

文章编号: 1003-0077 (2017) 00-0000-00

基于依存句法的交互注意力情感分析方法

钟建¹ 许文清² 陆启桐¹ 薛云¹ 胡晓晖¹

(1. 华南师范大学 物理与电信工程学院, 广东 广州 510006; 2. 华南师范大学 数学科学学院, 广东 广州 510631)

摘要: 近年来, 特定目标的情感分析受到了越来越多的关注, 其主要任务是判断句子在给定目标下的情感倾向。通常所给的目标是独立的词或者固定的词组, 所包含的信息量有限; 同时, 为了引入目标信息, 大部分的方法是将目标向量和句子中的词向量直接进行拼接, 但这种简单的处理方法不能充分地建立起两者之间的联系。因此该文提出了基于依存句法的交互注意力模型。首先通过对句子进行依存句法分析, 捕捉到和给定目标词存在直接关联的依存信息, 并根据依存关系类型筛选出含有情感信息的相关词, 再将其与原有的目标词整合成目标集合, 以表达更加完整的目标信息; 之后进一步将目标集合和句子分别进行建模得到初始特征表示, 然后设计交互注意力模型以考虑两者之间的整体相互影响, 从而充分地挖掘目标和句子的语义关系。在公开数据集 SemEval 2014 task4 上的实验表明, 我们的模型取得了比传统模型更好的效果, 验证了该文方法的有效性。

关键词: 特定目标; 情感分析; 依存关系; 注意力机制

中图分类号: TP391

文献标识码: A

A Dependency Parsing Based Interactive Attention Mechanism for Sentiment Analysis

ZHONG Jian¹, XU Wenqing², LU Qitong¹, XUE Yun¹, and HU Xiaohui¹

(1. School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou, Guangdong 510006, China; 2. School of Mathematical Sciences, South China Normal University, Guangzhou, Guangdong 510631, China)

Abstract: In recent years, increasing attention has been received for aspect based sentiment analysis whose main task is to identify the sentiment polarity of sentence based on the given target. In general, the given target is an isolated word or a fixed phrase, which contains finite information, and with the purpose of introducing the target information, most conventional approaches just concatenate the target embedding with word embedding of a sentence directly, but the relationship can't be fully established by those simple methods. In this paper, we firstly use dependency parsing to analyze the sentence, which could capture the directly relative information according to the given target word. Moreover, to express more complete information, we built a target word set combining the original target word with selected relevant words which convey sentiment information according to the dependency relationship. Then the target set and the sentence were modeled separately to obtain the initial feature representation and an interactive attention model was also designed to describe the mutual influence between them. Therefore, the semantic relationship between the target and the sentence is fully established. We carried out experiments on the SemEval task4 2014 dataset and the results demonstrate that our model is effective enough and prevail over the traditional models.

收稿日期: 2018-06-01; 定稿日期: 2018-06-?

基金项目: 基金名 (基金号); 基金名(基金号)

Key words: Aspect Based Sentiment Analysis; Sentiment Analysis; Dependency Relation; Attention Mechanism

0 引言

不同对象的情感倾向不一致

情感分析指应用自然语言处理技术将文本划分为积极倾向和消极倾向,近年来被广泛应用于电商评论分析、信息检索、社交网络分析等领域^[1-2]。早期的情感分析主要是为了得到句子的整体情感倾向,却忽略了句子中的评价目标及其情感倾向。而**特定目标的情感分析**作为一个细粒度情感分析任务,目的是要判断句子在特定目标下的情感倾向^[3-4],例如: I love the cell phone color, but the screen is too small. 通过分析可知对于给定的目标 cell phone color, 句子的情感倾向为积极; 对于给定的目标 screen, 句子的情感倾向为消极。

在特定目标的情感分析领域中, **目标词**对句子的最终情感分类具有非常重要的意义。传统方法主要通过基于情感词典和制定语言规则对目标和句子之间的关系进行建模,从而获得大量相关的特征,最终输入到分类器中得到情感分类结果^[5]。Jiang^[6]通过基于**语言规则**的方法,获得了特定目标在句子中的句法特征,同时人工制定多种规则以得到句子的上下文特征,然后将两者进行融合形成了更加全面的特征表示,最终输入到 SVM 分类器中进行预测,在 Twitter 数据集上获得了较高的准确率; Mitchell^[7]则将特定目标的情感分析视为**序列标注任务**,首先提取大量和目标词有关的人工特征,其中包括丰富的情感词典特征和统计特征,然后利用条件随机场模型(Conditional Random Field)进行预测以得到最终结果。虽然这些方法都考虑了目标对情感分类的重要性,但是它们都**高度依赖于所获得的特征质量**,并且**特征工程**的工作量也十分庞大。

为了克服传统特征工程的不足,学者们开始研究基于深度学习^[8-9]的情感分析方法,因为**神经网络可以高效地从目标和句子中自动提取特征**,并且拥有更加丰富的语义信息。因此,基于深度学习的方法逐渐成为目前解决特定目标情感分析任务的主流模型。Tang^[10]以**目标词为中心点**,将句子划分为左右两部分,并用两个独立的 LSTM 网络分别对这两部分的句子进行建模,从而能够充分地利用目标的语义信息; Huang^[11]则将**目标词和句子中每一个词语分别进行拼接**后输入

LSTM 网络,获得隐藏输出后,再使用注意力机制考虑给定目标对句子的影响; Wang^[12]则在 Huang 的基础上引入了依存信息,捕捉到句子中每个词语的长距离依存信息,并将其与隐藏层输出直接进行拼接,从而增加了目标特征信息; Ma^[13]则通过 LSTM 网络对目标词和句子分开进行建模,再利用各自隐藏层的输出进行注意力权重的计算。

尽管这些方法能够较好地解决特定目标的情感分析任务,但是它们仍然面临着两个挑战:一方面,任务中所给定的目标只是独立的词或者固定的词组,忽略了目标词与上下文之间的依存关系,从而使得目标包含的信息量变得很有限。另一方面,大部分基于拼接的简单方式不能充分建立起目标与上下文之间的关系。因此,受 Wang^[12]和 Ma^[13]的启发,我们提出**基于依存句法的交互注意力模型(Dependency Interactive Attention model, DIA)**来解决上述两个问题。我们的主要贡献如下:

(1) 本文对给定的句子进行**依存句法分析**,捕捉到和目标词直接关联的**长距离依存信息**,进一步根据依存关系类型**筛选出含有情感信息的相关项**,并将其与目标词整合成**目标集合**,以形成具有情感信息的目标集合。

(2) 对目标集合和句子分别进行**独立建模**以获取各自的初始特征表达,进一步设计**交互注意力模型**来考虑目标集合和句子之间的相互影响,从而使得最终的特征表达充分考虑了目标和句子的紧密联系。

1 相关工作

早期的情感分析任务主要是判断给定文本的整体情感倾向^[14]。但是在实际情况中,针对文本中的不同分析对象,往往会有不同的情感倾向,这种细粒度的分析能够为决策者提供更加全面的参考信息。因此作为一个细粒度的分析任务,特定目标的情感分析^[15]成为了当前 NLP 领域的研究热点。

1.1 基于依存句法的特定目标情感分析

依存句法分析是为了表示句子中词语之间

的依存关系^[16-17]。图 1 为一个标准的依存句法分析结构示例, 其中树的节点代表句子中的词语, 有向弧边代表两个词语之间句法关系, 弧边的起点对应的词为支配词, 弧边箭头对应的词为被支配词。

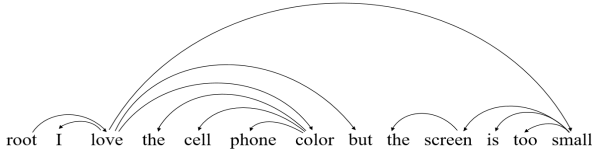


图 1 依存句法分析结构示例图

Jiang^[6]考虑到 Twitter 文本中字数少、歧义多的特点, 所以在解决此类特定目标的情感分析任务中, 引入了依存句法分析来获得更加丰富的句法特征, 从而扩充了目标信息; Wang^[13]则认为利用依存句法分析能够将语义相近但是间隔较远的词联系起来, 从而获得句子中每一个词的长距离依存信息以提升分类性能。

1.2 基于深度学习的特定目标情感分析

深度学习模型能够利用分布式表示来自动学习得到目标的相关特征, 近年来在特定目标的情感分析任务中取得了巨大的成功。Dong^[18]为了将目标信息整合到递归神经网络中, 提出一种自适应的深度学习方法, 能够提取句子中与目标相关的文本信息和句法结构特征; Ma 和 Peng^[19]则认为引入外部信息有助于学习外部信息中与目标词相关的特征, 因此将常识信息嵌入到传统的深度神经网络当中, 从而更好地预测特定目标下的句子情感倾向。

1.3 基于深度学习的特定目标情感分析

注意力机制^[20]通过模拟人脑, 能够对重要的信息给予更多的关注, 目前已经广泛应用于各大研究领域。Bahdanau^[21]首次在自然语言处理领域中引入注意力机制, 在机器翻译任务中可以自动地对句子中相关部分赋予较大权重, 使模型的翻译正确率得到显著的提升; 而在特定目标的情感分析领域中, 注意力机制可以关注句子中与给定目标存在关联的部分, Huang^[11]和 Ma^[13]在特定目标的情感分析任务中, 通过设计注意力机制模型找到特定目标下句子中关键部分, 使模型训练时可以更加关注这些重点部分, 以提升分类性能。

2 本文方法

本节主要介绍本文所提出的 DIA 模型, 整体

框架如图 2 所示。

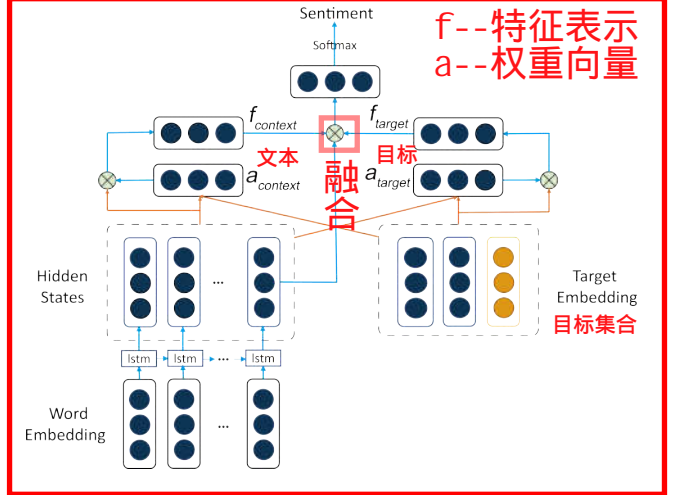


图 2 DIA 模型结构示例图

首先, 因为长短期记忆网络 (Long Short-Term Memory)^[22]能够有效地捕捉到句子中的语义特征, 非常适合解决序列型的问题, 所以本文将句子中的每个词依次输入到 LSTM 模型, 将得到的隐藏层输出作为句子的初始语义表达; 同时, 因为依存句法分析能够将语义相关但间隔较远的词语直接联系起来, 所以我们通过对句子进行依存句法分析, 找到目标词的依存信息, 并根据依存关系类型筛选出含有情感信息的相关项, 然后与原始目标词融合为目标集合, 形成目标的初始特征表示; 最终, 本文对目标集合与句子分别进行训练, 设计了交互注意力模型, 不仅使模型能够关注句子中与目标紧密相关的部分, 而且还考虑了彼此之间的整体相互影响, 从而充分建立起两者之间的联系。

2.1 LSTM 网络

普通的循环神经网络 (Recurrent Neural Network) 在对句子进行编码的时候存在梯度消失或者爆炸的情况^[23]。为了克服这个问题, Hochreiter^[22]等人提出了长短时记忆网络, 能够更好地学习句子中的长距离语义关系。LSTM 网络中分别设计了输入门、遗忘门、输出门来控制句子信息的流入和流出。具体的数学表达公式如下:

$$\text{输入 } i_t = \sigma(W_i \cdot [h_{t-1}; w_t] + b_i) \quad (1)$$

$$\text{遗忘 } f_t = \sigma(W_f \cdot [h_{t-1}; w_t] + b_f) \quad (2)$$

$$\text{输出 } o_t = \sigma(W_o \cdot [h_{t-1}; w_t] + b_o) \quad (3)$$

$$g_t = \tanh(W_r \cdot [h_{t-1}; w_t] + b_r) \quad (4)$$

$$\text{状态 } c_t = i_t \odot g_t + f_t \odot c_{t-1} \quad (5)$$

$$\text{隐藏层输出 } h_t = o_t \odot \tanh(c_t) \quad (6)$$

其中, i_t 、 f_t 、 o_t 分别为输入门、遗忘门和输出门; c_t 和 h_t 分别为当前细胞状态和隐藏层输出; W_i 、 W_f 、 W_o 为权重矩阵, b_i 、 b_f 、 b_o 为偏置项; σ 为 sigmoid 激活函数, \odot 为点乘操作。在本文模型中, 每一个句子都被映射成 LSTM 网络中的隐藏输出表示 $H=[h_1, h_2, \dots, h_n]$ 。

因此本文采用 LSTM 模型对句子进行建模, 将所得到的隐藏输出作为句子的初始特征表示, 为后续计算交互注意力权重奠定了基础。

2.2 基于依存句法分析的目标集合

传统特定目标的情感分析任务中所给定的目标通常是由单个词或者固定的词组构成, 所包含信息十分有限。因此本文希望能够从句子中挖掘出和给定目标相关的依存信息, 并从中筛选出具有情感的相关信息, 然后与原始目标进行融合, 从而得到包含情感信息的目标集合。

我们采用了 Stanford Parser^[24] 工具包对句子进行分析, 在相应句子中找出所有和目标词存在依存关系的词, 得到目标的长距离依存信息。但考虑到并非所有依存关系都对情感分类有重要贡献, 部分依存关系甚至会最终分类造成一定的干扰。所以我们制定了相应的语言规则对目标词的依存关系进行了筛选, 见表 1。

表 1 语言规则示例

依存关系	解释	举例
nsubj	名词性主语	(small, screen)
dobj	直接宾语	(love, color)
advmod	副词性修饰关系	(small, too)
neg	否定修饰关系	(like, don't)
conj	并列关系	(beautiful, clear)

得到扩充目标集合后, 我们通过语言模型将目标集合中的每一个词转化为词向量表示, 最终得到目标集合对应的词向量矩阵 $T=[t_1, t_2, \dots, t_n]$, 我们直接将其作为目标的特征表示以用于后续的交互注意力计算。

2.3 融合目标与句子内容的交互注意力机制

在解决特定目标的情感分析任务时, 传统的注意力模型中虽然考虑了特定目标对于句子不同成分的影响, 但是忽略了句子对特定目标的影响, 因此不能充分完整地捕捉到句子和目标集

当中的全部重要信息。为了克服这个问题, 本文基于传统注意力机制提出了交互注意力计算方式, 通过让目标和句子交互影响彼此的向量表示训练过程, 从而完整捕捉到目标和句子中的情感信息。

首先, 本文将句子中所有词对应的词向量输入到 LSTM 网络, 得到隐藏层输出 $H=[h_1, h_2, \dots, h_n]$, $H \in \mathbb{R}^{d_D \times N}$, 其中 d_D 为隐藏层的大小, N 为给定句子的长度。然后, 将目标词经过依存句法拓展后得到目标集合, 再将其中元素分别转换成词向量后得到目标集合矩阵表示

$T=[t_1, t_2, \dots, t_M]$, $T \in \mathbb{R}^{d_w \times M}$, 其中 d_w 为目标词向量的长度, M 为给定目标词拓展后的集合大小。

对于句子, 我们将目标集合的全部信息用于训练句子的特征表示, 以更充分地捕捉到句子中与目标相关的信息。通过式子 (7) 我们可以得到句子的最终特征表示:

$$\text{特征表示 } f_{\text{context}} = a_{\text{context}} H^T \quad (7)$$

$$\text{权重向量 } a_{\text{context}} = \text{softmax}(H \cdot \tanh(W_1 T W_2)) \quad (8)$$

其中, 句子对应的注意力权重向量 $a_{\text{context}} \in \mathbb{R}^N$, 代表分配给句子中每一个词权重的大小。 $W_1 \in \mathbb{R}^{N \times M}$, $W_2 \in \mathbb{R}^{d_w \times d_w}$ 均为注意力公式中的参数矩阵。

对于目标集合, 我们也同样基于注意力机制, 将句子输入 LSTM 模型得到的隐藏层信息 H 用于目标集合的注意力权重计算, 以充分利用句子与目标集合的相互关系, 得到目标集合更好的表示形式。通过式子 (9) 我们得到目标的最终特征表示:

$$f_{\text{target}} = a_{\text{target}} T^T \quad (9)$$

$$a_{\text{target}} = \text{softmax}(T \cdot \tanh(W_3 H W_4)) \quad (10)$$

其中, 目标词的注意力权重向量 $a_{\text{target}} \in \mathbb{R}^M$, 代表分配给目标集合中每一个词权重的大小。 $W_3 \in \mathbb{R}^{M \times N}$, $W_4 \in \mathbb{R}^{d_D \times d_D}$ 均为注意力公式中的参数矩阵。

为了形成最终的特征表示, 本文受到 Huang^[11] 的启发, 将上述得到的句子和目标的向量表示进行融合, 同时也考虑隐层中最后一个输出 h_N , 因为 h_N 包含句子最丰富的语义信息。具体计算方式如下: 最终特征表示

$$s = f_{\text{context}} + f_{\text{target}} + h_N W_5 \quad (11)$$

为了得到句子的情感分类结果, 我们将最终的特征表示 s 输入到 softmax 层, 从而计算出该句子在不同类别上的概率分布:

attention 其实用于权重计算的

$$\hat{y} = \text{softmax}(W_s s + b_s) \quad (12)$$

其中, W_s 和 b_s 为对应的权重矩阵和偏置项。概率最大的情感类别将作为最终的分类结果。

3 模型训练

本文采用了所有句子分类结果的交叉熵总和作为损失函数, 同时采用反向传播算法来对权重矩阵和参数进行更新。

$$\text{loss} = -\sum_{i \in S} \sum_{j=1}^C y_i^j \log \hat{y}_i^j + \lambda \|\theta\|^2 \quad (13)$$

其中, i 代表一个句子, j 代表一种情感倾向类别; y 为句子情感倾向的真实分布, \hat{y} 则为预测分布; S 代表训练数据, C 则代表所有的情感倾向类别; λ 和 θ 都为 L_2 正则项的参数。

4 实验

4.1 参数设置

本文实验中采用预训练好的 Glove¹ 词向量^[25]对句子中的词和目标词进行初始化, 词向量维度为 300; 权重参数初始值由均匀分布 $U(-0.01, 0.01)$ 产生; 隐藏层大小设置为 300; 采用 AdaGard 作为模型的优化器, 对应的学习率设置为 0.01; 同时为了防止过拟合, L_2 正则项的权重设置为 0.001。我们采用准确率作为评价指标, 用于评估本文模型和 baseline 模型的性能。准确率的具体计算方式如下:

$$\text{准确率} = \frac{\text{预测正确的样本数}}{\text{所有样本总数}} \quad (14)$$

4.2 数据集介绍

本文在 SemEval 2014 Task4^[26] 中的笔记本电脑评论数据和餐厅评论数据中分别进行了模型验证。其中, 每个评论由句子和目标共同组成, 句子的情感极性包含了积极、消极、中性三个类别, 关于数据集的统计描述详见表 2。本文模型目的就是要将给定目标下的句子划分到正确的类别中。

表 2 实验数据集

数据集	积极	消极	中性	共计
LPT-TR	987	866	460	2313
LPT-TE	341	128	169	638
RST-TR	2164	805	633	3602
RST-TE	728	196	196	1120

其中, LPT-TR 为笔记本电脑评论数据的训练集, LPT-TE 为测试集; RST-TR 为餐厅评论数据的训练集, RST-TE 为测试集。

4.3 实验结果及分析

为了更加全面的评估本文模型的性能, 我们选择了 4 个 baseline 模型进行比较。

LSTM: 标准的单层 LSTM 网络, 进行分类的时候没有利用目标的信息。

TD-LSTM: 以目标为中心, 分别从左右两个方向采用 LSTM 网络进行建模, 从而能够得到目标的上下文信息。

ATAE-LSTM: 在输入层中将每个词向量和目标词向量进行拼接, 从而形成上下文输入, 同时在隐藏层中将 LSTM 网络的输出和目标词向量进行拼接, 然后使用注意力机制得到最终的特征表示。

DAT-LSTM: 在 ATAE-LSTM 模型的基础上, 引入了句子词语之间的依存关系, 以获得词语之间的长距离依存信息。

同时我们也给出了本文模型 (DIA) 的实验结果。

表 3 对比实验结果

方法	Res	Lap
LSTM	74.3	66.5
TD-LSTM	75.6	68.1
ATAE-LSTM	77.2	68.7
DAT-LSTM	78.1	69.4
DIA	78.5	71.9

4.3.1 结果分析

实验结果如表 3 所示, 标准的单层 LSTM 网络在两个数据集上的准确率只有 74.3% 和 66.5%, 均为最低值。主要因为标准单层 LSTM 网络忽略了目标信息, 无法根据不同的目标判断句子对应的情感倾向。相比之下, 考虑了目标信息的 4 个模型平均准确率分别为 77.4% 和 69.5%, 领先标准单层 LSTM 模型 3.1% 和 3.0%, 表明了目标信息对提高分类精度起到了重要的作用。ATAE-LSTM

模型则在考虑目标信息的基础上引入了注意力机制,在两个数据集上分别实现了 77.2%和 68.7% 的准确率,比 TD-LSTM 模型高 1.6%和 0.6%,主要是因为 TD-LSTM 仅考虑了目标信息,同等地对待每一个词在最终结果中起到的作用,不能识别出句子中具有重要信息的词语。DAT-LSTM 模型在注意力机制上进一步引入了依存分析,并在餐厅评论数据和笔记本电脑评论上都取得了不错的分类效果,相比 ATAE-LSTM 分类准确率提升了 0.9%和 0.7%。

本文提出的 DIA 模型取得了当前实验中的最好结果,一方面是因为本模型扩展了目标信息,使得最终的目标集合包含了丰富的情感信息。同时另一方面通过引入了交互注意力机制,充分挖掘出目标集合和句子之间的语义关系。最终的实验结果证实了模型的有效性。

4.3.2 案例研究

我们分别选择了“The ambience was nice, but service was not so great.”和“Service was slow, but the people were friendly.”作为研究例子。对于第一句话中的给定目标“ambience”,句子的情感标签为积极的;而对于第二句话中的给定目标“service”,句子的情感标签则为消极的。

首先对句子进行依存句法分析,以获得目标词“ambience”和“service”的直接依存信息,并根据依存关系进行筛选,可以分别形成两个目标集合{ambience, nice}和{service, slow}。在此基础上使用 DIA 模型分别对句子和给定目标进行建模。为了更加直观地呈现结果,我们将句子和目标的注意力权重分布进行了可视化,结果如图 3 所示。



图 3 案例研究可视化图

图 3 中的第一个例句显示,句子中的“nice”和“ambience”词所分到的注意力权重较大,而其他词的注意权重则较小。在第二个例子中则更为明显,“slow”的注意力权重远远高于句子中其他词的权重值。这表明 DIA 模型在进行情感分类时能够重点关注句子中和目标相关的信息,而

其他相关程度较低的词所分配的权重则较小。同时,扩充后的目标集合中富含情感信息的词(“nice”和“slow”)也受到了较大的关注,符合人类的实际认知。以上结果进一步验证了 DIA 模型的有效性。

5 结语

在传统特定目标的情感分析任务中,给定目标所包含信息量十分有限。针对这个问题,本文对句子进行了依存句法分析,根据依存关系类型筛选出含有情感信息的依存项,然后将其与目标词一起扩充为目标集合,从而将情感信息引入了目标集合。同时,在引入目标信息的过程中,大部分基于拼接的方法不能充分地建立起目标和句子间联系。针对这个问题,本文对目标集合和句子分别进行独立建模以获取各自的初始特征表达,再进一步设计交互注意力模型以考虑目标集合和句子之间的相互影响,从而使得最终的特征表达能充分挖掘目标和句子之间的深层次语义关系。实验结果证明了本文提出的模型在解决特定目标情感分析任务中的有效性,取得了比传统方法更优的结果。

在特定目标的情感分析任务中,未来还有很多研究工作需要开展。首先,将目标信息有效地嵌入到神经网络单元中是一个比较可行的研究方向。其次,利用不同句子之间的交互信息训练模型以提升情感分类性能也会是我们未来的主要研究重点。

参考文献

- [1] PANG B, LEE L, VAITHYANATHAN S. Thumbs up? sentiment classification using machine learning techniques[J]. Proceedings of Emnlp, 2002:79-86.
- [2] QIAN Q, HUANG M, Lei J, et al. Linguistically regularized LSTM for sentiment classification[C]// ACL 2017: Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2017:1679-1689.
- [3] GAO Y, ZHANG Y, XIAO T. Implicit Syntactic Features for Target-dependent Sentiment Analysis[C]//Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers). 2017, 1: 516-524.
- [4] BARBOSA L, FENG J. Robust sentiment detection on Twitter from biased and noisy data[C]// COLING 2010: International Conference on Computational Linguistics: Posters. Association for Computational Linguistics, 2010:36-44.
- [5] VO D T, ZHANG Y. Target-dependent twitter sentiment classification with rich automatic features[C]// ICAI 2015: International Conference on Artificial Intelligence.

- AAAI Press, 2015:1347-1353.
- [6] JIANG, LONG, YU, et al. Target-dependent Twitter sentiment classification[C]// ACL 2011: Meeting of the Association for Computational Linguistics: Human Language Technologies. Association for Computational Linguistics. Stroudsburg, PA: ACL, 2011:151-160.
- [7] MITCHELL M, AGUILAR J, WILSON T, et al. Open domain targeted sentiment[C]//Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. 2013: 1643-1654.
- [8] BRADBURY J, MERITY S, XIONG C, et al. Quasi-Recurrent Neural Networks[J]. 2016.
- [9] TANG D, QIN B, LIU T. Aspect Level Sentiment Classification with Deep Memory Network[J]. 2016:214-224
- [10] TANG D, QIN B, FENG X, et al. Effective LSTMs for Target-Dependent Sentiment Classification[J]. Computer Science, 2015.
- [11] WANG Y, HUANG M, ZHU X, et al. Attention-based LSTM for Aspect-level Sentiment Classification[C]// EMNLP 2016: Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2016:606-615.
- [12] WANG X, CHEN G. Dependency-Attention-Based LSTM for Target-Dependent Sentiment Analysis[C]// SMP 2017: Chinese National Conference on Social Media Processing. Springer, Singapore, 2017:206-217.
- [13] MA D, Li S, ZHANG X, et al. Interactive Attention Networks for Aspect-Level Sentiment Classification[C]// IJCAI 2017: Twenty-Sixth International Joint Conference on Artificial Intelligence. 2017:4068-4074.
- [14] NASUKAWA T, YI J. Sentiment analysis: capturing favorability using natural language processing[C]// K-CAP 2003: International Conference on Knowledge Capture. DBLP, 2003:70-77.
- [15] LIU J, ZHANG Y. Attention modeling for targeted sentiment[C]//Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers. 2017, 2: 572-577.
- [16] CHEN W, ZHANG Y, ZHANG M. Feature embedding for dependency parsing[C]//Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers. 2014: 816-826.
- [17] MUKHERJEE S, BHATTACHARYYA P. Feature Specific Sentiment Analysis for Product Reviews[J]. 2012, 7181:475-487.
- [18] DONG L, WEI F, TAN C, et al. Adaptive Recursive Neural Network for Target-dependent Twitter Sentiment Classification[C]//ACL 2014: Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2014:49-54.
- [19] MA Y, PENG H, CAMBRIA E. Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM[C]//AAAI. 2018.
- [20] MNIH V, HEES N, GRAVES A, et al. Recurrent models of visual attention[J]. 2014, 3:2204-2212.
- [21] BAHDANAU D, CHO K, BENGIO Y. Neural Machine Translation by Jointly Learning to Align and Translate[J]. Computer Science, 2014.
- [22] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.
- [23] GRAVES A. Supervised Sequence Labelling with Recurrent Neural Networks[J]. Studies in Computational Intelligence, 2008, 385.
- [24] SCHUSTER S, MANNING C. Enhanced English Uni-

versal Dependencies: An Improved Representation for Natural Language Understanding Tasks[C]// Tenth International Conference on Language Resources and Evaluation. 2016.

- [25] PENNINGTON J, SOCHER R, MANNING C. Glove: Global Vectors for Word Representation[C]// EMNLP 2014: Conference on Empirical Methods in Natural Language Processing. 2014:1532-1543.

- [26] PONTIKI M, GALANIS D, PAVLOPOULOS J, et al. SemEval-2014 Task 4: Aspect Based Sentiment Analysis[J]. Proceedings of International Workshop on Semantic Evaluation at, 2014:27-35.



钟建 (1994—), 硕士研究生, 主要研究领域为自然语言处理、深度学习、情感分析。
E-mail: zhongjian.nlp@foxmail.com



许文清 (1995—), 硕士研究生, 主要研究领域为自然语言处理、数据挖掘、文本情感分析。
E-mail: manchingmark@outlook.com



薛云 (1975—), 通信作者, 博士, 教授, 主要研究领域为自然语言处理、文本挖掘、情感分析。
E-mail: xueyun@m.scnu.edu.cn