

■ 数据挖掘专题(主持人:谢娟英)

引用格式:龚乐君,刘晓林,高志宏,等.基于双向 GRU 和 CNN 的药物相互作用关系抽取[J].陕西师范大学学报(自然科学版),2020,48(6):108-113. [GONG L J, LIU X L, GAO Z H, et al. Extraction of drug-drug interaction based on bidirectional GRU and CNN[J]. Journal of Shaanxi Normal University (Natural Science Edition), 2020, 48(6): 108-113.] DOI: 10.15983/j.cnki.jsnu.2020.01.023

基于双向 GRU 和 CNN 的药物相互作用关系抽取

龚乐君^{1,2*}, 刘晓林^{1,2}, 高志宏³, 李华康^{1,4,5}

(1 南京邮电大学 计算机学院、软件学院、网络空间安全学院, 江苏 南京 210023;

2 江苏大数据安全与智能处理重点实验室, 江苏 南京 210023;

3 浙江省智慧医疗工程技术研究中心, 浙江 温州 325035;

4 自然资源部 城市国土资源监测与仿真重点实验室, 广东 深圳 518034;

5 苏州派维斯信息科技有限公司, 江苏 苏州 215011)

摘要: 不同药物由于药效动力学和药代动力学的差异可能会产生不可预知的副作用,甚至威胁患者的生命安全。在信息技术飞速发展及指数级生物医学文献增加的背景下,从文本中提取药物相互作用成为可能,为此本文提出一种基于双向门控循环单元(GRU)和卷积神经网络(CNN)相融合的双层药物关系抽取模型,使用 DDIEExtraction2013 作为数据集进行多组实验评估,实验结果获得最高 75% 的综合测评率;与其他方法相比较,基于双向 GRU 和 CNN 的双层模型可以有效地抽取文本中的药物相互作用关系。

关键词: 药物相互作用;生物医学关系抽取;药物关系抽取;门控循环单元;卷积神经网络

中图分类号: TP391 **文献标志码:** A

文章编号: 1672-4291(2020)06-0108-06

开放科学(资源服务)标识码(OSID):

Extraction of drug-drug interaction based on bidirectional GRU and CNN

GONG Lejun^{1,2*}, LIU Xiaolin^{1,2}, GAO Zhihong³, LI Huakang^{1,4,5}

(1 School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, Jiangsu, China;

2 Jiangsu Key Lab of Big Data Security & Intelligent Processing, Nanjing 210023, Jiangsu, China;

3 Zhejiang Engineering Research Center of Intelligent Medicine, Wenzhou 325035, Zhejiang, China;

4 Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources, Shenzhen 518034, Guangdong, China;

5 Suzhou Privacy Information Technology Company, Suzhou 215011, Jiangsu, China)

Abstract: Drug-drug interaction (DDI) is the difference between pharmacodynamics and pharmacokinetics of different drugs, which may produce unpredictable side effects or even threaten the life safety of patients. With the rapid development of information technology and the increase of exponential biomedical literature, it is possible to extract drug interactions from texts. A two-layer drug relationship extraction model based on the fusion of bidirectional GRU (bi-gated recurrent unit, BiGRU) and convolutional neural network (CNN) is proposed. DDIEExtraction 2013 is used as data set to evaluate multiple groups of experiments, and obtain the highest comprehensive evaluation rate of 75% is obtained. Compared with other's works, the two-layer

收稿日期: 2020-06-01

基金项目: 浙江省智慧医疗工程技术研究中心资助项目(2016E10011);苏州市姑苏科技创业天使计划项目(CYTS2018233);南京邮电大学引进人才科研启动基金(NY217136)

* **通信作者:** 龚乐君,男,博士,副教授,主要从事数据与文本挖掘、生物医学信息处理的研究。E-mail: glj98226@163.com

model based on bidirectional GRU and CNN can effectively extract the drug interaction relationship in the text.

Keywords: drug-drug interaction (DDI); biomedical relationship extraction; drug relationship extraction; GRU; CNN

随着信息网络的迅速发展,各类生物医学数据都在以指数级增加^[1]。这些无结构、分布散的数据中蕴藏着巨大有价值的信息,从这些海量的无规则数据中快速准确地获取知识具有重要的研究意义。实体关系抽取作为信息检索、知识库自动构建等领域的核心任务和重要环节,能够自动地从非结构化的文本中抽取出实体对之间的语义关系,这种实体间的关系可以形式化的定义为 $\langle \text{entity}_1, \text{relation}, \text{entity}_2 \rangle$, entity_1 和 entity_2 是实体, relation 是实体关系集^[2]。药物间的相互作用关系抽取旨在抽取药物间的不良反应。药物不良反应(adverse drug reactions, ADRs)^[3]是指患者在服用药物后出现的各种不适症状,临床上表现为胃肠道反应,如恶心、呕吐、上腹部不适或者出现腹泻等情况;以及当服用不应同时给药的两种或者多种药物时,患者可能出现的头晕、耳鸣、眼花等情况,还有人会出现过敏反应,表现为瘙痒,甚至能够导致过敏性休克。因此,药物相互作用关系的抽取对指导临床用药和保障患者用药安全有重要意义。

在现有的药物相互作用关系抽取方法中,主要有基于传统机器学习的方法和基于神经网络的方法^[4]。传统机器学习方法依赖手工提取特征信息,而特征提取费时费力且不同特征组合差异会很大^[5]。Huang 等^[6]使用支持向量机(SVM)和长短期记忆网络(LSTM)对 DDI 进行关系抽取,首先使用基于特征的二分类器来分类出正实例,然后使用长短期记忆网络分类器将正实例分类为特定类别,实验结果取得了当时的最高 F 值 69.0%。基于传统机器学习的方法无法从原始语料数据中自动学习特征,因此需要手工定义大量的特征,构建这些特征将花费大量的时间和人力且模型的泛化能力较差。针对传统机器学习存在的这些问题,研究者们尝试将深度学习应用到关系抽取任务中,因深度学习的优势在于能够自动学习句子的特征,无需复杂的特征工程。Wang 等^[7]提出了一种基于依赖的神经网络模型,在双向长短期记忆网络(Bi-LSTM)的基础上建立了 3 个通道,分别为线性通道、DFS 通道和 BFS 通道,取得了 72.0% 的 F 值。Zhang 等^[8]基于循环神经网络并添加最短路径依赖特征,实验取得了 72.9% 的 F 值。在循环神经网络中,当网络层数

很深时,各层在参数共享时很容易造成梯度消失的问题,且模型的计算复杂度普遍很高。针对循环神经网络存在的这一问题,本文提出一种基于双向 GRU 和 CNN 融合的双层模型,GRU 能有效处理句子的长距离依赖。与循环神经网络(RNN)相比,GRU 能够解决循环神经网络中梯度消失问题;与 LSTM 相比,GRU 在模型结构上比 LSTM 减少了一个遗忘门单元,因此 GRU 的计算复杂度更低,能弥补 CNN 在处理长距离依赖上的不足。本文在 DDIExtraction2013^[9]数据集上进行模型性能的验证并与目前的主流方法^[10-21]进行实验对比。

1 方法及模型

单一的神经网络模型有其自身的不足,如 CNN^[16, 20]对句子中连续词特征的抽取比较充分,但是对一个较长的句子无法表征非连续词之间的相关性。而循环神经网络与 CNN 恰恰相反,可以很好地表示一个句子中长距离词之间的相关性,但是对句子中连续词的特征提取不如卷积神经网络。针对这种情况,本文提出一种基于 CNN 和 GRU^[17]相结合的药物实体关系抽取方法。为了充分抽取句子的语义信息和更好的特征表示,通过正向和反向两个方面同时利用句子的历史和未来信息。将预训练的词向量矩阵和位置向量作为双向 GRU 的输入,在经过双向 GRU 模型的训练处理后,将模型的输出拼接再输入到 CNN 层,经过卷积和池化层的作用最终得到分类结果。

1.1 双向 GRU 层

门控循环单元(gated recurrent unit, GRU)^[18, 19]是循环神经网络的变种,GRU 和 LSTM 网络都是对传统循环神经网络的改进,解决了当网络层数很深时可能存在的梯度消失问题,梯度消失问题使得各层网络之间的依赖性变弱,导致难以捕捉句子的长距离依赖。GRU 模型在结构上比 LSTM 少了一个遗忘门,在保持了长短期记忆网络效果的同时又使结构更加简单,因此 GRU 模型的结构更加简单且模型计算更快。在 GRU 模型中包含两个门结构,分别是更新门和重置门。更新门的作用类似于 LSTM 中的遗忘门和输入门,它能决定丢弃哪些信息和添加哪些信息;重置门用于决定

丢弃前一个状态信息的程度。GRU 的门控单元结构如图 1 所示。

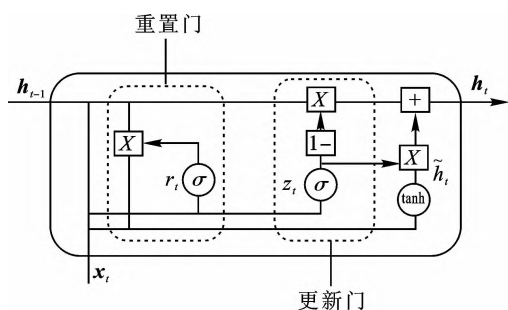


图 1 GRU 门控单元

Fig. 1 GRU gating unit

图 1 中的 GRU 门控单元包含两个模块,分别是重置门和更新门。重置门用于控制忽略前一时状态信息的程度,重置门的值越小说明忽略得越多。更新门用于控制当前状态前一时状态信息被带入到当前状态中的程度,更新门的值越大说明前一时状态信息带入越多。 h_{t-1} 对应前一个状态的输出, h_t 是当前状态的输出也是下一个状态的输入。更新门、重置门、节点状态和节点输出的计算过程如下:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]), \quad (1)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]), \quad (2)$$

$$\tilde{h}_t = \tanh(W \cdot [r_t \otimes h_{t-1}, x_t]), \quad (3)$$

$$h_t = (1 - z_t) \otimes h_{t-1} + z_t \otimes \tilde{h}_t. \quad (4)$$

其中: \otimes 表示矩阵对应元素相乘, σ 表示 sigmoid 函数, W_z 、 W_r 、 W 是 GRU 模型的共享参数。

对于一个输入的句子序列,假设序列 S_i 中有 T 个词,每个词为 $w_t^i, t \in [1, T]$,将句子 S_i 中的每个词作为句子序列的组成部分。那么,分别通过正向 GRU 和反向 GRU 模型就能得到句子的表达:

$$\vec{h}_t^i = \overrightarrow{\text{GRU}}(w_t^i), t \in [1, T], \quad (5)$$

$$\overleftarrow{h}_t^i = \overleftarrow{\text{GRU}}(w_t^i), t \in [1, T]. \quad (6)$$

1.2 CNN 层

卷积神经网络(convolutional neural network, CNN)被广泛地应用于图像处理领域^[20-21]。近年来,研究者开始将 CNN 应用于自然语言处理,例如情感分析、语义解析等任务。卷积神经网络的内部主要包含 3 个部分,分别是卷积层、池化层和 Drop-out。卷积层通过设置不同的卷积核大小对输入的向量进行卷积运算。例如对于输入的向量 $\mathbf{X} = [x_1, x_2, \dots, x_n]$,对于给定的滑动窗口和滤波器的尺寸 $k, t \in \mathbf{R}^{(d_w + 2d_p) \cdot k}$,其中 d_w 是词向量的维度, d_p 是位置向量的维度。局部特征 f_i 的计算公式如下:

$$f_i = \tanh(t \cdot \mathbf{X}_{i:i+k-1} + b). \quad (7)$$

式中, b 是偏差($b \in \mathbf{R}$), \tanh 是双曲正切函数,经过卷积运算最终会得到特征向量 $\mathbf{f} = [f_1, f_2, \dots, f_{n-k+1}]$ 。选取不同的滤波器尺寸 t 会得到不同粒度的特征向量。池化层从每个特征向量中提取最重要的特征,以减少后续层的计算复杂度,本文采用最大池化策略,即 $f_{\max} = \max(f_1, f_2, \dots, f_{n-k+1})$ 。Drop-out 用于在模型训练时随机让某些神经元的权重停止本次更新,同一层中每个神经元权重更新的概率相等,因此能防止模型过拟合,网络在提取训练集特征时,舍弃掉一部分特征来提高网络的泛化能力。

1.3 双向 GRU+CNN

基于双向 GRU 和 CNN 融合模型如图 2 所示,模型的输入是使用 word2vec 预训练好的词向量加上自定义的位置向量。通过双向 GRU 获取原始句子的正向和反向文本序列信息表示,双向 GRU 网络输出的特征向量学习了文本上下文的语序信息;接着将两个方向的文本序列信息向量拼接作为卷积神经网络的输入,通过卷积神经网络来进一步提取文本中的关键特征,在经过卷积和池化后通过 softmax 分类器得到分类结果。

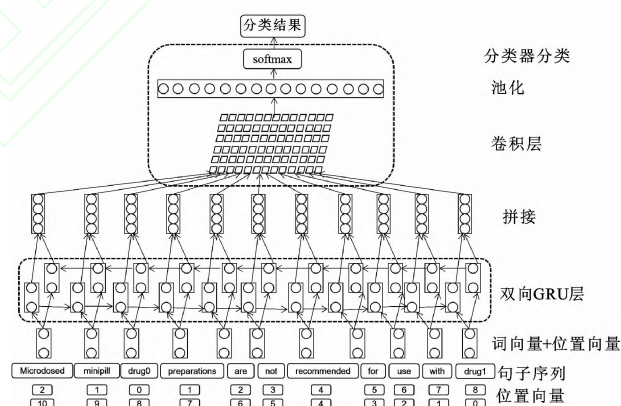


图 2 基于双向 GRU 和 CNN 的模型结构

Fig. 2 Model structure based on bidirectional GRU and CNN

2 实验结果和分析

DDIExtraction2013 数据集^[9-10]主要由两个部分构成,分别是 DrugBank 数据库和 MedLine 摘要,整个语料库包含 1 017 篇文本^[11]。在药物与药物相互作用关系中,两种药物之间的作用关系可以定义为 $R = r(e_1, e_2)$,其中: R 表示两个药物实体之间存在实体关系, r 表示药物实体之间关系的类型, e_1 表示第一个药物实体, e_2 表示第二个药物实体。在 DDIExtraction2013 数据集中,共标注定义了 5 类药物实体相互作用关系类型,分别为 Advice、Effect、Mechanism、Int 和 Negative。因此,药物相互作用

的关系类型可以表示为 $r \in \{\text{Advice}, \text{Effect}, \text{Mechanism}, \text{Int}, \text{Negative}\}$, 其中: Advice 表示在进行两种药物给药时的建议或推荐; Effect 表示药物实体间包含药效机制, 包括临床表现、体征或症状、毒性增加或治疗失败等; Mechanism 表示药物实体之间包含药代动力学机制, 包括实体水平或浓度的变化; Int 表示药物实体之间发生互动但是并不提供有关互动的任何信息; Negative 表示句子中未显示两种药物间存在相互作用。

2.1 数据预处理

DDIExtraction2013 数据集中 5 类药物相互作用关系的占比情况如下: Negative 类所占比重过高, 占比约为 85.11%; Int 类占比约为 0.84%; Effect 占比约为 6.10%; Mechanism 占比约为 4.83%; Advice 占比约为 3.12%。由于数据集中存在数据不平衡的现象, 因此采取过滤占比大的负例 Negative 类, 使用 Effect、Mechanism、Advice 和 Int 这 4 类和原始数据集中未过滤 Negative 的 5 种类型进行对比实验, 研究负例过滤对分类结果的影响。

原始数据集以 xml 的格式提供, 以自定义标签的形式标注文本、药物实体、实体关系对类型。具体的 5 类实例见表 1。

表 1 5 类药物相互作用关系实例
Tab.1 Examples of five DDI types

药物实体关系对	句子	实体类型
(progesterin, Soriatane)	Microdosed minipill progestin preparations are not recommended for use with Soriatane.	Advice
(reserpine, beta-blocking agents)	Catecholamine-depleting drugs, such as reserpine, may have an additive effect when given with beta-blocking agents.	Effect
(Acetazolamide, cyclosporine)	Acetazolamide may elevate cyclosporine levels.	Mechanism
(Campath, antibodies)	An immune response to Campath may interfere with subsequent diagnostic serum tests that utilize antibodies.	Int
(Imipramine, clonazepam)	Imipramine and clonazepam did not change fasting or overload glycemia.	Negative

在一个句子中会存在两个指代相同药物实体的情况, 比如 Propafenone hydrochloride 和 Propafenone 都是指普罗帕酮。因此, 当两种药物

实体有相同的名称、一种药物是另一种药物的缩写形式或一种药物是另一种药物的特例情况时都应该当成一种药物来处理。

词向量^[12-13]已经在自然语言处理中获得了巨大的成功, 被广泛用于机器翻译、文本分类和命名实体识别等领域。词向量提供了一个单词的分布式表示形式并可以通过两个向量间的余弦值来衡量这两个词之间的相似度。与独热编码(one-hot encoding)相比, 词向量的维度更低, 不会产生维度爆炸且模型的计算量更小, 在本文中使用 word2vec 来训练模型的输入。数据的预处理参考文献[14-15], 首先, 处理实例中相同指代的药物实体, 将两个相同指代的药物标记为一个药物实体, 接着将每个实例中的两个候选药物实体替换为“drug0”和“drug1”, 最后将处理后的文本转化为向量得到模型的输入。除了使用字嵌入, 还添加了位置向量(句子中非药物实体的其他词到药物实体词之间的绝对距离)。例如, 对于句子序列 Microdosed minipill [drug0] preparations are not recommended for use with [drug1], 相对于药物实体 drug0 的位置向量为[2,1,0,1,2,3,4,5,6,7,8], 相对于 drug1 的位置向量为[10,9,8,7,6,5,4,3,2,1,0]。使用位置向量能够通过非药物实体词与药物实体之间的距离来确定该词的重要性, 从而为距离药物实体更近的词分配更高的权重。本文基于 Keras 深度学习框架进行实验, 使用 word2vec 词向量训练工具得到词向量, 词向量的维度是 150。为了与预训练的词向量维度统一, 将位置向量的维度也设置为 150。因此对于句子长度不足 150 个单词的在空白位置使用 0 进行填充, 训练集和测试集样本分布情况如表 2 所示。

表 2 训练集和测试集样本分布情况

Tab.2 Sample distribution of training set and test set

数据集	Effect	Mechanism	Advice	Int	Negative
训练集	1 687	1 319	826	188	23 772
测试集	360	302	221	96	4 737

2.2 实验评测指标

为了综合考虑模型在整个数据集上的性能, 本文中使用精确率(Precision)、召回率(Recall)和综合测评率(F1)来衡量模型的性能, 具体如下:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\%, \tag{8}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\%, \tag{9}$$

$$\text{F1} = \frac{2\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \times 100\%. \tag{10}$$

式中,TP 是该类被正确分类的个数,FP 为错误的将其他类别分为该类的个数,FN 为属于该类但被错分为其他类别的个数。

2.3 实验结果和分析

为了研究数据不平衡对实验结果的影响,分别在过滤负例和未过滤负例的数据集上进行实验。过滤负例的性能结果如表 3 所示。

表 3 过滤 Negative 的实验性能

Tab. 3 Performances of filtering negative 单位: %			
关系类别	精确率	召回率	综合测评率
Effect	71	77	74
Mechanism	80	72	76
Advice	81	76	78
Int	79	47	59
平均值	77.75	68	71.75

未去除 Negative 类数据的实验结果如表 4 所示,可以看出 Int 类的召回率(46%)和综合测评率(58%)与其他类别相比都较低,这是由于该类药物相互作用关系对的样本数量较少,导致模型训练不足,因此召回率和综合测评率都较低。而 Negative 类则与之相反,精确率、召回率和综合测评率均获得了 5 种类别中的最高值。Effect、Mechanism 和 Advice 类的综合测评率分别为 73%、73%、75%,平均综合测评率为 75%。

表 4 未过滤 Negative 的实验性能

Tab. 4 Performances of unfiltered negative 单位: %			
关系类别	精确率	召回率	综合测评率
Effect	69	78	73
Mechanism	76	70	73
Advice	77	72	75
Int	77	46	58
Negative	95	96	96
平均值	78.8	72.4	75

与同样使用未过滤 Negative 类的数据集进行药物相互作用关系抽取的其他研究方法(CNN 和 GRU)进行对比,精确率、召回率和综合测评率的比较如表 5 所示。

表 5 与未过滤 Negative 的方法比较

Tab. 5 Comparisons of methods with negative class 单位: %			
模型	精确率	召回率	综合测评率
CNNs[文献 20]	75.72	64.66	69.75
GRUs[文献 19]	73.38	74.06	73.72
BiGRU+CNN	78.8	72.4	75

从表 5 的比较中可以看出,本文提出的双向 GRU 和 CNN 融合方法获得了精确率(78.8%)和

综合测评率(75%)。比单一的 CNN 模型的精确率提升了 3.08%,比单一 GRU 模型的精确率提升了 5.42%。在召回率上,BiGRU+CNN 模型的召回率比较均衡,结果为 72.4%。在综合测评率上,比单一 CNN 模型提升了 5.25%,比单一 GRU 模型提升了 1.28%。从以上分析中可以看出,通过 GRU 和 CNN 模型的融合显著地提升了单一 GRU 或者 CNN 模型的性能,有效地弥补了单一 CNN 或 GRU 模型特征提取不充分的问题。因此,基于双向 GRU 和 CNN 相融合的模型可以有效应用于药物关系的抽取。

数据集中样本分布的不平衡会对样本数较少的类别产生影响。去除 Negative 这类关系,使用 Effect、Mechanism、Advice 和 Int 这 4 类数据进行实验,和未过滤 Negative 类的数据进行实验对比(表 3、表 4),结果显示在过滤了 Negative 类后,除了 Effect 类的召回率有所下降外,其余各类别的精确率、召回率和综合测评率均有所提升,Effect、Mechanism、Advice 和 Int 的综合测评率分别为 74%、76%、78%和 59%,相比没有去除 Negative 类分别提升了 1%、3%、3%和 1%。由此可以得出,过滤 Negative 类可以使各类药物相互作用数据趋于平衡,在 4 类关系上的平均综合测评率为 71.75%。

与过滤 Negative 类的其他工作^[6]相比,实验的精确率、召回率和综合测评率比较如表 6 所示,结果显示使用双向 GRU 和 CNN 相融合的模型在精确率、召回率和综合测评率上均比单独的 LSTM 模型有所提升,精确率比 LSTM 模型高出了 2.45%,召回率比 LSTM 模型高出了 4.3%,综合测评率比 LSTM 模型高出了 2.75%。

表 6 与过滤 Negative 的工作比较

Tab. 6 Comparisons of filtering negative methods 单位: %			
模型	精确率	召回率	综合测评率
LSTM[文献 6]	75.30	63.70	69.00
BiGRU+CNN	77.75	68.00	71.75

综合上述对比分析可以得出,本文提出的双向 GRU 和 CNN 相融合的模型能有效地抽取药物实体的相互作用关系。

3 结语

本文针对单一神经网络模型在处理生物学关系抽取上存在特征提取不充分的问题,提出了一个双向 GRU 和 CNN 融合的模型用于药物的相互作用关系抽取;使用 DDIEExtraction2013 数据集进行模型的性能评估,与单一 GRU 或者 CNN 模型相

比,使用双向GRU和CNN融合模型可以改善单层模型存在的特征抽取不足的情况,能够提取句子中更多的特征信息。实验结果表明,双向GRU+CNN的融合模型比单一的GRU和CNN模型在平均综合测评率上分别提升1.28%和5.25%。未来工作中,尝试加入最短路径依赖和注意力机制来进一步提升模型的性能,以期更好地提高DDI抽取的性能。

参考文献:

- [1] 吴明智,崔雷. 生物医学实体关系抽取的研究[J]. 中华医学图书情报杂志,2010,19(5):5-10.
- [2] 李保利,陈玉忠,俞士汶. 信息抽取研究综述[J]. 计算机工程与应用,2003,39(10):1-5,66.
- [3] KUMAR N H, NARENDRA B, UPENDRA K, et al. A review on adverse drug reactions monitoring and reporting [J]. International Journal of Pharmacy Research & Technology, 2019, 9(2): 12-15.
- [4] SAHU S K, ANAND A. Drug-drug interaction extraction from biomedical texts using long short-term memory network[J]. Journal of Biomedical Informatics, 2018, 86:15-24.
- [5] 鄂海红,张文静,肖思琪,等. 深度学习实体关系抽取研究综述[J]. 软件学报,2019,30(6):1793-1818.
- [6] HUANG D G, JIANG Z C, ZOU L, et al. Drug-drug interaction extraction from biomedical literature using support vector machine and long short term memory networks[J]. Information Sciences, 2017, 415/416:100-109.
- [7] WANG W, YANG X, YANG C Q, et al. Dependency-based long short term memory network for drug-drug interaction extraction [J]. BMC Bioinformatics, 2017, 18(16):99-109.
- [8] ZHANG Y J, ZHENG W, LIN H F, et al. Drug-drug interaction extraction via hierarchical RNNs on sequence and shortest dependency paths[J]. Bioinformatics, 2018, 34(5):828-835.
- [9] SEGURABEDMAR I, MARTINEZ P, ZAZO M H, et al. SemEval-2013 task 9: extraction of drug-drug interactions from biomedical texts (DDIExtraction 2013) [C]//Joint Conference on Lexical and Computational Semantics. 2013: 341-350.
- [10] DEMNERFUSHMAN D, FUNG K W, DO P, et al. Overview of the TAC 2018 drug-drug interaction extraction from drug labels track[EB/OL]. [2020-06-01]. <https://www.researchgate.net/publication/337496164>.
- [11] HOU W, CEESAY B. Extraction of drug-drug interaction using neural embedding[J]. Journal of Bioinformatics and Computational Biology, 2018, 16(6):1840027.
- [12] SIEN S K. Adapting word2vec to named entity recognition[C]//Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA 2015), 2015 (Nodalida):239-243.
- [13] SEGER C. An investigation of categorical variable encoding techniques in machine learning: binary versus one-hot and feature hashing[EB/OL]. [2020-06-01]. <https://www.researchgate.net/publication/343709430>.
- [14] LIM S, LEE K, KANG J. Drug drug interaction extraction from the literature using a recursive neural network[J]. PLoS one, 2018, 13(1):e0190926.
- [15] ZHANG Y J, LIN H F, YANG Z H, et al. a hybrid model based on neural networks for biomedical relation extraction[J]. Journal of Biomedical Informatics, 2018, 81:83-92.
- [16] 李彦冬,郝宗波,雷航. 卷积神经网络研究综述[J]. 计算机应用,2016,36(9):2508-2515,2565.
- [17] DEY R, SALEM F M. Gate-variants of gated recurrent unit (GRU) neural networks[C]//2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS). Boston:IEEE, 2017:1597-1600.
- [18] ZHOU G B, WU J X, ZHANG C L, et al. Minimal gated unit for recurrent neural networks[J]. International Journal of Automation and Computing, 2016, 13(3):226-234.
- [19] 李丽双,钱爽,周安桥,等. 融合依存信息 Attention 机制的药物关系抽取研究[J]. 中文信息学报, 2019, 33(2):89-96.
- [20] LIU S Y, TANG B Z, CHEN Q C, et al. Drug-drug interaction extraction via convolutional neural networks [J]. Computational and Mathematical Methods in Medicine, 2016, 2016:1-8.
- [21] PUTELLI L, GEREVINI A E, LAVELLI A, et al. Applying self-interaction attention for extracting drug-drug interactions[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019: 445-460.

〔责任编辑 宋轶文〕