

DOI:10.13232/j.cnki.jnju.2020.01.014

融合依存信息和卷积神经网络的越南语新闻事件检测

王吉地^{1,2}, 郭军军^{1,2}, 黄于欣^{1,2}, 高盛祥^{1,2*}, 余正涛^{1,2}, 张亚飞^{1,2}

(1. 昆明理工大学信息工程与自动化学院, 昆明, 650500; 2. 云南省人工智能重点实验室, 昆明理工大学, 昆明, 650500)

摘要: 新闻事件检测是自动检测新闻文本中出现的相关事件, 需要大量人力设计模板, 而且难以获取句中隐含的语义信息, 识别触发词时多存在歧义。为解决以上问题, 利用融合依存句法信息的卷积神经网络(Dependency Parsing Convolutional Neural Networks, DPCNN), 针对句子级别越南语新闻事件进行检测。该模型在编码过程中融合了词义、位置信息、词性信息和命名实体信息, 利用传统卷积编码连续词之间的特征, 利用融合依存句法信息的卷积编码非连续词之间的特征, 再融合两部分特征作为事件编码, 进而实现事件检测。实验结果表明, 该方法在越南语新闻事件检测中取得了很好的效果。

关键词: 新闻事件检测, 依存句法信息, 卷积神经网络, 越南语

中图分类号: TP391 **文献标识码:** A

Vietnamese news event detection based on converge dependent information and convolutional neural networks

Wang Jidi^{1,2}, Guo Junjun^{1,2}, Huang Yuxin^{1,2}, Gao Shengxiang^{1,2*}, Yu Zhengtao^{1,2}, Zhang Yafei^{1,2}

(1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, 650500, China; 2. Yunnan Key Laboratory of Artificial Intelligence, Kunming University of Science and Technology, Kunming, 650500, China)

Abstract: News event detection automatically detects related events appearing in news. Currently, detecting news events requires a lot of manpower design templates. Implicit semantic information in sentences is difficult to obtain, and there are many ambiguities in identifying trigger words. In this paper, we use the convolutional neural network method of Dependency Parsing Convolutional Neural Networks (DPCNN) to detect sentence-level Vietnamese news events. Firstly, the model combines the words' semantics part of speech information, the entity information and the position information in the encoding process. Secondly, the traditional convolution is used to encode the continuous character features, and the convolution of fusion dependent information is used to encode the non-continuous character features, thereby implementing event detection. The experimental results show that the method has achieved good results in the detection of Vietnamese leaders travel events.

Key words: news event detection, dependency parsing information, convolutional neural network, Vietnamese

随着“一带一路”的提出, 中越两国共同关注的问题日益增多, 相关的新闻报道也随之增多。及时检测越南语新闻中的新闻事件有助于把握新

闻导向, 获取越南的舆情动态, 并作出有效应对。事件检测是自然语言处理的重要信息提取任务, 旨在识别文本中指定类型的事件。目前, 事件检

基金项目: 国家自然科学基金(61762056, 61732005, 61761026, 61472168, 61672271, 61972186), 国家重点研发计划(2018YFC0830105, 2018YFC0830100), 云南省高新计划专项(201606), 云南省自然科学基金(2018FB104)

收稿日期: 2019-08-20

* 通讯联系人, E-mail: gaoshengxiang.yn@foxmail.com

测研究大都在汉语、英语环境下展开,由于越南语属资源稀缺型语种,针对越南语的事件检测暂无人涉及。因此,本文针对越南语领导人出行领域新闻事件进行检测。

事件检测主要由两部分子任务组成:(1)触发词的检测,识别出句子中的触发词;(2)根据触发词进行分类,确定事件句所表示的事件类型^[1]。分析越南语的构词特点,每一个音节常常是一个有意义的单位,可以独立使用,每个单词又可作为构成多音节词的基础,绝大部分多音节词是双音节,在进行触发词识别时会存在歧义的问题,触发词在不同文本中属于不同事件类型。例如:

S1: Một triệu người tị nạn ở Việt Nam, chỉ có Trung Quốc đã đứng ra cứu trợ.

译文:越南百万难民,只有中国出面救济。

S1 中虽然包含出席活动事件的触发词“đứng ra(出面)”,但通过对事件句的分析可知,事件句中并不包含出席活动事件。事件检测的准确性主要依赖于对触发词的正确识别和事件句语义信息的充分理解。

目前大多是采用深度学习来进行事件检测,可以充分提取事件句的局部特征。但由于越南语独特的语言特征,如句法结构中主语在谓语之前、宾语和补语在动词之后、名词修饰语一般在名词之后,但数词、量词修饰语在名词之前、定语要放在所修饰的中心词之后等^[2],相隔较远的词之间的信息难以捕获。S1 中,通过“đứng ra(出面)”的前后词“Trung Quốc(中国)”“cứu trợ(救济)”以及“cứu trợ(救济)”和“người tị nạn(难民)”,这些语义信息就可以帮助判定“đứng ra(出面)”并非领导人出席活动的触发词。

随着对事件检测的研究,高源等^[3]和 Masung et al^[4]已经证明,依存句法信息对语义理解以及事件检测有重要的作用。本文提出融合依存句法信息的卷积神经网络(Dependency Parsing Convolutional Neural Networks, DPCNN)模型实现句子级越南语新闻事件检测,如 S1 中相隔较远的“cứu trợ(救济)”与“người tị nạn(难民)”两词的语义信息可以通过句法信息引入。同时,DPCNN 关注连续词之间的特征和有句法关系的非连续词之间的特征,挖掘事件句的深层语义信息,实现对

触发词的准确识别,进而实现越南新闻事件的高效检测。

1 相关工作

目前事件检测任务主要基于两类方法:(1)机器学习方法。张炫^[5]提出以狄利克雷过程事件混合模型(Dirichlet Process Event Mixture Model, DPEMM)为核心的事件抽取框架。裴东辉^[6]提出基于支持向量机模型的子事件类别自动识别。高永兵等^[7]针对微博的特征进行 TF-IDF(Term Frequency-Inverse Document Frequency)的改进得出事件提取结果。(2)深度学习方法。与机器学习方法相比,神经网络能自动学习构建特征,可以避免繁琐的特征工程。Nguyen et al^[8]在已有研究的基础上提出一种基于递归神经网络的联合方法进行英文事件抽取。Chen et al^[9]提出动态多池卷积神经网络(Dynamic Multi-Pooling Convolutional Neural Network, DMCNN),解决了句中多个事件的识别以及共享参数匹配的问题。Nguyen and Grishman^[10]使用卷积神经网络对句中的词进行卷积以获得句中隐含的语义信息。但这种连续的卷积无法解决信息依赖问题,有时相隔较远词之间的语义特征反而会有助于事件检测。Nguyen and Grishman^[11]提出 NCNN(N-Gram Convolutional Neural Network)模型对句子中所有非连续的 n 个词进行卷积来解决这个问题,但考虑所有的非连续 n 个单词可能会捕捉到不必要的信息。随着图卷积(Graph Convolutional Networks, GCNs)^[12-14]的提出,Nguyen and Grishman^[15]提出基于依存信息的图卷积神经网络和基于实体提及的池化方法用于事件检测,该方法对依存句法信息进行图卷积,解决句中非连续词之间的信息依赖问题。但 Nguyen 的方法只捕获有句法信息的词之间的关系,忽略了连续词之间的语义信息。

通过分析越南语的语言特性,本文提出融合依存句法信息的卷积神经网络(Dependency Parsing Convolutional Neural Networks, DPCNN)模型,同时采用传统卷积神经网络和融合依存信息的卷积神经网络获取连续词之间的语义信息和相隔较远的词之间的语义信息,既挖掘句子的深层语义信息又避免冗余信息的获取。

2 相关定义

本文对事件的定义如下:

(1)事件:在特定的时间和环境下发生、由若干个角色参与,表现出动作特征的一件事情,由事件触发词与事件参数构成.

(2)事件触发词:能清楚地表达一类事件发生,是触发事件的主要词,通常是单个动词或者名词.例如“Truy cập(出访)”“Tham dự(出席)”“Đề xuất(提出)”“Cuộc họp(会见)”.

(3)事件参数:描述事件发生的时间、地点、人物等信息.

(4)事件类型:将越南语领导人活动新闻事件分为“出访国家”“出席会议”“出席活动”“重要讲话”等五个事件类型和一个非事件类型,其中每个事件类型与多个触发词对应,例如“Đi thăm(访问)”和“Ghé thăm(拜访)”均为“Ra ngoài truy cập(外出访问)”的触发词.

3 方法

本文将事件检测问题转化为多分类问题,以此来判断句子中是否存在新闻事件,若存在则指出其事件类型.提出的DPCNN由三个部分组成:编码层、卷积层和池化层,如图1所示.

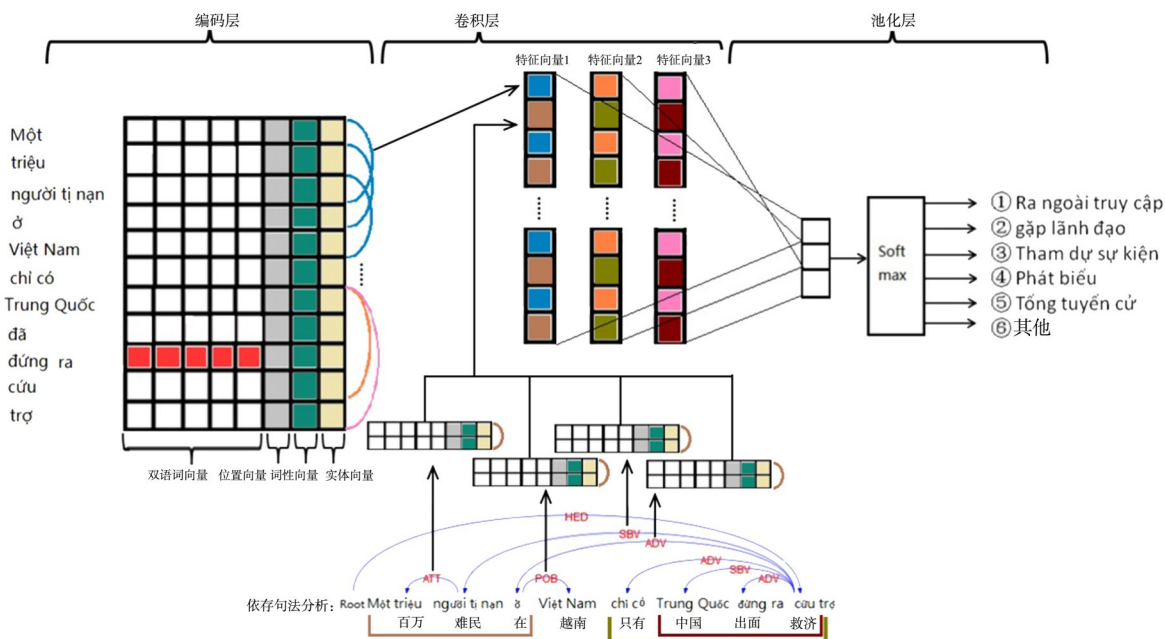


图1 DPCNN模型的结构

Fig. 1 Model structure of DPCNN

3.1 编码层 首先将句子中的词级信息转换成实值向量作为神经网络的输入. 设 $X = \{x_1, x_2, x_3, \dots, x_n\}$ 是一个长度为 n 的句子, 其中 x_i 是句子中的第 i 个词. 在自然语言处理任务中, 词的语义信息与其在句中的位置有关^[16], 词性和实体类型信息对触发词的识别和语义的理解有提升的作用. 本文融合词向量、位置向量、词性向量和实体向量作为模型输入.

词向量是一个实值向量, 目前有两种常用的表示方式: 一种是 one-hot 表示, 创建一个词表, 并

把每个词顺序编号; 另一种是基于分布式的词向量. 本文采用 Mikolov et al^[17] 提出的 word2vec 模型训练方法训练越南语词向量.

由于位置编码可以引入当前词的语义结构信息, 本文将位置编码作为编码的一部分, 指当前词与触发词的相对位置. 例如, S1 中 “đứng ra(出面)” 和 “người tị nạn(难民)” 之间的相对位置为 6.

由于词性和实体类型有助于获取当前的词语意信息, 依据侯中熙和杨蓓^[18] 提出的方法对越南语进行词性标注, 并定义词性嵌入表, 将 28 种词

性标签嵌入到词性向量中.

与词性向量类似,依据刘艳超等^[19]提出的方法对越南语进行命名实体识别,定义实体嵌入表,识别出句子中的人名、地名、组织机构名、时间等命名实体,将实体标签嵌入到实体向量中.共有十种实体类型,分为三大类(实体类、时间类和数字类)、七小类(人名、机构名、地名、时间、日期、货币和百分比).

3.2 卷积层 卷积层可以捕获整个句子的组合语义信息,并将这些有价值的语义压缩到特征映射中.卷积运算中最重要的一个部分就是滤波器 w ,它可以提取卷积窗内词之间的特征.当卷积核大小为 m 时,窗口内的 m 个词 $\{x_i, x_{i+1}, x_{i+2}, \dots, x_{i+m-1}\}$ 采用 $x_{i:i+m-1}$ 表示,得到的卷积特征用 c_i 表示,如式(1)所示:

$$c_i = f(wx_{i:i+m-1} + b) \tag{1}$$

其中, $b(b \in R)$ 是偏置项, f 是非线性激活函数,滤波器应用于句子中每个可能的窗口 $\{x_{1:m}, x_{2:m+1}, \dots, x_{n+m-1:n}\}$. 由于句中的特征并非单一,因此在卷积过程中使用多个滤波器来获取不同特征.当使用 k 个滤波器 $W = \{w_1, w_2, \dots, w_k\}$ 时,卷积运算如式(2)所示:

$$c_{ji} = f(w_j x_{i:i+m-1} + b_j), j \in [1, k] \tag{2}$$

采用李英等^[20]提出的方法构建越南语依存句法树,如图 2 所示.通过分析可知,“cứu trợ(救济)”和“người tị nạn(难民)”之间的句法关系为“SBV(主谓关系)”,这有助于判断“đứng ra(出面)”并非出席活动事件的触发词.

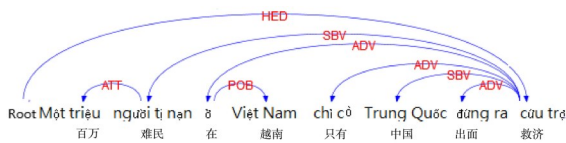


图 2 S1 的依存句法分析结果

Fig. 2 Analysis result of S1 by dependency parsing

卷积运算可以捕获窗口内连续词之间的语义信息,却不能捕获窗口外非连续词的特征.本文通过依存句法分析引入窗口外的信息,依存信息用 $D = \{N, E\}$ 表示,其中 $N = \{x_1, x_2, x_3, \dots, x_p\}$ ($p \leq n$) 表示句中存在依存关系的所有词节点,有

依存关系的两词节点用 x_s, x_t 表示, E 是两词节点之间边,每个边 (x_s, x_t) 代表从词节点 x_s 指向词节点 x_t , 并且有依存信息标签 $L(x_s, x_t)$. 例如,图 2 中节点“cứu trợ(救济)”和“người tị nạn(难民)”之间的有向边依存信息标签为 $L(x_s, x_t) = L(\text{cứu trợ}, \text{người tị nạn}) = \text{Datv}$. Kipf and Welling^[12]提出的方法表示信息流不只是按照标签指示的方向,因此这里添加了自循环 (x_s, x_s) 和反向边 (x_t, x_s) . 自循环的标签为 $L(x_s, x_s)$, 反向边的标签为 $L^-(x_t, x_s)$. 特定的依存信息标签具有固定的参数,依存特征的计算如下:

$$h_{ji} = f\left(\sum W_{jL(x_i, N)} x_{i, N} + b_{L(x_i, N)}\right) \tag{3}$$

其中, j 的范围是 1 到 k , f 为非线性激活函数, $W_L(x_i, N)$ 有三种形式,分别是原始边、反向边、自循环边, $b_L(x_i, N)$ 为偏置项. 最后,将卷积特征和融合依存句法信息的卷积特征拼接,作为当前句特征. 如式(4)所示:

$$E_{ji} = \lambda c_{ji} + (1 - \lambda) h_{ji} \tag{4}$$

$E \in R^{k(n-m+1)}$ 为卷积得到的结果矩阵, k 为滤波器的个数, n 为句子长度, m 为滤波器窗口大小.

3.3 池化层 池化层可以提取卷积特征中的最具代表性的特征. 本文选取最大池化的方法如式(5)所示:

$$E^* = \text{Each} - \max(E_1, E_2, E_3, \dots, E_k) \tag{5}$$

k 个滤波器,提取每个滤波器中最有价值的局部特征,其他特征值全部抛弃, k 个局部特征聚合成一个向量 E^* 作为事件编码. 最后,将事件编码送入全连接层,使用 *softmax* 激活函数对 E^* 进行分类,得到事件的分类概率,根据概率分布对事件的类型进行预测,如式(6)所示:

$$S_i = \frac{e^{E^*}}{\sum_i^C e^{E^*}} \tag{6}$$

其中, C 表示类别个数, i 表示类别索引 i 的范围是 1 到 6.

4 实验

4.1 实验设计 目前为止尚未有可供公开测评的越南语事件检测的语料库. 实验语料来源于东南亚舆情分析平台,参考 ACE (Automatic Content Extraction) 的事件标注体系对语料进行标

记,经过去重、筛选以后标注了领导人出行活动领域的越南语新闻文本 1233 篇,共 9576 条事件句。将 8069 条事件句作为训练数据,1507 条事件句作为测试集。本文构建的语料中划分了五种事件类型和一种非事件类型,分别为外出访问、领导会见、出席会议、出席活动、发表讲话和无事件类型,对应的触发词如表 1 所示。

表 1 触发词表

Table 1 Trigger vocabulary

事件类型	事件触发词
Ra ngoài truy cập 外出访问	Đi thăm, Ghé thăm, Truy cập, Nhiệm vụ, Kiểm tra 拜访, 访问, 出访, 走访, 探问, 考察, 探访...
gặp lãnh đạo 高层会见	Họp, Gặp gỡ, Nói chuyện, Tiếp kiến, Gặp nhau, Gặp mặt, Hội đàm 会见, 接见, 见面, 会晤, 会谈...
Tham dự sự kiện 出席活动	Tham dự, Đi ra, Tham gia, Sắp xếp, Đến rồi, Tại cuộc họp 出席, 出面, 参加, 列席, 到场, 到会...
Phát biểu 发表讲话	Bài giảng, xuất bản, nói, trình bày, phát biểu, đề nghị, nói chuyện 演讲, 发表, 发言, 提出, 讲话, 谈话...
Tổng tuyển cử 换届选举	Đề nghị, bỏ phiếu, Giới thiệu, bầu cử 推举, 选举, 推选, 投票竞选...

模型用 Tensorflow 深度学习框架实现。使用固定长度为 25 的句子作为输入,即句子长度大于 25 时截取前 25 个词,句子长度不足 25 时则使用特殊字符填充。使用 253 维的预训练词嵌入,位置嵌入 1 维,词性嵌入 1 维,实体类型嵌入 1 维。本模型采用两层卷积,每层卷积有三个滤波器,共三个卷积通道,卷积核大小分别为 3, 4, 5, 卷积步长为 1, 边界采用全 0 填充。初始学习率为 0.01, 使用交叉熵作为损失函数,采用反向传播和 Adam 优化算法训练模型,随机失活率 50%, batch-size 50, epoch 50, 衰减系数 0.96。为了防止过拟合,在分类时加入 L2 正则项,其系数为 3。

4.2 对比实验 采用准确率(P)、召回率(R)和 F 值(F)作为评价指标。

$$P = \frac{A}{A+B} \tag{7}$$

$$R = \frac{A}{A+C} \tag{8}$$

$$F = \frac{2PR}{P+R} \tag{9}$$

其中,A 为正确识别事件类型的数量,B 为错误识别事件类型的数量,C 为未被识别到的正确识别事件类型的数量。

4.2.1 卷积层数探究 为探究卷积层数对实验结果的影响,分别采用一层、两层和三层卷积的模型进行实验,找到最优层数。实验结果如表 2 所示。

表 2 卷积层数对实验结果的影响

Table 2 Influence of the number of convolution layers on experimental results

卷积层数	P	R	F
1	74.04%	62.63%	70.08%
2	76.78%	64.25%	71.45%
3	75.53%	59.01%	68.23%

由表 2 可见,当卷积层数为 2 时效果最佳,准确率、召回率和 F 值分别为 76.78%, 64.25%, 71.45%。而当卷积层数为 3 时,模型的性能却有所下降。因此后续实验中模型均采用两层卷积。

4.2.2 编码特征探究 针对词嵌入层融入的编码特征进行探究。在去掉某一项编码向量之后,剩余的两类编码向量和词向量融合作为模型的输入,探究不同编码特征组合对模型性能的影响。实验结果如表 3 所示。

表 3 编码特征对实验结果的影响

Table 3 Influence of combination coding characteristics on experimental results

编码特征	P	R	F
词向量、位置向量、词性向量和实体向量	76.78%	64.25%	71.45%
词向量、词性向量和实体向量	74.23%	62.3%	69.2%
词向量、位置向量和实体向量	71.88%	63.4%	69.3%
词向量、位置向量和词性向量	73.46%	64.02%	69.97%

由表 3 可知,去掉某一项编码向量之后,模型的准确率、召回率、F 值与本文模型相比均有所下降,所以同时使用三种编码向量可以提高事件检测的性能。

4.2.3 卷积核大小探究 为探究卷积核大小对模型的影响,分别采用不同大小的三个卷积核获取特征向量,实验结果如表 4 所示。

表 4 卷积核大小对实验结果的影响

Table 4 Influence of convolution kernel size on experimental results

卷积核大小	<i>P</i>	<i>R</i>	<i>F</i>
2,3,4	73.21%	63.59%	66.73%
3,4,5	76.78%	64.25%	71.45%
4,5,6	75.07%	61.22%	70.88%
5,6,7	73.12%	62.25%	67.54%

由表 4 可知,当卷积核为 3,4,5 时模型效果最佳。较小和较大的卷积核均会降低模型的性能。

4.2.4 不同模型的探究 为证明本文模型在越南语事件检测任务上的效果,将本文模型与传统卷积神经网络、Nguyen et al^[8]提出的循环神经网络模型(Recurrent Neural Networks, RNN)和 Nguyen and Grishman^[15]提出的融合依存句法信息的图卷积(GCNs)模型在越南语新闻事件检测任务上作对比,实验结果如表 5 所示。

表 5 不同模型的性能对比

Table 5 Performance of different models

模型	<i>P</i>	<i>R</i>	<i>F</i>
RNN	70.23%	65.89%	67.23%
CNN	73.23%	63.14%	69.23%
GCNs	75.00%	63.92%	70.24%
DPCNN	76.78%	64.25%	71.45%

由表 5 可知,卷积神经网络的模型均优于循环神经网络模型,其中 DPCNN 和 GCNs 的模型效果优于 CNN, DPCNN 优于 GCNs。由于循环神经网络存在长距离依赖的问题,会影响事件的准确检测,而引入依存句法信息可以捕获到 CNN 没有捕获到的信息。GCNs 和 DPCNN 相比,可以看到 DPCNN 的 *F* 值提升了 1.21%,说明大部分信息可以被 GCNs 所捕获,但同时使用连续的卷积神经网络和融合依存句法信息的卷积神经网络则可以捕获到句中更多的隐含信息。

5 结 论

本文提出一种越南语新闻事件检测的新型神经网络模型,该模型融合词向量、位置向量、词性向量和命名实体向量来捕捉词级别的语义信息,同时使用传统的卷积神经网络和融合依存句法信息的卷积神经网络获取句子级别语义信息。通过对模型设置不同的参数,将最佳模型与基线方法作比较,证明本文提出的方法在越南语新闻事件的检测中得到较好的效果。但由于越南语的语料尚不完善,本文提出的模型在越南语新闻事件检测任务上仍有提升空间。事件检测是事件抽取的重要环节,后期可在一个模型中同时识别出事件触发词和事件参数,从而实现事件抽取的任务。

参考文献

- [1] Linguistic D C. ACE (Automatic Content Extraction) Chinese annotation guidelines for events, version 5.5.1. <http://www ldc upenn edu/Projects/ACE>, 2005.
- [2] 黄敏中. 实用越南语语法. 北京: 北京大学出版社, 1997, 325.
- [3] 高源, 席耀一, 李弼程等. 基于依存句法分析与分类器融合的触发词抽取方法. 计算机应用研究, 2016, 33(5): 1407—1410. (Gao Y, Xi Y Y, Li B C, et al. Trigger extraction algorithm based on dependency parsing and classifier fusion. Application Research of Computers, 2016, 33(5): 1407—1410.)
- [4] Massung S, Zhai C, Hockenmaier J, et al. Structural parse tree features for text representation//The 7th International Conference on Semantic Computing. Irvine, CA, USA: IEEE, 2013: 9—16.
- [5] 张炫. 微博事件抽取. 硕士学位论文. 南京: 东南大学, 2017. (Zhang X. Event extraction from twitter. Master Dissertation. Nanjing: Southeast University, 2017.)
- [6] 裴东辉. 中文新闻事件抽取方法研究. 硕士学位论文. 昆明: 昆明理工大学, 2015. (Pei D H. Research on Chinese news event extraction method. Master Dissertation. Kunming: Kunming University of Science and Technology, 2015.)
- [7] 高永兵, 陈超, 熊振华等. 基于个人微博特征的事件提取研究. 计算机应用与软件, 2016, 33(7): 47—51. (Gao Y B, Chen C, Xiong Z H, et al. Research on

- event extraction based on personal microblog characteristics. *Computer Applications and Software*, 2016, 33(7):47–51.)
- [8] Nguyen T H, Cho K, Grishman R, et al. Joint Event extraction via Recurrent Neural Networks// *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Berlin, Germany: ACL Press, 2016: 300–309.
- [9] Chen Y, Xu L, Liu K, et al. Event Extraction via dynamic multi - pooling 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: ACL Press, 2015:167–176.
- [10] Nguyen T H, Grishman R. Event detection and domain adaptation 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: ACL Press, 2015:365–371.
- [11] Nguyen T H, Grishman R. Modeling Skip-Grams for event detection with convolutional neural networks// *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. Berlin, Germany: ACL Press, 2016:86–91.
- [12] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks// *The 5th International Conference on Learning Representations*. Vancouver, Canada: ACL Press, 2017:89–96.
- [13] Marcheggiani D, Titov I. Encoding sentences with graph convolutional networks for semantic role labeling// *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Vancouver, Canada: ACL Press, 2017:1506–1515.
- [14] Kearnes S, McCloskey K, Berndl M, et al. Molecular graph convolutions: moving beyond fingerprints. *Journal of Computer-Aided Molecular Design*, 2016, 30(8):595–608.
- [15] Nguyen T H, Grishman R. Graph convolutional networks with argument - aware pooling for event detection// *The 18th AAAI Conference on Artificial Intelligence*. New Orleans, LA, USA: AAAI Press, 2018:5900–5907.
- [16] Liu S, Chen Y, Liu K, et al. Exploiting argument information to improve event detection via supervised attention mechanisms// *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*. Vancouver, Canada: ACL Press, 2017:1789–1798.
- [17] Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space// *The 1st International Conference on Learning Representations*. Sofia, Bulgaria: ACL Press, 2013: 167–181.
- [18] 侯中熙, 杨蓓. 基于 SVMTool 的越南语词性标注. *价值工程*, 2016, 35(20):159–161. (Hou Z X, Yang P. Vietnamese word - of - speech tagging based on SVMTool. *Value Engineering*, 2016, 35(20): 159–161.)
- [19] 刘艳超, 郭剑毅, 余正涛等. 融合实体特性识别越南语复杂命名实体的混合方法. *智能系统学报*, 2016, 11(4):503–512. (Liu Y C, Gou J Y, Yu Z T, et al. A hybrid method for identifying Vietnamese complex named entities by merging entity features. *CAAI Transactions on Intelligent Systems*, 2016, 11(4): 503–512.)
- [20] 李英, 郭剑毅, 余正涛等. 越南语短语树到依存树的转换研究. *计算机科学与探索*, 2017, 11(4):599–607. (Li Y, Guo J Y, Yu Z T, et al. Constituent-to-dependency conversion for Vietnamese. *Journal of Frontiers of Computer Science & Technology*, 2017, 11(4):599–607.)

(责任编辑 杨可盛)