

实体关系抽取综述

王传栋, 徐 娇, 张 永

南京邮电大学 计算机学院, 南京 210023

摘 要: 实体关系抽取作为信息抽取任务的重要组成部分之一, 能够对更小粒度的信息进行语义分析, 为更多任务提供数据支持。关系抽取发展至今, 总体可分为基于传统机器学习和基于深度学习两种方式。基于传统机器学习的关系抽取研究主要以统计和基于规则相结合的方法为主。基于深度学习的框架通过引入远程监督、小样本学习、注意力机制、强化学习、多示例多标记学习等方法取得了丰富的研究成果。回顾实体关系抽取的发展历程, 对每种模型进行分析和讨论; 结合深度学习方法的最新动态, 对实体关系抽取未来的研究方向和趋势进行展望。

关键词: 实体关系抽取; 机器学习; 远程监督; 图卷积网络; 强化学习; 神经网络

文献标志码: A **中图分类号:** TP391.1 **doi:** 10.3778/j.issn.1002-8331.2003-0189

王传栋, 徐娇, 张永. 实体关系抽取综述. 计算机工程与应用, 2020, 56(12): 25-36.

WANG Chuandong, XU Jiao, ZHANG Yong. Survey of entity relation extraction. Computer Engineering and Applications, 2020, 56(12): 25-36.

Survey of Entity Relation Extraction

WANG Chuandong, XU Jiao, ZHANG Yong

School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China

Abstract: As an important part of information extraction, entity relation extraction can perform semantic analysis on smaller grained information and provide basic data support for more tasks. The development of relation extraction can be divided into two methods based on traditional machine learning and deep learning. In recent years, the research based on traditional machine learning has mainly focused on the combination of statistic-based and rule-based. The framework of deep learning has achieved abundant research results by introducing distant supervision, few-shot learning, attention mechanism, reinforcement learning and multi-instance multilabel. The development of entity relation extraction is reviewed and each model is analyzed. Combing the latest developments in deep learning methods, the development direction and trend of entity relationship extraction are prospected.

Key words: entity relation extraction; machine learning; distant supervision; graph convolutional network; reinforcement learning; neural network

1 引言

随着大数据的迅猛发展, 海量信息常以半结构化或者非结构化的形式呈现给用户, 如何通过文本深层分析模型向用户提供高质量、精准而有价值的信息成为学者们研究的热点问题。在这种背景下, 信息抽取的研究得到了快速发展, 实体关系抽取作为其重要子任务之一, 逐渐引起广大学者的关注。

关系抽取旨在已完成实体识别的基础上, 检索实体

间所存在的关系, 即在已标注出实体及实体类型的句子上确定实体间的关系类别。目前主流的实体关系抽取方法可细分为基于有监督的方式、基于半监督的方式、基于无监督的方式和面向开放域的关系抽取。随着近些年深度学习的不断发展和完善, 学者们开始尝试将一些基于深度学习的神经网络引入到关系抽取任务中^[1], 成为新的研究热点。

作为信息抽取的重要任务之一, 关系抽取能够对更

作者简介: 王传栋(1971—), 男, 副教授, 研究领域为数据挖掘、深度学习、自然语言处理; 徐娇(1994—), 女, 硕士研究生, 研究领域为实体关系抽取、机器学习, E-mail: xujiao_27@163.com; 张永(1994—), 男, 硕士研究生, 研究领域为自然语言处理、机器学习。

收稿日期: 2020-03-12 **修回日期:** 2020-04-29 **文章编号:** 1002-8331(2020)12-0025-12

CNKI网络出版: 2020-05-13, <http://kns.cnki.net/kcms/detail/11.2127.tp.20200512.1308.002.html>

小粒度样本数据中的信息进行语义关系分析,通过对海量信息进行关系抽取,可以将无结构文本转化为格式统一的关系数据,为知识图谱、推荐系统、信息检索等任务提供支持。同时,关系抽取的研究对篇章理解、自动摘要生成等研究领域也有深刻意义,具有广阔的应用场景。

2 关系抽取的发展

1998年美国国防高级研究计划局召开第七届消息理解会议(Message Understanding Conference, MUC),并首次提出模板抽取任务。MUC-7中关系抽取任务首次单独作为评测任务被提出,评测语料内容主要来自于纽约时报中涉及飞机失事、航天发射事件的相关新闻^[2],其中包含三类实体关系:Location_of、Employee_of和Product_of,并且设计了相应的评价体系。

1999年美国国家标准技术研究院召开自动内容抽取会议(Automatic Content Extraction, ACE),旨在研究新闻语料中的信息抽取任务。关系抽取任务属于ACE会议中定义的关系检测与识别(Relation Detection and Recognition, RDR)^[3],继MUC和ACE评测会议后, SemEval(Semantic Evaluation)也成为信息抽取领域的重要会议,其间补增评测任务中的实体关系类型,进一步促进了实体关系抽取问题的研究。

随着大数据时代的来临,异构数据呈指数级的增加,而上述测评会议所发布的依靠人工标注方式得到的语料集已经无法满足新的需求。人工标注虽可以获得高质量数据,但成本较高且语料的覆盖面窄,对于医疗等某些特定领域,需要更高昂的标注成本,导致模型可拓展性较差。一方面为了获得大规模、多领域的语料支持,许多学者开始关注面向开放域的信息抽取任务,另一方面随着涵盖更多领域信息的Freebases、维基百科和YAGO等知识库的建立和壮大,ACL 2009会议上Mintz首次提出将Distant Supervision应用到关系抽取任务中,并取得了一定效果^[4]。传统机器学习方法存在特征提取误差传播问题,极大影响关系抽取模型的性能。随着深度学习的崛起,学者们逐渐将深度学习引入关系抽取任务,大量基于CNN、RNN、LSTM、GRU、GCN等神经网络结构的关系抽取方法被提出。

3 关系抽取的研究现状

实体关系抽取是构建知识库的重要步骤,也是许多NLP下游任务的基础,根据对人工标注数据的依赖程度,主流的实体关系抽取方法主要分为三种:有监督学习方法、半监督学习方法和无监督学习方法^[5]。近年来,随着深度学习在联合学习、远程监督等方面的应用,使关系抽取任务相比此前基于传统机器学习的方法取得了更好的效果。

3.1 有监督的关系抽取

基于监督学习方法的实体关系抽取任务,通过在人工标注的数据上训练模型,然后将其应用在特定领域,具有较高准确率,主要包括基于规则的方法、基于特征向量的方法和基于核函数的方法^[6]。

基于规则的方法主要运用语言学的相关知识,对语料进行分析并归纳出关系表达式。Aone等^[7]通过人工制定的抽取规则,来对文本数据进行模式匹配,筛选出符合相应规则的关系样例。Miller等人^[8]通过使用与实体相关的语义信息来扩展语法树并联合表示句法和语义,生成规则进行实体关系抽取。Fundel等人^[9]使用斯坦福开发的句法分析器(Stanford Parser)构造句子的依存关系树(Dependency Tree),通过对依存关系树上两个实体间的路径进行分析来制定规则。然而制定关系规则时需要相关领域知识的支持,并且规则不具有通用性,无法有效进行跨领域迁移。

基于特征向量的方法首先从句子上下文中提取词性、实体位置等有用信息来构造特征向量,结合机器学习方法在特征向量上训练关系抽取模型。Kambhatla等人^[10]使用实体类型、实体词、句法分析树、依存关系等多种特征构造特征向量作为模型输入,并在特征向量上首次采用最大熵分类模型对实体关系抽取问题进行建模,实验证明:结合各层次的语言特征对关系抽取任务具有丰富的价值。Zhou等人^[11]在Kambhatla模型的基础上加入了WordNet、基本词组块和Name List信息来增强语义信息,采用SVM分类器在实体关系抽取上的F-measure达到55.5%。Jiang等人^[12]系统分析和比较了不同特征对关系抽取结果的影响,结果表明选取较基础的特征就能达到很好的效果,相反的,若所选取特征不相互独立时,会一定程度导致性能的下降。Bui等人^[13]、杨志豪^[14]以及Miwa^[15]使用词袋特征、词性特征以及依存关系特征等作为模型输入,并在生物医学领域的关系抽取上取得了一定成绩。

尽管基于特征向量的方法在关系抽取领域取得了很好的效果,但存在一定局限性,首先该方法很大程度上依赖大量特征工程的工作;其次当前使用的特征已基本覆盖大多数语言现象,性能上难以实现较大的提升。基于特征向量方法上述的局限性,更多的研究者尝试将核函数应用到实体关系抽取领域中,核方法不需要人为构造显性特征向量,而使用核函数的映射对多种信息进行融合来实现关系抽取。

采用基于核方法的关系抽取,对高维特征空间的样例只需计算其内积而无需得到具体的函数值,即使用隐性特征映射代替显性的特征映射,为基于特征向量的方法开拓了新的思路^[16]。Zelenko等人^[17]在2013年首次将核方法应用到关系抽取任务上,设计并提出了核函数及其计算方法,结合两个样本的浅层句法解析树来分析两

者间的相似性,结合SVM分类器在200篇新闻语料的数据上得到了不错的结果^[18]。Culotta等人^[19]改进Zlenko的方法,提出基于语法规则的依存树核进行关系抽取,通过使用词性、实体类型、WordNet上位词等特征来扩充树上的节点,提供了更丰富的句子表示形式,并在更大、包含更多关系类型的标注语料库上进行训练。Bunescu等人^[20]提出一种定义非常严格的最短路径依赖核,实现了比单纯依存树核更高的精度,但也因此导致召回率下降。Giuliano等人^[21]使用词袋(Bag-of-Words)代替稀疏子序列作为全局性信息结合词属性等相关特征作为局部信息,使用核函数的线性组合来整合两部分信息。实验表明:该方法在生物医学数据中提取实体间关系时,仅使用浅层语言特征就达到了较好的性能。

使用核函数方法可以学习文本的长距离特征而不需要构造特征向量,在关系抽取任务上的性能也超过了基于特征向量的方法,通过使用核方法复合不同核函数来表达高维特征空间时,会相应产生训练速度较慢的负面影响,无法很好适用于处理大规模语料下的关系抽取^[22]。

3.2 半监督的关系抽取

半监督学习只需通过对少量的种子标记样本和大量无标记的样本进行迭代训练就可以得到分类模型^[23],常用算法主要有Bootstrapping方法^[24]、协同训练方法^[25]和标注传播方法^[26]。

采用Bootstrapping技术的三个代表性半监督关系抽取系统即:DIPRE^[27](Dual Iterative Pattern Relation Expansion),Snowball^[28]和Zhang's method^[29]。其中DIPRE系统^[27]是最早被提出的基于Bootstrapping的半监督实体关系抽取方法,该方法在迭代初期使用少量书籍的(author,title)实体对作为种子,通过不断迭代能够自动从万维网获取新的书籍关系实例。Agichtein^[28]基于Bootstrapping采用和DIPRE相同的模式匹配方法从非结构化文本中抽取(organization,location)关系。Zhang等^[29]在结合SVM的Bootstrapping模型上,提出基于随机特征投影的BootProject算法,实验表明该方法可以显著降低对标记训练数据的依赖。Chen等人^[26]通过标注传播算法计算无标记样本中关系样例的最近邻来实现关系抽取,在ACE 2003语料上的实验结果表明,在仅有少量带标记的样本可使用时,基于标注传播(LP)算法的关系抽取模型性能明显优于SVM和Bootstrapping。张佳宏等人^[30]在Zhou^[31]的基础上通过引入无标注样本置信度衡量机制,在训练迭代次数得到显著减少的同时还取得了较高的F值。

半监督实体关系抽取一定程度上既能降低对标注语料的依赖,又可以得到较高的精度,所以能更好地适应于大规模训练语料的任务。目前基于半监督学习方法的实体关系抽取任务中应用最广泛的就是Bootstrapping算法,但该方法存在语义漂移等问题,还易受到初始关系种子质量的影响^[32]。

3.3 无监督的关系抽取

由于有监督和半监督实体关系抽取均依赖带标记的语料,虽然半监督方法仅使用少量的标记数据作为种子,但如何选择或选择多少数量种子的问题仍未得到有效解决。针对这些问题,一些研究者基于聚类的思想,通过学习对实体上下文进行抽取来刻画实体间的语义关系,实现无监督关系抽取。

Hasegawa等人^[33]在2004年ACL会议上首次提出无监督学习的命名实体关系抽取方法,为无监督实体关系抽取研究奠定了基础,但该方法一方面很难预先定义相似性的阈值,另一方面简单地按频率选择关系特征词时并没有考虑噪声特征。Zhang等人^[34]在进行聚类时,应用浅层句法树来表达实体间所存在的关系,通过衡量句法树的相似性来对实体间的相似性进行分析,充分考虑了低频实体间可能存在的语义关系。Chen等人^[35]使用DCM(Discriminative Category Matching)选取具有判别性的特征关系词,并通过提出的聚类质量评估公式得到最优聚类数目和最优特征子集,与Hasegawa等人^[33]的方法相比,在ACE语料库上的性能得到大幅度提高。秦兵等人^[36]提出一种基于无监督学习方法的中文实体关系抽取模型,在得到候选关系三元组和关系指示词后、结合句式规则对其进行筛选,实验在大规模无标记中文网络文本数据上的微平均准确率高于80%。

无监督的实体关系抽取方法不需要预定义任何关系类型,也不依赖标注数据,可以适应无规则内容文本,具有很好的领域移植性。但聚类出来的关系类别边界不够清晰、模型的整体准确率较低,并且低频关系实例使得模型的召回率较低^[18]。

3.4 开放域的关系抽取

开放域实体关系抽取是目前研究的热点,仅Freebase中就包含多达4 000万实体,上万种属性关系,这样数量级的关系抽取单纯依靠人工标注数据集是很难实现的,半监督和远程监督的学习方法一定程度上可以实现开放域的关系抽取。因其在数据规模、数据源类型、领域上的可拓展性,开放式信息抽取在处理大规模Web页面中异构信息时,具有其他关系抽取模型不可比拟的优势。

开放域的关系抽取目前主要有两种具有代表性的研究方向,一种是基于句法设计规则来对三元组进行过滤;另一种是基于知识监督的方法,这种思想后来慢慢发展出远程监督方法。开放式信息抽取的概念最早由Banko等人^[37]在2007年IJCAL会议上提出,同时提出第一个领域无关的OIE系统TextRunner,并可扩展到大规模Web语料库。Banko和Etzioni^[38]将抽取问题转换为在条件随机场模型上的序列标注任务,结合无监督同义词消解算法来处理候选关系和对象中的同义词。Zhu构建的Statsnowball系统^[39],可以看作是Snowball系统的开放域信息抽取版本,使用马尔科夫逻辑网络来抽取实

体间的关系。2010年Wu等人^[40]基于维基百科的信息构建WOE(Wikipedia-based Open Extractor)系统,实现了比TextRunner更高的召回率和准确率。Fader^[41]定义了两个由动词表达的二元关系的词法和语法约束,实验表明可以通过增加这些约束来改善Open IE系统的信息不连贯和信息不足的问题。

3.5 基于深度学习的关系抽取

基于传统机器学习的关系抽取方法在特征提取过程中存在误差传播问题,很大程度上限制了模型性能的提升。随着近些年深度学习的崛起,因其能够自动学习更高阶语义特征并具有较高的精确度,基于深度学习的方法成为了实体关系抽取领域新的研究热点^[42]。

根据语料集标注方式的差异性,目前国内外基于深度学习的实体关系抽取方法可被分为远程监督和有监督学习两类。有监督学习在训练过程中使用人工标注的数据集,而远程监督的学习方法通过对齐远程知识库自动对语料进行标注来获取带标记语料数据。

3.5.1 基于深度学习的有监督关系抽取

基于深度学习的有监督实体关系抽取根据实体识别和关系检测两个子任务完成顺序的不同,可细分为流水线方法和联合抽取方法。其中流水线学习方法是在实体已被标注的数据基础上进行实体间关系的抽取,联

合学习方法是同时进行实体识别和实体关系抽取任务^[42]。表1中整理了深度学习框架下有监督关系抽取经典方法,其中的模型主要是基于现有CNN、RNN、LSTM改进输入特征或网络结构,比如添加不同特征、结合多种Attention机制和引入依存树挖掘更深层次语义信息来提升模型的性能。随着图卷积神经网络的兴起,因其在处理异构数据所具备的天然优势,许多学者尝试引入GCN来学习依存树中蕴含的丰富信息。基于特征组合的关系抽取方法均依赖于其他特征工具包,另外很多模型将关系抽取建模为单标记问题,无法解决样本中关系重叠问题。下面对相关模型进行深入研究和分析。

(1) 基于RNN模型的关系抽取方法

递归神经网络(Recursive Neural Network, RNN)因为其既有内部反馈连接又有前馈连接,比前馈网络更适合处理序列化输入,基于RNN的实体关系抽取方法最早由Socher等人^[43]提出。模型中每个单词均由向量和矩阵组成,分别用来学习表示单词本身语义和对其他单词的修饰作用,可以自动学习到较长短语的深层语义,但模型需要学习的参数过多。Hashimoto等人^[44]提出一种基于解析树的RNN模型,通过对重要短语进行显式加权,结合短语类别、词性标注等特征实现性能提升。虽然RNN进行关系抽取已经表现出不错的性能,

表1 有监督实体关系抽取的经典模型

类别	模型名称	创新	优势	不足	提出年份
Pipeline	MV-RNN ^[43]	矩阵-向量表示	可以学习命题逻辑和自然语言中操作符的含义	需学习的参数过多	2012年 EMNLP
	基于解析树结构的RNN模型 ^[44]	短语显式加权 参数平均	引入额外特征提升性能 更稳定、泛化能力更高	依赖依存树的解析结果	2013年 EMNLP
	CDNN ^[45]	位置特征PF	无复杂预处理 突出实体影响	需人工构造词特征	2014年 COLING
	depLCNN+NS ^[46]	最短路径序列 负采样策略	更鲁棒 改善长距离关系	损失部分依存树信息	2015年 EMNLP
	Multi-Level Attention CNN ^[47]	两层注意力 合页损失函数	有效提取句子关键特征	结构复杂收敛困难	2016年 ACL
	TAC-CNN ^[48]	词级别注意力	引入词与关系的相关性	仅考虑单标记	2017年 ICONIP
	SDP-LSTM ^[49]	最短依赖路径 多通道LSTM	可集成异构信息	损失部分依存树信息	2015年 EMNLP
	BLSTM ^[50]	双向LSTM结构 融合多种特征	改善长距离关系 丰富单词语义特征	依赖特征工程	2015年 PACLIC
	C-GCN ^[51]	以路径为中心的修剪策略	保留关键信息的同时过滤 无关信息	需要指定修剪距离K	2018年 EMNLP
	AGGCN ^[52]	软修剪策略(Attention)	自动学习修剪依存树	全连接图破坏树原始结构	2019年 ACL
Joint	LST-AGCN ^[53]	句法传输图 注意力机制	引入依存关系嵌入	依赖依存树的解析结果	2020年 AAAI
	SPTree ^[54]	Bi-TreeLSTM	同时捕获单词序列和依存 树子结构信息	参数共享但过程依旧类似 Pipeline抽取形式	2016年 ACL
	Bi-LSTM+Attention ^[55]	BLSTM与注意力的组合	不依赖解析器	在PHYS和PARTWHOL关 系上的性能较差	2017年 ACL

但原始的RNN模型因其层数较多而更易出现梯度消失或梯度爆炸问题,无法有效解决文本数据中的长依赖问题。

(2)基于CNN模型的关系抽取方法

2014年Zeng等人^[45]首次提出使用卷积神经网络(Convolutional Neural Networks,CNN)进行关系抽取,不需要复杂的预处理仅将所有单词向量作为初始输入,结合卷积深度神经网络(CDNN)提取的词汇和句子特征进行关系分类。2015年Xu等人^[46]在Zeng工作的基础上,在依存关系树的最短依赖路径上通过卷积神经网络来学习更鲁棒的关系特征,同时提出一种简单负采样策略来改善实体距离较远所产生的问题。但模型性能很大程度受依存树解析结果的影响,并且仅在最短依赖路径上进行学习不能充分利用依存树所蕴含的丰富信息。Wang等人^[47]基于卷积神经网络,使用两层注意力机制提取句子特征并设计一个pair-wise合页损失函数,实验证明其性能优于标准损失函数。引入多级注意力机制捕获更高层语义信息的同时,也面临着结构复杂、参数较多、收敛困难等问题。基于单词的重要性与关系类型的高度相关性,Zhu等人^[48]在2017年提出TAC-CNN模型,结合词级别的注意力机制通过计算每个单词与关系类别的相关度计算相应权重。不引入二级注意力机制的情况下,在SemEval-2010 Task 8数据集上F1即达到87.3%。

(3)基于LSTM模型的关系抽取方法

Sundermeyer等人^[56]通过构建专门的记忆单元存储重要历史信息,以此获得远距离单词之间的关系,提出长短时记忆网络模型(Long Short-Term Memory network,LSTM)不但可以有效解决实体之间的长依赖问题,还可以结合聚类技术大幅度减少训练和测试用时。Xu等人^[49]结合实体对间的最短依存路径和长短时记忆网络模型,提出了用于关系分类的新型神经网络模型SDP-LSTM,它可以结合最短依存路径上的单词、POS标记、语法关系、WordNet上位词等信息迭代的学习与关系分类相关的特征。但SDP-LSTM在引入多种特征的同时,也使得模型更加依赖特征抽取工具,多特征的提取也加剧错误传播问题。Zhang等人^[50]基于每个时刻的输入不仅依赖文本中某个单词前面的单词,还依赖于后面单词的思想,提出结合前向和后向LSTM捕捉双向的语义依赖获取更多语序信息^[57]。在SemEval-2010语料库上的实验表明,模型仅使用单词嵌入作为输入特征就可以实现远优于CNN和CR-CNN方法的F1值。

(4)基于GCN模型的关系抽取方法

句法依存树包含句子中各词语间的依存关系,将其引入关系抽取任务可以挖掘更深层的语义信息。图卷积神经网络的提出^[58]实现了非欧式数据上的卷积操作,也为处理图结构数据提供了新思路。

Zhang等人^[51]提出一种基于修剪依存树的图卷积神

经网络并用于实体关系抽取问题,仅保留两个实体的最小公共祖先子树上 K 距离内的节点,并将修剪后的句法依存树引入图卷积网络进行实体关系抽取任务。实验表明,这种修剪方式过滤依存树中无关数据的同时,保留了对关系抽取任务有用的信息。但基于规则的硬性修剪策略却很容易产生过剪枝或欠剪枝,为了解决这个问题,Guo等人^[52]提出了注意力引导的图卷积网络AGGCN,可以理解为一种对句法依存树的软修剪策略,模型将完整依存树作为输入并结合注意力机制,在迭代训练中自动学习保留对关系抽取任务有用的子结构。AGGCN模型由 M 个相同模块组成,每个块包含注意力引导层、密集连接层和线性组合层,其中注意力引导层使用多头注意力机制构造 N 个注意力引导邻接矩阵,将输入依存树转换为 N 个不同的全连接边加权图,即每个注意力引导邻接矩阵对应一个全连接图。

Sun等人^[53]认为AGGCN模型使用的全连接图破坏了依存树原始结构,基于这个问题提出了可学习的句法传输注意力图卷积网络(LST-AGCN),通过引入连接节点的依存关系类型将树转换为加权图即句法传输图。通过词嵌入、依存关系嵌入和节点嵌入来建模可学习的传输矩阵 A ,并结合注意力机制学习合适权重来聚合所有图层输出的特征向量,得到最终句子表示再进行关系抽取,其中涉及的注意力机制主要用于整合每层GCN的输出。

表2将注意力机制按其结构分为单层自注意力、多层注意力和多头注意力机制。无论是基于CNN、RNN还是GCN衍生出的关系抽取模型,均可以通过引入不同Attention或其组合来提高性能。注意力机制良好的软性选择能力也可以有效缓解远程监督中的噪声问题,许多研究将其引入远程监督来过滤错误标记的样本。

表2 注意力机制相关方法统计

类别	模型名称
Self-Attention	TAC-CNN ^[48]
	LST-AGCN ^[53]
	BiLSTM+Attention ^[55]
	MIMLCNN ^[59]
	PCNN+Attention ^[60]
Multi-Level Attention	Multi-Level Attention CNN ^[47]
	Proto-HATT ^[61]
Multi-head Attention	AGGCN ^[52]

流水线方法虽然已取得不错的成绩,但仍存在局限性:顺序进行实体识别和关系抽取时,忽略了两个子任务间的内在联系;实体识别中产生的错误会降低关系抽取模型的性能即存在错误传播的问题。相比之下,联合学习方法可以充分利用实体和关系间的交互信息,有效地缓解了上述问题。

Miwa等人^[54]提出一种端对端模型来联合表示实体和关系,其中实体识别和关系抽取两个子任务共享

LSTM 编码层序列参数的思想,虽然考虑了两个子任务间的交互、缓解了错误传播问题,但模型学习过程仍然类似流水线方法,不属于真正意义上的联合抽取。Katiyar 等^[55]在深度 BiLSTM 序列标注方法的基础上引入注意力机制,使用多层双向 LSTM 将实体识别子任务建模成序列标注任务,结合序列标注结果和共享编码层表示进行实体关系抽取,模型可以拓展各种预定义的关系类型,是真正意义上第一个基于神经网络的关系联合抽取模型。Zheng 等人^[62]提出基于新标注策略的实体关系抽取方法,将联合抽取任务转化成端到端的序列标注问题而无需分别识别实体和关系,解决了流水线方法中实体冗余问题。但无论是上述的参数共享还是序列标注的联合抽取模型,均未有效解决重叠关系问题。

3.5.2 基于远程监督的关系抽取

Mintz 等人^[4]在 2009 年 ACL 上提出将远程监督应用到实体关系抽取任务上,基于“如果两个实体在已知知识库中存在某种关系,那么包含这两个实体的句子也表示该关系”的假设,通过将数据自动对齐远程知识库中的信息来对开放域中海量数据进行自动标注来获得标记样本。远程监督关系抽取较快得到大量标记样本的同时,大幅降低人工标注的工作量,但由于所基于的假设条件过于强烈,使得远程监督数据集中包含大量的错误标记样本。

如表 3 所示,针对远程监督中错误标签问题学者们提出了很可行的解决方法,比如引入多示例学习、结合 Attention 机制、对噪声进行拟合、建模为强化学习问题,下面对相关模型进行深入分析。

Zeng 等人^[63]通过多示例学习的方法来缓解远程监督中的噪音问题,在 2014 年 Zeng 等人^[45]的 CNN 基础上将句子根据实体位置切分为 3 段,分别进行池化来得到更多与实体相关的上下文信息。多示例学习是将包含同一实体对的所有样本看成一个包,模型只选择包中使得关系概率最大的一个示例作为实体对的表示,该方法一定程度上降低噪音数据的影响,但也因此丢失了大量有用的信息^[69],另外也无法处理关系重叠问题。Jiang 等人^[59]提出多示例多标记的卷积神经网络模型来松弛 at-least-once 假设,将关系抽取任务建模为多标记问题,解决了关系重叠问题。基于多示例学习的模型虽然可

以有效缓解远程监督中的噪声问题,但仅选取 bag 内最高置信度的样例可能会丢失大量有用信息。

Lin 等人^[60]通过引入 Attention 机制来自动学习如何筛选包内所有样本的信息,有效避免多示例模型中造成的信息损失。实验表明该模型能够学到合理权重来缓解远程监督中噪音问题的同时,还充分挖掘有用信息。Ji 等人^[64]从 Freebase 和 Wikipedia 页面中提取实体的描述作为额外信息来改善实体表示模块的性能,其中句级注意力模块与 Lin 等人^[60]类似,自动为包内的样本学得合理的权重。Feng 等人^[65]提出了一种基于强化学习框架的关系抽取模型,该模型由样本选择器和关系分类器两部分组成,其中将样本选择建模为强化学习问题。整个过程在没有明确的句子级标签的情况下,仅使用来自关系分类器的弱监督信息就可以有效过滤远程监督数据中的嘈杂句子。2018 年 ACL 会议上, Qin 等人^[66]提出一种基于深度强化学习的远程监督关系抽取模型,旨在使用深度强化学习框架不断训练得到一个正例、负例指示器,不但实现对负例的识别,还将其放至对应关系类别的负例集中。该模型不依赖于特定的关系分类器,是一种即插即用的技术,能被引入现有的任何一种远程监督关系抽取模型。但这两种结合强化学习的抽取模型均未解决重叠关系问题,而且属于流水线型抽取方法,不能有效学习子任务间的交互。2019 年 AAAI 会议上 Takanobu 等人^[67]上提出基于分层强化学习的关系抽取方法,将任务分解为实体检测和关系提取两个子任务,分层方法的性质可以对两个子任务间的交互进行有效建模,且擅长提取实体间的重叠关系。

Ren 等人^[70]提出基于远程监督的联合抽取模型 COTYPE 进行实体关系抽取,实验表明该方法不仅能扩展到不同领域,还有效减弱了错误的累积传播。Luo 等人^[68]认为可以根据数据中潜在的信息学习噪音的模式,在训练过程中通过动态转移矩阵对噪音进行建模来达到拟合真实分布的目的。在没有直接指导的情况下逐步学习对基础噪音模式进行建模,并灵活利用数据质量的先验知识来提高转移矩阵的有效性。2019 年 ACL 会议上, Fu 等人^[71]提出基于图卷积网络的联合关系抽取模型 GraphRel,对实体重叠问题提出了解决方案。模型通过堆叠 Bi-LSTM 句子编码器和 GCN 依存树编码器来自

表 3 远程监督中噪声问题的常见解决方法

解决方案	模型名称	优势
多示例	MIMLCNN ^[59]	将关系抽取任务建模为多标记问题,能提取重叠关系
	PCNNs+MIL ^[63]	松弛远程监督强假设,有效降低噪声影响
注意力机制	PCNN+ATT ^[60]	使用句子级注意力,自动为 bag 内数据学习合适权重
	APCNNs+D ^[64]	引入 KG 信息增加实体描述;模型可扩展性强
	RL ^[65]	无需显式的句子级注释
强化学习	Deep RL ^[66]	可以重新分配假正例样本;即插即用
	Hierarchical RL ^[67]	擅长建模两个子任务间的交互提取重叠关系
噪声拟合	显式建模噪声的关系抽取 ^[68]	基于课程学习,有效利用数据质量先验知识

动提取每个单词的隐含特征。第一阶段预测后计算实体损失和关系损失。为了考虑三元组之间的相互作用,在第二阶段添加了关系加权 GCN 来解决实体重叠问题。基于第二阶段提取的特征进行分类后可得到较准确的结果,训练时损失函数为两个阶段实体和关系预测损失的线性加和。GraphRel 基于关系加权的图卷积网络,考虑了命名实体与关系之间的相互作用。在 NYT 和 WebNLG 数据集上分别评估该模型,实验表明,此方法比以前的工作 F1 评测值分别提高了 3.2%和 5.8%。

4 关系抽取的数据集及评价指标

4.1 数据集

有监督学习关系抽取中常用的数据集主要包括 MUC 关系抽取数据集、ACE04、ACE05、TACRED、SemEval-2010 Task 8 和 FewRel 数据集,其中 FewRel 在有监督学习和小样本学习的关系抽取任务中均能应用。NYT-FB 是远程监督关系抽取任务应用最广泛的数据集,

语料来自于纽约时报,通过自动链接到 Freebase 知识库中的实体并经关系对齐等操作来标注实体间关系类别。

表 4 中模型 SPTree 和 Miwa 等人提出的关系抽取方法,在 ACE05 数据集上的性能均优于 ACE04, F1 评测值分别提高了 7.2%和 7.9%。在 SemEval-2010 Task 8 数据集上,模型 SDP-LSTM 的 F1 评测值仅比 C-AGGCN 高 2%,但两者在数据集 TACRED 上的差值却达到了 10.3%,分析可能因为 TACRED 数据集包含更多关系类别且“no_relation”类数据所占比重过大。而当基于有监督学习的关系抽取模型 CDNN 使用远程监督数据集进行训练时,由于远程监督中大量的噪声数据的影响,CDNN 方法的性能急剧下降。表 4 中数据说明模型的性能不仅受结构和初始化的影响,还取决于所使用的数据集,并且使用不同数据集进行训练得到的关系分类器性能往往有较大差距。

4.2 评价指标

关系抽取主要采用准确率(Precision)、召回率(Recall)

表 4 不同数据集及模型的性能

数据集	简述	模型	评测指标	评测值
ACE05	7 种实体类型和 6 种关系类别	SPTree ^[54]	F1	55.6
		Bi-LSTM+Attention ^[55]	F1	53.6
ACE04	7 种实体类型和 7 种关系类别	SPTree ^[54]	F1	48.4
		Bi-LSTM+Attention ^[55]	F1	45.7
SemEval-2010 Task 8	10 种关系类型(含“other”类),若考虑关系的方向则可细分为 19 种;训练集和测试集分别包含 8 000 和 2 717 条样本	MV-RNN ^[43]	F1	82.4
		CDNN ^[45]	F1	82.7
		Multi-Level attention CNN ^[47]	F1	88.0
		SDP-LSTM ^[49]	F1	83.7
		BLSTM ^[50]	F1	84.3
		C-GCN ^[51]	F1	84.8
		C-AGGCN ^[52]	F1	85.7
		LST-AGCN ^[53]	F1	86.0
TACRED	来自每年 TAC KBP 比赛语料库中新闻专线和网络文本;包含 106 264 个样本,42 种关系类型	SDP-LSTM ^[49]	F1	58.7
		C-GCN ^[51]	F1	66.4
		C-AGGCN ^[52]	F1	69.0
		LST-AGCN ^[53]	F1	68.8
FewRel	规模最大的精标关系抽取数据集和最大的小样本学习数据集;包含 100 种关系、70 000 条样本(每种关系均包含 700 条句子)	Proto-HATT ^[61]	Precision (5 Way 5 Shot)	90.12±0.04
			Precision (10 Way 5 Shot)	83.05±0.05
		MLMAN ^[72]	Precision (5 Way 5 Shot)	92.66±0.09
			Precision (10 Way 5 Shot)	87.29±0.15
NYT	应用最广泛的远程监督关系抽取数据集,包含 53 种关系类别,659 059 条数据(522 611 条训练样本,172 448 条测试样本)	Mintz ^[4]	Precision (Top100)	77.0
		PCNNs+MIL ^[63]	Precision (Top100)	86.0
			Precision (Top100)	87.0
		APCNNs+D ^[64]	Precision (Top100)	87.0
		CDNN ^[45]	F1	40.0
		CDNN+RL ^[65]	F1	42.0

和F值(F-Measure)3项作为基本评价指标,其中准确率和召回率是一对矛盾的度量,通常使用F值综合考虑两者来对系统性能进行总体评价。计算公式分别如下所示:

$$Precision_R = \frac{\text{被正确预测为关系}R\text{的样本个数}}{\text{被正确预测为关系}R\text{的样本总数}} \quad (1)$$

$$Recall_R = \frac{\text{被正确预测为关系}R\text{的样本个数}}{\text{真实标记为关系}R\text{的样本总数}} \quad (2)$$

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

F_β 中 β 是调节准确率和召回率比重的参数。当 $\beta=1$ 时,认为在评价模型性能时两者同等重要。由上式得到 F_1 表示为:

$$F_1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \quad (4)$$

远程监督兴起后,模型所需要处理的数据规模量级增加,在考察系统性能时,也将运行时间和内存占用作为评价指标的一部分进行考量。

5 关系抽取的挑战和趋势

实体关系抽取基于海量信息可以将无结构文本转化为格式统一的关系数据,为知识图谱、推荐系统、信息检索等任务提供基础的数据支持。同时,关系抽取研究对语义分析、篇章理解、自动问答等领域也具有重要意义。

目前,基于深度学习的实体关系抽取虽然已经取得了极大的成功,但在领域自适应性和召回率方面仍有很大的提升空间。其次,实体关系抽取任务中仍存在以下亟待解决的问题:很多主流的关系抽取方法均未有效解决关系类型的OOV(Out Of Vocabulary)问题,仅简单地将不属于预定义实体关系类型的数据归入Other类,而Other类中的实体对只能通过人工处理才能确定关系类型定义;为了缓解远程监督的错误标签问题,学者们分别结合多示例学习^[59, 63]、Attention机制^[60, 64]、强化学习框架^[65-67]、噪声建模^[68]等方法提出许多模型,但如何建立更有效的方法缓解远程监督中错误标签的影响仍是关系抽取中研究的重点问题。

5.1 未来研究方向

5.1.1 二元关系到多元关系抽取的扩展

目前大多数实体关系抽取的研究仅停留在二元关系的层面上,但实际中英文语料的多元实体关系却占据高达40%的比重^[73]。多元关系抽取能够获取到更多实体之间的关系,相较于二元实体关系抽取来说,多元关系抽取的研究具有更大挑战性。如何将二元抽取技术拓展至三元甚至多元层面是今后实体关系发展的一个热门研究方向。

5.1.2 基于远程监督的关系抽取

远程监督关系抽取技术通过外部知识库作为监督源,自动对语料库进行标注,能够以较低成本获取大量

带标记的样本。但由于远程监督所基于的假设过于肯定,难免含有大量的噪音数据,而如何使用有效的降噪方式来缓解远程监督中的错误标注问题一直是关系抽取研究的重要课题。

5.1.3 段落和篇章级关系抽取技术

现有关系抽取工作主要聚焦于句子级关系抽取而现实生活中实体对经常分别位于不同句子,根据从维基百科采样的人工标注数据的统计表明,至少40%的实体关系信息只能从多个句子中才能联合获取。研究段落级甚至篇章级的关系抽取要求模型具有更强大的逻辑推理、指代推理和常识推理能力,现有的跨句子 N 元关系抽取研究常通过引入指代消解任务的方法来提升模型的性能。未来融合改进指代消解和图结构的方法也许是解决段落级和篇章级实体关系抽取任务的有效方案。

5.2 基于深度学习的关系抽取新思路

5.2.1 融合图卷积网络的实体关系抽取

自第一个图卷积神经网络被Bruna等人^[58]提出以来,就受到了研究人员的大量关注,被广泛应用于推荐系统、交通流量预测、生物医学、计算机视觉等领域。图卷积神经网络在自然语言处理领域也有大量应用,涉及的常见图结构主要包括知识图谱、依存句法树、词共现图、文章引用网等。Liu等人^[74]和Nguyen等人^[75]使用基于依存句法树的图卷积神经网络来进行事件抽取任务。Marcheggiani等人^[76]基于句法依存树提出结合长短时记忆网络的图卷积模型,并成功应用于语法角色标注任务上。

图卷积神经网络也被引入关系抽取任务中,并表现出了较高的性能。Zhang等人^[51]提出一种基于修剪依存树的图卷积神经网络并用于实体关系抽取问题。2019年ACL会议上,Guo等人^[52]基于硬修剪策略可能存在的欠剪枝和过剪枝问题,结合多头注意力将全句法依存树作为图输入,使用图卷积神经网络自动学习依存树中对关系提取任务有用的子结构Sun等人^[53]使用句法传输图代替AGGC模型中的全连接图,结合自注意力机制提出一种新的图卷积关系抽取模型LST-AGCN。但以上模型一方面无法建模两个子任务间的交互,另一方面也不能处理重叠关系。Fu等人^[71]提出基于图卷积网络的联合关系抽取模型GraphRel则有效解决上述问题。

大量的研究已表明,在引入图卷积网络模型后,各项任务上的性能都出现了一定的提升。相对于传统自然语言处理中的序列化建模,使用图卷积网络能够挖掘更多非线性的复杂语义特征。对于关系抽取问题,在原LSTM的网络结构上引入基于句法依赖树的图卷积网络,模型不但能学习文本的语序性信息,还通过图卷积充分学习依赖树的空间结构信息,更好地挖掘潜在语义关系。

5.2.2 基于强化学习框架的实体关系抽取

强化学习是实体关系抽取问题中新的研究思路,2018年AAAI会议上Feng等人^[65]将强化学习与深度学习相结合提出了一种基于噪音数据的句子级实体关系抽取模型,仅使用来自关系分类器的弱监督信息就可以有效过滤远程监督数据中的嘈杂句子。Qin等人^[66]提出一种基于深度强化学习的远程监督关系抽取方法,不依赖于特定的关系分类器,是一种即插即用的技术,能被引入现有的任何一种远程监督关系抽取模型。Takanobu等人^[67]提出基于分层强化学习的关系抽取模型,将任务分解为实体检测和关系提取两个子任务,使用分层方法有效建模子任务间的交互,而且擅长提取实体间的重叠关系。无论是对于任务中远程监督的噪音问题还是重叠关系问题,强化学习的技术都提供了一种切实可行的解决办法。

5.2.3 基于小样本学习的实体关系抽取

小样本学习方法(few-shot learning)作为公认未来最具潜力的研究方向之一,过去的研究主要集中在计算机视觉领域,在自然语言处理领域的探索还较少。2018年Han等人^[77]首次将小样本学习引入到关系抽取任务中,提出小样本关系抽取数据集FewRel,希望推动自然语言处理特别是实体关系抽取任务中的小样本学习研究。2019年Gao等人^[61]提出一种基于小样本学习的关系抽取模型,该模型结合特征和实例两级注意力机制,实现降低噪声数据影响的同时保留对关系分类结果有用的特征信息。Ye等人^[72]在2019年ACL会议上提出用于一种解决小样本关系抽取问题的新方法,该模型结合多级匹配和整合结构对训练样本间的隐含关联进行学习,尽可能充分挖掘少量样本中对关系分类有用的潜在信息。Soares等人^[78]则采用预训练模型BERT来处理关系抽取任务中的小样本学习问题。因为基于海量数据训练的BERT包含丰富的语义特征,引入模型后能够有效缓解小样本学习中特征匮乏问题,实验表明其在FewRel数据集上的性能已经超过人工进行关系分类的水平。Gao等人^[79]进一步研究发现要将小样本学习模型用于生产环境中,则模型不仅要具备领域迁移性,还需要能判断出数据是不是属于“非以上关系”的关系类型。为了解决以上两个问题,Gao等人采集了大量医疗领域的数据并进行标注,并在FewRel数据集原 N -way K -shot的基础上添加了“以上都不是”选项,提出了FewRel2.0数据集。

总之,研究小样本学习的关系抽取方法,能使模型具备“举一反三”的高效学习能力,无论是对现有的小样本学习方法进行改进,还是提出新的小样本学习模型进行关系抽取,都将大大降低模型对标注样本数量上的依赖,对推动关系抽取技术的落地有着重大意义。

6 结束语

现阶段,基于监督的方式在关系抽取任务上具有较高的准确率和召回率,但模型却严重依赖准确的标注数据。当模型迁移到新的领域时,需要构建新的标注语料库并重新对模型进行训练。半监督方式的关系抽取降低了对标注数据的依赖,可以适应标注数据匮乏的情况,但却对初始种子的质量有很高的要求,并且经常存在语义漂移问题。无监督的关系抽取因其无需依赖标注数据,所以不但在领域可移植性上具有先天的优势,在处理开放域海量数据时也表现出良好的性能,但其聚类阈值的确定一直是较大的挑战。

关系抽取技术发展至今,在研究内容上逐渐由限定领域转向开放领域,关系类型的定义方式表现为由人工预先定义演变为关系类型自发现;在研究方法上,深度学习的方法在关系抽取任务的性能上大幅超越传统机器学习,基于深度学习框架的关系抽取技术越来越受到学者们的关注,在此基础上结合新兴的图卷积神经网络结构、注意力机制、强化学习和小样本学习的思想等均作为关系抽取问题提供了新的解决思路。

参考文献:

- [1] Kumar S.A survey of deep learning methods for relation extraction[J].arXiv:1705.03645,2017.
- [2] 谢德鹏,常青.关系抽取综述[J/OL].计算机应用研究,2019-06-10[2020-04-30].https://doi.org/10.19734/j.issn.1001-3695.2018.12.0923.
- [3] 张素香.信息抽取中关键技术的研究[D].北京:北京邮电大学,2007.
- [4] Mintz M,Bills S,Snow R,et al.Distant supervision for relation extraction without labeled data[C]//Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP,2009:1003-1011.
- [5] Zhou Zhihua.A brief introduction to weakly supervised learning[J].National Science Review,2017,5(1):44-53.
- [6] 唐敏.基于深度学习的中文实体关系抽取方法研究[D].成都:西南交通大学,2018.
- [7] Aone C,Halverson L,Hampton T,et al.SRA:description of the IE2 system used for MUC-7[C]//Proceedings of the 7th Message Understanding Conference,1998.
- [8] Miller S,Fox H,Ramshaw L,et al.A novel use of statistical parsing to extract information from text[C]//Proceedings of the 1st North American Chapter of the Association for Computational Linguistics Conference,2000:226-233.
- [9] Fundel K,Küffner R,Zimmer R.RelEx—relation extraction using dependency parse trees[J].Bioinformatics,2007,23(3):365-371.
- [10] Kambhatla N.Combining lexical, syntactic, and semantic features with maximum entropy models for extracting

- relations[C]//Proceedings of ACL on Interactive Poster and Demonstration Sessions, 2004: 22-26.
- [11] Zhou Guodong, Sun Jian, Zhang Jie, et al. Exploring various knowledge in relation extraction[C]//Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics, 2005: 427-434.
- [12] Jiang Jing, Zhai Chengxiang. A systematic exploration of the feature space for relation extraction[C]//Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics, 2007: 113-120.
- [13] Bui Q C, Katrenko S, Sloot P M A. A hybrid approach to extract protein-protein interactions[J]. *Bioinformatics*, 2011, 27(2): 259-265.
- [14] 杨志豪, 洪莉, 林鸿飞, 等. 基于支持向量机的生物医学文献蛋白质关系抽取[J]. *智能系统学报*, 2008(4): 361-369.
- [15] Miwa M, Sætne R, Miyao Y, et al. A rich feature vector for protein-protein interaction extraction from multiple corpora[C]//Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, 2009: 121-130.
- [16] 张晓峰. 基于核方法的实体关系抽取研究[D]. 南京: 东南大学, 2016.
- [17] Zelenko D, Aone C, Richardella A. Kernel methods for relation extraction[J]. *The Journal of Machine Learning Research*, 2003, 3: 1083-1106.
- [18] 孔兵. 中文文本实体关系抽取方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2016.
- [19] Culotta A, Sorensen J. Dependency tree kernels for relation extraction[C]//The 42nd Meeting of Association for Computational Linguistics, 2004.
- [20] Bunescu R C, Mooney R J. A shortest path dependency kernel for relation extraction[C]//Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, 2005: 724-731.
- [21] Giuliano C, Lavelli A, Romano L. Exploiting shallow linguistic information for relation extraction from biomedical literature[C]//Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics, 2006: 401-408.
- [22] 刘方驰, 钟志农, 雷霖, 等. 基于机器学习的实体关系抽取方法[J]. *兵工自动化*, 2013, 32(9): 57-62.
- [23] Chapelle O, Zien A. Semi-supervised learning[M]. Cambridge: MIT Press, 2006: 3-14.
- [24] Abe N, Mamitsuka H. Query learning strategies using boosting and bagging[C]//Proceedings of the Fifteenth International Conference on Machine Learning, 1998: 1-9.
- [25] Etzioni O, Cafarella M, Downey D, et al. Unsupervised named-entity extraction from the Web: an experimental study[J]. *Artificial Intelligence*, 2005, 165(1): 91-134.
- [26] Chen Jinxiu, Ji Donghong, Tan C L, et al. Relation extraction using label propagation based semi-supervised learning[C]//Proceeding of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the Association for Computational Linguistics, 2006: 129-136.
- [27] Brin S. Extracting patterns and relations from the world wide web[M]. Berlin: Springer Heidelberg, 1999: 172-183.
- [28] Agichtein E, Gravano L. Snowball: extracting relations from large plain-text collections[C]//Proceedings of ACM Conference on Digital Libraries, 2000: 85-94.
- [29] Zhang Z. Weakly-supervised relation classification for information[C]//Proceedings of the 13th ACM International Conference on Information and Knowledge Management, 2004: 581-588.
- [30] 张佳宏. 基于半监督学习的文本实体关系抽取研究[D]. 哈尔滨: 哈尔滨工程大学, 2017.
- [31] Zhou Zhihua, Li Ming. Tri-training: exploiting unlabeled data using three classifiers[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2005, 17(11): 1529-1541.
- [32] 台丽婷. 基于半监督机器学习的实体关系抽取算法研究[D]. 北京: 北京邮电大学, 2018.
- [33] Hasegawa T, Sekine S, Grishman R. Discovering relations among named entities from large corpora[C]//Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, 2004: 415-422.
- [34] Zhang M, Su J, Wang D, et al. Discovering relations between named entities from a large raw corpus using tree similarity-based clustering[C]//Proceedings of the 2nd International Joint Conference on Natural Language Processing, 2005: 378-389.
- [35] Chen Jinxiu, Ji Donghong, Tan C L, et al. Unsupervised feature selection for relation extraction[C]//Proceedings of the International Joint Conference on Natural Language Processing, 2005: 262-267.
- [36] 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.
- [37] Banko M, Cafarella M J, Soderland S, et al. Open information extraction from the Web[C]//Proceedings of the 20th International Joint Conference on Artificial Intelligence, 2007: 2670-2676.
- [38] Banko M, Etzioni O. The tradeoffs between open and traditional relation Extraction[C]//Proceedings of the Association for Computational Linguistics, 2008: 28-36.
- [39] Zhu Jun, Nie Zaiqing, Liu Xiaojing, et al. StatSnowball: a statistical approach to extracting entity relationships[C]//Proceedings of the 18th International Conference on World Wide Web, 2009.
- [40] Wu Fei, Weld D S. Open information extraction using

- Wikipedia[C]//Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics,2010:118-127.
- [41] Fader A, Soderland S, Etzioni O. Identifying relations for open information extraction[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing,2011:1535-1545.
- [42] 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.
- [43] Socher R, Huval B, Manning C D, et al. Semantic compositionality through recursive matrix-vector spaces[C]//Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning,2012:1201-1211.
- [44] Hashimoto K, Miwa M, Tsuruoka Y, et al. Simple customization of recursive neural networks for semantic relation classification[C]//Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing,2013:1372-1376.
- [45] Zeng Daojian, Liu Kang, Lai Siwei, et al. Relation classification via convolutional deep neural network[C]//Proceedings of the 25th International Conference on Computational Linguistics,2014:2335-2344.
- [46] Xu Kun, Feng Yansong, Huang Songfeng, et al. Semantic relation classification via convolutional neural networks with simple negative sampling[C]//Proceedings of the 15th Conference on Empirical Methods in Natural Language Processing,2015:536-540.
- [47] Wang Linlin, Cao Zhu, Melo G D, et al. Relation classification via multi-level attention CNNs[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics,2016:1298-1307.
- [48] Zhu Jizhao, Qiao Jianzhong, Dai Xinxiao, et al. Relation classification via target-concentrated attention CNNs[C]//Proceedings of the 24th International Conference ICOLNIP,2017:137-146.
- [49] Xu Yan, Mou Lili, Li Ge, et al. Classifying relations via long short term memory networks along shortest dependency paths[C]//Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing,2015:1785-1794.
- [50] Zhang Shu, Zheng Dequan, Hu Xinchun, et al. Bidirectional long short-term memory networks for relation classification[C]//Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation,2015:73-78.
- [51] Zhang Yuhao, Qi Peng, Manning C. Graph convolution over pruned dependency trees improves relation extraction[C]//Proceedings of the Empirical Methods in Natural Language Processing (EMNLP),2018:2205-2215.
- [52] Guo Zhijiang, Zhang Yan, Lu Wei. Attention guided graph convolutional networks for relation extraction[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics,2019:241-251.
- [53] Sun Kai, Zhang Richong, Mao Yongyi, et al. Relation extraction with convolutional network over learnable syntax-transport graph[C]//Proceedings of the 34th AAAI Conference on Artificial Intelligence,2020.
- [54] Miwa M, Bansal M. End-to-end relation extraction using LSTMs on sequences and tree structures[C]//Proceedings of the Meeting of the Association for Computational Linguistics,2016:1105-1116.
- [55] Katiyar A, Cardie C. Going out on a limb: joint extraction of entity mentions and relations without dependency trees[C]//Proceedings of the Meeting of the Association for Computational Linguistics,2017:917-928.
- [56] Sundermeyer M, Schlüter R, Ney H. LSTM neural networks for language modeling[C]//13th Annual Conference of the International Speech Communication Association,2012.
- [57] 李枫林,柯佳.基于深度学习框架的实体关系抽取研究进展[J].情报科学,2018,36(3):169-176.
- [58] Bruna J, Zaremba W, Szlam A, et al. Spectral networks and locally connected networks on graphs[C]//Proceedings of the International Conference on Learning Representations,2013:1-14.
- [59] Jiang Xiaotian, Wang Quan, Li Peng, et al. Relation extraction with multi-instance multilabel convolutional neural networks[C]//Proceedings of the 26th International Conference on Computational Linguistics,2016:1471-1480.
- [60] Lin Yankan, Shen Shiqi, Liu Zhiyuan, et al. Neural relation extraction with selective attention over instances[C]//Proceedings of the Meeting of the Association for Computational Linguistics,2016:2124-2133.
- [61] Gao Tianyu, Han Xu, Liu Zhiyuan, et al. Hybrid attention-based prototypical networks for noisy few-shot relation classification[C]//Proceedings of the 33rd AAAI Conference on Artificial Intelligence,2019:6407-6414.
- [62] Zheng Suncong, Wang Feng, Bao Hongyun, 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.
- [63] Zeng Daojian, Liu Kang, Chen Yubo, et al. Distant supervision for relation extraction via piecewise convolutional neural networks[C]//Proceedings of the Empirical Methods in Natural Language Processing,2015:1753-1762.
- [64] Ji Guoliang, Liu Kang, He Shizhu, et al. Distant supervision for relation extraction with sentence-level attention and entity descriptions[C]//Proceedings of the 31st

- AAAI Conference on Artificial Intelligence, 2017: 3060-3066.
- [65] Feng Jun, Huang Minlie, Zhao Li, et al. Reinforcement learning for relation classification from noisy data[C]// Proceedings of the 32nd AAAI Conference on Artificial Intelligence, 2018: 5779-5786.
- [66] Qin Pengda, Xu Weiran, Wang W. Robust distant supervision relation extraction via deep reinforcement learning[C]// Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, 2018: 2137-2147.
- [67] Takanobu R, Zhang Tianyang, Liu Jiexi, et al. A hierarchical framework for relation extraction with reinforcement learning[C]// Proceedings of the 33rd AAAI Conference on Artificial Intelligence, 2019: 7072-7079.
- [68] Luo Bingfeng, Feng Yansong, Wang Zheng, et al. Learning with noise: enhance distantly supervised relation extraction with dynamic transition matrix[C]// Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 2017: 430-439.
- [69] 余恒. 基于深度学习的中文文本实体关系抽取研究与实现[D]. 北京: 北京邮电大学, 2019.
- [70] Ren Xiang, Wu Zeqiu, He Weiqi, et al. CoType: joint extraction of typed entities and relations with knowledge bases[C]// Proceedings of the 26th International Conference on World Wide Web, 2017: 1015-1024.
- [71] Fu T, Li P, Xu Weiran, et al. 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.
- [72] Ye Zhixiu, Ling Zhenhua. Multi-level matching and aggregation network for few-shot relation classification[C]// Proceeding of the 57th Annual Meeting of the Association for Computational Linguistics, 2019: 2872-2881.
- [73] Christensen J, Mausam M, Soderland S, et al. An analysis of open information traction based on semantic role labeling[C]// Proceedings of the 6th International Conference on Knowledge Capture, Banff, 2011: 113-120.
- [74] Liu Xiao, Luo Zhunchen, Huang Heyan. Jointly multiple events extraction via attention-based graph information aggregation[C]// Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018: 1247-1256.
- [75] Nguyen T H, Grishman R. Graph convolutional networks with argument-aware pooling for event detection[C]// Proceedings of the 32nd AAAI Conference on Artificial Intelligence, 2018: 5900-5907.
- [76] Marcheggiani D M, Bastings J, Titov I. Exploiting semantics in neural machine translation with graph convolutional networks[C]// Proceeding of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2018: 486-492.
- [77] Han Xu, Zhu Hao, Yu Pengfei, et al. FewRel: a large-scale supervised few-shot relation classification dataset with state-of-the-art evaluation[C]// Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018: 4803-4809.
- [78] Soares L B, FitzGerald N, Ling J, et al. Matching the blanks: distributional similarity for relation learning[C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019: 2895-2905.
- [79] Gao Tianyu, Han Xu, Zhu Hao, et al. FewRel 2.0: towards more challenging few-shot relation classification[C]// Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language, 2019: 6251-6259.