

限定域关系抽取研究综述^①

袁清波, 杜晓明, 杨帆

(陆军工程大学 指挥控制工程学院, 南京 210007)

通讯作者: 袁清波, E-mail: 12661967@qq.com



摘要: 随着当前知识图谱的蓬勃发展, 关系抽取作为信息抽取的关键一环, 已受到越来越多研究者的关注. 关系抽取发展至今, 总体可以分为基于模板的抽取方法和基于机器学习的抽取方法; 之后随着深度学习抽取方法的广泛应用, 关系抽取的性能得到了较大提高. 本文利用时间顺序法对限定域条件下二元关系抽取方法进行归纳总结. 首先对关系抽取的概念定义、数据集以及评价指标等内容进行了简要介绍; 随后对关系抽取的相关方法进行了系统梳理, 重点分析了目前研究较热的深度学习关系抽取方法; 最后对关系抽取的未来研究方向及其应用进行了分析和展望.

关键词: 限定域; 关系抽取; 机器学习; 深度学习; 远程监督

引用格式: 袁清波, 杜晓明, 杨帆. 限定域关系抽取研究综述. 计算机系统应用, 2021, 30(9): 24-40. <http://www.c-s-a.org.cn/1003-3254/8079.html>

Survey on Relation Extraction in Restricted Domain

YUAN Qing-Bo, DU Xiao-Ming, YANG Fan

(College of Command and Control Engineering, Army Engineering University of PLA, Nanjing 210007, China)

Abstract: Amid the vigorous development of the knowledge graph, relation extraction, as a key part of information extraction, has attracted increasing attention from researchers. In general, relation extraction can be divided into template-based extraction and machine learning-based extraction. Later, with the extensive application of the extraction methods based on deep learning, the performance of relation extraction has been greatly improved. In this study, the time sequence method is employed to summarize the extraction methods of binary relations in a restricted domain. This study first briefly introduces the concept, data set, and evaluation indicators of relation extraction. Then it systematically sorts out the related extraction methods and highlights the current research on the relation extraction methods based on deep learning. Finally, it analyzes the future research direction and application of relation extraction.

Key words: restricted domain; relation extraction; machine learning; deep learning; distant supervision

随着计算机和互联网的快速发展, 人类产生、创造的数据量呈爆炸式增长. 如何对这些海量异构数据进行高效利用, 是当前亟需解决的问题, 也是信息抽取 (Information Extraction, IE) 研究的重要内容之一. 关系抽取 (Relation Extraction, RE) 作为信息抽取的重要子任务, 主要完成从非结构化数据中抽取出实体间的语义关系, 目前已广泛应用于搜索引擎、智能推荐、机

器翻译和问答服务等领域.

当前, 知识图谱技术研究较为热门, 关系抽取作为其中的一个重要方向, 引起了研究者们广泛关注. 对关系抽取的开拓性探索是 FASTUS 系统^[1] 的基于模板关系抽取方法, 该方法能够在小规模特定领域取得较好效果, 但是存在开发周期长、准确率较低、可移植性差等问题. 因此, 研究者们继而提出了有监督、弱监督

① 基金项目: 全军军事类研究生资助课题 (JY2019C078)

Foundation item: Military Postgraduate Research Project of PLA (JY2019C078)

收稿时间: 2020-12-07; 修改时间: 2021-01-11; 采用时间: 2021-01-20; csa 在线出版时间: 2021-09-02

和无监督等机器学习关系抽取方法. 有监督的关系抽取方法最初主要基于传统的机器学习, 包括 Miller 等人^[2]的基于特征向量关系抽取方法和 Zelenko 等人^[3]的基于核函数关系抽取方法. 基于传统的机器学习关系抽取方法比较依赖于人工构建的各种特征, 近年来基于深度学习的关系抽取方法开始被研究者们提出, 使关系抽取的性能取得了较大提高. 深度学习关系抽取方法不需要人工构建特征, 其输入一般包括句子中的词向量和位置向量表示, 分为流水线 (pipeline) 方法和联合 (joint) 方法两种, 其比较有代表性的是 Zeng 等人^[4]提出的卷积深度神经网络模型, 通过提取词汇和句子级别的特征以进行关系抽取. 有监督学习关系抽取方法的缺点是需要大量有标注的训练语料, 而人工标注这些语料则耗时耗力. 当训练语料较少时, 则可以利用弱监督学习方法来进行关系抽取, 该方法主要包括 Mintz 等人^[5]的基于远程监督 (distant supervision) 关系抽取方法和 Brin 等人^[6]基于 Bootstrapping 关系抽取方法. 无监督的关系抽取方法, 不需要人工标注数据, 通过聚类方法把相同关系类型聚合在起来.

近年来, 一些研究者对关系抽取方法进行过梳理, 如庄传志等人^[7]梳理了深度学习方法, 白龙等人^[8]梳理了远程监督方法, 但缺少对关系抽取方法整体发展脉络的系统归纳与总结.

本文主要利用时间顺序法对限定域条件下二元关系抽取方法进行归纳总结. 首先对关系抽取进行简要介绍; 其次对其相关方法进行详细归纳总结, 重点是是目前研究较热的深度学习关系抽取方法; 最后对其应用领域和未来研究方向进行分析展望.

1 概念定义

Bach 等人^[9]将关系以元组 $t=(e_1, e_2, \dots, e_n)$ 的形式定义, 其中实体 e_i 之间在文档 D 中具有某种预先定义关系 r . 二元关系可表示为 $\langle e_1, r, e_2 \rangle$ 的三元组形式, 其中 e_1 、 e_2 表示两个实体, r 为预定义目标关系类型. 根据实体数量的多少关系抽取可分为二元关系抽取和多元关系抽取, 目前大多数研究者都专注于二元关系抽取.

关系抽取的任务是从一段文本中提取出发生在两个实体或多个实体之间的语义关系. 完整的关系抽取包括两个子过程, 首先从非结构化数据中进行命名实体识别, 而后对所识别出的实体进行语义关系判别. 其中关系判别的过程, 也可称为关系分类 (relation

classification), 即判断实体之间属于哪种关系, 属于一种多分类问题.

如图 1 所示, 对于给定句子“姚明出生于上海的一个篮球世家”. 命名实体识别子过程识别出句子中具有“姚明”和“上海”两个实体; 语义关系判别子过程判断出“姚明”和“上海”两个实体间是“出生地”关系. 在阅读关系抽取相关论文时发现, 大多数的方法默认命名实体识别子过程已完成, 所做研究主要在于关系判别子过程, 即关系分类子过程.

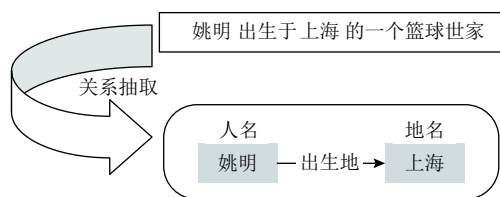


图 1 关系抽取示例

2 数据集

关系抽取的公开主流评测数据集主要有 ACE2005 数据集、SemEval-2010 任务 8 数据集、NYT2010 数据集等.

ACE 2005 数据集^[10]是一种多语种训练语料库, 包含完整的英语、阿拉伯语和汉语训练数据. ACE 2005 数据集中英文语料来源包括广播新闻、广播谈话、新闻专线、网站博客、新闻组和电话对话语音等共 6 个领域, 共包含 599 篇与上述领域内容相关的文档. ACE 2005 数据集里共定义了 7 大类实体和 7 大类、19 小类关系.

SemEval-2010 任务 8 数据集^[11]包含 10 717 个带注释的示例, 其中训练集包含 8 000 个示例, 测试集包含 2 717 个示例. 共定义有 9 种关系类型, 每种类型都有两个方向; 不属于这些关系类型的示例被标记为“Other”. 每个示例都包含标有两个名词 e_1 和 e_2 的句子, 任务是在考虑方向性的情况下预测 e_1 和 e_2 间关系. 在考虑方向性的情况下, SemEval-2010 任务 8 数据集共有 19 种关系类型.

纽约时报 (New York Times, NYT) 2010 数据集是用于远程监督关系抽取的标准语料库, 由 Riedel 等人^[12]发布于 2010 年. 该数据集是通过将 Freebase 知识库与纽约时报语料库对齐而形成, 使用斯坦福 NER 系统从文本中提取的命名实体, 并自动链接到

Freebase 知识库中的实体. 该数据集中共有 53 种可能的关系类型, 包括一种特殊的关系 NA, 它表示实体之间没有任何关系.

3 评价指标

关系抽取任务经常采用的评价指标有正确率 (Accuracy)、准确率 (Precision)、召回率 (Recall)、F 值 (F-Measure).

(1) 正确率

正确率是最常见的评价指标, 而且很容易理解, 就是被正确抽取的样本数除以所有的样本数, 通常来说, 正确率越高, 分类器越好. 其计算公式为:

$$Accuracy = \frac{\text{被正确抽取的样本数}}{\text{所有的样本数}} \quad (1)$$

(2) 准确率

准确率又称为查准率, 是针对抽取结果而言的, 它表示的是抽取结果为关系 R 的样本中有多少是对的. 把抽取结果集中正确抽取的记为 TP (True Positive), 错误抽取的记为 FP (False Positive). 其计算公式为:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

(3) 召回率

召回率又称为查全率, 是针对原来的样本而言的, 它表示的是存在关系 R 的样本中有多少被正确抽取

了. 把存在关系 R 的样本集中正确抽取的记为 TP , 错误抽取的记为 FN (False Negative). 其计算公式为:

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

(4) F 值

对于关系抽取来说, 准确率和召回率两个指标有时候会出现相互矛盾的情况, 二者实际上为互补关系. 这样就需要综合考虑它们, 最常见的方法就是 F 值, 又称为 F -Score. 其计算公式为:

$$F\text{-Measure} = \frac{(\beta^2 + 1) \times Precision \times Recall}{Precision + Recall} \quad (4)$$

其中, β 是用来平衡准确率和召回率在 F 值计算中的权重. 在关系抽取任务中, 一般 β 取 1, 认为两个指标一样重要. 此时 F 值计算公式为:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (5)$$

4 相关方法

关系抽取发展至今, 总体可以分为基于模板的关系抽取方法和基于机器学习的关系抽取方法, 具体如图 2 所示. 基于机器学习的关系抽取方法按照对语料的依赖程度分为 3 类: 有监督的关系抽取、弱监督的关系抽取和无监督的关系抽取. 在传统机器学习方法之后, 随着深度学习抽取方法的广泛应用, 关系抽取的性能得到了较大提高.

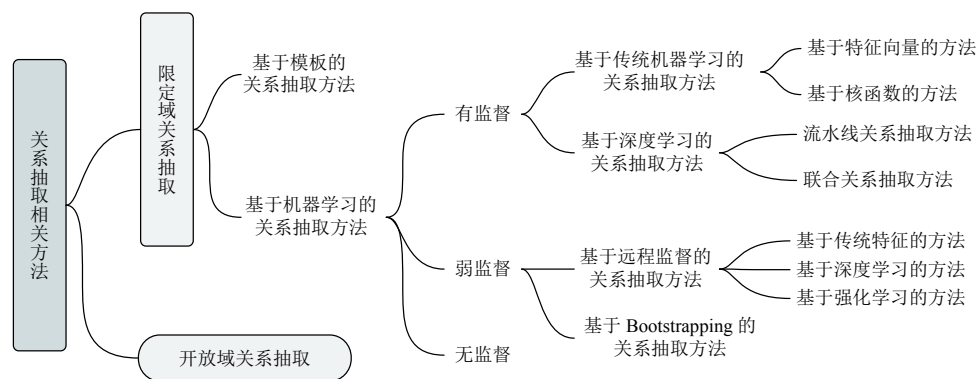


图2 关系抽取相关方法

4.1 基于模板的方法

早期的实体关系抽取方法大都采用基于模板匹配的方法实现. 该方法又称为基于规则或模式匹配的方法, 基于语言学知识和专业领域知识, 由专家手工编写

模板, 构造出基于词语、词性或语义的模板集合, 来实现特定关系实体对的抽取. 方法对比如表 1 所示.

在 1995 年 MUC-6 会议上, Appelt 等人^[1]提出了 FASTUS 抽取系统, 并引入了“宏”这一概念, 用户只需

在“宏”中修改相应参数,即可快速配置好特定领域实体间的关系抽取模板.在1998年MUC-7会议上,Yangarber等人^[13]提出了Proteus抽取系统,采用基于样本泛化的思想来构建关系抽取模板,用户通过分析含有关系的例句,进而对识别出的要素进行泛化后而形成关系抽取模板,这进一步改善了关系抽取的通用性.Aone等人^[14]提出了IE²抽取系统,通过人工编写关系抽取模板,从文本中抽取与模板匹配的关系实例.Fukumoto等人^[15]提出了Oki抽取系统,通过利用实体之间的谓语信息来判定两个实体间的语义关系.Humphreys等人^[16]提出了LaSIE-II抽取系统,通过对句子进行相应句法分析,从而人工构造出复杂的句法规则来识别出实体间的语义关系.

总体来说,基于模板的关系抽取方法能够在小规模特定领域取得较好效果,但是存在以下问题:第一,开发人员需在特定领域专家的指导下手工编写关系抽取模板集合,人工参与量大,系统开发周期长;第二,当抽取模板集合较小时,模板的覆盖范围不够,系统召回率不高;当抽取模板集合比较复杂时,不同模板之间容

易产生冲突,导致系统准确率降低;第三,当系统移植到其他领域时,需要重新编写抽取模板,系统可移植性较差,难以得到广泛使用.

表1 基于模板的关系抽取方法对比

序号	模型	年份	发表会议	F值评测结果
1	FASTUS ^[1]	1995	MUC-6	51
2	Proteus ^[13]	1998	MUC-7	76.5
3	IE ² ^[14]	1998	MUC-7	83.9
4	Oki ^[15]	1998	MUC-7	39.1
5	LaSIE-II ^[16]	1998	MUC-7	54.7

虽然基于模板的关系抽取方法存在一定缺陷,但却对后续基于机器学习的关系抽取研究起到了相应促进作用.

4.2 基于传统机器学习的方法

基于传统机器学习的关系抽取方法属于有监督的关系抽取方法,将二元关系抽取看作是分类问题,其主要工作在于如何抽取出表征两个实体间语义关系的有效特征.该方法通常分为基于特征向量(feature-based)的方法和基于核函数(kernel-based)的方法.

基于传统机器学习的关系抽取方法如表2所示.

表2 基于传统机器学习的关系抽取方法对比

抽取方法	序号	模型	年份	评测数据集	F值评测结果
特征向量	1	增强解析树 ^[2]	2000	MUC-7 Template Relation	71.23
	2	最大熵 ^[17]	2004	ACE	52.8
	3	支持向量机 ^[18]	2005	ACE	55.5
	4	条件随机场 ^[19]	2006	Wikipedia	67.91
核函数	5	树核函数 ^[3]	2003	Person-affiliation	86.8
				Organization-location	83.3
	6	依存树核函数 ^[20]	2004	ACE	45.8
	7	最短路徑依存核函数 ^[21]	2005	ACE	52.5
	8	子序列核函数 ^[22]	2006	ACE	47.7
	9	复合核函数 ^[23]	2006	ACE	72.1
	10	上下文相关解析树核函数 ^[24]	2007	ACE	75.8

(1) 基于特征向量的方法

基于特征向量的方法通过使用空间向量模型(vector space model),将自然句子转换为特征向量.该方法进行关系抽取通常分为以下3个步骤:第1步是特征提取,即从训练集句子中提取出如词汇、句法和语义等特征;第2步是模型训练,即利用第1步提取出的有效特征来训练我们的分类器模型;第3步是关系抽取,即利用训练好的模型对测试集中的句子进行分类,进而完成实体关系抽取.基于特征向量的方法根据分类器模型的不同可分为:基于句法解析增强的方

法、基于最大熵(maximum entropy)的方法、基于支持向量机(Support Vector Machine, SVM)的方法及基于条件随机场(Conditional Random Fields, CRF)的方法等.

Miller等人^[2]在2000年设计了一个联合模型(joint model)进行关系抽取,主要解决传统流水线模型(pipeline model)中前一步操作会传播到后一步操作的错误问题.Kambhatla^[17]在2004年基于最大熵模型并结合词汇、句法和语义特征进行关系抽取,只使用非常简单的词汇特征也取得了较高准确率,大大降低了

对句法分析树的依赖. Zhou 等人^[18]在2005年基于支持向量机模型并通过融合不同词汇、句法和语义知识进行关系抽取,该模型在 Kambhatla 的基础上,通过加入基本短语组块信息特征和 WordNet、Name List 等语义信息,来提升关系抽取效果. Culotta 等人^[19]在2006年基于条件随机场模型来进行关系抽取,提出了一种集成的有监督机器学习方法和构建了一个线性链条件随机场来提高关系抽取性能.

基于特征向量的关系抽取方法,虽然在一定程度上可以取得不错的效果,但其对关系实体对的上下文信息利用不够充分.此外,对于特征向量的选择和设置方面更多的依赖于模型构建者的经验知识,于是后续研究者们进而提出了基于核函数的关系抽取方法.

(2) 基于核函数的方法

基于核函数的关系抽取方法,不需要人为选择和设置特征向量,而是直接使用原始字符串作为输入,来计算任意两个实体间的核相似性(kernel similarity)函数.该方法进行关系抽取通常分为以下3个步骤^[25]:第1步是选择合适解析结构,即为了承载语句中隐含的特征信息,如使用语法树等剖析语句;第2步是选择合适基础核函数,即在解析结构基础上,来定量地计算解析结构中子成分的相似性;第3步是复合多个核函数,即为了充分利用各种特征,来提高分类精度.

Zelenko 等人^[3]在2003年首次将核函数应用于关系抽取任务中,提出一种使用核函数从非结构化自然语言文本中提取关系的方法. Culotta 等人^[20]在2004年提出使用依存树核函数方法进行关系抽取,该模型在文献[3]树核函数方法基础上进行改进,在依存关系树上定义树内核,并将该内核合并到 SVM 中,可以用来在 ACE 新闻语料库中进行实体关系的检测和抽取. Bunesu 等人^[21]在2005年提出使用最短路径依存核函数方法进行关系抽取,该方法优于文献[20]的依存树核函数方法.但由于其在计算两个实体间最短路径时要求依存树具有相同的节点数和高度,因此在一定程度上限制了其使用范围.于是两人^[22]在2006年提出使用子序列核函数方法以提升关系抽取的效果,新的模型使用了3种子序列核函数模式,用于抽取自然语言文本中实体之间的语义关系. Zhang 等人^[23]在2006年提出了复合核函数方法进行关系抽取,该模型中复合核函数由实体核函数和卷积解析树核函数构成,可以充分利用核函数方法的优良特性进行关系抽取. Zhou

等人^[24]在2007年提出了一种具有上下文相关结构化解析树信息的树核函数方法进行关系抽取,该模型通过扩展广泛使用的最短路径封闭树(SPT)来包含SPT之外的必要上下文信息,自动确定动态上下文相关树的跨度以进行关系抽取.

基于核函数方法的关系抽取方法,可以充分利用文本的长距离特征和结构化特征,实验结果表明其优于基于特征向量的关系抽取方法.但由于核函数是在高维的特征空间中隐式地计算对象间的距离,因此不可避免的会引入噪声,对特征向量的有效性会造成一定影响.此外,核函数的计算过程复杂度高,模型的训练过程相对比较慢,因此不适用于在大规模语料库中进行关系抽取.

4.3 基于深度学习的方法

基于传统的机器学习关系抽取方法比较依赖于人工构建的各种特征,近年来基于深度学习的关系抽取方法开始被研究者们提出.深度学习不需要人工构建特征,其输入一般包括句子中的词向量和位置向量表示.基于深度关系的关系抽取方法通常包括流水线(pipeline)关系抽取方法和联合(joint)抽取关系抽取方法两种.

基于深度学习的关系抽取典型方法如表3所示.

(1) 流水线方法

基于流水线的方法将命名实体识别和关系抽取作为两个独立的过程进行处理,关系抽取在命名实体识别完成的基础上进行.其过程可以描述为:把已经标注好命名实体对的句子作为模型输入,而后把实体关系三元组作为预测结果进行输出.基于流水线的方法通常包括基于卷积神经网络(Convolutional Neural Network, CNN)的方法、基于循环神经网络(Recurrent Neural Network, RNN)的方法、基于依存关系的方法和基于BERT的方法.

① 基于卷积神经网络模型的方法

基于卷积神经网络模型的关系抽取方法发展脉络如图3所示. Liu 等人^[26]在2013年提出一种结合词汇特征的卷积神经网络用于关系抽取,主要针对以前研究忽略词间语义信息的情况而引入同义词编码,通过同义词词典对输入词进行编码,将语义知识集成到神经网络中以进行关系抽取. Zeng 等人^[4]在2014年提出利用卷积深度神经网络模型(CNN+Softmax)提取词汇和句子级别的特征进行关系抽取,该模型无需复杂的预处理就可以将所有单词标记作为输入,就可以预测

两个标记名词之间的关系. 在 Zeng 等人^[4]的模型基础上, Nguyen 等人^[27]在 2015 年提出了一种新的改进模型, 其输入完全没有使用人工特征, 而是使用多尺寸卷积核进行 N-Gram 特征抽取. Dos Santos 等人^[28]在 2015 年提出了一种利用卷积神经网络排序进行关系分类的模型 (CR-CNN), 该模型为每个关系类型学习分布式矢量表示, 对于给定输入文本段, 使用卷积层生成文本的分布式矢量表示形式, 并将其与关系类别表示形式进行比较, 以便为每个关系类别生成分数. Shen 等人^[29]在 2016 年提出了一种基于注意力机制的 CNN 模型

(Attention-CNN) 进行关系抽取, 该模型使用 CNN 来提取句子级别特征, 将文本段编码为其语义表示, 并可以充分利用单词嵌入, 词性标签嵌入和位置嵌入信息. Wang 等人^[30]在 2016 年提出了一种基于多级注意力机制的 CNN 模型 (Att-Pooling-CNN) 进行关系抽取, 主要依赖两个层次的注意力机制, 以便更好地识别异构上下文中的模式. Zhu 等人^[31]在 2017 年提出了一种基于目标集中注意力机制的 CNN 模型 (TCA-CNN) 进行关系分类, 认为一个句子中不同的词具有不同的信息量, 并且词的重要性高度依赖于关系.

表 3 基于深度学习的关系抽取典型方法比较

抽取方法	序号	模型	年份	评测数据集	F1评测值
卷积神经网络	1	CNN ^[26]	2013	ACE 2005	83.9
	2	CNN+Softmax ^[4]	2014	SemEval-2010 Task 8	82.7
	3	CNN +N-Gram ^[27]	2015	SemEval-2010 Task 8	82.8
	4	CR-CNN ^[28]	2015	SemEval-2010 Task 8	84.1
	5	Attention-CNN ^[29]	2016	SemEval-2010 Task 8	85.9
	6	Att-Pooling-CNN ^[30]	2016	SemEval-2010 Task 8	88.0
	7	TCA-CNN ^[31]	2017	SemEval-2010 Task 8	87.3
循环神经网络	8	RNN+PI ^[32]	2015	SemEval-2010 Task 8	79.6
	9	BLSTM ^[33]	2015	SemEval-2010 Task 8	84.3
	10	Hier-BLSTM ^[34]	2016	SemEval-2010 Task 8	84.3
	11	Att-BLSTM ^[35]	2016	SemEval-2010 Task 8	84.0
	12	EAtt-BiGRU ^[36]	2017	SemEval-2010 Task 8	84.7
	13	LET-BLSTM ^[37]	2019	SemEval-2010 Task 8	85.2
流水线Pipeline	14	MV-RNN ^[38]	2012	SemEval-2010 Task 8	82.4
	15	FCM ^[39]	2014	SemEval-2010 Task 8	83.0
	16	DepNN ^[40]	2015	SemEval-2010 Task 8	83.6
	17	depLCNN+NS ^[41]	2015	SemEval-2010 Task 8	85.6
	18	SDP-LSTM ^[42]	2015	SemEval-2010 Task 8	83.7
	19	BRCNN ^[43]	2016	SemEval-2010 Task 8	86.3
	20	DRNNs ^[44]	2016	SemEval-2010 Task 8	86.1
依存关系	21	SDP-CFN ^[45]	2018	SemEval-2010 Task 8	86.4
	22	R-BERT ^[46]	2019	SemEval-2010 Task 8	89.25
	23	BERT _{EM} +MTB ^[47]	2019	SemEval-2010 Task 8	89.5
BERT	24	LSTM-RNN ^[48]	2016	SemEval-2010 Task 8	85.5
	25	LSTM+SLL+RLL ^[49]	2016	MPQA 2.0 corpus	54.98 (IS-ABOUT) 58.22 (IS-FROM)
	26	Bi-LSTM+ Attention ^[50]	2017	ACE 2004 ACE 2005	45.7 53.6
	27	BiLSTM-ED+CNN ^[51]	2017	ACE 2005	52.1
	28	LSTM-LSTM-Bias ^[52]	2017	NYT	52.0
	29	MRT ^[53]	2018	NYT	51.7
	30	Novel Graph Scheme ^[54]	2018	NYT	50.9
联合Joint	27	BiLSTM-ED+CNN ^[51]	2017	ACE 2005	52.1
	28	LSTM-LSTM-Bias ^[52]	2017	NYT	52.0
	29	MRT ^[53]	2018	NYT	51.7
	30	Novel Graph Scheme ^[54]	2018	NYT	50.9

② 基于循环神经网络模型的方法

基于循环神经网络模型的关系抽取方法发展脉络

如图 4 所示. Zhang 等人^[32]在 2015 年首次提出使用循环神经网络模型来进行关系分类, 认为关系分类问题

本质上可以看作是时间序列学习的任务问题,因此应通过时间模型进行建模.与CNN等其他深度学习模型相比,RNN模型可以处理远距离模式,因此特别适合于在较长上下文中学习关系.Zhang等人^[33]在2015年提出了双向长短期记忆网络(BLSTM)模型来解决关系分类问题,对于给定句子中的每个单词,BLSTM都有关于其前后所有单词的完整的顺序信息,在一定程度上可以解决长距离关系.Xiao等人^[34]在2016年提出了基于注意力机制的分层循环神经网络模型(Hier-BLSTM)进行语义关系分类,采用两个注意机制RNN从原始句子中学习有用的特征以进行关系分类.Zhou等人^[35]在2016年提出了基于注意力机制的双向长短期记忆网络(Att-BLSTM)模型进行语义关系分类,不依赖于NLP系统或词汇资源派生的任何特征,而是使用带有位置指示符的原始文本作为输入.Qin等人^[36]在2017年提出了一种基于实体对的双向注意力机制门控循环单元(Gated Recurrent Unit, GRU)模型(EAtt-BiGRU),该模型利用双向GRU捕获有价值的字符级信息,针对具体实例,将对应的实体对信息作为先验知识.Lee等人^[37]在2019年提出了一种结合实体感知注意机制和潜在实体类型(Latent Entity Type, LET)的端到端循环神经网络模型(LET-BLSTM)进行关系分类,该模型为了捕捉句子的上下文,通过自我注意力机制来获得单词的表示,并用双向长短期记忆网络来构建循环神经结构.

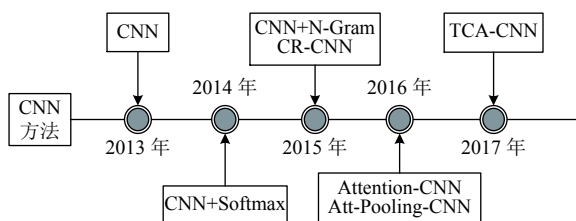


图3 基于CNN模型的方法发展时间轴

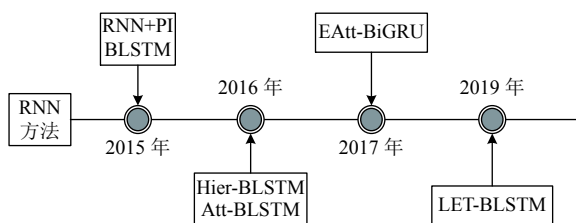


图4 基于RNN模型的方法发展时间轴

③ 基于依存关系模型的方法

基于依存关系模型的关系抽取方法发展脉络如图5

所示. Socher等人^[38]在2012年提出了一种矩阵向量递归神经网络模型(MV-RNN), 其为解析树中的每个节点分配向量和矩阵, 并通过根据解析树的句法结构递归地组合单词, 可以自底向上计算较长短语的表示形式. Yu等人^[39]在2014年提出了基于因子的组合嵌入模型(FCM)用于关系分类, 将带注释的句子分解为子结构因子后, 结合子结构特征与单词嵌入形成子结构嵌入, 再用Softmax层来预测这个句子级嵌入的输出标签. Liu等人^[40]在2015年提出了一种基于依存关系的神经网络模型(DepNN)进行关系分类, 提出了增强依存路径(Augmented Dependency Path, ADP)结构, 利用递归神经网络对子树进行建模和卷积神经网络捕获最短路径上的重要特征. Xu等人^[41]在2015年提出了一种简单负采样的卷积神经网络模型(depLCNN+NS)进行语义关系分类, 利用CNN从最短依存路径中学习更鲁棒的关系表示, 避免来自其它不相关语块或从句的负面影响, 并引入一种负采样策略来解决关系的方向性. Xu等人^[42]在2015年提出了一种基于最短依存路径的长短期记忆网络模型(SDP-LSTM)进行关系分类, 利用具有长期短期记忆单元的多通道递归神经网络沿最短依存路径来收集异构信息. Cai等人^[43]在2016年提出了一种双向循环卷积神经网络模型(BRCNN)进行关系分类, 利用基于双通道LSTM的循环神经网络对SDP中的全局模式进行编码, 并利用卷积层捕获由依存关系连接的每两个相邻单词的局部特征. Xu等人^[44]在2016年提出了一种基于数据扩充的深度循环神经网络模型(DRNNs)来进行关系分类, 通过多个RNN层来探索不同抽象级别和粒度下的表示空间, 同时提出了一种利用关系方向性进行数据扩充的方法. Sun等人^[45]在2018年提出了一种结合最短依存路径监督下关键词选择的粗粒度和细粒度网络模型(SDP-CFN)进行关系分类, 提出了粗粒度和细粒度两种关系分类网络、SDP监督下的选词网络和一种新的相反损失函数.

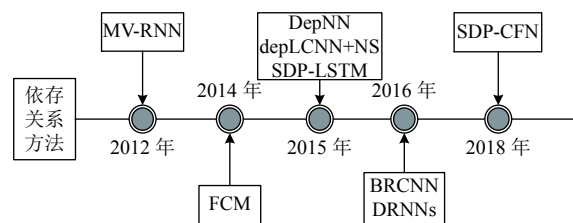


图5 基于依存关系模型的方法发展时间轴

④ 基于 BERT 模型的方法

Wu 等人^[46]在 2019 年提出了一种基于实体信息来丰富 BERT 预训练语言模型的方法 (R-BERT) 来进行关系分类, 结合预训练的 BERT 模型和目标实体信息来解决关系分类问题, 通过预先训练的体系结构定位目标实体并传递信息, 并合并两个实体的相应编码. Soares 等人^[47]在 2019 年提出了一种通用关系提取器 (BERTEM+MTB), 主要基于 Harris 的分布假设对关系的扩展以及在学习文本表示形式 (特别是 BERT) 方面的最新进展, 以完全从实体链接的文本构建与任务无关的关系表示形式. 通过实验证明, BERTEM+MTB 大大优于 SemEval 2010 Task 8 上的先前方法, 取得了目前最高 F1 值 89.5.

以上基于深度学习的流水线关系抽取方法, 其关系抽取的结果过度依赖于命名实体识别的结果, 忽略了两个过程之间存在的关系, 一定程度上影响了关系抽取的效果; 同时, 命名实体识别过程的错误也会传播到关系抽取过程中, 即流水线方法通常存在的错误累积传播问题.

(2) 联合方法

为了避免流水线方法所带来的错误累积传播问题, 基于深度学习的联合关系抽取方法开始被提出. 联合关系抽取方法使用单个模型将命名实体识别和关系抽取两个过程结合在一起, 以便在统一的模型中进行共同优化. 联合关系抽取方法发展脉络如图 6 所示.

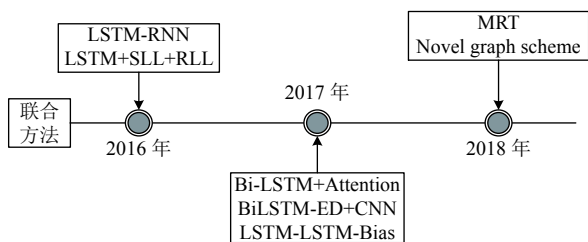


图 6 联合方法发展时间轴

Miwa 等人^[48]在 2016 年提出了一种基于序列和树结构的 LSTMs 端到端模型进行实体和关系联合抽取. 该模型在循环神经网络的基础上通过在双向序列 LSTM-RNNs 上叠加双向树结构 LSTM-RNNs 来同时捕获字序列和依存树子结构信息, 这允许模型在单个模型中使用共享参数共同表示实体和关系. Katiyar 等人^[49]在 2016 年提出了基于双向 LSTMs 的实体与关

系联合抽取模型. 这是第一次尝试使用深度学习方法处理完整的意见实体和关系抽取任务. 但该模型只能用于联合抽取意见实体和 IS-FROM 和 ISABOUT 关系, 而无法用于抽取其他实体关系类型, 不具有通用性. Katiyar 等人^[50]为了解决上述问题, 又在 2017 年提出了一种基于注意力机制的循环神经网络模型, 在不使用任何依存树信息的情况下进行实体标记和关系的联合提取. Zheng 等人^[51]在 2017 年提出了一种基于混合神经网络的实体与关系联合抽取模型, 可以在不需要任何人工特征的情况下提取实体及其语义关系. 该模型包含一个用于命名实体识别的双向编解码模块 (BiLSTM-ED) 和一个用于关系分类的 CNN 模块, 可以用来捕获实体标签之间的长距离关系. Zheng 等人^[52]在 2017 年又提出了一种基于标记方案的实体与关系联合抽取模型 (LSTM-LSTM-Bias). 该模型设计了一种带有端到端模型的标记方案, 其中包含实体及其所具有的关系的信息, 从而将实体和关系的联合抽取转换为了标记问题. Sun 等人^[53]在 2018 年提出了一种基于最小风险训练 (MRT) 方法的轻量级实体与关系联合抽取模型. 基于 MRT 的方法的优点是可以显式地优化全局句子级损失 (如 F1 值), 而不是局部标记级损失, 从而模型可以在训练时间内捕捉更多的句子级信息, 在测试时间内更好地匹配评价指标. Wang 等人^[54]在 2018 年提出了一种基于图形方案的实体与关系联合抽取模型. 该模型用直观的图形方案来共同表示实体和关系, 从而将端到端的关系抽取很容易地转换成类似解析的任务.

以上的实体与关系联合抽取模型所采用的方法基本上可以分为两大阵营, 一种是以 Miwa 等人^[48]为代表基于参数共享的实体关系抽取方法, 另外一种是以 Zheng 等人^[51]为代表基于序列标注的实体关系抽取方法. 基于参数共享的方法很好地缓解了流水线方法所带来的错误累积传播问题以及两个子过程间关系被忽视的问题. 而基于序列标注的方法在上两个问题之外, 还解决了流水线方法中的实体冗余的问题.

4.4 基于弱监督的方法

基于有监督学习的关系抽取方法需要大量有标注的训练语料, 而人工标注这些语料则耗时耗力, 尤其是面对海量非结构化网络数据时, 问题则更加突出. 当训练语料较少时, 则可以利用弱监督学习方法来进行关系抽取. 弱监督关系抽取主要包括基于远程监督 (distant supervision) 的方法和基于 Bootstrapping 的方法.

(1) 基于远程监督的方法

远程监督方法所基于的假设为: 若两个实体间存在某种关系, 则所有包含这两个实体的句子都可能以某种方式表达这种关系^[5]. 该方法通过将非结构化文本

中的实体对与知识图谱进行对齐来自动标注训练样本, 以构建大量的训练数据, 从而减少对人工的依赖, 增强了模型的跨领域适应能力.

基于远程监督的关系抽取典型方法如表4所示.

表4 基于远程监督的关系抽取典型方法对比

抽取方法	序号	模型	年份	评测数据集	指标	评测值
传统特征	1	Mintz ^[5]	2009	Wikipedia-FB	Precision (Top100)	77
	2	Riedel ^[12]	2010	NYT-FB	Precision (Top1000)	91
	3	MultiR ^[55]	2011	NYT-FB	Precision (Top100)	83
	4	MIML ^[56]	2012	NYT-FB	Precision (Top100)	85
	5	PROP ^[57]	2012	NYT-FB	Precision (Top100)	89
深度学习	6	PCNNs+MIL ^[58]	2015	NYT-FB	Precision (Top100)	86
	7	PCNN+ATT ^[59]	2016	NYT-FB	Precision (Top100)	76.2
	8	MIMLCNN ^[60]	2016	NYT-FB	Precision (Top100)	69
	9	APCNNs+D ^[61]	2017	NYT-FB	Precision (Top100)	87
	10	COTPYE ^[62]	2016	NYT	F1	66
	11	BGRU+STP+EWA+TL ^[63]	2018	NYT-10	Precision (Top100)	87
	12	RESIDE ^[64]	2018	Riedel dataset.	Precision (Top100)	84
	13	HRERE ^[65]	2019	NYT	Precision (Top100)	88
强化学习	13	CNN+RL ^[66]	2018	NYT	F1	42
	14	PE+REINF ^[67]	2018	NYT	Accuracy	49.2

① 基于传统特征的远程监督方法

基于传统特征的远程监督方法发展脉络如图7所示. Mintz等人^[5]在2009年首次将远程监督的方法运用到实体关系抽取任务中. 该模型使用一个具有数千个关系的大型语义数据库 Freebase 来提供远程监督, 对于出现在 Freebase 关系库的每一对实体, 在一个大的未标记语料库 Wikipedia 中找出包含这些实体的所有句子, 并提取文本特征以训练一个关系分类器. Riedel等人^[12]在2010年提出了一种基于无向图模型和远程监督框架的多实例学习模型来进行关系抽取. 该模型认为 Mintz等人^[5]先前的假设过于绝对, 进而提出了“at least one sentence”假设, 即如果将所有包含两个相关实体的句子看成是一个包, 那么这个包中至少有一个句子表达了相应的关系. Hoffmann等人^[55]在2011年提出了一种可以处理多重关系的多实例学习概率图模型 MultiR 来进行关系抽取. Mintz及Riedel等人的模型假设实体间只存在一种关系, 事实上部分实体间往往会存在多重关系. Surdeanu等人^[56]在2012年提出了一种多实例多标签学习模型 MIML 用于关系抽取, 认为一个句子中一对实体所表达的关系是未知的, 有可能有关系或根本没有关系. 该模型利用具有潜

在变量的概率图模型, 将文本中含有实体对的所有实例及其所有关系标签联合起来进行建模. Takamatsu等人^[57]在2012年提出了一种减少使用远程监督假设创建的错误关系标签数量的方法. 该模型直接模拟了远程监督的启发式标记过程, 并使用预测模式来判断指定的关系标签是正确的还是错误的. 如果模式能够成功预测关系标签, 那么标注样本应当保留; 如果不能, 则标注样本应该抛弃.

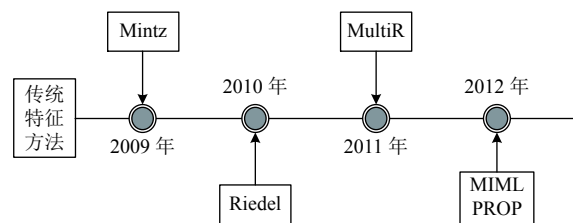


图7 基于传统特征的远程监督方法发展时间轴

以上模型方法都是基于传统特征的, 然而传统特征的设计周期相对较长, 而且模型应用范围受限, 扩展性较差. 同时, 在提取特征时会使用自然语言处理工具, 导致误差不断向下传播, 进而影响关系抽取的效果.

② 基于深度学习的远程监督方法

近年来, 随着深度学习神经网络的快速发展, 目前基

于深度学习的远程监督方法相关研究已经占据主导地位. 基于深度学习的远程监督方法发展脉络如图8所示.

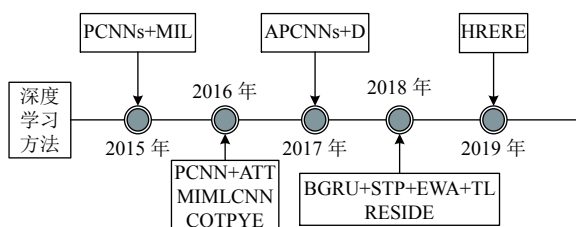


图8 基于深度学习的远程监督方法发展时间轴

Zeng 等人^[58]在2015年提出了一种多实例学习分段卷积神经网络(PCNNs+MIL)关系抽取模型. 该模型将远程监督关系抽取问题看作是一个多实例问题, 考虑了实例标签的不确定性; 其无需复杂的NLP预处理, 而是采用了带分段最大池化操作的卷积神经网络来自动学习相关特征. Lin 等人^[59]在2016年提出了一种基于句子级选择性注意力机制的卷积神经网络(PCNN+ATT)关系抽取模型, 是对Zeng 等人^[58]模型的改进, 主要还是解决远程监督错误标注的问题. Jiang 等人^[60]在2016年提出了一种多实例多标签卷积神经网络(MIMLCNN)关系抽取模型. 该模型首先放宽了“至少一次表达”的假设, 并采用了跨句最大池化操作, 以使信息可以在不同句子之间共享; 然后, 使用神经网络分类器通过多标签学习处理多重关系. Ji 等人^[61]在2017年提出使用句子级注意力机制结合实体描述信息的远程监督模型(APCNNs+D)进行关系抽取. 其中, 注意力机制可以通过为有效实例分配较高的权重, 为无效实例分配较低的权重, 来选择包中的多个有效实例; 而实体描述信息可以提供更多背景知识. 除此之外, Ren 等人^[62]在2016年提出了一种联合抽取模型(COTPYE)用于关系抽取, 以解决将实体抽取和关系抽取两项工作分别进行所带来的错误累积传播问题. Liu 等人^[63]在2018年提出了一种基于语句内降噪和迁移学习模型(BGRU+STP+EWA+TL)进行关系抽取, 通过建立子树解析(STP)来去除与关系无关的噪声词, 运用实体注意力机制来识别每一个实例中关系词的重要语义特征, 通过转移学习从实体分类的相关任务中学习先验知识使模型对噪声具有更强的鲁棒性. Vashishth 等人^[64]在2018年提出了一种基于边信息的图形卷积网络(GCN)模型(RESIDE)进行关系抽取, 一方面利用知识库中的附加边信息来改进远程监督关系提取,

另一方面利用GCN从文本中对语法信息进行编码. Xu 等人^[65]在2019年提出了一种将语言和知识与异构表示联系起来以进行神经关系抽取的模型(HRERE), 使用知识库嵌入(KBE)进行链路预测来改进关系抽取, 通过一个统一学习关系抽取和知识库嵌入模型的框架帮助缩小差距, 从而显著提高关系抽取的效果.

③ 基于强化学习的远程监督方法

自从AlphaGo在围棋领域打败了人类专业棋手后, 强化学习就进入了众多研究者的视野. 将强化学习应用到远程监督关系抽取领域的研究也开始不断出现.

Feng 等人^[66]在2018年提出了一种使用强化学习框架来解决远程监督关系抽取中噪声的模型(CNN+RL). CNN+RL模型主要用于在句子层面上降噪, 由一个实例选择器和一个关系分类器组成. 实例选择器为关系分类器选择高质量的句子, 关系分类器在句子级别预测关系, 并向选择器提供奖励, 以作为监督实例选择过程的微弱信号. 这将实例选择实际转化为一个强化学习问题. Zeng 等人^[67]同样在2018年提出了使用强化学习方法在远程监督数据集上进行关系抽取的模型(PE+REINF). PE+REINF模型遵循“至少表达一次”的假设来预测包关系, 但从预测的角度重新表述: 当预测包的关系时, 当且仅当包中的所有句子都表示NA关系(无关系)时, 包才是NA关系, 否则, 包是用它的句子表示的真实关系. 模型利用实体对的关系作为远程监督, 借助强化学习方法指导关系抽取器的训练.

基于远程监督的关系抽取方法作为弱监督方法的一种, 其优点在于可以通过较低成本获得大量训练数据, 克服了有监督方法需要大量标注数据的弊端, 其在面对大量无标注数据时会显现出特有的优势, 在一些缺乏标注数据集的垂直领域中具有很好的应用前景; 其缺点在于当所采用的知识库不完备及有错误标注时会产生噪声问题, 将导致抽取效果比有监督方法差很多, 这给关系抽取准确率的提升带来了较大挑战. 因此, 在远程监督关系抽取方法研究中, 如何克服噪声问题已成为研究关注的重点.

(2) 基于 Bootstrapping 的方法

自动化模板抽取通常采用自举法(Bootstrapping)算法来实现, 如图9所示. 针对某个特定类型的关系实例抽取任务, 自举法的基本思想是: ① 实体标注, 即为该关系类型标注少量的初始种子实体对; ② 句子查找,

即找到实体对在数据集中所出现的句子集合;③模板提取,即基于上述句子集合提取表达关系的模板;④实例抽取,即使用经筛选和评估后的新模板去数据集中抽取新的实体对.上述“模板提取+实例抽取”的过程循环迭代,直至不再发现新的关系实例.这个过程也被称为“滚雪球(snowball)”^[68].

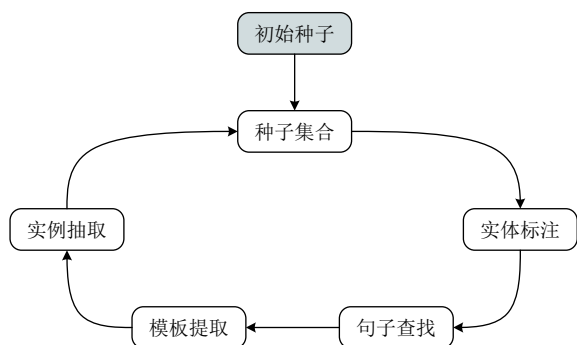


图9 基于 Bootstrapping 的关系抽取流程

基于自举法的关系抽取方法得到了广泛应用,出现了很多具有代表性的系统,主要有 Brin 等人^[6]在 1998 年提出的 DIPRE (Dual Iterative Parttern Relation Expansion) 抽取系统, Agichtein 等人^[69]在 2000 年提出的 Snowball 抽取系统, Etzioni 等人^[70]在 2005 年提出的 KnowItAll 抽取系统,以及卡内基梅隆大学 (CMU) 的 Tom Mitchell 教授领导的团队在 2010 年开发的 NELL (Never-Ending Language Learner) 抽取系统^[71].

自举法的优点是关系抽取系统构建成本低,不需要过多的人工标记数据,适合大规模的关系抽取任务.但是,自举法也存在不足之处,包括对初始种子集较为敏感、存在语义漂移问题、抽取结果准确率较低等.

4.5 基于无监督的方法

无监督关系抽取方法主要基于分布式假设理论:如果两个词语出现在相同上下文中且用法相似,那么它们意思相近.相应的,在关系抽取任务中,具有相同语义关系的实体对也倾向于具有相似的上下文语境,其上下文可作为表征该语义关系的特征.该方法进行关系抽取通常分为以下两个步骤:第一步是实体对聚类,即采用某种聚类方法将语义相似度高的实体对聚为一类;第二步是关系标记,即在上下方中选择具有代表性的词语来标记这种关系.

Hasegawa 等人^[72]在 2004 年首次提出了一种基于无监督的大型语料库关系发现方法,其核心思想是根

据命名实体之间上下文词语的相似度对命名实体进行聚类. Chen 等人^[73]在 2005 年对 Hasegawa 等人^[72]的方法进行了改进,该方法将每个实体对的上下文,而不是所有相同实体对的上下文,作为实体之间的语义关系特征. Rozenfeld 等人^[74]在 2006 年提出了一种无监督的关系识别和提取系统 URIES,该系统使用一种基于模式的上下文表示来代替实体对的上下文,使关系抽取结果取得了较大提高. Shinyama 等人^[75]在 2006 年提出了一种多层级聚类的无监督关系抽取方法,该方法试图在一个文档中发现多个实体之间的并行对应关系,并使用基本模式作为特征进行聚类. Bollegala 等人^[76]在 2010 年提出了一种用于针对 Web 上实体对的无监督关系抽取方法,该方法利用关系的对偶性,使用协同聚类来发现实体对其关系模板的聚类簇,并从中选择具有代表性的模板作为对应的关系. Yao 等人^[77]在 2012 年提出了一种基于语义消歧的无监督关系发现方法,该方法使用主题模型将实体对及其对应的关系模板分配到不同的语义类别上,然后再使用聚类方法将这些语义类别映射到语义关系上. Simon 等人^[78]在 2019 年提出了一种具有关系分布损失的正则化判别方法来进行无监督信息抽取,该模型通过引入偏度损失函数和分布距离损失函数来提高判别模型的性能. Tran 等人^[79]在 2020 年提出了一个简单的无监督关系抽取方法,该方法仅使用命名实体来推导关系类型,与现有方法相比可以获得更好的性能.

无监督关系抽取方法的优点是无需预先定义关系类型,并可以发现新的关系类型,适用领域范围广.但也存在缺点,由于发现的关系往往是相似模板的聚类,因此关系往往不具有语义信息,很难被用来构建知识库.如果要得到具有语义信息的关系,需要人工方式添加语义信息或将其与现有知识库的关系进行对齐.

4.6 有监督深度学习方法与远程监督方法对比

有监督的深度学习关系抽取方法所采用的数据集规模一般相对较小,以人工标为主,特点是噪声小、准确率高,但花费成本较高;训练出的关系抽取模型抽取效果较好,但领域可迁移性较差.

无监督的远程监督关系抽取方法的数据集主要采用远程知识库方式,数据集规模较大,特点是噪声大、准确率低,但花费成本较低;训练出的关系抽取模型抽取效果比有监督的方法相比差一些,但领域可迁移性相对较好.

有监督深度学习方法与远程监督方法对比具体如表5所示^[80]。

表5 有监督深度学习方法与远程监督方法对比

抽取方法	有监督深度学习	远程监督
数据集规模	相对较小	较大
数据集标注	人工标注为主	采用远程知识库
数据集特点	噪声小, 准确率高	噪声大, 准确率低
花费成本	较高	较低
抽取效果	较好	较差
领域迁移性	较差	较好

5 未来研究方向及应用分析

5.1 未来研究方向

尽管关系抽取在近20年间已得到了学术界的广泛关注和研究, 各种关系抽取方法也在不断地得到尝试, 但关系抽取在实际应用过程中仍然面临着很多挑战, 从理论研究转变为工业实践还有很长的路要走。下面对关系抽取的未来研究方法进行分析和展望。

(1) 基于模板和深度学习相融合的关系抽取研究

在早期的时候, 基于模板的关系抽取方法研究得相对比较多。基于模板的关系抽取方法优点是抽取准确率比较高, 但召回率较低。同时, 该方法还存在“完全匹配”或“硬匹配”问题, 即无法应用于语义相同而表述不同的短语中。而基于深度学习的关系抽取则能够匹配表述不同而语义相近的短语。因此, 如果能将两者融合在一起, 则有可能提高关系抽取的性能^[81]。

(2) 基于深度学习新进展的关系抽取研究

随着前些年深度神经网络在其他领域的成熟应用, 学者们已将各种神经网络模型(如CNN、RNN)相互结合应用于了关系抽取任务中, 获得了丰富的研究成果。近几年, 随着强化学习^[82,83]、生成对抗学习^[84,85]、图卷积神经网络^[86,87]、预训练模型^[88-90]等深度学习新技术的提出, 又有很多学者开始研究如何将这些方法应用于关系抽取中。因此, 未来可以尝试将这些新的深度学习技术应用于关系抽取任务中, 从而来提升关系抽取的效果。

(3) 段落级和篇章级的关系抽取研究

如今的关系抽取研究大多集中在词汇级和语句级层面, 很少有学者将其扩展到段落级和篇章级层面。而仅仅根据给定实体对和句子来抽取关系, 往往会缺乏必要的背景知识, 从而造成关系抽取的困难。如果能够根据整个段落或篇章文字来获取实体的背景知识, 则

有可能带来实体关系抽取性能的大幅提升。因此, 如何结合背景知识进行联合实体关系抽取以及进行段落级、篇章级的联合抽取研究具有重要的研究价值。

(4) 多元实体关系抽取研究

目前文献中关于关系抽取的研究, 多数集中于从单一句子中抽取二元关系。这些关系抽取模型基于的假设都是句子中给定的两个标注实体间只存在一种关系。而在实际生活中, 我们所面对的句子中的实体间存在着不止一种关系, 三元甚至是多元关系。简单的二元关系抽取模型已经无法满足现实任务的要求。因此, 针对多元实体关系的抽取研究将具有重要的现实意义。

5.2 应用分析

关系抽取作为信息抽取的一项重要内容和知识图谱构建中的一个关键环节, 具有十分广阔的应用前景。具体应用表现在以下几个方面:

(1) 有助于深层自然语言理解

目前的机器语言理解系统只能理解浅层次语言, 在深层次语言理解上正确率较低、效果难以令人满意。关系抽取作为句子、段落和篇章中关系内容抽取的一项关键技术, 对文本中核心内容的理解具有重要意义。因此, 文本语义关系抽取的研究对实现机器真正理解人类语言具有重要推动作用, 对机器翻译等自然语言处理领域的任务性能提升也将具有重要意义。

(2) 自动构建大规模知识图谱

当前很多互联网应用都离不开底层通用知识图谱和领域知识图谱的支撑。如何有效利用互联网上海量多源异构数据以构建大规模知识图谱, 则会对提升互联网应用性能起到重要作用。如果利用人工构建知识图谱特别是构建领域知识图谱的话, 则将面临构建成本高、开发周期长、知识覆盖率低和领域数据稀疏等一系列问题。利用关系抽取技术则可以很好地解决上述问题, 根据需求自动抽取结果以形成知识图谱。

(3) 为其它应用提供技术支持

关系抽取对问答系统和信息检索等具体应用均可提供技术支持。在问答系统中, 关键问题就是要构建一个与领域无关的问答类型体系, 而后找出与问答类型体系中每个问答类型相对应的答案模式, 这就需要关系抽取技术的支持。在信息检索中, 有了关系抽取技术的支持, 可以构建出大规模的知识图谱, 而后通过对检索信息进行关联搜索和推理, 可以为用户提供更加智能化的检索服务。

参考文献

- 1 Appelt DE, Hobbs JR, Bear J, *et al.* SRI international FASTUS system: MUC-6 test results and analysis. Proceedings of the 6th Conference on Message Understanding. Columbia, OH, USA. 1995. 237–248.
- 2 Miller S, Fox H, Ramshaw L, *et al.* A novel use of statistical parsing to extract information from text. Proceedings of the 1st Meeting of the North American Chapter of the Association for Computational Linguistics. Seattle, WA, USA. 2000. 226–233.
- 3 Zelenko D, Aone C, Richardella A. Kernel methods for relation extraction. The Journal of Machine Learning Research, 2003, 3: 1083–1106.
- 4 Zeng DJ, Liu K, Lai SW, *et al.* Relation classification via convolutional deep neural network. Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers. Dublin, Ireland. 2014. 2335–2344.
- 5 Mintz M, Bills S, Snow R, *et al.* Distant supervision for relation extraction without labeled data. 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. Suntec, Japan. 2009. 1003–1011.
- 6 Brin S. Extracting patterns and relations from the world wide web. Proceedings of International Workshop on The World Wide Web and Databases. Valencia, UK. 1998. 172–183. [doi: [10.1007/10704656_11](https://doi.org/10.1007/10704656_11)]
- 7 庄传志, 靳小龙, 朱伟建, 等. 基于深度学习的关系抽取研究综述. 中文信息学报, 2019, 33(12): 1–18. [doi: [10.3969/j.issn.1003-0077.2019.12.001](https://doi.org/10.3969/j.issn.1003-0077.2019.12.001)]
- 8 白龙, 靳小龙, 席鹏弼, 等. 基于远程监督的关系抽取研究综述. 中文信息学报, 2019, 33(10): 10–17. [doi: [10.3969/j.issn.1003-0077.2019.10.002](https://doi.org/10.3969/j.issn.1003-0077.2019.10.002)]
- 9 Bach N, Badaskar S. A review of relation extraction. https://www.researchgate.net/publication/265006408_A_Review_of_Relation_Extraction, (2011-05-12).
- 10 Walker C, Strassel S, Medero J, *et al.* ACE 2005 multilingual training corpus. Progress of Theoretical Physics Supplement, 2006, 110(110): 261–276. [doi: [10.35111/mwxc-vh88](https://doi.org/10.35111/mwxc-vh88)]
- 11 Hendrickx I, Kim S N, Kozareva Z, *et al.* Semeval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals. Proceedings of the Workshop on Semantic Evaluations: Recent Achievements and Future Directions. Uppsala, Sweden. 2009. 94–99.
- 12 Riedel S, Yao LM, McCallum A. Modeling relations and their mentions without labeled text. Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Barcelona, Spain. 2010. 148–163. [doi: [10.1007/978-3-642-15939-8_10](https://doi.org/10.1007/978-3-642-15939-8_10)]
- 13 Yangarber R, Grishman R. NYU: Description of the Proteus/PET system as used for MUC-7 ST. Seventh Message Understanding Conference (MUC-7). Fairfax, VA, USA. 1998.
- 14 Aone C, Halverson L, Hampton T, *et al.* SRA: Description of the IE2 system used for USAMUC-7. Seventh Message Understanding Conference (MUC-7). Fairfax, VA, USA. 1998.
- 15 Fukumoto J, Masui F, Shimohata M, *et al.* Oki electric industry: Description of the Oki system as used for MUC-7. Proceedings of the 7th Message Understanding Conference. Fairfax, VA, USA. 1998.
- 16 Humphreys K, Gaizauskas R, Azzam S, *et al.* University of Sheffield: Description of the LaSIE-II system as used for MUC-7. Seventh Message Understanding Conference (MUC-7). Fairfax, VA, USA. 1998.
- 17 Kambhatla N. Combining lexical, syntactic, and semantic features with maximum entropy models for extracting relations. Proceedings of the ACL 2004 on Interactive Poster and Demonstration Sessions. Barcelona, Spain. 2004. 22-es. [doi: [10.3115/1219044.1219066](https://doi.org/10.3115/1219044.1219066)]
- 18 Zhou GD, Su J, Zhang J, *et al.* Exploring various knowledge in relation extraction. Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics. Ann Arbor, MI, USA. 2005. 427–434. [doi: [10.3115/1219840.1219893](https://doi.org/10.3115/1219840.1219893)]
- 19 Culotta A, McCallum A, Betz J. Integrating probabilistic extraction models and data mining to discover relations and patterns in text. Proceedings of the Main Conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics. New York City, NY, USA. 2006. 296–303. [doi: [10.3115/1220835.1220873](https://doi.org/10.3115/1220835.1220873)]
- 20 Culotta A, Sorensen J. Dependency tree kernels for relation extraction. Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics. Barcelona, Spain. 2004. 423-es. [doi: [10.3115/1218955.1219009](https://doi.org/10.3115/1218955.1219009)]
- 21 Bunescu RC, Mooney RJ. A shortest path dependency kernel for relation extraction. Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing. Vancouver, BC, Canada. 2005. 724–731. [doi: [10.3115/1220575.1220666](https://doi.org/10.3115/1220575.1220666)]
- 22 Bunescu RC, Mooney RJ. Subsequence kernels for relation

- extraction. Advances in Neural Information Processing Systems. Vancouver, BC, Canada. 2006. 171–178.
- 23 Zhang M, Zhang J, Su J, *et al.* A composite kernel to extract relations between entities with both flat and structured features. Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the Association for Computational Linguistics. Sydney, Australia. 2006. 825–832. [doi: [10.3115/1220175.1220279](https://doi.org/10.3115/1220175.1220279)]
- 24 Zhou GD, Zhang M, Ji DH, *et al.* Tree kernel-based relation extraction with context-sensitive structured parse tree information. Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. Prague, Czech Republic. 2007. 728–736.
- 25 Zhang XF, Gao ZQ, Zhu M. Kernel methods and its application in relation extraction. Proceedings of 2011 International Conference on Computer Science and Service System. Nanjing, China. 2011. 1362–1365. [doi: [10.1109/CSSS.2011.5972181](https://doi.org/10.1109/CSSS.2011.5972181)]
- 26 Liu CY, Sun WB, Chao WH, *et al.* Convolution neural network for relation extraction. Proceedings of the 9th International Conference on Advanced Data Mining and Applications. Hangzhou, China. 2013. 231–242. [doi: [10.1007/978-3-642-53917-6_21](https://doi.org/10.1007/978-3-642-53917-6_21)]
- 27 Nguyen TH, Grishman R. Relation extraction: Perspective from convolutional neural networks. Proceedings of the 1st Workshop on Vector Space Modeling for Natural Language Processing. Denver, CO, USA. 2015. 39–48. [doi: [10.3115/v1/W15-1506](https://doi.org/10.3115/v1/W15-1506)]
- 28 Dos Santos C, Xiang B, Zhou BW. Classifying relations by ranking with convolutional neural networks. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Beijing, China. 2015. 626–634. [doi: [10.3115/v1/P15-1061](https://doi.org/10.3115/v1/P15-1061)]
- 29 Shen YT, Huang XJ. Attention-based convolutional neural network for semantic relation extraction. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. Osaka, Japan. 2016. 2526–2536.
- 30 Wang LL, Cao Z, De Melo G, *et al.* Relation classification via multi-level attention CNNs. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Berlin, Germany. 2016. 1298–1307. [doi: [10.18653/v1/P16-1123](https://doi.org/10.18653/v1/P16-1123)]
- 31 Zhu JZ, Qiao JZ, Dai XX, *et al.* Relation classification via target-concentrated attention CNNs. Proceedings of the 24th International Conference on Neural Information Processing. Guangzhou, China. 2017. 137–146. [doi: [10.1007/978-3-319-70096-0_15](https://doi.org/10.1007/978-3-319-70096-0_15)]
- 32 Zhang DX, Wang D. Relation classification via recurrent neural network. arXiv: 1508.01006, 2015.
- 33 Zhang S, Zheng DQ, Hu XC, *et al.* Bidirectional long short-term memory networks for relation classification. Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation. Shanghai, China. 2015. 73–78.
- 34 Xiao MG, Liu C. Semantic relation classification via hierarchical recurrent neural network with attention. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. Osaka, Japan. 2016. 1254–1263.
- 35 Zhou P, Shi W, Tian J, *et al.* Attention-based bidirectional long short-term memory networks for relation classification. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). Berlin, Germany. 2016. 207–212. [doi: [10.18653/v1/P16-2034](https://doi.org/10.18653/v1/P16-2034)]
- 36 Qin PD, Xu WR, Guo J. Designing an adaptive attention mechanism for relation classification. Proceedings of 2017 International Joint Conference on Neural Networks. Anchorage, AK, USA. 2017. 4356–4362. [doi: [10.1109/IJCNN.2017.7966407](https://doi.org/10.1109/IJCNN.2017.7966407)]
- 37 Lee J, Seo S, Choi YS. Semantic relation classification via bidirectional LSTM networks with entity-aware attention using latent entity typing. Symmetry, 2019, 11(6): 785. [doi: [10.3390/sym11060785](https://doi.org/10.3390/sym11060785)]
- 38 Socher R, Huval B, Manning CD, *et al.* Semantic compositionality through recursive matrix-vector spaces. Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. Jeju Island, Republic of Korea. 2012. 1201–1211.
- 39 Yu M, Gormley M, Dredze M. Factor-based compositional embedding models. NIPS Workshop on Learning Semantics. Montreal, QC, Canada. 2014. 95–101.
- 40 Liu Y, Wei FR, Li SJ, *et al.* A dependency-based neural network for relation classification. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Beijing, China. 2015. 285–290. [doi: [10.3115/v1/P15-2047](https://doi.org/10.3115/v1/P15-2047)]
- 41 Xu K, Feng YS, Huang SF, *et al.* Semantic relation

- classification via convolutional neural networks with simple negative sampling. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, Portugal. 2015. 536–540. [doi: [10.18653/v1/D15-1062](https://doi.org/10.18653/v1/D15-1062)]
- 42 Xu Y, Mou LL, Li G, *et al.* Classifying relations via long short term memory networks along shortest dependency paths. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, Portugal. 2015. 1785–1794. [doi: [10.18653/v1/D15-1206](https://doi.org/10.18653/v1/D15-1206)]
- 43 Cai R, Zhang XD, Wang HF. Bidirectional recurrent convolutional neural network for relation classification. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Berlin, Germany. 2016. 756–765. [doi: [10.18653/v1/P16-1072](https://doi.org/10.18653/v1/P16-1072)]
- 44 Xu Y, Jia R, Mou LL, *et al.* Improved relation classification by deep recurrent neural networks with data augmentation. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. Osaka, Japan. 2016. 1461–1470.
- 45 Sun YP, Cui Y, Hu JL, *et al.* Relation classification using coarse and fine-grained networks with SDP supervised key words selection. Proceedings of the 11th International Conference on Knowledge Science, Engineering and Management. Changchun, China. 2018. 514–522. [doi: [10.1007/978-3-319-99365-2_46](https://doi.org/10.1007/978-3-319-99365-2_46)]
- 46 Wu SC, He YF. Enriching pre-trained language model with entity information for relation classification. Proceedings of the 28th ACM International Conference on Information and Knowledge Management. Beijing, China. 2019. 2361–2364. [doi: [10.1145/3357384.3358119](https://doi.org/10.1145/3357384.3358119)]
- 47 Soares LB, FitzGerald N, Ling J, *et al.* Matching the blanks: Distributional similarity for relation learning. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy. 2019. 2895–2905. [doi: [10.18653/v1/P19-1279](https://doi.org/10.18653/v1/P19-1279)]
- 48 Miwa M, Bansal M. End-to-end relation extraction using LSTMs on sequences and tree structures. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin, Germany. 2016. 1105–1116. [doi: [10.18653/v1/P16-1105](https://doi.org/10.18653/v1/P16-1105)]
- 49 Katiyar A, Cardie C. Investigating LSTMs for joint extraction of opinion entities and relations. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Berlin, Germany. 2016. 919–929. [doi: [10.18653/v1/P16-1087](https://doi.org/10.18653/v1/P16-1087)]
- 50 Katiyar A, Cardie C. Going out on a limb: Joint extraction of entity mentions and relations without dependency trees. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Vancouver, BC, Canada. 2017. 917–928. [doi: [10.18653/v1/P17-1085](https://doi.org/10.18653/v1/P17-1085)]
- 51 Zheng SC, Wang F, Bao HY, *et al.* Joint extraction of entities and relations based on a novel tagging scheme. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. Vancouver, BC, Canada. 2017. 1227–1236. [doi: [10.18653/v1/P17-1113](https://doi.org/10.18653/v1/P17-1113)]
- 52 Zheng SC, Hao YX, Lu DY, *et al.* Joint entity and relation extraction based on a hybrid neural network. Neurocomputing, 2017, 257: 59–66. [doi: [10.1016/j.neucom.2016.12.075](https://doi.org/10.1016/j.neucom.2016.12.075)]
- 53 Sun CZ, Wu YB, Lan M, *et al.* Extracting entities and relations with joint minimum risk training. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels, Belgium. 2018. 2256–2265. [doi: [10.18653/v1/D18-1249](https://doi.org/10.18653/v1/D18-1249)]
- 54 Wang SL, Zhang Y, Che WX, *et al.* Joint extraction of entities and relations based on a novel graph scheme. Proceedings of the 27th International Joint Conference on Artificial Intelligence. Stockholm, Sweden. 2018. 4461–4467. [doi: [10.24963/ijcai.2018/620](https://doi.org/10.24963/ijcai.2018/620)]
- 55 Hoffmann R, Zhang CL, Ling X, *et al.* Knowledge-based weak supervision for information extraction of overlapping relations. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. Portland, OR, USA. 2011. 541–550.
- 56 Surdeanu M, Tibshirani J, Nallapati R, *et al.* Multi-instance multi-label learning for relation extraction. Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. Jeju Island, Republic of Korea. 2012. 455–465.
- 57 Takamatsu S, Sato I, Nakagawa H. Reducing wrong labels in distant supervision for relation extraction. Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics. Jeju Island, Republic of Korea. 2012. 721–729.
- 58 Zeng DJ, Liu K, Chen YB, *et al.* Distant supervision for relation extraction via piecewise convolutional neural networks. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon, Portugal. 2015. 1753–1762. [doi: [10.18653/v1/D15-1203](https://doi.org/10.18653/v1/D15-1203)]
- 59 Lin YK, Shen SQ, Liu ZY, *et al.* Neural relation extraction with selective attention over instances. Proceedings of the

- 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Berlin, Germany. 2016. 2124–2133. [doi: [10.18653/v1/P16-1200](https://doi.org/10.18653/v1/P16-1200)]
- 60 Jiang XT, Wang Q, Li P, *et al.* Relation extraction with multi-instance multi-label convolutional neural networks. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. Osaka, Japan. 2016. 1471–1480.
- 61 Ji GL, Liu K, He SZ, *et al.* Distant supervision for relation extraction with sentence-level attention and entity descriptions. Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco, CA, USA. 2017. 3060–3066.
- 62 Ren X, Wu ZQ, He WQ, *et al.* CoType: Joint extraction of typed entities and relations with knowledge bases. Proceedings of the 26th International Conference on World Wide Web. Perth, WA, Australia. 2017. 1015–1024. [doi: [10.1145/3038912.3052708](https://doi.org/10.1145/3038912.3052708)]
- 63 Liu TY, Zhang XS, Zhou WH, *et al.* Neural relation extraction via inner-sentence noise reduction and transfer learning. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels, Belgium. 2018. 2195–2204. [doi: [10.18653/v1/D18-1243](https://doi.org/10.18653/v1/D18-1243)]
- 64 Vashishth S, Joshi R, Prayaga S S, *et al.* RESIDE: Improving distantly-supervised neural relation extraction using side information. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels, Belgium. 2018. 1257–1266. [doi: [10.18653/v1/D18-1157](https://doi.org/10.18653/v1/D18-1157)]
- 65 Xu P, Barbosa D. Connecting language and knowledge with heterogeneous representations for neural relation extraction. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis, MN, USA. 2019. 3201–3206. [doi: [10.18653/v1/N19-1323](https://doi.org/10.18653/v1/N19-1323)]
- 66 Feng J, Huang ML, Zhao L, *et al.* Reinforcement learning for relation classification from noisy data. Proceedings of the 32nd AAAI Conference on Artificial Intelligence. New Orleans, LA, USA. 2018. 5779–5786.
- 67 Zeng XR, He SZ, Liu K, *et al.* Large scaled relation extraction with reinforcement learning. Proceedings of the 32nd AAAI Conference on Artificial Intelligence. New Orleans, LA, USA. 2018. 5658–5665.
- 68 王昊奋, 漆桂林, 陈华钧. 知识图谱: 方法、实践与应用. 北京: 电子工业出版社, 2020. 136–138.
- 69 Agichtein E, Gravano L. *Snowball*: Extracting relations from large plain-text collections. Proceedings of the 5th ACM Conference on Digital Libraries. San Antonio, TX, USA. 2000. 85–94. [doi: [10.1145/336597.336644](https://doi.org/10.1145/336597.336644)]
- 70 Etzioni O, Cafarella M, Downey D, *et al.* Unsupervised named-entity extraction from the web: An experimental study. Artificial Intelligence, 2005, 165(1): 91–134. [doi: [10.1016/j.artint.2005.03.001](https://doi.org/10.1016/j.artint.2005.03.001)]
- 71 Carlson A, Betteridge J, Kisiel B, *et al.* Toward an architecture for never-ending language learning. Proceedings of the 24th AAAI Conference on Artificial Intelligence. Atlanta, GA, USA. 2010.
- 72 Hasegawa T, Sekine S, Grishman R. Discovering relations among named entities from large corpora. Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-04). Barcelona, Spain. 2004. 415–422. [doi: [10.3115/1218955.1219008](https://doi.org/10.3115/1218955.1219008)]
- 73 Chen JX, Ji DH, Tan CL, *et al.* Unsupervised feature selection for relation extraction. Companion Volume to the Proceedings of Conference Including Posters/Demos and Tutorial Abstracts. Jeju Island, Republic of Korea. 2005.
- 74 Rozenfeld B, Feldman R. High-performance unsupervised relation extraction from large corpora. Proceedings of the 6th International Conference on Data Mining. Hong Kong, China. 2006. 1032–1037. [doi: [10.1109/ICDM.2006.82](https://doi.org/10.1109/ICDM.2006.82)]
- 75 Shinyama Y, Sekine S. Preemptive information extraction using unrestricted relation discovery. Proceedings of the Human Language Technology Conference of the NAACL, Main Conference. New York City, NY, USA. 2006. 304–311. [doi: [10.3115/1220835.1220874](https://doi.org/10.3115/1220835.1220874)]
- 76 Bollegala DT, Matsuo Y, Ishizuka M. Relational duality: Unsupervised extraction of semantic relations between entities on the Web. Proceedings of the 19th International Conference on World Wide Web. Raleigh, NC, USA. 2010. 151–160. [doi: [10.1145/1772690.1772707](https://doi.org/10.1145/1772690.1772707)]
- 77 Yao LM, Riedel S, McCallum A. Unsupervised relation discovery with sense disambiguation. Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Jeju Island, Republic of Korea. 2012. 712–720.
- 78 Simon É, Guigue V, Piwowarski B. Unsupervised information extraction: Regularizing discriminative approaches with relation distribution losses. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy. 2019. 1378–1387. [doi: [10.18653/v1/P19-1133](https://doi.org/10.18653/v1/P19-1133)]
- 79 Tran TT, Le P, Ananiadou S. Revisiting unsupervised

- relation extraction. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020. 7498–7505. [doi: [10.18653/v1/2020.acl-main.669](https://doi.org/10.18653/v1/2020.acl-main.669)]
- 80 鄂海红, 张文静, 肖思琪, 等. 深度学习实体关系抽取研究综述. 软件学报, 2019, 30(6): 1793–1818. [doi: [10.13328/j.cnki.jos.005817](https://doi.org/10.13328/j.cnki.jos.005817)]
- 81 肖仰华, 徐波, 林欣, 等. 知识图谱: 概念与技术. 北京: 电子工业出版社, 2020. 155–156.
- 82 Qin PD, Xu WR, Wang WY. Robust distant supervision relation extraction via deep reinforcement learning. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne, VIC, Australia. 2137–2147. [doi: [10.18653/v1/P18-1199](https://doi.org/10.18653/v1/P18-1199)]
- 83 Takanobu R, Zhang TY, Liu JX, *et al.* A hierarchical framework for relation extraction with reinforcement learning. Proceedings of the 33rd AAAI Conference on Artificial Intelligence. Honolulu, HI, USA. 2019. 7072–7079. [doi: [10.1609/aaai.v33i01.33017072](https://doi.org/10.1609/aaai.v33i01.33017072)]
- 84 Wu Y, Bamman D, Russell S. Adversarial training for relation extraction. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Copenhagen, Denmark. 2017. 1778–1783. [doi: [10.18653/v1/D17-1187](https://doi.org/10.18653/v1/D17-1187)]
- 85 Qin PD, Xu WR, Wang WY. Dsgan: Generative adversarial training for distant supervision relation extraction. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne, VIC, Australia. 2018. 496–505. [doi: [10.18653/v1/P18-1046](https://doi.org/10.18653/v1/P18-1046)]
- 86 Zhang YH, Qi P, Manning CD. Graph convolution over pruned dependency trees improves relation extraction. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels, Belgium. 2018. 2205–2215. [doi: [10.18653/v1/D18-1244](https://doi.org/10.18653/v1/D18-1244)]
- 87 Guo ZJ, Zhang Y, Lu W. Attention guided graph convolutional networks for relation extraction. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy. 2019. 241–251. [doi: [10.18653/v1/P19-1024](https://doi.org/10.18653/v1/P19-1024)]
- 88 Liu YH, Ott M, Goyal N, *et al.* RoBERTa: A robustly optimized bert pretraining approach. arXiv: 1907. 11692, 2019.
- 89 Lan ZZ, Chen MD, Goodman S, *et al.* ALBERT: A lite bert for self-supervised learning of language representations. Proceedings of the 8th International Conference on Learning Representations. Addis Ababa, Ethiopia. 2019.
- 90 Sun Y, Wang SH, Li YK, *et al.* ERNIE 2.0: A continual pre-training framework for language understanding. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(5): 8968–8975. [doi: [10.1609/aaai.v34i05.6428](https://doi.org/10.1609/aaai.v34i05.6428)]