

基于门控卷积网络的篇章级事件可信度识别方法

张 贇 李培峰 朱巧明

苏州大学计算机科学与技术学院 江苏 苏州 215006

(20175227048@stu.suda.edu.cn)



摘 要 事件可信度表示文本中事件的真实程度,描述了事件是否是一个事实,或是一种可能性,又或者是一种不可能的情况。事件可信度识别是问答系统、篇章理解等诸多相关任务的重要基础。目前,事件可信度识别的研究基本上还停留在句子级,很少涉及篇章级。因此,文中提出了一个基于门控卷积网络的篇章级事件可信度识别方法 DEFI(Document-level Event Factuality Identification)。该方法首先使用门控卷积网络从句子和句法路径中抽取篇章中事件的语义和句法信息,然后通过自注意力(Self-Attention)层获取每个序列相对于自身更重要的整体信息的特征表示,从而识别出篇章级事件可信度。在中英文语料上的实验显示,与基准系统相比,DEFI的宏平均 $F1$ 值和微平均 $F1$ 值均得到了提高,其中在中英文语料上宏平均 $F1$ 值分别提高了 2.3% 和 4.4%,微平均 $F1$ 值分别提升了 2.0% 和 2.8%;同时,所提方法在训练速度上也提升了 3 倍。

关键词: 篇章理解,事件可信度识别,门控卷积

中图法分类号 TP391.1

Document-level Event Factuality Identification Method with Gated Convolution Networks

ZHANG Yun, LI Pei-feng and ZHU Qiao-ming

School of Computer Sciences and Technology, Soochow University, Suzhou, Jiangsu 215006, China

Abstract Event factuality represents the factual nature of events in texts, it describes whether an event is a fact, a possibility, or an impossible situation. Event factuality identification is the basis of many relative tasks, such as question-answer system and discourse understanding. However, most of the current researches of event factuality identification focus on the sentences level, and only a few aim at the document-level. Therefore, this paper proposed an approach of document-level event factuality identification (DEFI) with gated convolution network. It first uses gated convolution network to capture both the semantic information and the syntactic information from event sentences and syntactic path, and then uses the self-attention layer to capture the feature representation of the overall information that is more important for each sequence itself. Finally, it uses the above information to identify the document-level event factuality. Experimental results on both the Chinese and English corpus show that the proposed DEFI outperforms the baselines both on macro-F1 and micro-F1. In Chinese and English corpus, the macro-average $F1$ value increased by 2.3% and 4.4%, while the micro-average $F1$ value increased by 2.0% and 2.8%, respectively. The training speed of this method is also increased by three times.

Keywords Discourse understanding, Event factuality identification, Gated convolution network

1 引言

事件可信度识别用于判断一个事件是否为一个事实,或一种可能性,又或一种不可能的情况。事件可信度识别是自然语言处理的重要基础,可用于情感分析^[1]、观点检测^[2]及问答系统^[3]等。事件可信度识别任务主要分为句子级和篇章级两类。其中,句子级事件可信度识别用于识别一个事件实例的可信度;篇章级事件可信度识别则是识别一篇文档中频繁谈及的某个事件的可信度。如例 1 所示的文档中多次谈到“发现海盗沉船”这一事件。在各个事件句(包含事件的句子)

中该事件实例的可信度值就是其句子级的可信度。如例 1 的首句“报告称在非洲发现(CT-)著名海盗沉船,被联合国否认。”中,事件“发现”被否定线索词“否认”修饰,该句中事件“发现”的可信度值为“一定不可信/CT-”。在其他谈到该事件的句子中,其句子级可信度值分别是“可能可信/PS+”或“一定可信/CT+”等。但从全文的角度综合考虑,该事件的篇章级可信度值为“一定不可信/CT-”。一个事件的篇章级可信度值唯一,而它在各个事件实例中的可信度值可以不同。

例 1 报告称在非洲发现(CT-)著名海盗沉船,被联合国否认。

收稿日期:2019-02-05 返修日期:2019-06-09 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金(61836007,61772354,61773276)

This work was supported by the National Natural Science Foundation of China (61836007,61772354,61773276).

通信作者:李培峰(pfli@suda.edu.cn)

多年来人们一直希望能够发现(PS+)基德船长掠夺的宝藏。

中新网7月15日电,据外媒报道,联合国教科文组织发表声明,否认了先前有关发现(CT-)一艘17世纪著名海盜沉船的报告。

一个搜寻小组此前声称在非洲马达加斯加海岸发现(CT+)了苏格兰著名海盜基德船长的沉船“加里冒险号”,并且发现了一个50公斤的银锭。

今年5月,搜寻小组在一次全球瞩目的仪式上把这个银锭送给了马达加斯加总统。

联合国教科文组织的声明说,这个从海里捞上来的“银锭”事实上含有95%的铅。

受命就这项发现(CT+)展开调查的联合国教科文组织小组说,发现(CT-)“银锭”的“残骸”并不是基德船长的“加里冒险号”沉船残骸,而是圣玛利港口设施的破损部分。

基德船长的故事曾经激励英国小说家罗伯特·路易斯·史蒂文森创作了著名的小说《金银岛》。

许多专家认为,人们300年来都在寻找基德船长的宝藏,但他的宝藏只是一个传说。

句子级事件可信度识别只需要考虑一个句子;而篇章级可信度识别需要综合全文内容,更加具有挑战性。目前,由于缺乏相关语料库资源,篇章级事件可信度识别的研究工作还处于起步阶段,仅有的工作是Qian^[4]构建了一个篇章级事件可信度识别语料库,并提出了一种基于对抗网络的篇章级事件可信度识别模型。但是,该模型中的BiLSTM神经网络存储单元有限,并且会给输入序列两端的信息赋予更大的权重,在输入序列加长时,这会削弱事件词与其邻近词语间的相关性;同时,该模型难以实现并行操作,无法充分利用计算资源,会耗费更多的时间来进行模型训练。

本文首次将门控卷积网络(Gated Convolution Neural Network,GCNN)用于篇章级事件可信度识别任务。首先通过GCNN和自注意力机制层抽取篇章语义和句法信息,通过这些信息综合判断篇章级事件可信度;再通过多层叠加进行多次卷积操作,以获取序列的全局信息;同时,每个层次的卷积操作是相对独立的过程,可避免Qian^[4]的模型无法进行并行化训练的弊端。

2 相关工作

本节将从事件可信度语料库和相关识别方法两方面来介绍相关工作。

在句子级事件可信度语料库方面,Sauri等^[5]构建了基于TimeBank^[6]的英文事件可信度语料库FactBank^[7],多数英文句子级事件可信度识别方法基于此语料库。Cao等^[8]在ACE 2005中文事件语料库的基础上,标注了各个事件的可信度及其相关属性。在篇章级事件可信度语料库方面,Qian^[4]从中国日报网和新浪双语新闻上选取了1730篇英文文档和4650篇中文文档,构建了一个篇章级事件可信度语料库。标注中所用的英文线索词来自BioScope语料库^[9],中文线索词来自CNeUn语料库^[10]。标注的事件可信度标签分别为“不清楚/Uu”“一定不发生/CT-”“可能不发生/PS-”“可能发生/PS+”“一定发生/CT+”。表1给出了该语料库

中各类别样本的分布情况。

表1 语料库的分布统计

| | Chinese | English |
|------------|-------------|-------------|
| CT- | 1342/28.8% | 279/16.13% |
| PS+ | 848/18.24% | 274/15.84% |
| CT+ | 2403/51.68% | 1150/66.47% |
| 其他(Uu+PS-) | 57/1.22% | 27/1.56% |

在句子级事件可信度识别研究方面,Sauri^[11]和Werner等^[12]均采取了基于支持向量机和特征工程的方法。Qian等^[13]提出了一个统计和规则相结合的模型;随后,其又提出了一个基于生成式对抗网络的方法^[14]。在中文方面,He^[15]提出了一种基于CNN模型的中文事件可信度识别方法。

在篇章级事件可信度方面的研究还处于起步阶段。Qian^[4]在自主标注的语料库上提出了一种基于对抗网络的篇章级事件可信度识别模型。该模型通过使用BiLSTM对包含事件的句子信息和句法路径进行编码,然后通过注意力机制将这些编码后的信息融合在一起,得到综合整篇文章的事件信息的特征表示,并将这些特征表示输入softmax层对事件可信度进行识别。本文的工作是在Qian^[4]工作的基础上进行的扩展。

3 基于GCNN的篇章级事件可信度识别方法

在篇章级事件可信度识别任务上,Qian^[4]提出了一种基于对抗网络的篇章级事件可信度识别模型,采用BiLSTM分别抽取文章中事件的语义信息和句法信息,并利用注意力机制^[16]得到综合的特征表示,最后识别篇章级事件可信度。Qian^[4]认为事件句的前后邻句对判断事件可信度有重要作用,并根据当前句子的邻句中是否含有不确定或否定线索词,将句法路径定义为:1)先计算当前事件到达当前句子根节点的依存句法路径,再计算邻句线索词到达其所在句的根节点的依存句法路径,最后将二者的根节点合并为一个结点;2)计算当前句中不确定或否定线索词到达事件的依存句法路径。Qian将这种包含邻句的句子和句法路径作为该模型所用特征。

Qian^[4]将事件句与其邻句共同作为句子特征,这样可以使模型抽取到更多有用的信息,但同时也会带来更多的干扰信息。由于BiLSTM神经网络的记忆单元容量有限,虽然相对于传统的RNN网络而言可以改善梯度弥散问题,但还是会使得序列两端的信息占据更大的比重。在输入中包含事件句的邻句会使得输入序列增长,从而倾向于把序列两端的词语作为更有用的信息,而忽略了事件词与相临近词语间的相关性;而且循环神经网络每一时刻的输入都依赖于前一时刻的输出,因此无法进行并行化操作,难以充分利用计算资源。

本文首次把GCNN应用于篇章级事件可信度识别任务,实现了一个基于门控卷积的篇章级事件可信度识别模型DEFI。由于多层叠GCNN可以通过多层次卷积操作来获取序列的全局信息,而每个层次的卷积操作都是一个独立的过程,因此DEFI很大程度上解决了Qian^[4]模型无法进行并行化操作的问题,同时缓解了其在处理较长输入序列时对两端信息过度信赖的问题。本文模型包含输入层、特征提取层和事件可信度识别层,其结构如图1所示。

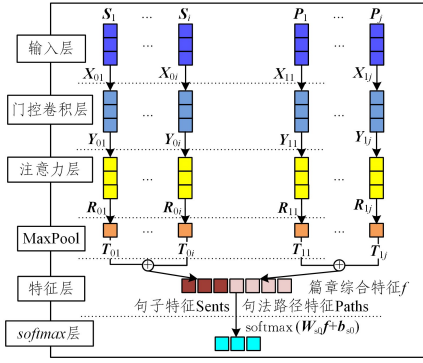


图1 DEFI模型框架图

Fig. 1 Structure diagram of DEFI

3.1 输入层

与 Qian^[4]提出的模型类似,本文将一个篇章中包含该事件的句子(本文假设每个篇章中的事件已知,但每个事件的可信度未知)和句法路径分别作为输入的序列。其中,每个句子或句法路径都是由词组成的序列。把一个含有 n 个词的序列记作:

$$\mathbf{X} = (x_1, x_2, \dots, x_n)$$

其中, x_i 表示该序列中第 i 个词在字典中的 id。通过 Word2vec 工具,对本文中语料库的相关领域语料进行预训练,得到 Embedding 矩阵,将序列中的每一个词 x_i 映射为一个低维度的稠密词向量, $x_i \in \mathbb{R}^d$,从而将序列转换成一个向量表示 $\mathbf{X}, \mathbf{X} \in \mathbb{R}^{d \times n}$ 。

本文与 Qian^[4]模型选取同样的特征输入,即选取事件句的句子序列和依存句法路径序列作为输入,具体输入 i 个句子序列 $S_1 - S_i, j$ 个句法路径序列 $P_1 - P_j$ 。经过输入层之后,由句子序列 $S_1 - S_i$ 得到 $\mathbf{X}_{01} - \mathbf{X}_{0i}, \mathbf{X}_{0i} \in \mathbb{R}^{d \times n}$ 表示第 i 个句子的词向量序列。由句法路径序列 $P_1 - P_j$ 得到 $\mathbf{X}_{11} - \mathbf{X}_{1j}, \mathbf{X}_{1j} \in \mathbb{R}^{d \times n}$ 表示第 j 个句法路径的词向量序列。

3.2 特征提取层

特征提取层由门控卷积层、注意力层和 Max-Pooling 层组成。

门控卷积层采用门控卷积网络^[17]。本文所用的门控卷积层不同于一般的 GCNN,是在卷积层 CNN 的基础上引入了门控机制,并且堆叠多层 CNN 结构。其优点如下:1)RNN 模型在下一时刻的输出依赖前一个时刻的隐藏层状态,因此无法实现并行运算;而 CNN 模型每次操作都只考虑当前局部特征,因此可以进行并行运算,从而能够充分利用计算资源,提高计算速度。2)在引入了门控机制后,可以减缓梯度消失,加快模型的收敛。3)通过叠加多层 CNN 结构来学习词序列的上下文依赖关系,使其在长文本学习中取得不错的效果。

在输入层得到句子序列和句法路径序列的向量表示之后,将各个序列输入到门控卷积层中,经过叠加三层的 GCNN 网络,学习到句子序列和句法路径序列的特征表示。本文把第一层 GCNN 的输出记作 Y_i^1 :

$$Y_i^1 = (\mathbf{X}_i * \mathbf{W}_1 + b_1) \otimes \sigma(\mathbf{X}_i * \mathbf{V}_1 + c_1) \quad (1)$$

其中, $\mathbf{W}_1, \mathbf{V}_1 \in \mathbb{R}^{d \times h}, b_1, c_1 \in \mathbb{R}^h$ 是参数矩阵, \otimes 表示逐元素相乘, σ 表示 sigmoid 函数。

普通 CNN 模型仅捕捉局部信息,而通过叠加卷积层和加入门控机制可以使得模型学习到词序列的前后依赖关系,

也可以得到更多复杂的语义特征和句法特征,能更加均衡地抽取全局信息。

从本质上讲,深度学习中的注意力机制也是从全局信息中获取对当前任务更重要的信息。最近的一些工作^[18]通过注意力机制对输入序列进行编码,如式(2)所示:

$$Attention(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}(\mathbf{Q}\mathbf{K}^T / \sqrt{d_k}) \quad (2)$$

其中, $\mathbf{Q}, \mathbf{K}, \mathbf{V}$ 表示输入序列, $\mathbf{Q} \in \mathbb{R}^{m \times d_q}, \mathbf{K} \in \mathbb{R}^{m \times d_k}, \mathbf{V} \in \mathbb{R}^{m \times d_v}; \sqrt{d_k}$ 用来限制内积,防止内积值过大。编码过程是将 \mathbf{Q} 序列的每一个词向量与 \mathbf{K} 序列的每个词向量进行匹配得到相似度,然后与 \mathbf{V} 序列相乘,以更好地保留原始输入信息和表示整体信息。

通过门控卷积层对每个序列抽取语义特征和句法特征后,将每个序列依次单独输入注意力层。自注意力机制是指将式(2)中的 $\mathbf{Q}, \mathbf{K}, \mathbf{V}$ 3 个序列选取为相同序列,然后抽取出 Y_i^3 相对于自身更重要的全局信息。注意力如式(3)所示:

$$\mathbf{R}_i = \text{Attention}(Y_i^3 \mathbf{W}_s^Q, Y_i^3 \mathbf{W}_s^K, Y_i^3 \mathbf{W}_s^V) \quad (3)$$

其中, Y_i^3 为序列 i 经过第三层 GCNN 后的输出; $\mathbf{W}_s^Q, \mathbf{W}_s^K, \mathbf{W}_s^V \in \mathbb{R}^{\text{exd}}$; \mathbf{R}_i 为经过 Self-Attention 层抽取的特征向量。

经过注意力层抽取特征之后,每个输出序列 \mathbf{R}_i 中的各个词向量都包含了该序列自身上下文的信息。将 \mathbf{R}_i 通过 Max-Pooling 层,选取同一维度上最活跃的信息,得到更抽象的表示 T_i ,即句子序列或句法路径序列的特征表示。将句子序列的每个特征表示 T_{0i} 拼接起来,得到整篇文章综合的句子特征表示 $Sents$;将句法路径序列的每个特征表示 T_{1j} 拼接起来,得到整篇文章综合的句法路径的特征表示 $Paths$;最后再将二者连接起来,得到事件的语义信息和句法信息的篇章综合特征表示 f 。

$$Sents = \text{Concat}(T_{01}, T_{02}, \dots, T_{0i}) \quad (4)$$

$$Paths = \text{Concat}(T_{11}, T_{12}, \dots, T_{1j}) \quad (5)$$

$$f = \text{Concat}(Sents, Paths) \quad (6)$$

3.3 事件可信度识别层

根据 S_i 和 P_j 所包含的事件可信度信息,将句子中的词序列转换为词向量矩阵,再输入特征提取层,最终得到事件所包含的语义信息和句法信息的篇章综合特征表示 f 。将 f 输入一层前馈神经网络实现篇章级事件可信度识别,如式(7)所示:

$$y = \text{softmax}(\mathbf{W}_{s0} * f + b_{s0}) \quad (7)$$

其中, $\mathbf{W}_{s0} \in \mathbb{R}^{n \times 3}, b_{s0} \in \mathbb{R}^3$ 是 softmax 层参数矩阵。

3.4 对抗训练

为了提高模型的鲁棒性,本文对模型增加了小幅扰动,并采用 Fast Gradient Sign Method (FGSM)^[19]来计算扰动,将其添加到词向量上,得到扰动后的词向量。类似于前人的工作,本文在对抗训练中加入小幅扰动至词嵌入向量中,然后再次将其输入模型中得到扰动后的交叉熵损失 (Cross-Entropy Loss_{adv}),最后将扰动后的交叉熵损失与原有的交叉熵损失 (Cross-Entropy Loss) 相加得到新的 Loss,通过优化 Loss 来得到鲁棒性更好的模型。

4 实验

4.1 实验语料和设置

本文实验所用语料及其划分与文献^[4]一致,在中文和英

文语料库上都进行了10折交叉验证。在每折中,按8:1:1来划分训练集、开发集和测试集。使用准确率 P(Precision)、召回率 R(Recall)以及 F1 值(F1-measure)来描述其性能;同时,通过微平均(Micro-averaging)和宏平均(Macro-averaging)来描述本实验在各个分类上的综合性能。本文把 GCNN 的维度设置为 100,句法路径和句子中词语的词嵌入向量维度均为 100,并且通过 Word2Vec 进行初始化。

4.2 实验结果

在事件可信度识别任务中,前人的相关工作主要关注句子级事件可信度识别,只有 Qian^[4]进行过篇章级事件可信度识别研究。

在该模型所用语料库的中英文语料中,PS-和 Uu 的文

档分别占 1.22% 和 1.56%。因此,Qian^[4]主要考虑实验中 CT-、PS+ 和 CT+ 的性能。为了验证本文模型的性能,把 Qian^[4]的系统作为基准系统。表 2 列出实验中宏平均值和微平均值的各项性能指标对比,表 3 列出各类别的性能指标对比。

表 2 微平均和宏平均值结果

Table 2 Micro-average and macro-average results

| Language | System | Micro-ave | | | Macro-ave | | |
|----------|---------------------|-----------|-------|-------|-----------|-------|-------|
| | | P | R | F1 | P | R | F1 |
| Chinese | Qian ^[4] | 83.86 | 84.19 | 84.03 | 82.70 | 80.92 | 81.64 |
| | DEFI | 85.65 | 86.41 | 86.03 | 83.80 | 84.08 | 83.93 |
| English | Qian ^[4] | 83.56 | 83.56 | 83.56 | 77.62 | 75.76 | 76.28 |
| | DEFI | 85.71 | 87.03 | 86.37 | 82.13 | 80.70 | 80.74 |

表 3 各类别实验结果

Table 3 Experimental results for each category

| Language | System | CT- | | | PS+ | | | CT+ | | |
|----------|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | | P | R | F1 | P | R | F1 | P | R | F1 |
| Chinese | Qian | 83.68 | 83.17 | 83.35 | 79.05 | 69.74 | 74.06 | 85.36 | 89.86 | 87.52 |
| | DEFI | 86.59 | 84.20 | 85.38 | 76.49 | 77.12 | 76.81 | 88.32 | 90.93 | 89.60 |
| English | Qian | 77.21 | 77.78 | 76.87 | 66.89 | 58.46 | 62.14 | 88.76 | 91.04 | 89.84 |
| | DEFI | 79.75 | 82.45 | 80.28 | 77.33 | 66.41 | 70.76 | 89.32 | 93.24 | 91.19 |

从表 2 可以看出,在中文语料实验中,DEFI 在 3 个分类的宏平均 F1 值和微平均 F1 值上分别比基准系统提升了 2.29% 和 2.00%;在英文语料实验中,DEFI 在 3 个类别的宏平均 F1 值和微平均 F1 值上分别比基准系统提升了 4.46% 和 2.81%。这是由于本文模型不仅注意到事件词相邻词语,同时考虑了相距较远的词语。Qian^[4]的模型则比较重视较远距离词语的作用。但新闻文档往往会更直接地表达对事件的态度和判断,尤其是在英文文档中,直接报道事件的文档比例更高。因此,本文模型在英文语料方面的性能提升更大。

从表 3 可以看出,在中文语料的实验中,DEFI 在 CT-, PS+ 和 CT+ 3 个类别上的 F1 值分别提升了 2.03%, 2.75% 和 2.08%。在英文语料的实验中,CT-, PS+ 和 CT+ 的 F1 值分别提升了 4.41%, 8.62% 和 1.35%。从各个类别的结果对比中可以发现,DEFI 在 PS+ 类别上带来的提升最显著,这是由于对于 PS+ 这类不确定的事件来说,在事件句中表达其可信度之后,通常会在邻句中给出不同的看法和态度,过多关注邻句则容易被邻句信息所误导。综合上述性能对比来看,DEFI 明显优于 Qian^[4]所提模型。

DEFI 模型的性能优于 Qian^[4]的模型,其主要归因于以下几点。

DEFI 使用了三层 GCNN 网络,可抽取篇章中包含事件的句子的语义和句法信息,并且使模型在抽取局部特征的同时也能抽取全局特征,能够保留词序列的上下文依赖关系。Qian^[4]的模型用 BiLSTM 抽取句子的语义信息和句法信息,虽然考虑了上下文依赖关系,但是容易忽略相邻词语间的相关性,更加倾向于信赖句子两端的信息。如例 2 所示,此篇文章中事件“签署”的篇章级可信度为 CT-。DEFI 根据语义可判断出否定线索词“未”作用于事件“签署”,因此此句中事件可信度值应为 CT-, 与这篇文档的篇章级可信度一致。而 Qian^[4]的模型比较偏重于整个句子两端的信息。该例句

的后半句中多次出现不确定线索词,导致其将此句的可信度值误判为 PS+, 影响了整个篇章级事件可信度的识别。

例 2 对于媒体报道的与宝马合资建厂的传闻,长城汽车 13 日在港交所发布公告称,未(neg_cue)与宝马签署任何关于在国内成立合资公司的法律文件。截至本公告发布之日,本公司与宝马就 MINI 品牌汽车合作可行性仍在初步探讨(ps_cue)和评估(ps_cue)之中,双方是否(ps_cue)开展正式合作尚有重大不确定性(ps_cue),不会对本公司 2017 年的财务状况和经营业绩构成重大影响。

BiLSTM 网络的记忆容量有限,当输入的序列长度较大时,其存储信息的能力会下降。本文采用的中文语料或英文语料中包含邻句的事件句的序列变长,导致 Qian^[4]的模型在对这类文档进行特征抽取时会大大减小事件词相邻词语所占比重。而本模型通过多层的门控卷积,可以在重视局部特征的同时注意到全局特征,通过增加门控机制过滤掉无用的信息。

在模型训练速度上,DEFI 每训练一轮所用时间不足 Qian^[4]模型的 1/3,充分证明了本文模型对训练速率带来的提升。

4.3 错误分析

表 4 和表 5 分别为 DEFI 模型在篇章事件可信度识别任务的中英文语料库上错误识别的关系类型分布。从表中数据可以看出,在中文任务中有 10.7% 的 CT- 类样例和 15.7% 的 PS+ 类样例被误分类为 CT+ 类样例。在英文任务中有 15.8% 的 CT- 类样例和 26.2% 的 PS+ 类样例被误分类为 CT+ 类样例。其主要原因有以下两点:1) 在 Qian^[4]所标注的篇章级事件可信度识别中英文语料中,CT+ 类样例的比例占到了 1/2 以上;2) 一些句子中多次出现否定线索词如“不过”“搁置”“并未”等,以及不确定线索词如“研究”“提议”“讨论”等,并且语义结构复杂,模型难以准确识别哪些线索词作

用于事件“登月”(如例3所示),从而影响到对篇章级事件可信度值的判断。

表4 中文误分类样例所占比例

Table 4 Proportion of Chinese misclassified samples
(单位:%)

| Label | CT- | PS+ | CT+ |
|-------|-----|-----|------|
| CT- | — | 4.9 | 10.7 |
| PS+ | 7.1 | — | 15.7 |
| CT+ | 4.2 | 4.8 | — |

表5 英文误分类样例所占比例

Table 5 Proportion of English misclassified samples
(单位:%)

| Label | CT- | PS+ | CT+ |
|-------|-----|-----|------|
| CT- | — | 2.2 | 15.8 |
| PS+ | 7.3 | — | 26.2 |
| CT+ | 3.0 | 3.8 | — |

例3 不过,10月11日,日本文部科学省的小委员会的汇总报告称,日本宇宙航空研究开发机构(JAXA)提议的日本宇航员登月构想的相关讨论被搁置,且报告中并未提及如何具体推进宇航员的活动。

结束语 本文提出了一个基于门控卷积的篇章级事件可信度识别方法DEFI,通过使用层叠的门控卷积来抽取文章中事件的语义信息和句法信息,在抽取局部信息的同时也能通过层叠过程捕获到全局特征,然后通过自注意力层来获取每个序列相对于自身更重要的全局信息,最后将事件的语义信息和句法信息的特征表示结合起来,从而判别出该事件在篇章级别的事件可信度值。实验结果表明,本文模型DEFI在宏平均F1值和微平均F1值方面优于Qian^[4]提出的方法。在接下来的工作中,将着重考虑同一个事件在多篇文章中多次出现时,对该事件的可信度的综合判断。

参考文献

- [1] KLENNER M, CLEMATIDE S. How factuality determines sentiment inferences[C]// Meeting of the Joint Conference on Lexical and Computational Semantics. 2016:75-84.
- [2] QAZVINIAN V, ROSENGREN E, RADEV D R, et al. Rumor has it: Identifying Misinformation in Microblogs [C]// Meeting of the Empirical Methods in Natural Language Processing. 2011:1589-1599.
- [3] SAURÍ R, VERHAGEN M, PUSTEJOVSKY J. Annotating and recognizing event modality in text[C]// Meeting of the International FLAIRS. 2006:333-339.
- [4] QIAN Z. Research on Methods of Event Factuality Identification [D]. Suzhou: Soochow University, 2018.
- [5] SAURÍ R, PUSTEJOVSKY J. FactBank: a corpus annotated with event factuality[J]. Language Resources and Evaluation, 2009, 43(3): 227-268.
- [6] PUSTEJOVSKY J, PATRICK H, SAURÍ R, et al. The time-bank corpus[C]// Meeting of the Corpus Linguistics. 2003: 647-656.
- [7] SAURÍ R. A factuality profiler for eventualities in text[D]. Massachusetts: Brandeis University, 2008.
- [8] CAO Y, ZHU Q M, LI P F. The Construction of Chinese Event Factuality Corpus[J]. Journal of Chinese Information Processing, 2012, 27(6): 38-44.
- [9] VINCZE V, SZARVAS G, FARKAS R, et al. The bioscope corpus: biomedical texts annotated for uncertainty, negation and their scopes [J]. Bmc Bioinformatics, 2008, 9(S2): S9-S9.
- [10] ZOU B W, ZHU Q M, ZHOU G D. Research on natural language text oriented negation and uncertainty extraction [J]. Frontiers of Computer Science, 2016, 10(6): 1039-1051.
- [11] SAURÍ R, PUSTEJOVSKY J. Are you sure that this happened? Assessing the factuality degree of events in text[J]. Computational Linguistics, 2012, 38(2): 1-39.
- [12] WERNER G, VINODKUMAR P, MONA D, et al. Committed belief tagging on the factbank and lu corpora: A comparative study[C]// Meeting of the Second Workshop on Extra-Propositional Aspects of Meaning in Computational Semantics. 2015: 32-40.
- [13] QIAN Z, LI P F, ZHU Q M. A two-step approach for event factuality identification[C]// Meeting of the International Conference on Asian Language Processing. 2015: 103-106.
- [14] QIAN Z, LI P F, ZHANG Y, et al. Event Factuality Identification via Generative Adversarial Networks with Auxiliary Classification[C]// Meeting of the Joint Conference on Artificial Intelligence. 2018: 4293-4300.
- [15] HE T X, LI P F, ZHU Q M. Identifying Chinese event factuality with convolutional neural networks[C]// Meeting of the Chinese Lexical Semantic Workshop. 2017: 284-292.
- [16] ZHOU X J, WAN X J, XIAO J G. Attention-based lstm network for cross-lingual sentiment classification[C]// Meeting of the Empirical Methods in Natural Language Processing. 2016: 1650-1659.
- [17] DAUPHIN Y N, ANGELA F, MICHAEL A, et al. Language modeling with gated convolutional networks[C]// Meeting of the Machine Learning. 2017: 933-941.
- [18] PETER S, JAKOB U, ASHISH V. Self-Attention with relative position representations[J]. arXiv: 1803. 02155.
- [19] MIYATO T, MAEDA S, KOYAMA M, et al. Distributional smoothing with virtual adversarial training [J]. arXiv: 1507. 00677.



ZHANG Yun, born in 1993, postgraduate, is member of China Computer Federation. His main research interests include natural language processing.



LI Pei-feng, born in 1971, Ph.D, professor, Ph.D supervisor, is member of China Computer Federation. His main research interests include natural language processing, and machine learning.