其他传统模型,图结构能更好地挖掘句子中的语义特征,但是根据已有模型的总结,大多数模型采用同质图建模。传统的同质图只含有同种类型的边和节点,实际任务中的实体类型多样、关系复杂。在这种情况下,同质图很难学习到关键信息,会导致模型在EPO问题上表现不佳。相较于同质图,异质图可以容纳更多的节点和边的类型,在建模过程中学习到潜在的异构特征,提升模型性能,所以将同质图扩展异质图是必然选择。常用的异质图网络包括HAN(heterogeneous graph attention network)^[48]、HetGNN(heterogeneous graph neural net-

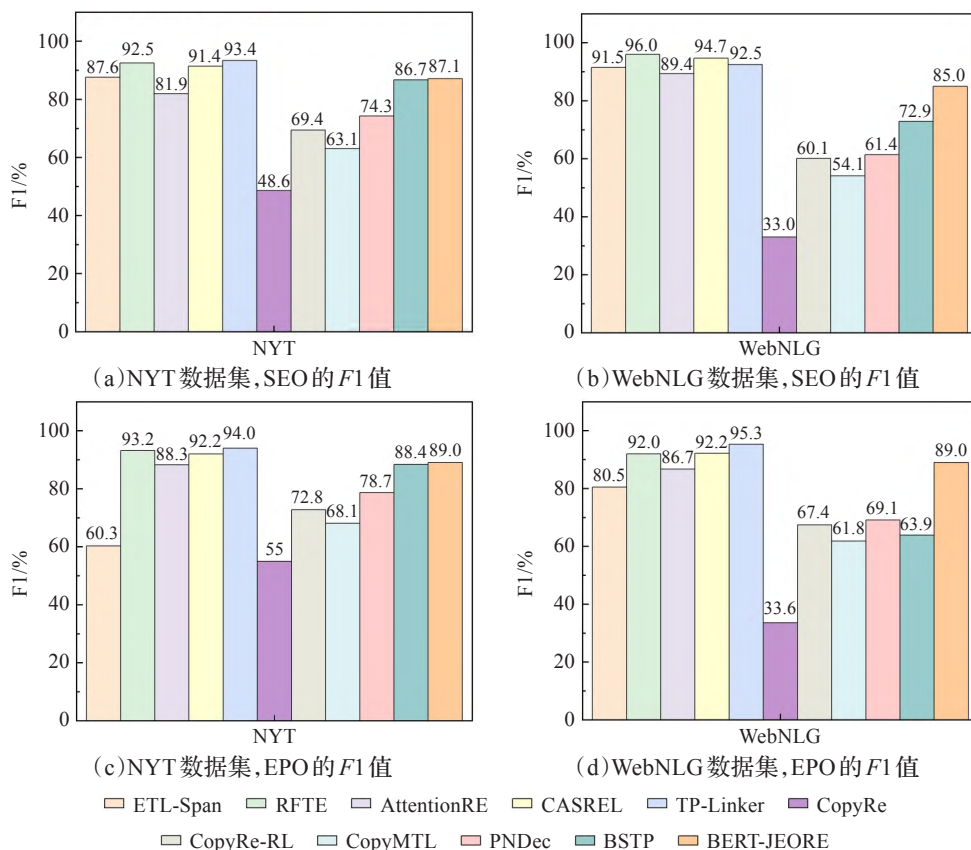


图5 不同重叠类型句子的F1值

Fig.5 F1 score on sentences with different overlapping types

work)^[49]。Zhao 等人^[50]率先引入异质图神经网络,将关系和单词同时视为图表上的节点,然后通过消息传递机制迭代融合来自不同节点的信息,在获得单词和关系的节点表示后,进行关系提取步骤生成最终的三元组。但是异质图动态地构建图结构更具有挑战性,因为更多类型特征需要从数据中获取。因此,图模型是未来热门的研究方向之一。

3.3 预训练语言模型

BERT 模型在重叠实体关系抽取领域中已经取得了惊人的成绩,但仍有着许多可以研究的方向。其中包括:模型精细化调参、模型结合知识图谱和改进模型。

模型精细化调参:BERT 模型在通用领域语料库进行预训练,如果让 BERT 参与到特定领域的下游任务,需要对其进行精细化的调参,让其发挥更好的性能。现有的调参策略包含:长文本处理、层数选择、层间学习率设置^[51]。

模型结合知识图谱:BERT 通过掩码策略经过预训练可以学习到丰富的语义,但是模型并未学会如何表示知识,而结合知识图谱使模型可以学习到知识,并且对知识的表示更加贴合真实世界^[52]。

改进模型:提升预训练语言模型规模,其中最为典型的模型是 RoBERTa, RoBERTa 相较于 BERT 具有更大的训练数据,每批次数据量更大,并且预训练过程采用动态掩码的方式,与之对应的训练时间也更大,但是模型最终也取得了不错的提升^[53]。

总之,现阶段预训练语言模型改进方法,已经有学者进行了初步的探索,但是在重叠实体关系抽取领域还未被广泛利用,改进预训练语言模型仍然是未来几年的研究热点。

3.4 知识推理

知识推理的方法是非常具备潜力的研究方向之一,但目前对于重叠关系抽取任务的探索还比较少, NYT 和 WebNLG 数据集集中都具备着十分复杂的关系以及隐式三元组存在,十分符合知识推理的应用条件。而现有抽取方法主要专注于可以直接发现的显式三元组,但通常忽略缺乏明确表达的隐式三元组,这将导致构建的知识图谱严重不完整。实际上,句子中的其他三元组提供了用于发现具有隐式关系的实体对的补充信息。此外,可以在现实世界中以关系推理模式识别隐式实体对之间的关系类型。在 2021 年 ACL 会议上, Chen 等人^[54]提出了一个统一的框架,共同提取显式和隐式关系三元组。引入了一个二进制指针网络,探索可以通过关系隐式连接的实体对,以顺序地提取与每个单词相关的重叠三元组,并保留在外部存储器中先前提取的三元组的信息。为了推断隐式三元组的关系类型,在模型中引入实际关系推理模式,并通过关系网络捕获这些模式。

总之,研究知识推理在重叠关系抽取的应用,可以使模型发现更多三元组的存在,让知识图谱构建的完整度更高,同时大大降低模型对标注样本的依赖程度,在实际应用方面具备重大意义。

4 结束语

本文详细介绍了重叠实体关系抽取研究的最新进展,包括基于 Seq2Seq、基于图和基于预训练语言模型的方法,详细描述了其方法以及核心思想,阐述了模型中出现的问题。此外,本文在总结了现有模型的优缺点的同时,详细对比了模型在不同问题上表现的性能。虽然现阶段重叠实体关系抽取任务已经具备了很高的 F1 值,但是仍然存在面临的挑战;本文进一步总结了未来的研究方向来推进这一任务。

参考文献:

- [1] 鄂海红,张文静,肖思琪,等.深度学习实体关系抽取研究综述[J].软件学报,2019,30(6):1793-1818.
E H H, ZHANG W J, XIAO S Q, et al. Survey of entity relationship extraction based on deep learning[J]. Journal of Software, 2019, 30(6): 1793-1818.
- [2] 李冬梅,张扬,李东远,等.实体关系抽取方法研究综述[J].计算机研究与发展,2020,57(7):1424-1448.
LI D M, ZHANG Y, LI D Y, et al. Review of entity relation extraction methods[J]. Journal of Computer Research and Development, 2020, 57(7): 1424-1448.
- [3] 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.
- [4] ZELENKO D, AONE C, RICHARDELLA A. Kernel methods for relation extraction[J]. Journal of Machine Learning Research, 2003, 3: 1083-1106.
- [5] CHAPPELLE O, ZIEN A. Semi-supervised learning[M]. Cambridge: MIT Press, 2006: 3-14.
- [6] ZHU J, NIE Z, LIU X, et al. StatSnowball: A statistical approach to extracting entity relationships[C]//Proceedings of the 18th International Conference on World Wide Web, 2009: 101-110.
- [7] ZHANG Z. Weakly-supervised relation classification for information extraction[C]//Proceedings of the 13th ACM International Conference on Information and Knowledge Management, 2004: 581-588.
- [8] HOFFMANN R, ZHANG C, LING X, et al. Knowledge-based weak supervision for information extraction of overlapping relations[C]//Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, 2011: 541-550.

- [9] 秦兵, 刘安安, 刘挺. 无指导的中文开放式实体关系抽取[J]. 计算机研究与发展, 2015, 52(5): 1029-1035.
- QIN B, LIU A A, LIU T. Unsupervised Chinese open entity relation extraction[J]. Journal of Computer Research and Development, 2015, 52(5): 1029-1035.
- [10] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313: 504-507.
- [11] BAI F, RITTER A. Structured minimally supervised learning for neural relation extraction[C]// Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2019: 3057-3069.
- [12] ZHENG S, HAO Y, LU D, et al. Joint entity and relation extraction based on a hybrid neural network[J]. Neurocomputing, 2017, 257: 1083-1106.
- [13] MINTZ M, BILLS S, SNOW R, et al. Distant supervision for relation extraction without labeled data[C]// Proceedings of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the AFNLP, 2009: 1003-1011.
- [14] SUTSKEVER I, VINYALS O, LE Q V. Sequence to sequence learning with neural networks[C]// Proceedings of Annual Conference on Neural Information Processing Systems, 2014: 3104-3112.
- [15] ZHENG S, WANG F, BAO H, et al. Joint extraction of entities and relations based on a novel tagging scheme[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 2017: 1227-1236.
- [16] YU B W, ZHANG Z Y, SHU X B, et al. Joint extraction of entities and relations based on a novel decomposition strategy[C]// Proceedings of the 24th European Conference on Artificial Intelligence, 2020: 2282-2289.
- [17] 田佳来, 吕学强, 游新冬, 等. 基于分层序列标注的实体关系联合抽取方法[J]. 北京大学学报(自然科学版), 2021, 57(1): 53-60.
- TIAN J L, LYU X Q, YOU X D, et al. Joint extraction of entities and relations based on hierarchical sequence labeling[J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2021, 57(1): 53-60.
- [18] 赵敏钧, 赵亚伟, 赵雅捷, 等. 一种新的基于深度学习的重叠关系联合抽取模型[J]. 中国科学院大学学报, DOI: 10.7523/j.ucas.2020.0026.
- ZHAO M J, ZHAO Y W, ZHAO Y J, et al. A new joint model for extracting overlapping relations based on deep learning[J]. Journal of University of Chinese Academy of Sciences, DOI: 10.7523/j.ucas.2020.0026.
- [19] LIN K, MIAO K, HONG W, et al. Relation extraction based on relation label constraints[C]// Proceedings of 2020 IEEE 6th International Conference on Computer and Communications (ICCC), 2020: 2166-2170.
- [20] ZHUANG C, ZHANG N, JIN X, et al. Joint extraction of triple knowledge based on relation priority[C]// 2020 IEEE International Conference on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking, 2020: 562-569.
- [21] YUAN Y, X ZHOU, PAN S, et al. A relation-specific attention network for joint entity and relation extraction[C]// Proceedings of the 29th International Joint Conference on Artificial Intelligence, 2020: 4054-4060.
- [22] LIU J, CHEN S, WANG B, et al. Attention as relation: Learning supervised multi-head self-attention for relation extraction[C]// Proceedings of the 29th International Joint Conference on Artificial Intelligence, 2020: 3787-3793.
- [23] LUO X, LIU W, MA M, et al. BiTT: Bidirectional tree tagging for joint extraction of overlapping entities and relations[J]. arXiv:2008.13339, 2020.
- [24] MA L, REN H, ZHANG X. Effective cascade dual-decoder model for joint entity and relation extraction[J]. arXiv:2106.14163, 2021.
- [25] WEI Z, SU J, WANG Y, et al. A novel cascade binary tagging framework for relational triple extraction[C]// Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020: 1476-1488.
- [26] WANG L, XIONG C, DENG N. A research on overlapping relationship extraction based on multi-objective dependency[C]// Proceedings of the 15th International Conference on Computer Science & Education, 2020: 618-622.
- [27] WANG Y, YU B, ZHANG Y, et al. TPLinker: Single-stage joint extraction of entities and relations through token pair linking[C]// Proceedings of the 28th International Conference on Computational Linguistics, 2020: 1572-1582.
- [28] 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, 2018: 506-514.
- [29] 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, 2019: 367-377.
- [30] BAI C, PAN L, LUO S, et al. Joint extraction of entities and relations by a novel end-to-end model with a double-pointer module[J]. Neurocomputing, 2020, 377: 325-333.
- [31] ZENG D, ZHANG H, LIU Q. Copymtl: Copy mechanism

- for joint extraction of entities and relations with multi-task learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence,2020:9507-9514.
- [32] WANG X, LI Q, DING X, et al. A new method for complex triplet extraction of biomedical texts[C]//Proceedings of the 12th International Conference on Knowledge Science, Engineering and Management, 2019:146-158.
- [33] NAYAK T, NG H T. Effective modeling of encoder-decoder architecture for joint entity and relation extraction[C]//Proceedings of the 34th Conference on Artificial Intelligence, 2020:8528-8535.
- [34] WANG S, ZHANG Y, CHE W, et al. Joint extraction of entities and relations based on a novel graph scheme[C]//Proceedings of the 27th International Joint Conference on Artificial Intelligence, 2018:4461-4467.
- [35] 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.
- [36] WANG X, WANG D, JI F. A span-based model for joint entity and relation extraction with relational graphs[C]//Proceedings of 2020 IEEE International Conference on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking, 2020: 513-520.
- [37] 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, 2019:1409-1418.
- [38] FEI H, REN Y, JI D. Boundaries and edges rethinking: An end-to-end neural model for overlapping entity relation extraction[J]. Information Processing & Management, 2020, 57(6):102311.
- [39] DUAN G, MIAO J, HUANG T, et al. A relational adaptive neural model for joint entity and relation extraction[J]. Frontiers in Neurobotics, 2021, 15:635492.
- [40] 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, 2019:4171-4186.
- [41] LI C, TIAN Y. Downstream model design of pre-trained language model for relation extraction task[J]. arXiv: 2004.03786, 2020.
- [42] SUI D, CHEN Y, LIU K, et al. Joint entity and relation extraction with set prediction networks[J]. arXiv: 2011.01675, 2020.
- [43] LIU L, WANG M, HE X, et al. Extracting relational facts based on hybrid Syntax-Guided transformer and pointer network[J]. Journal of Intelligent & Fuzzy Systems, 2021, 40(6):12167-12183.
- [44] YE H, ZHANG N, DENG S, et al. Contrastive triple extraction with generative transformer[C]//Proceedings of the 35th AAAI Conference on Artificial Intelligence, 2021:14257-14265.
- [45] HANG T, FENG J, WU Y, et al. Joint extraction of entities and overlapping relations using source-target entity labeling[J]. Expert Systems with Applications, 2021, 177:114853.
- [46] FENG J, HUANG M, ZHAO L, et al. Reinforcement learning for relation classification from noisy data[C]//Proceedings of the 32nd AAAI Conference on Artificial Intelligence, 2018:5779-5786.
- [47] TAKANOBU R, ZHANG T, LIU J, et al. A hierarchical framework for relation extraction with reinforcement learning[C]//Proceedings of the 33rd AAAI Conference on Artificial Intelligence, 2019:7072-7079.
- [48] WANG X, JI H, SHI C, et al. Heterogeneous graph attention network[C]//Proceedings of World Wide Web Conference, 2019:2022-2032.
- [49] GAO H, PEI J, HUANG H. ProGAN: Network embedding via proximity generative adversarial network[C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019:1308-1316.
- [50] 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.
- [51] SUN C, QIU X, XU Y, et al. How to fine-tune bert for text classification?[C]//Proceedings of China National Conference on Chinese Computational Linguistics, 2019: 194-206.
- [52] 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.
- [53] LIU Y, OTT M, GOYAL N, et al. Roberta: A robustly optimized BERT pretraining approach[J]. arXiv: 1907.11692, 2019.
- [54] CHEN Y, ZHANG Y, HU C, et al. Jointly extracting explicit and implicit relational triples with reasoning pattern enhanced binary pointer network[C]//Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2021:5694-5703.