

文章编号: 2095-2163(2023)08-0186-07

中图分类号: TP391

文献标志码: A

融合数据增强和注意力机制的中医实体及关系联合抽取

杨延云¹, 杜建强², 聂斌², 罗计根², 贺佳²

(1 江西工程职业学院, 南昌 330025; 2 江西中医药大学 计算机学院, 南昌 330004)

摘要: 中医药领域高质量标注数据缺乏, 可用作实体及关系联合抽取任务的语料数量偏少。BiLSTM-CRF 方法在序列标注上有优势, 但对上下文信息与当前信息的关联关注不够, 影响模型的性能。针对上述问题, 提出一种融合数据增强和注意力机制的中医实体及关系联合抽取方法(DA-BiLSTM-Attention-CRF)。该方法首先采用 EDA 数据增强方法对未标注训练集进行多倍增强, 然后将增强后的伪标注数据加入训练集, 采用自训练策略进行多次实验。最后, 测试集数据通过模型预测得到标注序列并进行三元组提取。在中医语料库上进行实验评估, 实体及关系联合抽取 F1 值为 82.43%。

关键词: 中医文本; 实体及关系联合抽取; 数据增强; 注意力机制

Entity and relationship joint extraction method for traditional medical text integrating data augmentation and attention mechanism

YANG Yanyun¹, DU Jianqiang², NIE Bin², LUO Jigen², HE Jia²

(1 Jiangxi Engineering Vocational College, Nanchang 330025, China;

2 School of Computer Science, Jiangxi University of Chinese Medicine, Nanchang 330004, China)

[Abstract] There is a lack of high-quality annotation data in the field of traditional medicine, and the number of corpora that can be used for joint entity and relationship extraction tasks is small. The BiLSTM-CRF method has advantages in sequence labeling, but it does not pay enough attention to the correlation between context information and current information, which affects the performance of the model. In response to the above problems, a joint extraction method of traditional medicine entities and relationships (DA-BiLSTM-Attention-CRF) that integrates data enhancement and attention mechanisms is proposed. In this method, the EDA data enhancement method is used to enhance the unlabeled training set in multiple times, then the enhanced pseudo-labeled data is added to the training set, and a self-training strategy is used for multiple experiments. Finally, the test set data is predicted by the model to get the annotation sequence and extract the triples. The experimental evaluation is carried out on the Chinese medicine corpus, and the F1 value of the joint extraction of entities and relationships is 82.43%.

[Key words] TM text; joint extraction of entities and relationships; data enhancement; attention mechanism

0 引言

命名实体识别^[1] (Named Entity Recognition, NER) 和关系抽取^[2] (Relation Extraction, RE) 作为信息抽取的基础任务^[3], 旨在从半结构化、非结构化的文本中提取实体, 并识别实体间的语义关系。中医药领域的实体及关系联合抽取就是提取中医文本中包含的中药名、方剂名、证型名、症状名等, 并识别 2 个实体间的关系类别。高质量、精准的实体及关系抽取结果为后续知识图谱的构建、信息检索效

率、智能问答系统的搭建等提供良好的数据基础^[4-5]。

实现命名实体识别和关系抽取两个任务传统是基于流水线方法, 分 2 步执行。第一步是实体识别, 第二步将上一步结果作为关系抽取的输入进行后续关系抽取操作。基于流水线方法易于建模且操作更加灵活, 但也有局限性, 会导致错误累积, 忽略 2 个子任务间的相关性, 并产生冗余信息。为解决流水线方法存在的问题, 实体及关系联合抽取方法应运而生, 充分考虑二者的相关性, 同时提取实体及关

基金项目: 国家重点研发计划(2019YFC1712301); 江西省自然科学基金资助项目(20202BAB202019); 江西省教育厅科技项目(GJJ215702); 江西省教育厅科技项目(GJJ190863); 国家自然科学基金(62141202)。

作者简介: 杨延云(1995-), 女, 硕士研究生, 主要研究方向: 自然语言处理、数据挖掘。

通讯作者: 杨延云 Email: 1441910806@qq.com

收稿日期: 2022-09-02

系,组成形如<实体 1,关系,实体 2>的三元组^[6]。本文的研究主要针对将联合抽取任务转化为端对端的序列标注任务,并频繁应用循环神经网络(Recurrent Neural Network, RNN)^[7]及其变体解决该问题。2017年,Zheng等学者^[8]首次将实体及关系联合抽取转化为序列标注任务,提出了一种基于新序列标注的联合抽取方法,还设计了带有偏置损失函数的端到端模型,实现了真正意义上的实体及关系联合抽取。曹明宇等学者^[9]借鉴 Zheng等学者^[8]的研究思路,通过改进标注策略在 DDI(drug-drug interactions) 2013 语料上有效缓解了同一实体参与多个关系的重叠问题。张军莲等学者^[10]利用图卷积神经网络的编码局部特征和先验词间关系的能力,提升抽取实体和关系三元组的性能。

循环神经网络及其变体在序列标注任务上显示出巨大的优势,能够全面捕捉上下文序列语义特征,但却没有突出与当前信息有关上下文信息的关联性和依赖性。Attention 机制旨在对输入序列依据其对目标任务结果的重要程度赋予相应不同的权重,强化重点,进而强调上下文信息与当前信息的关联和依赖,捕获更有效的上下文语义信息,Attention 机制既可以单独使用,也可以与其他神经网络模型融合训练。2014年,Google Deep-mind 团队^[11]使用 RNN 与 Attention 机制的混合模型做图像识别,验证了 Attention 机制在图像处理任务上的有效性。随后,Bahdanau 等学者^[12]实现了 Attention 机制在自然语言处理领域的首次应用,在英语到法语翻译任务上效果可与当时最先进的方法相媲美。文献^[13]针对传统流水线方法的不足以及实体重叠问题提出了一种新颖的标记策略,并融入注意力机制为包含 n 个单词的句子针对每个查询位置生成对应的 n 条标注序列。该模型在 NYT 数据集上较 Zheng 等学者^[8]提出的模型 F_1 值提高 4.3%。模型 LSTM-LSTM-2AT-Bias^[14]引入对抗训练增强模型的鲁棒性,并加入自注意力机制增强对文本信息的编码能力,在 NYT 数据集上实验的 F 值为 0.521 ± 0.006 。王勇超等学者^[15]利用指针标注机制实现实体及其关系的联合抽取,引入注意力机制有助于捕捉文本内部信息的相互关联。

针对中医文本的实体及关系联合抽取任务而言,语料中存在着大量无标签数据,基于序列标注的实体及关系联合抽取方法虽克服了传统流水线的不足,但由于上述方法均是以标注语料作数据基础,数据打标消耗大量的人力、物力、财力,以及 BiLSTM-

CRF 无法强调上下文信息与当前信息关联的问题。由此,本文提出了一种融合数据增强和注意力机制的中医实体及关系联合抽取方法(DA-BiLSTM-Attention-CRF),利用数据增强改善领域标注数据缺乏问题,在 BiLSTM-CRF 模型中加入注意力机制强化上下文信息与当前信息的关联性。

1 方法描述

融合数据增强和注意力机制的中医药领域的实体及关系联合抽取方法采用自训练方法多次实验,模型构建流程具体如下。

(1) 第一阶段,模型训练。训练步骤分述如下:

Step 1 将已标注的中医语料按照 7:3 划分为训练集和测试集,并转化为 200 维字词并联拼接向量作为序列标注模型 BiLSTM-Attention-CRF 的输入,学习文本深层次特征,将已训练完成的模型记为 BiLSTM-Attention-CRF_{origin}。

Step 2 采用数据增强方法 EDA 对无标签的训练集数据进行多倍增强。

Step 3 使用 BiLSTM-Attention-CRF_{origin} 预测增强后数据所对应标签,得到伪标注数据。

(2) 第二阶段,模型训练。训练步骤分述如下:

Step 4 研究将训练集和伪标注数据作为训练数据重新训练模型,已训练完成的模型记为 DA-BiLSTM-Attention-CRF。

Step 5 无标签测试集数据通过模型预测打上相应的标签,结合提取策略进行三元组提取。

Step 6 进行模型可靠性和有效性验证与评价。

需要注意的是,该方法中训练集指带有人工标注标签且用作模型训练的中医文本语料;测试集同理;无标签的训练集指只包含训练集中的文本数据,不包含对应的标签;无标签测试集同理。该方法涉及的主要深度学习模型 BiLSTM-Attention-CRF,其网络模型结构如图 1 所示。

该模型结构主要包含 3 个模块,分别为输入模块、神经网络模块和输出模块。其中,输入模块由文本输入层和向量嵌入层组成;神经网络模块包括双向 LSTM 层和注意力机制层;输出模块有 CRF 层对中医文本的序列标注操作和结合提取策略进行三元组抽取处理。下面对上述 3 个模块所包含层级结构的作用和功能进行详细介绍。

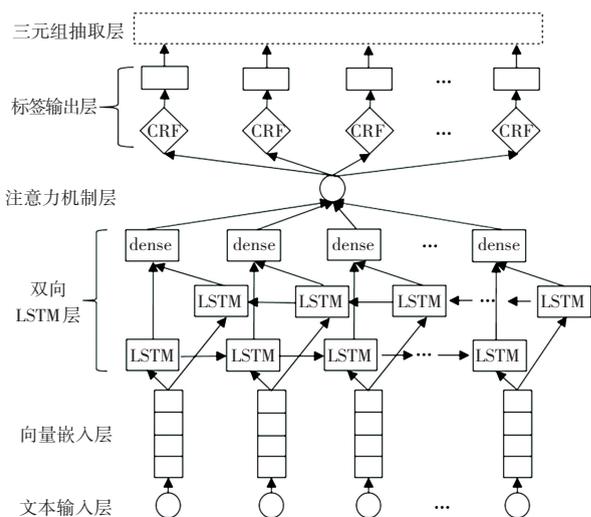


图1 BiLSTM-Attention-CRF 结构图

Fig. 1 BiLSTM-Attention-CRF structure diagram

(1) 输入模块

① 文本输入层: 文本输入为中医文本序列及每个字符所对应的标签。

② 向量嵌入层: 对输入的文本序列进行向量表示。向量所用语料来源于《中医证候鉴别诊断学》^[16]、《中医 150 证候辨证论治辑要(何晓晖)》^[17]和《中医学概论》^[18] 三本中医相关书籍。中医语料利用 jieba 分词^[19] 工具, 并加载自定义的中医领域自定义词典进行分词, 自定义词典主要包含大量的证型、方剂等信息, 通过 Word2Vec^[20] 训练得到 100 维词向量。中医语料使用 Word2Vec 训练得到 100 维字向量, 并以字为基本语义单元进行字词向量并联拼接, 丰富语义特征信息^[21]。

(2) 神经网络模块

① 双向 LSTM 层: 以字向量为基本语义单元与该字所在词的词向量进行并联拼接, 得到 200 维字向量作为 BiLSTM 的输入, BiLSTM 由前向的 LSTM 与后向的 LSTM 结合而成, 得到一个前向 t 时刻的隐藏层输出 \vec{h}_t 和一个后向 t 时刻的隐藏层输出 \overleftarrow{h}_t 拼接而成 $[\vec{h}_t, \overleftarrow{h}_t]$, 更加充分地利用了上下文信息, 进一步获取文本深层次特征表达。

② 注意力机制层: 注意力机制层的输入为上一层 BiLSTM 的特征向量输出, 通过计算概率权重, 为不同字向量分配不同注意力权重, 区分不同信息的重要程度, 达到强化与当前信息有关联的上下文信息的目的。

(3) 输出模块

① CRF 层: 其作用是输出每个字符的最大概率

标签, 使用 CRF 代替 *softmax* 函数, 是因为 CRF 层可以从训练数据中获得标签约束性规则, 进而保证预测标签的准确性。

② 三元组抽取层: CRF 层的序列标注输出作为该层的输入, 根据预先定义的提取策略进行实体关系三元组抽取, 其结果呈现类似于 <杏苏散, 方剂/中药, 杏仁>。

1.1 标注模式及提取策略

本文实验语料来源于《中医证候鉴别诊断学》^[16]、《中医 150 证候辨证论治辑要(何晓晖)》^[17]和《中医学概论》^[18] 三本中医相关文献, 累计 2 968 个句子。标注模式共涉及 34 类标签, 表 1 展示了部分标签的内容及含义。“BIES”表示字在词的位置; 方剂/中药、证型/方剂、证型/症状、病因/证型和 M 表示关系类别, M 表示该实体与多个实体组成关系不同的三元组; 1、2 和 M 表示关系角色, 1 和 2 分别表示该实体在所属三元组中处于实体 1 或实体 2 的位置, M 表示该实体与多个实体组成三元组且处于不同位置。

通过标注模型标注文本得到标签后, 将有相同关系类别(包括“M”类型)的 2 个实体标签组合, 再根据关系角色标签确定其在三元组中的位置。

表 1 部分标签及含义

Tab. 1 Part of the labels and meanings

标签	含义
O	无关词汇
B-方剂/中药-1	方剂/中药类别三元组实体 1 的首部
I-方剂/中药-1	方剂/中药类别三元组实体 1 的中部
E-方剂/中药-1	方剂/中药类别三元组实体 1 的尾部
B-M-2	与多个实体构成不同关系类别三元组的实体 2 的首部
I-M-2	与多个实体构成不同关系类别三元组的实体 2 的中部
E-M-2	与多个实体构成不同关系类别三元组的实体 2 的尾部
B-M-M	与多个实体构成不同关系类别三元组的实体 1 或实体 2 的首部
I-M-M	与多个实体构成不同关系类别三元组的实体 1 或实体 2 的中部
E-M-M	与多个实体构成不同关系类别三元组的实体 1 或实体 2 的尾部

需要强调的是, 标签 M 既可以扮演实体 1, 也可以扮演实体 2。标注实例和三元组抽取结果展示如图 2 所示。

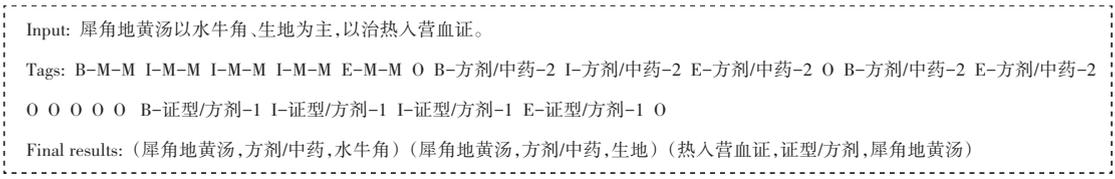


图 2 标注实例

Fig. 2 An example of labeling

1.2 自训练策略

当使用神经网络模型时,用于训练模型的数据是越多越好,而在有监督的任务中,所需训练数据必须根据目标任务进行打标,而想要构建这样的标注数据集依赖于大量的人力、物力和财力。对于这个问题,半监督学习可以利用少量已标注数据进行模型训练,再使用已经训练好的模型对无标签数据进行预测,得到准确率较高的伪标签数据,不断迭代上述步骤将所有数据打标。其中,自训练就是一个标准的半监督过程^[22]。自训练的算法流程主要步骤见如下:

Step 1 将已标注的数据按比例划分训练集和测试集传送给模型,对模型进行训练。

Step 2 使用训练好的模型预测无标签的数据,所得结果记作伪标签数据。

Step 3 将伪标记数据与原训练数据组合,组合的“伪标记”和正确标记训练数据共同来重新训练模型。

Step 4 使用经过训练的模型来预测已标注的测试集,根据所采用的评价指标评估模型性能。

1.3 数据增强

数据增强主要用于数据集较小的时候,用来防止模型过拟合,提升模型泛化能力。对于深度学习模型来说,大的数据集是实验效果好的基本要求,而对于特定的领域任务,已标注的数据却是非常缺乏的。数据增强是一种有效的扩充数据集的方法,在计算机视觉领域取得较好的效果,主要涉及的技术有随机旋转、随机裁剪、色彩抖动、高斯噪声、翻转等,通过增加图像的多样性扩充数据量。然而在自然语言处理领域数据增强的方法相对较少,本文采用文献^[23]提出的数据增强方法(下文简称 EDA)对训练集进行数据增强处理。EDA 采用 4 种技术进行数据增强,分别是同义词替换、随机插入、随机交换、随机删除。

接下来将详细介绍这 4 种技术,并以原本“主治肝郁气滞证的方剂是黄芩散。”进行举例说明:

(1)同义词替换:从句子中随机选择非停用词,并用其同义词进行替换,则可以得到“主治肝郁气

滞证的方药是黄芩散”。

(2)随机插入:找句子中一个随机词(非停用词)的同义词,插入到句子的随机位置,则可以得到“主治肝郁气滞证的方剂是黄芩散疏肝”。

(3)随机交换:随机选择句子中的 2 个词,交换二者的位置,则可以得到“主治肝郁气滞证的方剂黄芩是散”。

(4)随机删除:随机删除句子中的词语,则可以得到“主治肝郁气滞的方剂是黄芩散”。

1.4 Attention 机制

Attention 机制^[12]是一种模拟人脑注意力资源分配的机制,通过有限的注意力选择性地分配给更重要的信息,而弱化其他无用信息。例如,给出中医文本:大承气汤由大黄、枳实各 12 g,炙厚朴 24 g,芒硝 6 g。根据文本回答问题“大承气汤由哪几味中药药?”,那么答题人便只会专注去数大承气汤的中药组成个数;而当问题改为“大承气汤中枳实的用量是多少?”,那么答题人便只聚焦枳实的克数。注意力机制就是随着任务的变化,变化注意力区域,找到对于任务要求最有价值的信息。注意力机制的基本结构如图 3 所示。

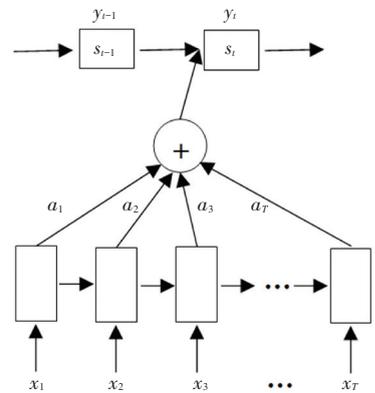


图 3 注意力机制结构图

Fig. 3 Structure diagram of attention mechanism

具体计算公式如下:

$$u_t = \tanh(h_t w_t + b_t) \tag{1}$$

$$a_t = \frac{\exp(u_t u_w)}{\sum_{t=1}^T \exp(u_t u_w)} \tag{2}$$

$$s_t = \sum_{i=1}^T a_i h_i \quad (3)$$

其中, h_t 表示 t 时刻上一层神经网络的输出向量; w_t 是第 t 时刻权重系数; b_t 是第 t 时刻偏置系数; u_t 表示第 t 时刻输入到全连接层获得其隐藏表示; u_w 表示随机初始化的注意力权重矩阵; a_t 表示在 t 时刻通过归一化操作计算权重向量; s_t 表示在 t 时刻最终编码后的句子向量。

2 实验分析

2.1 实验设置

本文实验环境是 Windows 64 位操作系统, 处理器是 Inter(R) Celeron(R) CPU G530 @ 2.40 GHz, 内存 8 GB。采用 JetBrains PyCharm, Python 编程语言, 搭建 Tensorflow 框架, 并运用 Jieba 分词工具。评价指标选取精确率 (Precision, P)、召回率 (Recall, R)、 F_1 值^[24]。模型涉及多个超参数, 根据所得精确率、召回率和 F_1 值进行参数调整, 实验参数设置见表 2。

表 2 实验参数设置

Tab. 2 Experimental parameter settings

超参数	值
<i>learning rate</i>	0.01
<i>dropout</i>	0.4
<i>optimizer</i>	RMSProp
<i>batch - size</i>	64
<i>hidden - dim</i>	300
<i>epoch</i>	40

2.2 模型对比实验

本节在使用相同字词向量并联拼接作为模型输入特征向量, 相同标注模式和提取策略的情况下, 对 BiLSTM_CRF、BiLSTM-Attention-CRF_origin、DA-BiLSTM-Attention-CRF 进行对比实验, 结果见表 3。

表 3 模型对比实验结果

Tab. 3 Experimental results comparison of models %

模型	实体及关系联合抽取		
	P	R	F_1
BiLSTM_CRF	85.32	76.90	80.89
BiLSTM-Attention-CRF_origin	87.45	77.03	81.91
DA-BiLSTM-Attention-CRF	87.75	77.71	82.43

注: 黑色加粗字体表明该方法效果优于其他方法

由表 3 可以得出结论如下:

(1) BiLSTM-Attention-CRF_origin 与 BiLSTM_CRF 模型相比增加注意力机制结构, 能够对 BiLSTM

输出的特征向量针对目标任务赋予不同的注意力权重, 最大限度地强调上下文信息对目标词的重要性。研究中, 2 个模型在相同数据集上做实验, BiLSTM-Attention-CRF_origin 均有更好的效果。

(2) 使用数据增强对训练集进行扩充, 结合自训练策略对 BiLSTM-Attention-CRF 模型进行多次训练学习, 不仅没有降低模型性能, 反而在领域标注语料不足的情况使模型性能得到一定的提升, F_1 分别较原模型提升 0.52%。

2.3 DA 部分参数对比实验

为了确保新方法中数据增强部分参数 α 和 $navg$ 的选取对于中医语料的可靠性, 选取文献[23]推荐数据量为 2 000 时采用的参数组合 0.05、8 和本实验选取的参数组合 0.05、4 分别进行对比, 以实体及关系联合抽取实验的精确率、召回率和 F_1 值作为评价指标, 其结果见表 4。

表 4 DA 部分参数对比实验结果

Tab. 4 Experimental results comparison of some parameters of DA %

模型	参数组合	实体及关系联合抽取		
		P	R	F_1
BiLSTM-Attention-CRF_origin	/ /	87.45	77.03	81.91
DA-BiLSTM-Attention-CRF	0.05 8	85.05	77.24	80.96
DA-BiLSTM-Attention-CRF	0.05 4	87.75	77.71	82.43

注: 黑色加粗字体表明该方法效果优于其他方法

由表 4 可以分析得出结论如下:

在 DA-BiLSTM-Attention-CRF 模型上实验, 利用 EDA 方法进行数据扩充, 当选取参数组合 0.05、8 时, 实体识别和联合抽取综合评价指标 F_1 均有所下降, 这是因为由模型 BiLSTM-Attention-CRF_origin 预测所得伪标注数据存在一定概率的错误标签数; 而参数组合 0.05、4 的 P 、 R 、 F_1 值都高于未经数据增强处理的情况, 与前者对比, 这说明相对多倍数据增强后的伪标注语料存在的噪音数据对于原标注语料稀释作用较大, 实验效果变差, 而适当倍数的数据增强有助于提升模型性能。

2.4 模型成分消减实验

本文实验在数据处理部分采用了数据增强操作, 模型输入采用字词向量并联拼接, 核心算法结构使用双向 LSTM、注意力机制和条件随机场。为探究这些结构对于目标任务的影响, 本节在相同标注模式和提取策略的前提下, 分别对模型输入、数据增强、Attention、CRF 结构进行成分消减实验。模型成分消减实验结果见表 5。

表 5 模型成分消减实验结果

Tab. 5 Model component reduction experiment results %

模型	实体及关系联合抽取		
	P	R	F1
DA-BiLSTM-Attention-CRF	87.75	77.71	82.43
-word embedding	88.30	72.16	79.42
-DA	87.45	77.03	81.91
-Attention	85.83	77.33	81.36
-CRF	88.20	74.62	81.60

注: 黑色加粗字体表明该方法效果优于其他方法。

由表 5 可以得到如下结论:

(1) 使用字词向量并联拼接作为模型输入, 与采用单独字向量相比, 在召回率和 F_1 值有显著的提升, 证明了字词向量并联拼接将字和词的信息有效地结合起来, 更有利于提取有效特征。

(2) 删减数据增强部分, 即只有本文方法的第一阶段, 实验的 3 个评价指标均有所降低, 证明了采用 EDA 做数据扩充和自训练策略的有效性。

(3) 注意力机制的作用是区分不同信息的重要程度, 强化与当前信息有关联的上下文信息, 而弱化其他无用信息。增加 Attention 层, 显著提升了中医文本实体及关系联合抽取的精确率 (提升接近 2%), F_1 值增加 1.07%。

(4) CRF 对于所有预测标签序列进行全局归一化操作, 去掉 CRF 层模型较原模型降低了联合抽取的 F_1 值。

3 结束语

本文提出了一种融合数据增强和注意力机制的中医实体及关系联合抽取方法, 选用 EDA 对中医语料进行数据增强, 使用自训练策略将原数据集和预测所得的伪标注数据共同学习, 在中医语料上取得较优的效果。将本文提出的方法与目前主流的序列标注模型进行对比, 证明了本文方法的优越性; 然后进行数据增强部分的参数组合对比实验, 结果表明本实验选取参数取得更佳的实验效果; 最后通过模型成分消减实验, 验证了模型各个部分的必要性。

未来工作将寻求不同数据增强方法用于解决领域标注数据缺少的问题, 对于本文采用的自训练方式进一步优化, 提升模型性能, 也可尝试探索预训练和自训练的融合策略对于实体及关系联合抽取任务的影响也是此后的研究工作。

参考文献

[1] GRISHMAN R, SUNDHEIM B M. Message understanding conference-

6: A brief history [C]//COLING 1996 Volume 1: The 16th International Conference on Computational Linguistics. 1996, 96: 466-471.

- [2] KUMAR S. A survey of deep learning methods for relation extraction[J]. arXiv preprint arXiv:1705.03645, 2017.
- [3] GOLSHAN P N, DASHTI H A R, AZIZI S, et al. A study of recent contributions on information extraction[J]. arXiv preprint arXiv:1803.05667, 2018.
- [4] CAO Yixuan, CHEN Dian, XU Zhengqi, et al. Nested relation extraction with iterative neural network[J]. Frontiers of Computer Science, 2021, 15(3).
- [5] 王传栋, 徐娇, 张永. 实体关系抽取综述[J]. 计算机工程与应用, 2020, 56(12): 25-36.
- [6] 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(06): 1793-1818.
- [7] MIKOLOV T, KARAFIÁT M, BURGET L, et al. Recurrent neural network based language model [C]// ESEC/FSE 2018: Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering. Makuhari, Chiba, Japan: ACM, 2015: 896-899.
- [8] ZHENG Suncong, WANG Feng, BAO Hongyun, et al. Joint extraction of entities and relations based on a novel tagging scheme [J]. arXiv preprint arXiv:1706.05075, 2017.
- [9] 曹明宇, 杨志豪, 罗凌, 等. 基于神经网络的药物实体与关系联合抽取[J]. 计算机研究与发展, 2019, 56(07): 1432-1440.
- [10] 张军莲, 张一帆, 汪鸣泉, 等. 基于图卷积神经网络的中文实体关系联合抽取[J/OL]. 计算机工程: 1-10 [2021-03-15]. <https://doi.org/10.19678/j.issn.1000-3428.0059574>.
- [11] MNIH V, HEES N, GRAVES A. Recurrent models of visual attention[J]. Advances in Neural Information Processing Systems, 2014, 27: 2204-2212.
- [12] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate [J]. arXiv preprint arXiv:1409.0473, 2014.
- [13] DAI Dai, XIAO Xinyan, LYU Y, et al. Joint extraction of entities and overlapping relations using position-attentive sequence labeling [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Honolulu; AAAI, 2019, 33: 6300-6308.
- [14] 黄培馨, 赵翔, 方阳, 等. 融合对抗训练的端到端知识三元组联合抽取[J]. 计算机研究与发展, 2019, 56(12): 2536-2548.
- [15] 王勇超, 穆华岭, 周灵智, 等. 基于指针网络的实体与关系联合抽取方法 [J/OL]. 计算机应用研究: 1-5 [2021-03-15]. <https://doi.org/10.19734/j.issn.1001-3695.2020.04.0113>.
- [16] 姚乃礼. 中医证候鉴别诊断学[M]. 北京: 人民卫生出版社, 2002.
- [17] 何晓晖, 陈建章. 中医 150 证候辩证论治辑要[M]. 北京: 北京师范大学出版社, 2003.
- [18] 郝丽莉, 傅南琳. 中医学概论[M]. 2 版. 北京: 科学出版社, 2016.
- [19] 董洋溢, 李伟华, 陈世亮. 中文领域知识半自动化 OWL 本体构建方法研究[J]. 计算机应用与软件, 2016, 33(05): 23-25, 71.
- [20] MIKOLOV T, SUTSKEVER I, CHEN Kai, et al. Distributed representations of words and phrases and their compositionality[J]. Advances in Neural Information Processing Systems, 2013, 26: 3111-3119.

(下转第 196 页)