

# 基于注意力机制和 BiLSTM-CRF 的消极情绪意见目标抽取



丁锋 孙晓

合肥工业大学计算机与信息学院 合肥 230601

合肥工业大学情感计算与先进智能机器安徽省重点实验室 合肥 230601

(fengdf10@163.com)

**摘要** 基于方面情感分析 (Aspect-Based Sentiment Analysis, ABSA) 是自然语言处理的热门课题,其中意见目标抽取和意见目标情感极性分类是 ABSA 的基本子任务之一。而很少有研究直接抽取特定情感极性的意见目标,尤其是抽取更有潜在价值的消极情绪意见目标。文中提出了一种全新的 ABSA 子任务—抽取消极情绪意见目标 (Negative-Emotion Opinion Target Extraction, NE-OTE),并提出了基于注意力机制和单词与字符混合嵌入的 BiLSTM-CRF 模型 (Attention-based BiLSTM-CRF with Word Embedding and Character Embedding, AB-CE),在双向长短期记忆网络 (Bi-directional Long Short-Term Memory, BiLSTM) 学习文本语义信息和捕获长距离双向语义依赖关系的基础上,通过注意力机制使模型更好地关注输入序列中的关键部分和捕获与意见目标及其情感倾向相关的隐含特征,最终通过 CRF 层预测句子级别的全局最佳标签序列,实现对消极情绪意见目标的抽取。文中基于主流 ABSA 任务基准数据集构建了 3 个 NE-OTE 任务数据集,并在这些数据集上进行了广泛的实验,实验结果显示,所提模型能够有效识别消极情绪意见目标,且识别效果明显优于其他基线模型,验证了所提方法的有效性。

**关键词:** 情感分析;消极情绪意见目标;注意力机制;双向长短期记忆网络;条件随机场

**中图法分类号** TP391

## Negative-emotion Opinion Target Extraction Based on Attention and BiLSTM-CRF

DING Feng and SUN Xiao

School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230601, China

Anhui Province Key Laboratory of Affective Computing and Advanced Intelligent Machine, Hefei University of Technology, Hefei 230601, China

**Abstract** Aspect-based sentiment analysis (ABSA) is a popular topic for natural language processing, in which opinion target extraction and sentiment polarity classification of opinion target are one of the basic subtasks of ABSA. However, few studies directly extract the opinion targets of specific emotional polarity, especially the negative emotion opinion targets with more potential value. A new ABSA subtask—negative emotion opinion target extraction (NE-OTE) is proposed, and a BiLSTM-CRF model based on attention mechanism and character and word mixture embedding (AB-CE) is proposed. On the basis of bi-directional long short-term memory (BiLSTM) learning textual semantic information and capturing long distance bi-directional semantic dependency, through the attention mechanism, the model can better pay attention to the key parts in the input sequence and capture the implied characteristics related to the opinion target and its emotional tendency. Finally, the CRF layer can be used to predict the optimal tag sequence at the sentence level, so as to extract the negative emotional opinion target. This paper builds three NE-OTE task datasets based on the mainstream ABSA task baseline datasets and conducts extensive experiments on these datasets. Experimental results show that the model proposed in this paper can effectively identify the target of negative emotional opinions, and is significantly better than other baseline models, which verifies the effectiveness of the method proposed in this paper.

**Keywords** Sentiment analysis, Negative emotion opinion target, Attention, BiLSTM, CRF

## 1 引言

文本情感分析<sup>[1]</sup>又称意见挖掘<sup>[2]</sup>,可以基于文本挖掘出人们对服务、事件、产品等实体对象的态度、情绪、意见等主观感受,是近年来备受关注的自然语言处理任务。按照文

本的粒度,文本情感分析可以划分为 3 种级别:篇章级别情感分析、句子级别情感分析和基于方面情感分析。例如,对电商评论进行“消极”“积极”“中立”的情感倾向的 3 种分类就属于句子级别的情感分析。而更加细粒度的基于方面情感分析 (ABSA) 可以识别文本中特定目标或方面的情感倾向。

收稿日期:2021-01-06 返修日期:2021-05-25 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(61976078)

This work was supported by the National Natural Science Foundation of China(61976078).

通信作者:孙晓(sunx@hfut.edu.cn)

例如,挖掘出用户在某家餐厅的评论“The sandwich was delicious, but the table was a bit dirty.”中用户对“sandwich”和“table”两个意见目标及其对应的餐厅口味、环境两个方面分别表达了积极和消极的情绪,而篇章/句子级文本情感分析并不能挖掘出用户评论中方面级别的用户情感。近年来,基于方面情感分析在电商领域、舆情分析等领域应用广泛,随着神经网络模型的发展,ABSA任务也不断取得新进展。Wang等<sup>[3]</sup>提出了基于注意力机制和LSTM的ATAE-LSTM模型,用于方面级别的情感分类,使方面信息在注意力权重的计算中发挥作用,当涉及不同方面时,该模型能够关注句子的不同部分。相比ATAE-LSTM中简单地连接上下文的单词信息和方面信息,Tay等<sup>[4]</sup>提出了一种融合方面信息的注意力学习方法,通过循环卷积进行更有效的方面-词融合。随着预训练模型和微调思想的发展,预训练模型和微调思想在ABSA任务中也取得了进展,Sun等<sup>[5]</sup>提出了一种新的ABSA解决方案,通过构造辅助句子,并将ABSA任务转换成句子对分类任务,最终通过微调预训练模型BERT来完成当前的句子对分类。Li等<sup>[6]</sup>探索了将BERT嵌入分量与各种神经网络模型相结合,验证了基于BERT的模型捕捉基于方面的情感具有优越性。

ABSA任务包含方面类别检测、意见目标抽取(Opinion Target Extraction, OTE)、意见目标/方面情感极性分类(Opinion Target Sentiment Classification, OTSC)等多个子任务。其中,意见目标抽取和意见目标情感极性分类是两个重要的ABSA子任务。意见目标指评论中表示用户态度的特征或实体的词,意见目标情感极性指评论中用户对特定意见目标的情感倾向,包括消极、积极和中性。近年来,在这两个子任务上已经开展了大量基于神经网络的工作<sup>[7-9]</sup>,并取得了很大的进展。但是目前仍没有同时抽取意见目标及其情感极性的研究,以上述的用户评论为例,〈table:消极〉,〈sandwich:积极〉是两对〈意见目标:意见目标的情感极性〉搭配,这种搭配模式本身有着潜在的价值,而且对ABSA任务具有重要的意义。

典型的ABSA任务可以帮助分析文本中具体方面的情绪极性,但不能面向特定的情绪极性去识别意见目标,尤其是被消极评价的意见目标,被消极评价的意见目标往往有着很高的潜在价值,例如,它表达了用户的深层消费情绪,有助于商家更好地了解用户的真实消费体验,发现自身服务的不足,辅助商家提升自身服务质量。同时,NE-OTE任务也有助于提高ABSA任务的解释性。

为此,本文提出了一个全新的ABSA子任务——抽取消极情绪的意见目标(NE-OTE)。在NE-OTE任务中,给定用户评论文本,任务的目的是抽取用户评论中被消极评价的意见目标——消极情绪意见目标(Negative-Emotion Opinion Target, NE-OT)。以上述的用户评论为例,用户评论中共有“table”和“sandwich”两个意见目标,但NE-OTE任务最终只会抽取“table”这个意见目标,而不会抽取“sandwich”,这是因为“table”被消极评价,满足消极情绪意见目标的条件,而“sandwich”被积极评价,不满足条件,因此不需要被抽取。更多的NE-OTE任务案例如图1所示。

#### Reviews:

- The **hostess** was rude and I got a distinct feeling that they did not want to serve us.
- As BFC doesn't take reservations you almost always have to wait by the bar and be abused by the **front** of house staff until you are seated, which can be over an hour later!
- bad **staff**.
- I took one look at the **chicken** and I was appalled.
- I must say I am surprised by the bad reviews of the **restaurant** earlier in the year, though.

注:加粗字体代表消极情绪意见目标

图1 NE-OTE任务案例

Fig. 1 NE-OTE task case

消极情绪意见目标抽取的难点如下。

#### (1) 在线评论的文本质量较差

在线评论的文本质量一般较差,经常包含拼写错误的单词、新奇的单词、罕见的单词,未登录词等。因此,我们假设在输入序列的单词表示中加入字符级特征是有用的,字符级特征有助于提升模型对拼写错误情况的健壮性和对未登录词的泛化能力。实验结果表明,字符级特征对NE-OTE任务是有效的,其中F1值在3个数据集上分别提升了4.59%,4.45%,3.14%。

#### (2) 捕获在线评论中双向长距离的语义信息比较困难

对于NE-OTE任务来说,在线评论的文本质量较差,同时在线评论中意见目标与包含意见目标情感倾向信息的意见词的相对位置通常没有规律,但是模型捕获在线评论中双向长距离的语义信息的能力对NE-OTE任务来说十分重要,我们基于神经网络在自然语言处理方面的成功应用,提出使用循环神经网络(Recurrent Neural Network, RNN)来负责捕获双向长距离的语义信息。循环神经网络的结构既可以接收不定长的输入序列,还可以进行从前到后的线性传播,天然适配解决自然语言处理的问题,尤其是循环神经网络的变种之一——双向长短时记忆网络(BiLSTM)进一步改善了RNN中梯度弥散/爆炸和捕获双向语义信息等问题,大大提升了双向长距离上下文语义信息的捕获能力。

#### (3) 如何捕获意见目标的情感倾向信息

对于NE-OTE任务,仅仅捕获意见目标的信息是不够的,意见目标的情感倾向的信息也同样重要。本文提出利用注意力机制辅助模型来捕获输入序列中意见目标及其情感倾向的关键信息。注意力机制已被证明是一种有效的机制,并在很多领域取得了显著效果,如图像识别<sup>[10]</sup>、机器翻译<sup>[11]</sup>和文本摘要<sup>[12]</sup>。针对NE-OTE任务的特点,我们在模型中加入注意力机制,以辅助模型更好地捕获输入序列中与意见目标及其情感倾向相关的隐含特征,提高模型提取文本表征的质量。

#### (4) 样本不均衡问题

在基于ABSA任务主流数据集构建的NE-OTE任务语料中存在样本不均衡问题,语料中的语句只有小部分语句包含NE-OTE任务的抽取目标——消极情绪意见目标,影响了模型的性能和鲁棒性。为了改善样本不均衡问题,我们进行了大量的数据增强实验,实验结果表明,随机同义词替换的数据增强效果最为明显。

本文的主要贡献如下:

(1)本文提出了一个全新的 ABSA 子任务——抽取消极情绪的意见目标(NE-OTE),任务本身有潜在的应用价值,同时可以提升 ABSA 任务的解释性。

(2)本文通过序列标注的方式来完成 NE-OTE 任务,并提出了一种基于注意力机制和单词与字符混合嵌入的 BiLSTM-CRF 序列标注神经网络模型来解决 NE-OTE 任务。

(3)本文针对在线评论文本质量较差的缺点,提出了将单词的字符级特征加入单词的表示中,实验结果表明,字符级特征大大提升了模型的健壮性和泛化能力。

(4)为了更好地捕获输入序列中与意见目标及其情感倾向相关的隐含特征,本文提出使用注意力机制来辅助模型的文本表征提取。实验结果表明,注意力机制是有效的。同时,注意力机制也提高了模型的可解释性和透明度。

(5)本文基于主流的 ABSA 基准数据集构建了 3 个 NE-OTE 任务数据集,并在这些数据集上进行了广泛的实验,结果表明,本文模型的性能显著超过各种基线模型。

## 2 相关工作

### 2.1 意见目标抽取任务

目前已有大量针对意见目标抽取任务的研究,有无监督/半监督和监督等传统方法<sup>[13-14]</sup>。近年来,基于深度学习的方法在意见目标抽取任务方面也取得了巨大进展。Liu 等<sup>[15]</sup>使用基于预先训练过的单词向量的循环神经网络来解决这个任务。Poria 等<sup>[16]</sup>使用深度卷积神经网络(Convolutional Neural Networks, CNN)来解决意见目标抽取任务。而 Xu 等<sup>[17]</sup>提出了一个双嵌入的 CNN 模型。Li 等<sup>[18]</sup>针对汉字的特点,提出了基于字向量的模型来解决中文意见目标的抽取。但是上述方法都不能很好地适用于抽取特定情感极性意见目标的任务。

相比之下,本文方法是一种监督学习方法,通过序列标注的方式抽取消极情绪的意见目标。同时,本文针对 NE-OTE 任务和在线评论文本的特点设计了一些辅助特征来提升模型的性能。

### 2.2 任务计划

本文将 NE-OTE 任务转换成序列标注问题来解决。序列标注是一种经典的自然语言处理问题,几十年来,它一直受到研究人员的关注,包括命名实体识别<sup>[19]</sup>、词性标注<sup>[20]</sup>、关键词抽取<sup>[21]</sup>等。同时,将自然语言处理任务转换成序列标注问题也是一种常见的解决方案。序列标注问题的本质是根据上下文表征对输入序列中的每个元素进行标签分类,其中,元素指输入序列中的字或者单词。在不同的任务中,可能需要使用代表不同含义的标签集合,但是最终的目的都是为输入序列中的每一个元素预测标签<sup>[22]</sup>。

对于一个由长度为  $m$  的句子中的单词组成的输入序列  $S=(\text{单词 } 1, \text{单词 } 2, \dots, \text{单词 } m)$  来说,当前 NE-OTE 任务的目的就是对输入序列  $S$  进行序列标注,预测每个单词的标签,最终根据预测的标签序列抽取消极情绪意见目标。我们针对 NE-OTE 任务设计了  $\{B, O\}$  标签集合,每一个句子中的单词都被标记上  $\{B, O\}$  标签集合中的一个标签。其中,标签  $B$  表示当前单词为消极情绪意见目标词,标签  $O$  则表示

其他单词。以上述的用户评论“The sandwich was delicious, but the table was a bit dirty.”为例,用户评论中只有单词“table”是消极情绪意见目标词,其他单词都不是。因此,最终的序列标注结果为:

The/O sandwich/O was/O delicious/O,/ but/O the/O table/Bwas/O a/O bit/O dirty/O./O

## 3 Attention-based BiLSTM CRF with Word Embedding and Character Embedding(AB-CE)

### 3.1 BiLSTM-CRF (W-BiLSTM-CRF)

在 BiLSTM-CRF 模型中,首先通过 Embedding 层将初始输入的单词序列  $W=(w_1, w_2, \dots, w_m)$  转换成向量表示  $X=(x_1, x_2, \dots, x_m)$ ,其中  $x_i$  表示维度为  $d$  的单词词向量,接着通过 Word-BiLSTM 层学习输入序列的上下文特征和表示,将 Word-BiLSTM 层输出的文本特征  $H=(h_1, h_2, \dots, h_m)$ (其中  $h_i$  是当前单词的上下文表征)作为 CRF 层的输入,最终 CRF 层结合学习的相邻标签之间的依赖关系,从所有可能的标签序列中选择出句子级别的全局最佳标签序列作为最终输出。

#### 3.1.1 双向长短期记忆网络

长短期记忆网络(Long Short-Term Memory, LSTM)是循环神经网络(RNN)的变种之一,由 Hochreiter 等于 1997 年首次提出<sup>[23]</sup>,相比 RNN, LSTM 增加了能够控制丢弃或者增加信息的“门限”结构来控制信息的遗忘和记忆,增强了捕获长距离语义依赖和对历史信息的记忆能力,因此该模型在进行预测时可以依赖远距离的“早期”信息,同时改善 RNN 长距离传播时梯度弥散和爆炸的问题。此外,与单向 LSTM 相比, BiLSTM 可以更好地利用输入序列的双向上下文信息。因此, BiLSTM 在文本分类<sup>[24]</sup>和命名实体识别<sup>[25]</sup>等自然语言处理任务中都得到了广泛的应用。LSTM 的单元结构如图 2 所示。

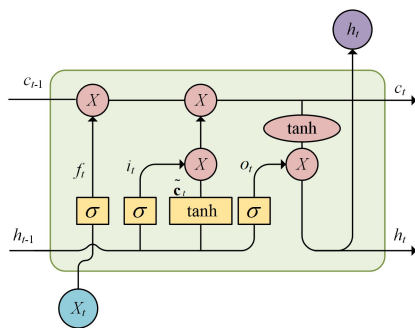


图 2 LSTM 的单元结构<sup>[23]</sup>

Fig. 2 LSTM unit structure<sup>[23]</sup>

LSTM 单元包含遗忘门(forget gate)、输入门(input gate)、输出门(output gate)。LSTM 单元可以根据当前时刻网络的输入  $x_t$ 、前一个隐藏状态  $h_{t-1}$  和前一个单元状态  $c_{t-1}$  计算出当前隐藏状态  $h_t$ ,其中输入门  $i_t$  控制计算当前时刻网络的输入  $x_t$  记忆到当前单元状态  $c_t$  中(如式(1)所示),遗忘门  $f_t$  控制计算上一个单元状态  $c_{t-1}$  记忆到当前单元状态  $c_t$  中(如式(2)所示),当前单元根据上个时刻单元状态  $c_{t-1}$  的遗忘程度和当前输入的单元状态  $\tilde{c}_t$ (见式(3))的保留程度来更新  $c_t$ (见式(4))。输出门控制计算当前单元状态  $c_t$  输出作为

当前的隐藏状态  $h_t - o_t$  (见式(5)), 3 个门限和单元状态更新的计算式如下:

$$i_t = \delta(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}x_{t-1} + b_i) \quad (1)$$

$$f_t = \delta(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (2)$$

$$\tilde{c}_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (3)$$

$$c_t = f_t * x_t + i_t * \tilde{c}_t \quad (4)$$

$$o_t = \delta(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_{t-1} + b_o) \quad (5)$$

$$h_t = o_t * \tanh(c_t) \quad (6)$$

其中,  $\delta$  为 sigmoid 函数,  $W_{xf}, W_{hf}, W_{cf}, W_{xc}, W_{hc}, W_{xo}, W_{ho}, W_{co}$  和  $b_i, b_f, b_c, b_o$  分别表示权重参数和偏置项。

但是, LSTM 是一种前向传播算法, 对于当前任务中的在线评论文本来说, 除了序列前值, 序列后值也是非常重要的, 需要同时学习序列前值和序列后值。因此, 本文采用双向长短时记忆网络 (BiLSTM) 来捕获双向上下文语义信息, 最终拼接单词上下文表征  $\vec{h}_t, \overleftarrow{h}_t$  作为最终单词的上下文表征, 如式(7)所示:

$$h_t = \vec{h}_t \oplus \overleftarrow{h}_t \quad (7)$$

### 3.1.2 条件随机场

条件随机场 (Conditional Random Field, CRF) 与 BiLSTM 和其他模型不同, 其并不能考虑长距离双向上下文信息, 相反, 它更多考虑的是输入序列局部特征的线性加权组合。因此, Lafferty 等<sup>[26]</sup> 采用条件随机场联合预测整个输入序列的全局最佳标签序列。对于序列标注任务, 考虑到相邻标签之间的依赖性是非常重要的<sup>[27]</sup>, 因此 CRF 引入了转移矩阵  $A$ , 其中  $A_{ij}$  表示从第  $i$  个标签到第  $j$  个标签的转移得分。假设输入序列  $X = (x_1, x_2, \dots, x_m)$  对应的预测的标签序列  $Y = (y_1, y_2, \dots, y_m)$ , 其中  $y_i$  表示当前单词的标签。本文模型对整个序列的打分  $score(X, Y)$  等于各个位置的分数之和, 如式(8)所示:

$$score(X, Y) = \sum_{i=1}^n P_{i, y_i} + \sum_{i=1}^{n+1} A_{y_{i-1}, y_i} \quad (8)$$

其中,  $P_{i, y_i}$  表示序列元素  $x_i$  被打标签  $y_i$  的得分。经过最大化对数似然函数的训练后, CRF 会计算所有可能标签序列路径的得分, 并选择出全局最佳标签序列路径, 也就是从所有可能的标签序列路径中选择出得分最高的路径, 如图 3 所示。

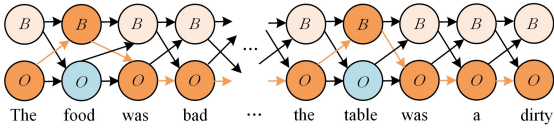


图 3 CRF 从所有可能的标签序列路径中选择出得分最高的路径

Fig. 3 CRF selects the path with the highest score from all possible tag sequence paths

### 3.2 BiLSTM-CRF with Word Embedding and Character Embedding (WC-BiLSTM-CRF)

除了单词级别特征之外, 单词中的字符也包含了丰富的结构信息。这些特征 (如字符、单词前缀和单词后缀) 在目前的序列标注问题中也是比较常用的<sup>[28]</sup>。

为了获得单词的字符特征, 本文将长度为  $m$  的输入序列中单词的所有字符的 One Hot 向量  $O = (o_1, o_2, \dots, o_m)$  ( $o_i$

表示当前字符的 One Hot 向量) 送入 Character-BiLSTM 层学习单词中的前后方向字符表征  $\vec{h}_{c_1}, \overleftarrow{h}_{c_m}$ 。最终, 将 Character-BiLSTM 层输出单词的前后方向字符表征拼接作为当前单词的字符特征向量  $c$ 。其中, Character-BiLSTM 层的结构如图 4 所示。

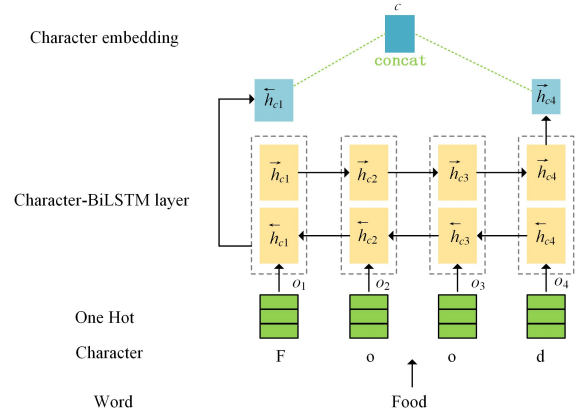


图 4 Character-BiLSTM 层获得单词的字符特征

Fig. 4 Character-BiLSTM layer obtains the character characteristics of the word

为了提升模型对 NE-OTE 任务中文本质量较差的在线评论的健壮性和泛化能力, 我们将每个单词的字符特征向量  $c_i$  都拼接到单词的单词向量  $x_i$  中, 共同作为单词的表示送入 Word-BiLSTM 层, Word-BiLSTM 层输出的上下文表征就同时包含了单词的字符级特征和单词级特征。

### 3.3 Attention-based BiLSTM-CRF (AB)

NE-OTE 作为意见目标级别任务, 标准的 BiLSTM-CRF 模型并不擅长这样的细粒度文本情感分析任务, 尤其是捕获输入序列中意见目标的情感倾向信息。为了解决这个问题, 我们提出在 BiLSTM-CRF 的基础上增加注意力计算层。其中, 意见目标的情感倾向是 NE-OTE 任务中的重要信息, 注意力机制有助于在意见目标及其情感倾向相关的隐含特征之间建立联系, 辅助模型更好地区分意见目标的情感倾向。Attention-based BiLSTM-CRF 模型的结构如图 5 所示。

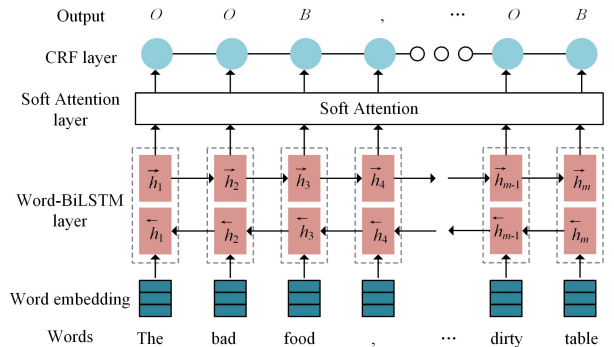


图 5 Attention-based BiLSTM-CRF 模型结构

Fig. 5 Attention-based BiLSTM-CRF model structure

Xu 等<sup>[29]</sup> 受到机器翻译领域的注意力机制的启发, 提出了 Hard Attention 和 Soft Attention 两种注意机制。两者的不同之处在于, Hard Attention 的注意力焦点集中在当下, 集中在一点, 而 Soft Attention 的注意力焦点则更加发散。对于

自然语言处理领域来说 Soft Attention 发散的注意力焦点更有优势,应用也更广泛<sup>[30]</sup>。

本文将 Word-BiLSTM 层抽取的输入序列文本特征向量  $\mathbf{H}=(\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_m)$  送入 Soft Attention 层进行计算,并输出上下文的 Attention 向量  $\mathbf{K}=(\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_m)$  ( $\mathbf{k}_i$  表示当前单词的 Attention 向量)作为输入序列新的上下文表征,并将其送入 CRF 层进行标签预测。Soft Attention 层结构如图 6 所示,其中 Attention 向量的计算步骤如下。

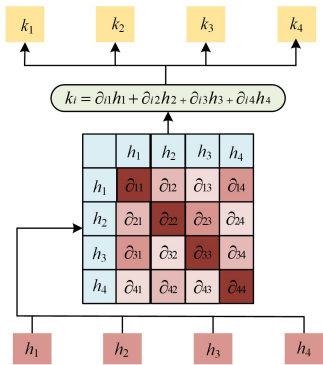


图 6 Soft Attention 层结构

Fig. 6 Soft Attention layer structure

Step1 计算输入序列元素之间的相似程度  $e_{ij}$ ,也就是计算输入序列中,位置为  $j$  的单词对位置为  $i$  的单词的影响程度,如式(9)所示:

$$e_{ij} = f(h_i, h_j), i, j = 1, 2, \dots, m \quad (9)$$

其中,  $f$  是相似度计算方法,包含点乘、拼接权重、感知器、co-sine 这 4 种基本计算方式<sup>[31]</sup>,其中  $h_i, h_j$  是 BiLSTM 层输出的文本特征向量  $\mathbf{H}=(\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_m)$  中位置为  $i, j$  的单词的上下文表征。

Step2 对得到的相似度  $e_{ij}$  进行归一化,得到计算句输入序列中位置为  $j$  的单词对位置为  $i$  的单词的影响权重  $\partial_{ij}$ ,如式(10)所示:

$$\partial_{ij} = \text{softmax}(e_{ij}), i, j = 1, 2, \dots, m \quad (10)$$

Step3 根据输入序列所有位置的单词对当前单词的影响权重进行加权计算,最终得到当前单词的 Attention 向量  $\mathbf{k}_i$ ,并将其作为当前单词的表示,如式(11)所示:

$$\mathbf{k}_i = \partial_{i1}h_1 + \partial_{i2}h_2 + \dots + \partial_{im}h_m, m = 1, 2, \dots, m \quad (11)$$

### 3.4 Attention-based BiLSTM-CRF with Word Embedding and Character Embedding (AB-CE)

在 W-BiLSTM-CRF 模型的基础上,模型 WC-BiLSTM-CRF 在单词表示中加入字符级特征是为了提高模型的健壮性和泛化能力,为了充分利用字符级特征,我们将字符特征向量加入到 AB-CE 模型中的单词向量中。最终,AB-CE 模型中的 Word-BiLSTM 层输入的单词特征向量中同时包含了单词的单词特征和字符特征,它们共同作为单词的表示送入 Word-BiLSTM 层,并在注意力计算层的权重计算中共同发挥作用。然后 Soft Attention 层输出注意力向量并将其作为输入序列新的文本表征送入 CRF 层,CRF 层将结合相邻标签的依赖信息联合预测出句子级别的全局最佳标签序列。AB-CE 模型结构如图 7 所示。

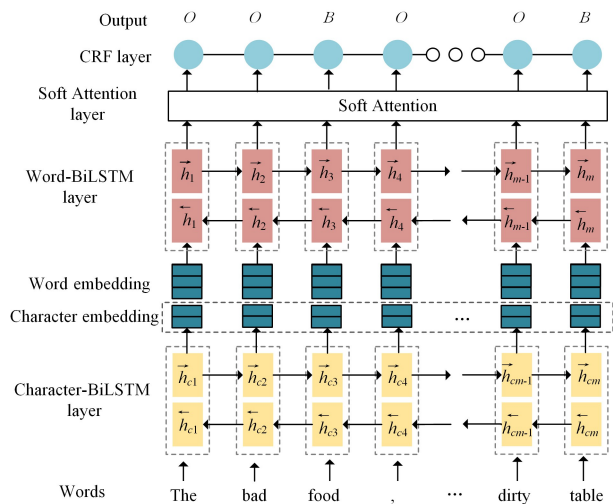


图 7 AB-CE 模型结构

Fig. 7 AB-CE model structure

## 4 实验及结果分析

### 4.1 实验数据集

本文基于 SemEval Challenge 2014 Task4, SemEval Challenge 2015 task 12 和 SemEval Challenge 2016 task 5<sup>[32-34]</sup> 的数据构造了 NE-OTE 任务语料。这些数据集是 ABSA 任务的主流数据集。我们对上述语料中每条句子的意见目标进行情感倾向的人工筛选和判定,标注出消极情绪的意见目标,每个句子由 3 个人标注,若发生冲突,则进行复审。我们只保留包含意见目标的句子,同时,删除有歧义或者无情绪表达的句子。

NE-OTE 任务语料存在样本不平衡的问题(语料中的语句只有小部分语句包含了 NE-OTE 任务的抽取目标——消极情绪意见目标),影响了模型的性能和鲁棒性。为了改善样本不平衡的问题,我们选择使用数据增强的方式改善样本不平衡问题,常见的自然语言处理数据增强方式<sup>[35]</sup>包括随机替换、随机插入、随机交换、随机删除。实验结果表明,随机同义词替换的数据增强效果最明显。数据增强的实验参数如表 1 所列。

表 1 实验参数列表

Table 1 List of experimental parameters

实验参数	说明	取值
optimizer	优化器	Adam
Word_LSTM_hidden_size	Word_LSTM 层 LSTM 单元隐藏层单元数量	200
Character_LSTM_hidden_size	Character_LSTM 层 LSTM 单元隐藏层单元数量	100
attention_type	Attention 相似度计算方法	感知器
word_embedding_dim	单词词向量维度	300
batch_size	一次训练所选取的样本数	8
learning_rate	学习率	0.005
aug_type	数据增强方式	同义词替换
exchange_target	同义词替换对象	只针对负面情感词
exchange_frequency	同义词替换频率	Top2 单词相似度的同义词

我们构造了 3 个数据集,其中 14res 来自 SemEval2014, 15res 来自 SemEval 2015, 16res 来自 SemEval 2016。数据集统计结果如表 2 所列。

表2 数据集统计结果

Table 2 Statistical results of datasets

数据集	句子数量	包含 NE-OT 的 句子数量	NE-OT 数量
14res	训练集	1301	215
	测试集	324	58
15res	训练集	604	101
	测试集	150	21
16res	训练集	864	147
	测试集	215	41

## 4.2 实验设置

本文采用手工调整结合网格搜索法<sup>[36]</sup>进行神经网络模型的参数调节,最终获得模型的最优参数集合。本文模型的部分超参数来自一些先前研究命名实体识别(NER)和意见目标抽取(OTE)论文中的经验,如优化器选择、单词向量维度等。其他参数则根据当前 NE-OTE 任务的特点和在线评论文本的特点来设置,如最长序列长度、Word-LSTM 层 LSTM 单元隐藏层单元数量、Character-LSTM 层 LSTM 单元隐藏层单元数量等。本文的主要模型实验参数与数据增强实验参数的取值及其说明如表 1 所列。

## 4.3 评价指标

我们使用序列标注任务常用的评价指标,即准确率(Precision, P)、召回率(Recall, R)和 F1(F1-score, F1)值来评估本文模型的效果,如式(12)~式(14)所示:

$$P = \frac{\text{正确识别的 NE-OT 的数量}}{\text{识别的 NE-OT 的数量}} \times 100\% \quad (12)$$

$$R = \frac{\text{正确识别的 NE-OT 的数量}}{\text{NE-OT 的数量}} \times 100\% \quad (13)$$

表3 对照实验结果

Table 3 Results of control experiments

模型	14res			15res			16res		
	P	R	F1	P	R	F1	P	R	F1
W-RNN	63.63	36.20	46.15	57.14	19.04	28.57	66.66	14.63	24.0
W-BiRNN	86.11	53.44	65.95	75.0	42.85	54.54	71.42	36.58	48.38
W-LSTM	70.96	37.93	49.43	85.71	28.57	42.85	44.44	19.51	27.11
W-BiLSTM	67.85	65.51	66.66	76.92	47.61	58.82	71.42	48.78	57.97
W-BiLSTM-CRF	69.09	65.51	67.25	62.5	71.42	66.66	61.36	65.85	63.52
C-BiRNN	55.17	27.58	36.78	43.47	17.24	24.69	42.85	14.63	24.69
C-BiLSTM	65.95	53.44	59.04	60.86	66.66	63.63	70.96	53.65	61.11
C-BiLSTM-CRF	66.66	55.17	60.37	61.90	61.90	61.90	60.46	63.41	61.90
WC-BiRNN	85.0	58.62	69.38	72.22	61.90	66.66	72.0	43.90	54.54
WC-BiLSTM	77.08	63.79	69.81	64.0	76.19	69.56	75.0	58.53	65.75
WC-BiLSTM-CRF	82.2	63.79	71.84	66.66	76.19	71.11	70.27	63.41	66.66
AB-CE	81.25	67.24	73.58	69.56	76.19	72.72	72.22	63.41	67.53

首先,综合表 3 中的结果来看,对于抽取负面情绪意见目标任务来说,基于单词嵌入的模型优于基于字符嵌入的模型,以单词嵌入为输入比以字符嵌入为输入的召回率更高。其中,模型 W-RNN 的效果不理想,是所有方法中最差的,它的召回率特别低,相比模型 C-RNN,模型 AB-CE 的 F1 值提高了 40%左右。

其次,在基于单词嵌入的模型的基础上增加字符嵌入,大大提升了模型的性能。其中,相比基于单词嵌入为输入的 W-BiRNN, W-BiLSTM, W-BiLSTM-CRF 模型,基于单词与字符混合嵌入的 WC-BiRNN, WC-BiLSTM, WC-BiLSTM-CRF 模型的准确率和召回率均得到大幅提升,消极情绪意见目标被正确预测的比例和模型预测结果的质量也大大提升,验

$$F1 = \frac{2 \times P \times R}{P + R} \times 100\% \quad (14)$$

## 4.4 对照方法

由于我们是首次研究 NE-OTE 任务,因此无对比模型,虽然意见目标抽取(OTE)任务中有很多复杂的模型,但是与当前 NE-OTE 任务是不同的,这些 OTE 模型侧重于整体语句的语义信息,对于 NE-OTE 任务,更关键的是情感信息与意见目标的依赖关系,故单纯地将 OTE 模型移植到 NE-OTE 是不合适的。因此我们只能自己设计和实现 NE-OTE 的基线模型。我们将与以下基线模型进行比较。

(1)W-RNN/W-BiRNN:基于单词的 W-RNN/W-Bi-RNN,以词嵌入为输入,如 Ali 等<sup>[37]</sup>的研究所示。

(2)W-LSTM/W-BiLSTM:基于单词嵌入的 LSTM/BiLSTM 与 W-RNN/W-BiRNN 具有相同的输入,如 Liu 等<sup>[15]</sup>的研究所示。

(3)W-BiLSTM-CRF: Huang 等<sup>[38]</sup>提出以单词嵌入为输入,用 BiLSTM-CRF 模型来完成序列标注任务。

(4)C-BiRNN/C-BiLSTM/C-BiLSTM-CRF:基于字符嵌入的 BiRNN/BiLSTM/BiLSTM 模型,以字符嵌入为输入。

(5)WC-BiRNN/WC-BiLSTM/WC-BiLSTM-CRF:基于单词和字符混合嵌入的 BiRNN/BiLSTM/BiLSTM 模型,以单词与字符混合嵌入为输入。

## 4.5 实验结果

表 3 列出了 11 种基线模型与本文 AB-CE 模型的实验结果。我们从实验结果中得出了 4 个结论。

证了增加字符嵌入对提升模型鲁棒性和泛化能力的有效性。

然后,捕获输入序列双向上下文语义对于抽取负面情绪意见目标任务来说是有必要的。相比基于单向语义的 W-RNN, W-LSTM 模型,基于双向上下文语义的 W-BiRNN, W-BiLSTM 模型的准确率、召回率和 F1 值均得到大幅提升,这表明我们之前的分析是正确的。

最后,结合特征提取层提取的上下文表征和 CRF 层可以提高模型的性能。如表 3 所列, W-BiLSTM-CRF, C-BiLSTM-CRF 和 WC-BiLSTM-CRF 模型分别在 W-BiLSTM, C-BiLSTM, WC-BiLSTM 模型的基础上增加了 CRF 层,使 F1 值均得到提升。

相比其他基线模型, AB-CE 模型在各个数据集上均取得

了最好的表现。AB-CE 模型在 W-BiLSTM-CRF 模型的基础上增加了字符嵌入和注意力计算层,其准确率、召回率、F1 值均得到了大幅提升,其中 AB-CE 的 F1 值比 W-BiLSTM-CRF 提升了 5%左右,准确率提升了 10%左右。

#### 4.6 模型分析

本文分析了在基线模型 W-BiLSTM-CRF 的基础上增加

字符嵌入和注意力计算层的有效性,如表 4 所列。表 4 第一行为 W-BiLSTM-CRF 模型,它在所有以单词嵌入为输入的基线模型中性能最优。接下来的两行显示了在 W-BiLSTM-CRF 的基础上分别增加字符嵌入、注意力计算层的 WC-BiLSTM-CRF 模型和 AB 模型的实验结果。最后一行显示了本文提出的 AB-CE 模型的实验结果。

表 4 消融实验结果

Table 4 Results of ablation experiments

模型	14res			15res			16res		
	P	R	F1	P	R	F1	P	R	F1
W-BiLSTM-CRF	69.09	65.51	67.25	62.5	71.42	66.66	61.36	65.85	63.52
WC-BiLSTM-CRF	82.2	63.79	71.84	66.66	76.19	71.11	70.27	63.41	66.66
AB	74.07	68.96	71.42	62.96	80.95	70.83	69.44	60.97	64.93
AB-CE	81.25	67.24	73.58	69.56	76.19	72.72	72.22	63.41	67.53

如表 4 第一行和第二行所列,相比 W-BiLSTM-CRF 模型,基于单词与字符混合嵌入的 WC-BiLSTM-CRF 模型的性能得到了明显的改善。同时,如表 3 所列,在基于单词嵌入模型的基础上增加字符嵌入能够普遍提高模型的召回率,对于抽取负面情绪意见目标任务和在线评论文本的特点来说,字符嵌入的加入,使模型的健壮性和泛化能力均得到提升,这与我们之前对增加字符嵌入的分析是一致的。

如表 4 第一行和第三行所列,相比 W-BiLSTM-CRF 模型,AB 模型在 W-BiLSTM-CRF 的基础上增加了注意力计算层,注意力计算层利用 W-BiLSTM 层输出的上下文表征和输入序列中所有元素之间的影响权重来重新计算出新的上下文表征,这有利于模型区分意见目标的情感倾向。与我们之前的分析一致,注意力计算层的加入显著提升了模型的准确率和 F1 值,其中 AB 模型在 3 个数据集上的准确率分别提升了 4.98%,0.46%,8.08%,其 F1 值分别提升了 4.17%,4.17%,1.31%。

**结束语** 本文提出了一个全新的 ABSA 子任务——消极情绪意见目标抽取(NE-OTE),其目的是识别出在线评论中被用户消极评论的意见目标。为此,本文设计了一个新的神经网络模型 AB-CE,该模型利用注意力机制能更好地捕获输入序列中的意见目标及其情感倾向相关的隐含特征,这对于 NE-OTE 任务来说十分重要,同时本文在基于单词嵌入的基础上增加了字符嵌入,显著提升了模型在质量较差的在线评论文本中的健壮性和泛化能力。实验结果表明,AB-CE 在所有数据集上都取得了最佳的表现。

在未来的工作中,NE-OTE 任务可以用来提高 ABSA 等文本情感分析任务的可解释性和性能。NE-OTE 任务本身仍有多个方面值得进一步的研究:如何实现对隐含消极情绪意见目标的识别,可以尝试使用 Seq2Seq 架构,根据上下文文本表征在解码端生成消极情绪意见目标;模型的跨领域鲁棒性仍待研究,可以尝试使用基于预训练和微调思想的词向量,从而获得更好的单词表示,提高模型跨领域的鲁棒性。

#### 参考文献

[1] NASUKAWA T, YI J. Sentiment analysis: Capturing favorabili-

ty using natural language processing[C]// Proceedings of the 2nd International Conference on Knowledge Capture. 2003: 70-77.

[2] LIU B. Sentiment analysis and opinion mining [J]. Synthesis Lectures on Human Language Technologies, 2012, 5(1): 1-167.

[3] WANG Y, HUANG M, ZHU X, et al. Attention-based LSTM for aspect-level sentiment classification[C]// Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016: 606-615.

[4] TAY Y, TUAN L A, HUI S C. Learning to attend via word-aspect associative fusion for aspect-based sentiment analysis[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2018: 5956-5963.

[5] SUN C, HUANG L, QIU X. Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence[J]. arXiv: 1903.09588, 2019.

[6] LI X, BING L, ZHANG W, et al. Exploiting BERT for end-to-end aspect-based sentiment analysis[J]. arXiv: 1910.00883, 2019.

[7] ZHANG J, DUAN L G, LI A P, et al. Fine-grained Sentiment Analysis Based on Combination of Attention and Gated Mechanism[J]. Computer Science, 2021, 48(8): 226-233.

[8] PORIA S, CAMBRIA E, GELBUKH A. Aspect extraction for opinion mining with a deep convolutional neural network[J]. Knowledge-Based Systems, 2016: 42-49.

[9] XU H, LIU B, SHU L, et al. Double embeddings and cnn-based sequence labeling for aspect extraction[J]. arXiv: 1805.04601, 2018.

[10] MNH V, HEES N, GRAVES A. Recurrent models of visual attention[J]. Advances in Neural Information Processing Systems, 2014, 27: 2204-2212.

[11] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[J]. arXiv: 1409.0473, 2014.

[12] RUSH A M, CHOPRA S, WESTON J. A neural attention model for abstractive sentence summarization[J]. arXiv: 1509.00685, 2015.

[13] HU M, LIU B. Mining and summarizing customer reviews[C]// Proceedings of the Tenth ACM SIGKDD International Confe-

- rence on Knowledge Discovery and Data Mining. 2004;168-177.
- [14] ZHUANG L, JING F, ZHU X Y. Movie review mining and summarization[C]// Proceedings of the 15th ACM International Conference on Information and Knowledge Management. 2006; 43-50.
- [15] LIU P, JOTY S, MENG H. Fine-grained opinion mining with recurrent neural networks and word embeddings[C]// Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. 2015;1433-1443.
- [16] PORIA S, CAMBRIA E, GELBUKH A. Aspect extraction for opinion mining with a deep convolutional neural network[J]. Knowledge-Based Systems, 2016(108), 108:42-49.
- [17] XU H, LIU B, SHU L, et al. Double embeddings and cnn-based sequence labeling for aspect extraction[J]. arXiv:1805.04601, 2018.
- [18] LI Y, LIU T, LI D, et al. Character-based BiLSTM-CRF Incorporating POS and Dictionaries for Chinese Opinion Target Extraction[C]// Asian Conference on Machine Learning. 2018; 518-533.
- [19] LI J, SUN A, HAN J, et al. A survey on deep learning for named entity recognition[J]. IEEE Transactions on Knowledge and Data Engineering, 2020, 3(17):1-20.
- [20] PLANK B, SØGAARD A, GOLDBERG Y. Multilingual part-of-speech tagging with bidirectional long short-term memory models and auxiliary loss[J]. arXiv:1604.05529, 2016.
- [21] MARTINC M, ŠKRLJ B, POLLAK S. TNT-KID: Transformer-based Neural Tagger for Keyword Identification [J]. arXiv: 2003.09166, 2020.
- [22] HE Z, WANG Z, WEI W, et al. A Survey on Recent Advances in Sequence Labeling from Deep Learning Models[J]. arXiv:2011.06727, 2020.
- [23] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8):1735-1780.
- [24] BAI X. Text classification based on LSTM and attention[C]// 2018 Thirteenth International Conference on Digital Information Management (ICDIM). IEEE, 2018;29-32.
- [25] CHIU J P C, NICHOLS E. Named entity recognition with bidirectional LSTM-CNNs[J]. Transactions of the Association for Computational Linguistics, 2016, 4:357-370.
- [26] LAFFERTY J, MCCALLUM A, PEREIRA F C N. Conditional random fields: Probabilistic models for segmenting and labeling sequence data[C]// Proc. of the 18th Int. Conf. on Machine Learning. New York: ACM, 2021;282-289.
- [27] HUANG Z, XU W, YU K. Bidirectional LSTM-CRF models for sequence tagging[J]. arXiv:1508.01991, 2015.
- [28] LIU S, TANG B, CHEN Q, et al. Drug name recognition: approaches and resources[J]. Information, 2015, 6(4):790-810.
- [29] XU K, BA J, KIROS R, et al. Show, attend and tell: Neural image caption generation with visual attention[C]// International Conference on Machine Learning. 2015;2048-2057.
- [30] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[J]. arXiv:1409.0473, 2014.
- [31] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// Advances in Neural Information Processing Systems. 2017;5998-6008.
- [32] PONTIKI M, PAPAGEORGIOU H, GALANIS D, et al. SemEval-2014 Task 4: Aspect Based Sentiment Analysis[C]// Proc. 8th Int. Workshop Semantic Eval(SemEval). 2014;27-35.
- [33] PONTIKI M, GALANIS D, PAPAGEORGIOU H, et al. SemEval-2015 task 12: Aspect based sentiment analysis[C]// Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015). 2015;486-495.
- [34] PORIA S, CAMBRIA E, GELBUKH A. Aspect extraction for opinion mining with a deep convolutional neural network[J]. Knowledge-Based Systems, 2016, 108:42-49.
- [35] WEI J, ZOU K. Eda: Easy data augmentation techniques for boosting performance on text classification tasks[J]. arXiv: 1901.11196, 2019.
- [36] LAMESKI P, ZDRAVEVSKI E, MINGOV R, et al. SVM parameter tuning with grid search and its impact on reduction of model over-fitting[M]// Rough sets, fuzzy sets, data mining, and granular computing. Cham: Springer, 2015;464-474.
- [37] ALI M N A, TAN G, HUSSAIN A. Bidirectional recurrent neural network approach for Arabic named entity recognition[J]. Future Internet, 2018, 10(12):123-135.
- [38] HUANG Z, XU W, YU K. Bidirectional LSTM-CRF models for sequence tagging[J]. arXiv:1508.01991, 2015.



**DING Feng**, born in 1996, postgraduate. His main research interests include natural language processing, machine learning and textual sentiment analysis.



**SUN Xiao**, born in 1980, Ph.D, professor, is a member of China Computer Federation. His main research interests include affective computing, natural language processing, machine learning and human-machine interactions.

(责任编辑:柯颖)