

DOI:10.3969/j.issn.1671-0673.2021.05.012

基于深度学习的中文命名实体识别综述

郑洪浩, 宋旭晖, 于洪涛, 李邵梅, 郝一诺

(信息工程大学, 河南 郑州 450001)

摘要:命名实体识别作为信息抽取的核心任务,能够从文本中识别出各类命名实体。近年来,深度学习技术在字词表示、特征提取等方面上的应用,使中文命名实体识别任务取得了较为丰富的研究成果。目前,基于深度学习的中文命名实体识别技术,在特征提取的深度和模型的精确度上已逐渐超过了传统的基于规则的方法、基于特征工程的有监督方法和基于无监督的方法。围绕深度学习的识别框架,将现有基于深度学习的中文命名实体识别方法分嵌入层、编码层和标签解码层三部分进行介绍,并对未来可能的研究方向进行探讨和展望。

关键词:中文命名实体识别;深度学习;嵌入层;编码层;标签解码层

中图分类号:TP391

文献标识码:A

文章编号:1671-0673(2021)05-0590-07

Survey of Chinese Named Entity Recognition Based on Deep Learning

ZHENG Honghao, SONG Xuhui, YU Hongtao, LI Shaomei, HAO Yinuo

(Information Engineering University, Zhengzhou 450001, China)

Abstract: As the core task of information extraction, named entity recognition can identify various named entities in text. In recent years, deep learning technology has been applied to Chinese named entity recognition, and relatively rich research results have been achieved. At present, Chinese named entity recognition technology based on deep learning has gradually surpassed traditional rule-based methods, supervised methods based on feature engineering, and unsupervised methods in terms of feature extraction depth and model accuracy. Focusing on the recognition framework of deep learning, this paper introduces the existing deep learning-based Chinese named entity recognition methods into three parts: embedding layer, coding layer and label decoding layer. Finally, it discusses and looks forward to possible future research directions.

Key words: Chinese named entity recognition; deep learning; embedded layer; encoder layer; tag decoder layer

随着互联网技术的不断普及和发展,人们需要处理的数据爆炸式增长。如何从非结构化的文本中提取有效信息,已经成为当下亟须解决的问题。命名实体识别作为信息抽取的核心模块,广泛用于智能问答^[1]、机器翻译^[2]、信息检索^[3]等复杂的自然语言处理任务。命名实体识别是指识别文本中提到的每个命名实体并标记类型,实体的通用类型

包括实体名称(人名、地名、机构名)、时间表示(日期、时间)和数字表示(度量衡、百分比表达式)^[4]。近年来,随着知识图谱技术的兴起,命名实体识别得到广泛关注,其最新研究成果已发表于 ACL^[5]、NAACL^[6]、EMNLP^[7]等自然语言处理顶级会议。

传统的中文命名实体识别方法分为基于规则的方法^[8-9]、基于特征工程的有监督方法^[10-12]、基

收稿日期:2021-03-11;修回日期:2021-04-09

基金项目:国家自然科学基金资助项目(61601513);郑州市协同创新重大专项资助项目(162/32410218)

作者简介:郑洪浩(1992-),男,硕士生,主要研究方向为命名实体识别、关系抽取。

于无监督的方法^[13-14]。传统的中文命名实体识别方法或依赖于特征模板的制定或领域专家制定规则,在面对大规模多领域的复杂文本数据时,传统方法暴露出许多不足。

随着深度学习技术的发展及其在自然语言处理领域的广泛应用,研究人员陆续将深度学习技术应用于中文命名实体识别,取得了显著提升^[15-16]。基于深度学习的命名实体识别的方法论通过序列标注方法,在对每个字词进行标注的基础上,预测实体的边界和类型。笔者查阅了近年的综述文献^[15-16],发现基于深度学习的中文命名实体识别方法和经典的中文命名实体识别方法相比,其主要优势在于:①深度学习技术拥有强大的非线性转化能力、向量表示能力和计算能力;②深度学习无需复杂的特征工程,能够学习高维潜在语义信息;③深度学习模型是端到端的,避免了误差传递的问题^[17]。因此,本文重点围绕深度学习技术来深入探讨中文命名实体识别方法。

1 基于深度学习的识别框架

基于深度学习的识别框架如图1所示,分为嵌入层、编码层和标签解码层。其中嵌入层旨在将字词级别的特征转化为向量表示,编码层旨在进一步获取文本上下文特征,标签解码层旨在获取序列之间的规则特征并对编码层输出的特征向量进行分类。

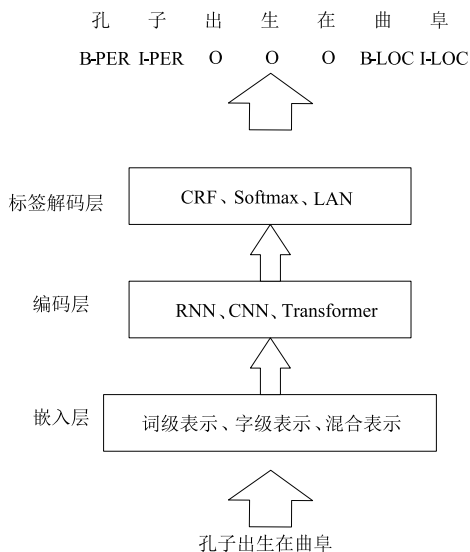


图1 深度学习的识别框架

1.1 嵌入层

嵌入层的功能是将字词级别的输入通过嵌入层模型得到对应的向量表示。相较于其他语言,中

文没有明显的词边界信息。按照输出向量的不同,将嵌入层的输出表示分为词级表示、字级表示和混合表示。下面对3种表示方式依次进行介绍。

1.1.1 词级表示

中文命名实体识别起步较晚,早期研究大多借鉴英文命名实体识别的思路,采用词级表示的方式。词级表示通过连续的CBOW和连续的skip-gram模型等无监督学习算法对大量无标签文本进行预训练获得^[18]。使用词级表示作为中文命名实体识别可以在模型训练期间进行固定或者进一步微调。

文献[19]提出一种SelfAttention-BiLSTM-CRF神经网络模型进行短文本的命名实体识别,SelfAttention-BiLSTM-CRF利用CBOW模型在大规模无监督的训练预料上学习词向量,每个词由300维向量表示。文献[20]面向司法领域的命名实体识别提出了深度学习模型BiLSTM-CRF。BiLSTM-CRF采用Skip-gram在CAIL2018_Small和CJNER_AI司法语料上训练词向量,语料内容包含20.8万份判决书,词向量由300维向量表示。

词级表示可以携带语义信息,但也存在以下两个缺点:①齐普夫定律表明:大部分的词都会被认定为未登录词,未登录词是指训练时未出现,测试时出现了的单词。词级表示会造成语料的稀疏,引起过拟合现象;②分词模型和词向量模型是分离的,并不是端到端的训练,存在误差传播的问题^[21]。

1.1.2 字级表示

字级表示是中文命名实体识别最常用的方法。和词级表示相比,已经有研究表明字级表示一般在中文命名实体识别任务上表现更优^[22-23]。其主要原因在于:①字级表示可以有效解决未登录词的问题;②字级表示不需要分词,避免了错误传播;③字级表示可以获取更多细粒度的特征,如“天安门”、“地安门”等词中的“门”字,携带明显的地名实体特征。

文献[24]提出用Transformer模型抽取出字级表示,并进一步输入到Transformer编码器中获取上下文信息。字级表示在中文命名实体识别任务上较词级表示有显著提升,但也存在不能携带语义信息的缺点^[25]。

1.1.3 混合表示

除了词级表示和字级表示,研究者们还提出了字词融合表示以及添加额外特征的表示,本文统一将该类表示称为混合表示。

字词融合表示一直是中文命名实体识别的重

点研究方向,主要可分结构融合与数据融合两类。结构融合是指通过调整模型结构将词信息融入到字级表示中,数据融合主要通过多任务学习实现。利用结构融合的研究有:文献[5]提出用 Lattice-LSTM 结构将词信息融合到字级表示之中,它能允许节点接收长距离依赖信息,并将模型从链式结构拓展为图式结构。文献[6]在 Lattice-LSTM 的基础上提出 WC-LSTM 模型,其重点改进了 Lattice-LSTM 模型不支持批量处理的缺点,利用 4 种策略为每个节点都添加一个且仅一个远程词信息。利用数据融合的研究有:文献[25]提出一种对抗迁移学习模型实现数据融合,它利用 3 个 LSTM 结构分别建模字信息、词信息和字词融合信息,本质上是一种多任务学习。相对于结构融合的方式,该模型优势在于能够缓解新词的问题。文献[26]利用 CNN 模型同时处理字信息提取和分词,二者共享参数,本质上也是一种多任务学习。

添加额外的特征会提升中文命名识别的性能,但同时也会损失模型的通用性。中文不同于英语等语言,除了字和词其语义信息存在于多维特征之中,如笔画、拼音等。文献[27]提出了一种基于部件和字向量的 LSTM-CRF 模型,额外获取了形旁和声旁部首信息。此外,近年来提出 BERT、XL-net^[28-29]、ERNIE^[30]等预训练语言模型通常通过大规模语料库进行训练,且在训练过程中需要加入位置编码等信息。鉴于此,本文将该类表示统一归纳为混合表示。

1.2 编码层

编码层的功能是对嵌入层输出的字词表示编码,进一步获取上下文信息。编码主要使用循环神经网络、卷积神经网络和 Transformer^[31]。

1.2.1 循环神经网络

循环神经网络及其变体(RNNs),如长短期记忆(Long Short-Term Memory, LSTM)、门控循环单元(Gated Recurrent Unit, GRU)以及上述模型的双向结构,在处理序列类时间数据方面有显著的优势^[32-33]。特别是双向循环神经网络可以有效地利用文本的前向信息和后向信息,获得文本的深层上下文语义。

文献[34]首先提出将 LSTM 模型应用于命名实体识别任务,LSTM 模型展现了良好的序列建模能力。跟随他们的工作,大量研究人员将 RNNs 应用于中文命名实体识别。其中最主流有效的方法就是使用双向 LSTM 模型对上下文信息进行编码。文献[35]将双向 LSTM 模型应用于网络电子病历

命名实体识别,文献[36]首先在渔业领域的命名实体识别上采用双向 LSTM 模型。为了缓解 LSTM 模型梯度消失(爆炸)的问题,文献[37]使用双向 GRU 模型进行通用实体的命名实体识别,实验表明双向 GRU 模型可以获取更长距离的文本信息。

1.2.2 卷积神经网络

卷积神经网络(Convolutional Neural Network, CNN)是一种深层的前馈神经网络,用于处理类似网络结构数据的网络。CNN 有 3 个结构上的特性:局部连接、权重共享以及汇聚,这些特性使得卷积神经网络具有一定程度上的平移、缩放和旋转不变性。

文献[38]最早提出用 CNN 获取上下文信息,其具体做法是利用卷积层提取每个词的局部特征构建全局特征向量,然后利用条件随机场模型计算对应标签的分数。其后,文献[39]在此基础上提出了一种利用膨胀卷积网络对上下文进行编码的方法,该方法在经典的卷积之上增加了一个膨胀宽度以增大感受视野,从而获取远距离的文本信息。和循环神经网络相比,卷积神经网络可以有效利用计算机的并行计算资源,减少训练时间。

1.2.3 Transformer

文献[31]提出的 Transformer 完全不同于循环神经网络和卷积神经网络,它利用自注意力机制和全连接层构建编码结构。Transformer 在多种自然语言处理任务上都实现了突破,且将训练时间大幅度压缩。

Transformer 可以有效解决 RNNs 存在的 3 点问题,具体解决方案如下:①Transformer 不同于 RNNs 的串行计算结构,采用并行计算结构以充分利用计算机的并行计算资源;②Transformer 采用自注意力机制,在结构上消除了梯度消失和梯度爆炸的问题,可以获取长文本的依赖信息;③Transformer 不同于 RNNs 的双向拼接,可以实现双向参数的统一更新,不会割裂上下文关系。相比于卷积神经网络,Transformer 不受卷积核感受野的影响,可以获得文本的长距离信息。

文献[40]采用 Transformer 对输入的电子病历文本中选择性学习,即通过注意力机制大量电子病历语料中筛选出有用的信息,且模型不受输入序列长度的限制。文献[24]对 Transformer 进行嵌入相对位置信息的改进,实验表明,基于改进 Transformer 的中文命名实体识别方法达到了领域最高效果。

1.3 标签解码层

标签解码层的功能是将编码层输出的向量进

行解码和分类。标签解码层的输入为编码层输出的向量表示,输出与文本对应的标签序列。标签解码层的主要网络有:多层感知机+Softmax 层和条件随机场(CRF)层。

1.3.1 多层感知机+Softmax 层

多层感知机+Softmax 层将命名实体识别转化为一个多分类问题,即将编码层输出的包含上下文语义的向量转化为对应的标签序列。该方法的优势在于具有较强的非线性表示能力,能够学习潜在的信息。多层感知机+Softmax 层的缺陷在于:该方法假设标签序列是相互独立的,然而标签之间存在一定的依赖关系和规则。标签的独立假设会带来信息损失,影响分类结果。

1.3.2 CRF 层

CRF 层是一种标签联合预测的方法,通过加入标签转移得分矩阵参数和定义序列的预测得分来进行全局优化。

文献[41]首次将 CRF 作为输出层加入到命名实体识别模型,在 4 种语言的数据集上获得了最优异的表现。此后,大量基于 CRF 输出结构的中文命名实体识别方法涌现而出。文献[42]提出一种对中文电子病历的命名实体识别方法,该方法利用 CRF 层对症状、检查、疾病、药物、治疗五类实体联合预测,避免了标签独立假设带来得到信息损失。

相比于多层感知机+Softmax 层,CRF 层可以进行全局优化以避免标签独立假设带来的信息损失。同时 CRF 层存在以下两个缺点:①CRF 层受到马尔可夫假设的限制,对标签相关信息较弱的数据处理效果较差;②CRF 层采用维特比解码技术,计算开销和时间成本较高。

2 基于深度学习的中文命名实体识别方法总结

表 1 总结了近年来的基于深度学习的中文命名实体识别方法。可以看出:BiLSTM-CRF 是目前最为主流的方法,其他方法大多在该方法的基础上少量改进。文献[43]提出的 Flat 模型在 Onto notes、Weibo、MSRA、Resume 四类主流数据集上都达到了最先进的性能。当下提高中文命名实体识别效果最有效的方式在于增强嵌入层的表示,BERT 等预训练语言模型和字词融合表示结合的方式已经成了一种新的范式。

对于嵌入层,字词融合表示以及增加额外特征表示的方式提高了中文命名实体识别的效果,也带来许多问题。其中字词融合表示存在以下两点问题:①给系统带来较高的时间复杂度;②不能解决新词的问题。其中增加额外特征表示的方式缺点在于:①人工特征表示会带来较高的成本;②人工选择的特征会影响深度学习端到端的学习方法。

对于编码层,RNNs 可以有效地利用文本的前向信息和后向信息,获得文本的深层上下文语义,仍是目前最主流的编码层方法。另一方面,Transformer 在结构上有效解决了 RNNs 的串行计算、梯度消失(爆炸)和单向建模的问题,更多地应用于编码层。

对于标签解码层,在标签数量较小时,CRF 仍是最佳的选择。在标签数量较大时,CRF 的性能会明显下降且时间复杂度很高。利用注意力机制的标签解码层可以有效解决这一问题。

在具体应用中,应根据数据量、标签数、时间成本、计算成本和通用性选择最适合的方法。

表 1 基于深度学习的中文命名实体识别方法

词级表示	嵌入层		编码层	标签解码层	实验结果
	字级表示	混合表示			
-	-	字词融合表示 (Lattice-LSTM)	LSTM	CRF	Onto notes: 73.88%, Weibo: 53.13% MSRA: 93.18 %, Resume: 94.46%
-	-	字词融合表示 (WC-LSTM)	LSTM	CRF	Onto notes: 74.43%, Weibo: 49.86% MSRA: 93.36 %, Resume: 94.96%
-	-	字词融合表示 (Muti-digraph)	LSTM	CRF	Onto notes: 76.0% ,MSRA: 94.40 %
-	-	字词融合表示 (LR-CNN)	CNN	CRF	Onto notes: 74.40 % ,Weibo: 56.54% MSRA: 93.71% , Resume: 95.11%
-	-	字词融合表示 (PLTE)	Transformer	CRF	Onto notes: 74.60%, Weibo: 55.15% MSRA: 93.26 %, Resume: 95.40%
-	注意力机制	-	CNN	LAN	MSRA: 93.40 % ,Weibo: 59.23% 军事文本数据集: 65.97%
-	Glove, CNN	-	LSTM	CRF	CCKS2017: 90.97% 人民日报: 92.19%

续表 1 基于深度学习的中文命名实体识别方法

词级表示	嵌入层		编码层	标签解码层	实验结果
	字级表示	混合表示			
-	CNN	五笔特征	LSTM	CRF	Ontonotes:75.84%, Weibo:63.31%
-	-	BERT	IDCNN	CRF	MSRA:94.41%
Word2vec	-	-	Transformer	CRF	电子病历数据集:95.02%
-	-	字词融合表示 (BGRU)	GRU	CRF	MSRA:94.26%, Onto Notes:76.56% 人民日报:97.40%
-	-	ERNIE	GRU	CRF	人民日报:94.07%
-	-	字词融合表示 (Flat)、BERT	Transformer	CRF	Onto notes:81.82%, Weibo:68.55% MSRA:96.09%, Resume:95.86%

3 未来展望

目前,基于深度学习的中文命名实体识别已成为业内的主流框架,但仍值得研究人员不断探索。通过对现有的中文命名实体识别工作进行总结,未来可以从以下几个方面展开相关研究。

3.1 标签数据稀疏问题

深度学习模型的特点之一是需要的大量的标签语料进行监督学习。然而,实体识别领域标签语料获取需要耗费较高的人力成本和时间成本。

标签语料稀疏问题是制约基于深度学习的中文命名实体识别模型效果的重要因素。针对该问题,研究者们已使用不同的方法。例如,利用迁移学习的方式,通过资源丰富的源任务或自监督方式获取更有效的语言表示以提升标签数据稀疏领域的实体识别效果,利用数据增强的方式,扩展训练数据集。其中效果最为显著地有两类方法,一类是文献[44]在2020年提出将传统的数据增强方法应用于NER中,将训练数据通过同标签替换、同义词替换、实体提及替换和顺序打乱四种方式进行扩充。另一类是文献[45]提出的语言模型生成方法,该方法将原始的序列标注标签和句子序列进行混合,丰富了实体的上下文信息。

然而,命名实体识别领域对于标签语料稀疏数据集的性能还普遍低于70%的F1值。命名实体识别是一项精细的序列标注任务,在进行全局结构化预测时,增强数据中的噪声会使模型产生错误倾向,导致指标下降。文献[46]提出一种负采样的方法对增强数据进行去噪,通过实验证实了数据去噪的有效性。因此,对数据去噪是解决标签数据稀疏问题的突破口。

3.2 非正式文本问题

如表1所示,现有基于深度学习的中文命名实体识别方法在新闻类、电子简历类等规范文本上的

识别F1值已经达到了较高的水平,但在社交类等非正式文本上的F1值最高仅为68.55%。社交类等非正式文本相比于规范文本存在篇幅较短、用语不规范、特征稀疏等特点。

针对该问题,研究者们普遍采用结合特定领域的外部知识的方式提高对非正式文本的命名实体识别性能。领域专家介入、构建专业词典等方法已经证实了融合外部知识的有效性^[47]。下一步,如何实现人机协同、提高外部知识的构建,是对非正式文本领域命名实体识别的重要研究方向。

3.3 嵌套实体问题

现今,大多数的基于深度学习的中文命名实体识别方法只处理扁平化的实体,嵌套实体的识别仍然是一大难题。尤其在生物医学等领域,嵌套实体的比例较大,严重限制了识别效果。

针对该问题,研究者提出基于超图、跨度、边界信息、状态转换和阅读理解等方法提高对嵌套实体的识别性能,实验验证了上述方法的有效性。然而,现阶段嵌套实体的语料库还较为稀少,无法有效评估模型的泛化能力。下一步,构建嵌套实体库仍是该领域值得进一步探索的方向。

参考文献:

- [1] PIZZATO L A, MOLLA D, PARIS C. Pseudo relevance feedback using named entities for question answering [C]//Proceedings of the 2006 Australasian Language Technology Workshop. Sydney, 2006: 83-90.
- [2] BABYCH B, HARTLEY A. Improving machine translation quality with automatic named entity recognition [C]//Proceedings of the 7th International EAMT Workshop on MT and other Language Technology Tools. Budapest, Hungary, 2003: 1-8.
- [3] MANDL T, WOMSER-HACKER C. The effect of named entities on effectiveness in cross-language information retrieval evaluation [C]//Proceedings of the 2005 ACM Symposium on Applied Computing. Santa Fe, New

- Mexico, 2005: 1059-1064.
- [4] GRISHMAN R, SUNDHEIM B. Message understanding conference-6: a brief history [C]//Proceedings of the 16th Conference on Computational Linguistics. Copenhagen, Denmark, 1996: 466-471.
- [5] ZHANG Y, YANG J. Chinese NER using lattice LSTM [DB/OL]. (2018-05-05) [2021-04-09]. <https://arxiv.org/abs/1805.02023>.
- [6] LIU W, XU T G, XU Q H, et al. An encoding strategy based word-character LSTM for Chinese NER [C]// Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics. Minneapolis, Minnesota, 2019: 2379-2389.
- [7] GUI T, ZOU Y C, ZHANG Q, et al. A lexicon-based graph neural network for Chinese NER [C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). Hong Kong, China, 2019: 1040-1050.
- [8] 向晓雯, 史晓东, 曾华琳. 一个统计与规则相结合的中文命名实体识别系统 [J]. 计算机应用, 2005, 25 (10): 2404-2406.
- [9] 李楠, 郑荣廷, 吉久明, 等. 基于启发式规则的中文化学物质命名识别研究 [J]. 现代图书情报技术, 2010(5): 13-17.
- [10] LUO W C, YANG F. An empirical study of automatic Chinese word segmentation for spoken language understanding and named entity recognition [C]//Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics. San Diego, California, 2016: 238-248.
- [11] 张玥杰, 徐智婷, 薛向阳. 融合多特征的最大熵汉语命名实体识别模型 [J]. 计算机研究与发展, 2008, 45(6): 1004-1010.
- [12] MILLER S, GUINNESS J, ZAMANIAN A. Name tagging with word clusters and discriminative training [C]// Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics. Boston, Massachusetts, USA, 2004: 337-342.
- [13] ALFONSECA E, MANANDHAR S. An unsupervised method for general named entity recognition and automated concept discovery [C]//Proceedings of the 1st International Conference on General Word Net Mysore India. 2002: 1-9.
- [14] ETZIONI O, CAFARELLA M, DOWNEY D, et al. Un-supervised named-entity extraction from the web: an experimental study [J]. Artificial Intelligence, 2005, 165 (1): 91-134.
- [15] 刘浏, 王东波. 命名实体识别研究综述 [J]. 情报学报, 2018, 37(3): 329-340.
- [16] 孙镇, 王惠临. 命名实体识别研究进展综述 [J]. 现代图书情报技术, 2010(6): 42-47.
- [17] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521(7553): 436-444.
- [18] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space [DB/OL]. (2013-01-16) [2021-04-09]. <https://arxiv.org/abs/1301.3781>.
- [19] 徐啸, 朱艳辉, 冀相冰. 基于自注意力深度学习的微博实体识别研究 [J]. 湖南工业大学学报, 2019, 33 (2): 48-52.
- [20] 林义孟. 面向司法领域的命名实体识别研究 [D]. 昆明: 云南财经大学, 2019.
- [21] LI X Y, MENG Y X, SUN X F, et al. Is word segmentation necessary for deep learning of Chinese representations [DB/OL]. (2019-05-14) [2021-04-09]. <https://arxiv.org/abs/1905.05526>.
- [22] HE J Z, WANG H F. Chinese named entity recognition and word segmentation based on character [C]//Proceedings of the Sixth SIGHAN Workshop on Chinese Language Processing. 2008.
- [23] LI H B, HAGIWARA M, LI Q, et al. Comparison of the impact of word segmentation on name tagging for Chinese and Japanese [C]//LREC. 2014: 2532-2536.
- [24] YAN H, DENG B, LI X N, et al. Tener: adapting transformer encoder for named entity recognition [DB/OL]. (2019-11-10) [2021-04-09]. <https://arxiv.org/abs/1911.04474>.
- [25] CAO P F, CHEN Y B, LIU K, et al. Adversarial transfer learning for Chinese named entity recognition with self-attention mechanism [C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. Brussels, Belgium, 2018: 182-192.
- [26] WU F Z, LIU J X, WU C H, et al. Neural Chinese named entity recognition via CNN-LSTM-CRF and joint training with word segmentation [C]//The World Wide Web Conference, San Francisco, CA, USA, 2019: 3342-3348.
- [27] DONG C H, ZHANG J J, ZONG C Q, et al. Character-based LSTM-CRF with radical-level features for chinese named entity recognition [M]//LIN C Y, XUE N W, ZHAO D Y, et al. Natural language understanding and intelligent applications. Cham: Springer, Cham, 2016: 239-250.
- [28] YANG Z L, DAI Z H, YANG Y M, et al. Flat: Chinese NER using flat-lattice transformer [DB/OL]. (2019-06-19) [2021-04-09]. <https://arxiv.org/abs/>

- 1906.08237? source=techstories.org.
- [29] 郑洪浩, 郝一诺, 于洪涛. 基于 Xlnet 嵌入的中文命名实体识别[J]. 信息工程大学学报, 2021, 22(4): 473-477.
- [30] ZHANG Z Y, HAN X, LIU Z Y, et al. ERNIE: enhanced language representation with informative entities [DB/OL]. (2019-05-17) [2021-04-09]. <https://arxiv.org/abs/1905.07129v3>.
- [31] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[DB/OL]. (2017-06-12) [2021-04-09]. <https://arxiv.org/abs/1706.03762>.
- [32] MA X Z, HOVY E. End-to-end sequence labeling via bi-directional LSTM-CNNs-CRF [DB/OL]. (2016-03-04) [2021-04-09]. <https://arxiv.org/abs/1603.01354>.
- [33] HUANG Z H, XU W, YU K. Bidirectional LSTM-CRF models for sequence tagging[DB/OL]. (2015-08-09) [2021-04-09]. <https://arxiv.org/abs/1508.01991>.
- [34] HAMMERTON J. Named entity recognition with long short-term memory [C]//Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL. Edmonton, Canada, 2003: 172-175.
- [35] 杨红梅, 李琳, 杨日东, 等. 基于双向 LSTM 神经网络电子病历命名实体的识别模型[J]. 中国组织工程研究, 2018, 22(20): 3237-3242.
- [36] 孙娟娟, 于红, 冯艳红, 等. 基于深度学习的渔业领域命名实体识别[J]. 大连海洋大学学报, 2018, 33(2): 265-269.
- [37] 杨飘, 董文永. 基于 BERT 嵌入的中文命名实体识别方法[J]. 计算机工程, 2020, 46(4): 40-45, 52.
- [38] COLLOBERT R, WESTON J, BOTTOU L, et al. Natural language processing (almost) from scratch[J]. The Journal of Machine Learning Research, 2011, 12(1/2): 2493-2537.
- [39] 李妮, 关焕梅, 杨飘, 等. 基于 BERT-IDCNN-CRF 的中文命名实体识别方法[J]. 山东大学学报(理学版), 2020, 55(1): 102-109.
- [40] 李博, 康晓东, 张华丽, 等. 采用 Transformer-CRF 的中文电子病历命名实体识别[J]. 计算机工程与应用, 2020, 56(5): 153-159.
- [41] LAMPLE G, BALLESTEROS M, SUBRAMANIAN S, et al. Neural architectures for named entity recognition [DB/OL]. (2016-03-04) [2021-04-09]. <https://arxiv.org/abs/1603.01360v3>.
- [42] 王若佳, 魏思仪, 王继民. BiLSTM-CRF 模型在中文电子病历命名实体识别中的应用研究[J]. 文献与数据学报, 2019, 1(2): 53-66.
- [43] LI X N, YAN H, QIU X P, et al. Flat: Chinese NER using flat-lattice transformer[DB/OL]. (2020-04-24) [2021-04-09]. <https://arxiv.org/abs/2004.11795v2>.
- [44] DAI X, ADEL H. An analysis of simple data augmentation for named entity recognition [DB/OL]. (2020-10-22) [2021-04-09]. <https://arxiv.org/abs/2010.11683>.
- [45] DING B S, LIU L L, BING L D, et al. DAGA: data augmentation with a generation approach for low-resource tagging tasks[DB/OL]. (2020-11-03) [2021-04-09]. <https://arxiv.org/abs/2011.01549>.
- [46] PENG M L, XING X Y, ZHANG Q, et al. Distantly supervised named entity recognition using positive-unlabeled learning[DB/OL]. (2019-06-04) [2021-04-09]. <https://arxiv.org/abs/1906.01378>.
- [47] SHANG J B, LIU L Y, REN X, et al. Learning named entity tagger using domain-specific dictionary[DB/OL]. (2018-09-10) [2021-04-09]. <https://arxiv.org/abs/1809.03599v1>.

(编辑: 颜峻)