

# 基于多层注意力机制的农业病虫害远程监督关系抽取研究

乐毅<sup>1,2</sup>, 王文宇<sup>1</sup>, 张凯<sup>1</sup>, 梁振京<sup>2,3</sup>, 刘飞<sup>1,2</sup>, 陈祎琼<sup>1,2</sup>, 吴云志<sup>1,2</sup>, 张友华<sup>1,2</sup>

(1. 安徽农业大学信息与计算机学院, 合肥 230036; 2. 安徽农业大学, 安徽省北斗精准农业信息工程实验室, 合肥 230036; 3. 安徽农业大学植物保护学院, 合肥 230036)

**摘要:** 针对大多数现有关系抽取模型存在对语义特征提取不充分、速度慢且数据集匮乏的缺点, 提出一种 PCNN(piecewise convolutional neural network)模型和多层注意力机制相结合的远程监督关系抽取方法进行农业病虫害领域的关系抽取。模型由两个实体把句子分成三段, 对卷积后的每一段进行最大池化获得特征, 同时在实例和池化特征层面上分别引入注意力机制有效降低信息噪声。在 F1 评价指标上比传统方法提高了 5.75%, 在耗时上是传统方法的 10.93%, 且减少了手工标注数据集的成本。

**关键词:** 关系抽取; 农业病虫害; 注意力机制; 卷积神经网络; 远程监督

中图分类号: S126

文献标识码: A

文章编号: 1672-352X(2020)04-0682-05

## Agricultural pest and disease relation extraction based on multi-attention mechanism and distant supervision

YUE Yi<sup>1,2</sup>, WANG Wenyu<sup>1</sup>, ZHANG Kai<sup>1</sup>, LIANG Zhenjing<sup>2,3</sup>,  
LIU Fei<sup>1,2</sup>, CHEN Yiqiong<sup>1,2</sup>, WU Yunzhi<sup>1,2</sup>, ZHANG Youhua<sup>1,2</sup>

(1. School of Information and Computer Science, Anhui Agricultural University, Hefei 230036;

2. Anhui Agricultural University, Engineering Laboratory of Beidou Precision Agriculture, Hefei 230036;

3. School of Plant Protection, Anhui Agricultural University, Hefei 230036)

**Abstract:** Aiming at the shortcomings of most existing relation extraction models: insufficient semantic feature extraction, slow speed, and lack of datasets, we propose a method based on distant supervision to do relation extraction in the agricultural field. The model combines PCNN (Piecewise Convolutional Neural Network) model and multilayer attention mechanisms. The model divides the sentence into three segments by two entities, and takes the max-pooling result as the feature for each segment after convolution. Then attention mechanism is introduced at the instance feature and the pooling feature to effectively reduce the noise of information. In the F1 score, our model is 5.75% higher than the traditional model. In time consuming, our model just takes 10.93% of the traditional model. The cost of manually labeling the dataset is reduced at the same time.

**Key words:** relation extraction; agricultural pest and disease; attention mechanism; convolutional neural network; distant supervision

当今大数据时代, 互联网信息每天都在爆炸式增长, 并且其中绝大部分都是非结构化的、杂乱无章的、机器不可直接使用的非结构化信息。关系抽取通过识别句子中两个实体之间存在的某种或多种关系, 将非结构化数据对齐为结构化数据, 为构建知识图谱、自动问答系统等自然语言处理任务提供

数据基础<sup>[1-2]</sup>。而在农业领域, 病虫害之间存在着靠传统方法<sup>[3-4]</sup>进行关系抽取效果较差的问题, 本研究提出了一种 PCNN (piecewise convolutional neural network) 分段池化卷积神经网络模型和多层注意力机制相结合的远程监督<sup>[5]</sup>关系抽取方法进行农业病虫害领域的关系抽取, 首先由两个实体把句子分成

收稿日期: 2019-10-28

基金项目: 安徽农业大学大创项目[XJDC2019206], 安徽省大学生创新创业教育训练计划项目[201910364206], 安徽省级教学团队计算机科学与技术教学团队项目[2018JXTD114]共同资助。

共同第一作者简介: 乐毅, 副教授。E-mail: yyyue@ahau.edu.cn 王文宇, 本科生。E-mail: black938@ahau.edu.cn

\* 通信作者: 张友华, 教授, 博士生导师。E-mail: zhangyh@ahau.edu.cn

三段, 再进行卷积后取分段最大池化操作, 同时在实例和池化特征层面上引入多层注意力机制有效降低信息噪声, 并且模型中远程监督的方法有效的减少了手工标注数据集的工作量。在 F1 值和算法耗时方面均相对于传统方法有所提高。

## 1 研究方法

传统的关系抽取主要有基于传统机器学习<sup>[3]</sup>和基于模板<sup>[4]</sup>的方法, 需要大量领域专家来构建模板, 人工及时间成本过高。传统机器学习又难以充分提取文本潜在的语义特征。而深度学习方法能够在反向传播的过程中进行学习, 自动提取文本的语义特征, 获得了较好的效果, 成为了近几年的研究热点。

Liu 等<sup>[6]</sup>第一次将卷积神经网络模型运用到自然语言处理任务中, 进行关系分类, 模型结构简单, 没有池化层。Zeng 等<sup>[7]</sup>设计了词法级别和句级别的特征提取, 使用了多个卷积核, 效果略优于传统方法, 但是使用不变的窗口大小参数, 提取到的特征比较单一。Zhang 等<sup>[8]</sup>尝试使将双向循环神经网络模型运用到关系抽取任务中, 由于循环神经网络的特性, 模型能捕获更多信息, 为解决循环神经网络的长距离依赖问题, 长短时记忆单元 LSTM (long short term memory network) 被提出, 通过在隐藏层增加三个门控记忆单元来存储历史信息, 对于长句子的处理拥有很好的效果。Zhou 等<sup>[9]</sup>开始将当时在图像领域流行的注意力机制应用到自然语言处理中, 但单层的注意力机制没有取得可观的降噪效果。且由于 RNN 模型是基于时间序列的, 因此无法并行计算, 其训练耗时长、开销大的特点为人诟病。

由于上述大多数受监督的关系抽取方法需要大量手工标记的数据, 费时费力, Zeng 等<sup>[10]</sup>提出了 PCNN 模型以及用于关系抽取任务的远程监督的思想, 通过对齐 FreeBase、百度百科、互动百科等知识库自动生成数据集。原理是利用已有的知识库中的两个实体  $e_1$ 、 $e_2$  以及对应的某个关系  $r$ , 在自制大规模语料库中进行回标, 如果某个句子  $s$  中同时包含了  $e_1$ 、 $e_2$ , 则假定句子  $s$  包含了上述的关系  $r$ 。由此即可获得一对三元组  $\langle e_1, r, e_2 \rangle$  和一个语料句  $s$ 。优点是通过此种假定可以获得大量标注数据用以训练, 免去了大量的手工标注成本, 而且可以针对领域知识, 对特定词进行的数据回标。这种方法的缺点也是很明显的, 假设性过于强, 引入了许多噪声数据。因为包含了实体  $e_1$ 、 $e_2$  的两个句子不一定包含关系  $r$ 。因此采用远程监督获取数据集的模型必须进行有效的降噪工作, 突出关键信息对模型分类输出结果的影响, 而注意力机制是一种模拟人脑取信息的模型,

通过计算概率分布提高重要特征的权重, 从而减少噪声, 近年来国内外学者对注意力机制做了大量研究, Yang 等<sup>[11]</sup>第一次将注意力机制应用在自然语言处理中, 进行机器翻译, 在关系抽取领域, Lin 等<sup>[12]</sup>通过实例级的注意力机制, 有效降低了远程监督模型中错误标签带来的噪声影响, 提高了关系抽取的效果。近年来, 在特定领域进行基于注意力机制的关系抽取的成果层出不穷<sup>[13]</sup>, 吴粤敏等<sup>[14]</sup>在农业金融文本领域结合注意力机制和 GRU 并用于农业企业投资领域知识可视化, 朱珊珊等<sup>[15]</sup>在军事领域利用基于注意力机制的 LSTM 网络抽取关系, 构建军队军事结构资料库, Gu 等<sup>[16]</sup>在生物医学领域, 利用注意力机制从生物医学文献中提取化学疾病的关系, 加快化学药品的开发。

基于上述研究, 本研究提出了一种在 PCNN 网络上叠加多层注意力机制的关系抽取模型, 利用远程监督的思想构建农业病虫害领域的关系抽取数据集, 同时加入实例级的注意力机制和池化特征级的注意力机制, 对噪音数据进行降噪。实验结果表明, 该模型应用于农业病虫害的关系抽取中, 避免了大量繁琐的手工标注工作, 且相较于之前的传统方法有着更好的抽取效果, 验证了本研究模型的有效性。

## 2 模型设计

整个模型如图 1 所示, 分段池化和多实例注意力参考文献<sup>[10]</sup>的设计。包含输入映射层、分段卷积池化层、池化特征注意力层、多实例注意力层和分类层。

### 2.1 输入映射层

**2.1.1 词向量化** 给定一个由  $n$  个单词组成的句子  $S = \{w_1, w_2, w_3, \dots, w_n\}$  通过公式将每一个单词  $w_i$  映射到一个向量,

$$w_i = w^{word} + V^w$$

其中,  $w^{word} \in R^{n \times |m|}$  是采用 Glove<sup>[17]</sup>训练得来的词向量矩阵,  $d^w$  表示词向量的维数,  $m$  表示句子中所有词语的个数。  $V^w$  是词语  $w$  的 one-hot 向量表示。由此可以将句子转化为一个维度为  $n \times d^w$  的矩阵。

**2.1.2 位置向量化** 为了更加准确获取句子含义, 将句子中每一个单词到两个实体之间的距离作为位置向量特征, 拼接到词向量特征之后, 从而得到整个句子的向量表达。以“稻飞虱是危害水稻的害虫”为例, 句中“稻飞虱”和“水稻”作为两个实体, 则“危害”到“稻飞虱”的距离为-2, “危害”到“水稻1”的距离为1。则“危害”的位置向量为[-2,1]。

设位置向量化的向量维度为  $d^p$ , 词向量维度为  $d^w$ , 则句子中每个单词的维度为  $d^s = d^w + 2d^p$ , 最终将句  $S$  转换为一个  $n \times d^s$  维的矩阵。

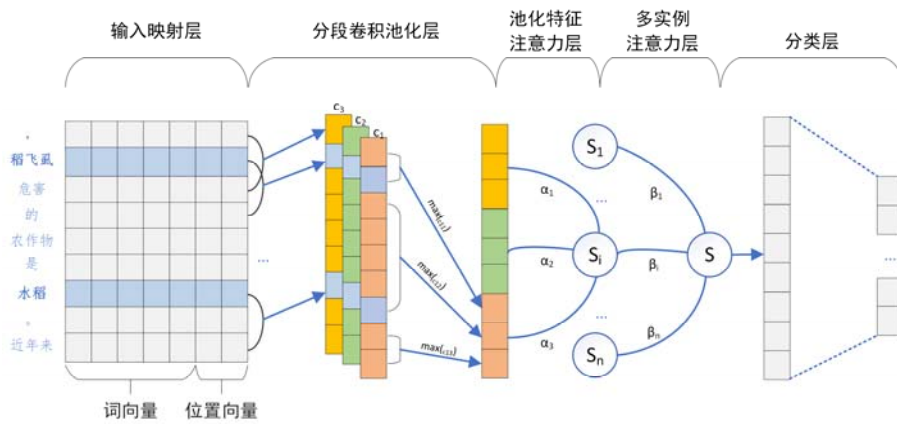


图 1 PCNN 多注意力机制关系抽取模型

Figure 1 PCNN and multi-attention model for relation extraction

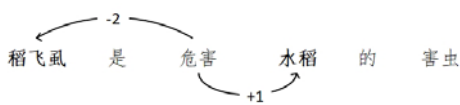


图 2 位置向量化

Figure 2 Position embedding

2.2 分段卷积池化层

2.2.1 卷积层 卷积的本质是一种线性运算，卷积层的目的是提取句子中邻近几个单词间的局部特征，最后拼接局部特征得到全局特征。并且卷积层通过提取高层特征，限制传统神经网络输入层到隐藏层之间的连接数来减少模型参数量，提高模型收敛效率。经过输入映射层后，句  $s$  的维度为  $n \times d^s$ ，由于本模型采用的是词向量，每一个句子的最小组成元素为词级别，所以卷积核的宽度与每一个词的维度统一，也就是  $d^s$ ，而卷积核的高度根据感受野的不同通常取 2~5 个。假设卷积核的大小为  $k \times d^s$ ，包含有  $n$  个单词的句子  $S = \{w_1, w_2, w_3, \dots, w_n\}$  其中  $w \in d^s$ 。则一个卷积核生成的局部特征计算公式为：

$$c_i = F(W \cdot w_{i:i+k-1} + b)$$

其中  $w \in n \times d^s$  是卷积核参数矩阵， $b$  是偏置项。 $F$  为激活函数，本研究使用的是 ReLu 函数。计算公式为：

$$y = \max(0, x)$$

针对一个特定的卷积核  $W$ ，完成一次卷积后捕捉到的句级向量表达为  $c = [c_i, c_{i+1}, \dots, c_{n-k+1}]$ 。

2.2.2 分段池化层 典型的卷积神经网络在卷积层后接入最大池化来捕捉更高层的语义特征，但也会丢失诸多如实体与实体间关系这种信息，为了更好的捕捉实体和实体之间的联系，Zeng<sup>[10]</sup>等人提出了分段池化的思想，将卷积核  $\omega_i$  捕捉到的句级向量  $c_i$  根据两个实体的位置切分成 3 个片段  $\{c_{i1}, c_{i2}, c_{i3}\}$  则分段最大池化的结果为：

$$p_{ij} = \max(c_{ij}), 1 \leq i \leq n, 1 \leq j \leq 3$$

其中  $n$  为卷积核的个数，对于每一个卷积核进行分段最大池化后的输出  $p_i = [p_{i1}, p_{i2}, p_{i3}]$ ，进行拼接得到分段卷积池化层的最后输出  $P = [p_i, p_i, p_i]$ ， $P \in 3 \times n$ 。

2.3 池化特征注意力层 本层在池化特征上进一步进行高层语义的特征提取。将上一层最终的输出  $P$  先通过 tanh 函数压缩至  $(-1, 1)$  区间，向量  $V$  表示各池化特征经注意力机制后的加权值：

$$M = \tanh(P), \alpha = \text{softmax}(W^T M), V = P \alpha^T$$

其中  $P \in 3 \times n$ ， $n$  为卷积核个数， $W$  为待训练的参数矩阵，跟随反向传播算法更新。

2.4 多实例注意力层 远程监督通过已有的实体对和关系数据在互联网文本中进行自动回标，由于缺乏人工监督，一些数据不可避免的并不能表达实体对所对应的关系，对模型性能有较大影响。为了在一定程度上降低噪声数据的影响，参考文献[12]引入多实例注意力层来计算包含拥有同一实体对的所有句子和关系标签的相关程度，赋予与关系标签相关性最高的句子更高的权重，从而降低噪声句对关系预测的影响。对于蕴含同一实体对  $(e_1, e_2)$  的实例集合  $(V_1, V_2, \dots, V_n)$ ，设  $\beta$  为注意力权重矩阵，一种计算方式为  $\beta_i = 1/n$ ，认为集合中所有句子对关系的贡献都是等同的，把集合向量的表达平均到每一个句子上。但是平均后无法使得注意力参数进行更新，本研究通过句子集合中句子向量和对应关系的关系向量内积再经过一层 softmax 函数得到  $\beta$ ， $\beta$  体现句子  $V_i$  和关系  $r$  之间的相关程度，实现注意力机制。计算公式如下：

$$\beta_i = \frac{\exp(V_i A r)}{\sum_{k=1}^n \exp(V_k A r)}$$

其中  $A$  为加权对角矩阵，随机初始化后在反向传播算法中不断迭代更新。给出该集合的向量标示

$S$  为:

$$S = \sum_{i=1}^n \frac{\exp(V_i Ar)}{\sum_{k=1}^n \exp(V_k Ar)} V_i$$

接下来就可以定义一个线性函数计算每一个关系对应的得分:

$$score = MS + b$$

其中  $M$  为关系矩阵,  $b$  为偏置项。

## 2.5 分类层

得到前面所有层的所有输出  $score$  后, 通过 softmax 函数计算预测关系的条件概率和最有可能的关系,  $n_r$  为所有事先预订的关系数量个数,  $\hat{r}$  作为模型最后的输出, 计算公式如下:

$$\hat{p}(r | S) = \frac{\exp(score_r)}{\sum_{i=1}^{n_r} \exp(score_i)}$$

$$\hat{r} = \operatorname{argmax}(\hat{p}(r | S))$$

针对这种多分类的问题, 本研究将最小化交叉熵损失函数作为优化目标, 损失函数定义如下:

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m \log \hat{p}(r_i | S_i, \theta) + \frac{\lambda}{2} \|\theta\|_2^2$$

其中  $\theta$  为模型中所有可训练参数, 采用了 L2 正则化, 在每次更新时乘以系数  $\lambda$ , 进行权重衰减, 避免过拟合现象, 采用 Adam 算法最小化损失函数进行参数更新, 并且本研究还在模型中加入 dropout 层<sup>[18]</sup>, 使隐藏层节点按照一定比例随机舍弃, 进一步增强模型鲁棒性。

## 3 结果与分析

### 3.1 实验数据与评价标准

实验采用的数据是通过上文提及的远程监督思想, 调用思知知识图谱、复旦知识工场提供的接口获得实体对信息, 并对齐到由爬虫获取的百度百科、中国农业信息网等互联网语料库后形成的。数据集上包含 5 种关系, 包含危害作物、危害部位、分布区域、颜色属性和病虫害别称。数据集按照 8:2 的比例来切分, 其中训练集包括实例数为 5 668, 测试集包含实例数为 1 417。以危害作物关系为例, 数据集的一条实例为 (“<稻飞虱,危害作物,水稻>”, “稻飞虱繁殖快、繁殖率高, 是危害水稻的一种主要害虫”)。本研究通过计算准确率 (precision)、召回率 (recall) 和 F1 值作为性能评价指标。公式如下:

$$precision = \frac{out_{right}}{out}$$

$$recall = \frac{out_{right}}{all}$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall}$$

其中  $out_{right}$  为正确预测关系个数,  $out$  表示所有预测关系个数,  $all$  表示测试集中所有的关系个数。

### 3.2 实验环境与参数设置

实验环境: 操作系统 Ubuntu16.04; 处理器 Intel Xeon E5-2620 v4@2.10GHz; 内存 16GB; Python 版本: 3.6.4。本研究参照文献<sup>[14]</sup>的经验对实验参数进行设置, 具体参数设置由表 1 所示。

表 1 参数设置

Table 1 Parameter setting

| 参数 Parameter                  | 值 Value |
|-------------------------------|---------|
| 词向量维度 WE Dimension            | 50      |
| 位置向量维度 PE Dimension           | 5       |
| 句向量维度 Sentence Dimension      | 230     |
| 学习率 Learning Rate             | 0.001   |
| 批大小 Batch Size                | 64      |
| Dropout 率 Dropout Probability | 0.5     |
| L2 正则项参数 L2-norm Parameter    | 0.001   |

表 2 关系提取效果 F1 值

Table 2 F1 score for relation extraction %

| 关系类型 Relation type            | F1 值 F1 Value |
|-------------------------------|---------------|
| 危害作物 Harmfal to corps         | 60.04         |
| 危害部位 Harmfal to Part of corps | 61.44         |
| 分布区域 Distribution area        | 57.70         |
| 颜色属性 Color attributes         | 59.21         |
| 别称 Another name               | 54.40         |

表 3 不同提取方法结果

Table 3 Comparison of extraction results by different methods %

| 方法 Method | 准确率 Accuracy | 召回率 Recall | F1 值 F1 Value |
|-----------|--------------|------------|---------------|
| BiLSTM    | 53.29        | 53.77      | 53.53         |
| PCNN      | 55.87        | 60.73      | 58.20         |
| PCNN+1ATT | 56.82        | 59.26      | 58.01         |
| PCNN+2ATT | 57.23        | 61.48      | 59.28         |

表 4 不同提取方法耗时

Table 4 Comparison of time consuming by different methods

| 方法 Method | 训练耗时/s Train time consuming | 预测耗时/s Predict time consuming |
|-----------|-----------------------------|-------------------------------|
| BiLSTM    | 115.46                      | 1.89                          |
| PCNN      | 32.74                       | 0.19                          |
| PCNN+1ATT | 33.04                       | 0.21                          |
| PCNN+2ATT | 33.71                       | 0.22                          |

### 3.3 实验结果及分析

在上述实验设置下, 对五类关系进行提取。具体结果如表 2 所示。

由实验结果可以看出, 危害作物、危害部位等关系抽取的效果相对于病虫害别称和分布区域较好, 原因是在数据采集时别称属性以及分布区域的

实体关系数量较少, 样本数量的不均衡导致了关系抽取的结果存在一定差异。

将文献[15]中基于双向 LSTM 的模型、无注意力 PCNN 模型和本研究提出的 PCNN 叠加不同层级注意力机制模型从抽取结果和运行时间两方面做比较, 实验结果如表 3、表 4 所示。实验结果表明, 采用 PCNN 的方法在 F1 值上均超过了 LSTM 模型, 其原因是数据集的单句字符数大多小于 30, 在这种短文本上 PCNN 利用不同大小的卷积核窗口捕捉单词之间的局部语义关系, 会有很大优势。而 LSTM 网络的优势在于处理较长序列时, 可以通过门结构捕捉长距离依赖, 但线性处理的方式增加了反向传播的长度, 且碍于结构无法像 PCNN 一样并行计算。平均每 epoch 使得模型训练耗时变长了 3.48 倍, 在模型预测时 LSTM 模型与 PCNN 模型相差了一个数量级, 后者只需前者 10.93% 的耗时即可完成所有预测。另外可以注意到, 叠加了注意力机制后模型的表现更优, 其中加了单层实例级注意力因为可以去掉数据集中的错误标注文本噪声而相对无注意力 PCNN 模型效果高, 证明了多实例层注意力机制的作用。再经过添加池化特征注意力层后, 对池化后高层语义特征有了更高的权重分配, 模型效果进一步提升。上述实验表明本研究提出的基于多层注意力机制的远程监督方法提高了在农业病虫害领域数据集上关系抽取的性能, 可对农业病虫害领域关系的自动抽取提供新的方法支持。

#### 4 结论

本研究提出了一种分段池化卷积神经网络模型和多层注意力机制相结合的远程监督关系抽取方法, 并应用于农业病虫害领域进行关系抽取。分别在实例级和池化特征级加入注意力机制, 实例级注意力机制能够减少远程监督获取的数据集的噪音问题, 池化特征级注意力机制则能够综合考量高层语义信息。二者结合后的模型实验结果相较于传统模型有一定的提升。但该方法只能抽取事先预定好的关系集合, 下一步研究的方向是针对开放领域自动发现关系, 并进行关系抽取。

#### 参考文献:

- [1] CHINCHOR N, MARSH E. Muc-7 information extraction task definition[C]//Proceeding of the seventh message understanding conference (MUC-7). Association for Computational Linguistics, Stroudsburg, 1998: 359-367.
- [2] 黄勋, 游宏梁, 于洋. 关系抽取技术研究综述[J]. 现代图书情报技术, 2013(11): 30-39.
- [3] HADDOUD M, MOKHTARI A, LECROQ T, et al. Combining supervised term-weighting metrics for SVM text classification with extended term representation[J]. Knowl Inf Syst, 2016, 49(3): 909-931.
- [4] 韩红旗, 徐硕, 桂婕, 等. 基于词形规则模板的术语层次关系抽取方法[J]. 情报学报, 2013(7): 708-715.
- [5] 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: Volume 2-ACL-IJCNLP '09. Suntec, Singapore. Morristown, NJ, USA: Association for Computational Linguistics, 2009: 1003-1011.
- [6] LIU C Y, SUN W B, CHAO W H, et al. Convolution neural network for relation extraction[C]//International conference on advanced data mining and applications. Springer, Berlin, Heidelberg, 2013: 231-242.
- [7] ZENG D, LIU K, LAI S, et al. Relation classification via convolutional deep neural network[C]//Proceedings of COLING 2014, the 25th international conference on computational linguistics: technical papers, Dublin, 2014.
- [8] ZHANG D X, WANG D. Relation classification via recurrent neural network[EB/OL]. 2015: arXiv:1508.01006 [cs.CL]. <https://arxiv.org/abs/1508.01006>.
- [9] ZHOU P, SHI W, TIAN J, et al. Attention-based bidirectional long short-term memory networks for relation classification[C]//Proceedings of the 54th annual meeting of the association for computational linguistics (Volume 2: Short Papers). Berlin, Stroudsburg: Association for Computational Linguistics, 2016: 207-212.
- [10] ZENG D J, LIU K, CHEN Y B, et al. Distant supervision for relation extraction via piecewise convolutional neural networks[C]//Proceedings of the 2015 conference on empirical methods in natural language processing. Lisbon, Portugal. Stroudsburg, PA, USA: Association for Computational Linguistics, 2015: 1753-1762.
- [11] YANG Z C, YANG D Y, DYER C, et al. Hierarchical attention networks for document classification[C]//Proceedings of the 2016 conference of the north american chapter of the association for computational linguistics: human language technologies. San Diego, California. Stroudsburg, PA, USA: Association for Computational Linguistics, 2016: 1480-1489.
- [12] LIN Y K, SHEN S Q, LIU Z Y, et al. Neural relation extraction with selective attention over instances[C]// Proceedings of the 54th annual meeting of the association for computational linguistics (Volume 1: Long Papers). Berlin, Germany. Stroudsburg, PA, USA: Association for Computational Linguistics, 2016: 2124-2133.
- [13] 白龙, 靳小龙, 席鹏弼, 等. 基于远程监督的关系抽取研究综述[J]. 中文信息学报, 2019, 33(10): 10-17.
- [14] 吴粤敏, 丁港归, 胡滨. 基于注意力机制的农业金融文本关系抽取研究[J]. 数据分析与知识发现, 2019, 3(5): 86-92.
- [15] 朱珊珊, 唐慧丰. 基于 BiLSTM\_Att 的军事领域实体关系抽取研究[J]. 智能计算机与应用, 2019, 9(4): 96-99.
- [16] GU J H, SUN F Q, QIAN L H, et al. Chemical-induced disease relation extraction via attention-based distant supervision[J]. BMC Bioinform., 2019, 20(1): 1-14.
- [17] PENNINGTON J, SOCHER R, MANNING C. Glove: global vectors for word representation[C]//Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). Doha, Qatar. Stroudsburg, PA, USA: Association for Computational Linguistics, 2014: 1532-1543.
- [18] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors[EB/OL]. 2012: arXiv:1207.0580[cs.NE]. <https://arxiv.org/abs/1207.0580>.