

一种融合位置特征的方面级情感分析方法

翟社平^{1,2}, 成大宝¹, 张文晴¹, 刘园彪¹

(1. 西安邮电大学 计算机学院, 陕西 西安 710121;

2. 陕西省网络数据分析与智能处理重点实验室, 陕西 西安 710121)

摘要:方面级情感分析是如今情感分析领域的重要研究任务之一,旨在计算文本中多个方面词的情感极性。现有的方面级情感分析方法通常将整个句子直接输入复杂的神经网络,尽管此类方法能够有效捕捉到词与词之间的依赖关系,却忽略了方面词与其上下文之间所隐含的位置特征。因此,提出了一种融合位置特征的方面级情感分析方法。将文本分别采用基于方面词间隔的上下文划分方式与基于单词距离的上下文划分方式,通过两个微调后的 BERT 模型,完成词向量的表达;将两种词向量送入多头注意力机制,计算其文本特征;使用平均池化将语义信息进行融合,在输出层完成方面词的情感极性分类。在 SemEval2014 Task4 数据集和 Twitter 数据集上的实验表明,提出的融合位置特征的方面级情感分析方法能够充分利用方面词上下文之间的位置特征,有效提升了准确率和 F1 值。

关键词:方面级情感分析;位置特征;注意力机制;BERT;深度学习

中图分类号:TP391.1

文献标识码:A

文章编号:1673-629X(2023)05-0167-06

doi:10.3969/j.issn.1673-629X.2023.05.025

An Aspect-level Sentiment Analysis Method with Location Features

ZHAI She-ping^{1,2}, CHENG Da-bao¹, ZHANG Wen-qing¹, LIU Yuan-biao¹

(1. School of Computer Science & Technology, Xi'an University of Posts and Telecommunications,
Xi'an 710121, China;

2. Shaanxi Key Laboratory of Network Data Analysis and Intelligent Processing, Xi'an University of
Posts and Telecommunications, Xi'an 710121, China)

Abstract: Aspect-level sentiment analysis is one of the important research tasks in the field of sentiment analysis, aiming at calculating the sentiment polarity of various aspect words in the text. The existing aspect-level sentiment analysis methods usually input the whole sentence directly into complex neural networks. Although this kind of method can effectively capture the dependency between words, it ignores the implicit location features between aspect words and its contexts. Therefore, we propose an aspect-level sentiment analysis method with location features. The aspect words and their contexts are divided into context based on aspect word interval and context based on word distance, respectively, and the expression of word vectors is completed by two fine-tuned BERT models. Two kinds of word vectors are sent into the multi-head attention mechanism, and their text features are calculated. Average pooling is used to fuse semantic information, and the emotional polarity classification of aspect words is completed at the output level. Experiments on SemEval2014 Task4 data set and Twitter data set show that the proposed aspect-level sentiment analysis method with location features can make full use of the location features between the contexts of aspect words, and effectively improve the accuracy and F1-Measure.

Key words: aspect-level sentiment analysis; location features; attention mechanism; BERT; deep learning

0 引言

情感分析,又称为观点挖掘、意见挖掘,是自然语言处理领域中的一项基本任务,其基本目标是挖掘文本中的情感信息,分析其情感倾向性^[1]。情感分析人

物根据粒度的不同,被划分为文档级情感分析、句子级情感分析以及细粒度的方面级情感分析。

方面级情感分析旨在判断文本中具体方面所对应的情感极性^[2]。在文本“In addition, the food is very

收稿日期:2022-05-06

修回日期:2022-09-08

基金项目:工业和信息化部通信软科学项目(2018R26);陕西省重点研发计划项目(2022GY-038);国家级大学生创新创业训练计划项目(202111664004)

作者简介:翟社平(1971-),男,博士,教授,CCF高级会员(77328S),通讯作者,研究方向为语义web、自然语言处理;成大宝(1997-),男,硕士研究生,研究方向为情感分析。

good but the prices are unreasonable.”中,“food”和“prices”是对同一个餐厅进行描述的两个方面,而他们的极性却是相反的,使用文档级情感分析或句子级情感分析无法判别文本真实的情感极性,而使用方面级情感分析能够分别计算文本中多个方面的情感极性。目前随着文本信息的复杂,方面级情感分析成为了一项重要的研究内容。

随着深度学习的发展,在目前现有的方面级情感分析方法中,主要包括构建复杂的神经网络模型以及将不同的神经网络模型与注意力机制等进行结合。这类模型研究能够模仿人们阅读时的注意力,从而能够捕获方面词的极性信息。但此类研究均只在模型构建层面进行迭代与累积,并未考虑方面词在文本中的位置关系,从而导致在不同数据集上的鲁棒性较差。为解决以上问题,本文着重研究方面词及其上下文之间的位置特征对情感极性的影响,将方面词与上下文分别输入模型,同时利用注意力机制进行计算,提出一种融合位置特征的方面词情感分析方法。本文的主要工作有以下三点:

(1)对 BERT 模型进行两种不同的改进,将方面词及其上下文采用两种文本划分方式,并独立进行权重词向量的表示。

(2)将权重词向量送入多头注意力机制中,计算其文本特征,最后在输出层将两种特征进行融合和池化。

(3)在 SemEval2014 数据集和 Twitter 数据集上进行实验,实验证明方面词情感极性经过融合位置特征后,相较于其他对比实验在准确率和 F1 值上均有不错表现。

1 相关工作

目前针对方面级情感分析任务,大部分研究者利用深度学习中卷积神经网络及其变体进行建模。Vo 等^[3]通过利用基于分布式单词表示的特征,研究了基于目标的 Twitter 情感分类,提出了一种全新的上下文表示方法,同时采用自动句法分析器减少了噪声的影响。Xue 等^[4]提出了一种融合卷积神经网络与门控机制的模型,卷积网络对方面词语情感信息分别建模,门控单元能够根据给定的方面有选择地输出对应的情感特征。注意力机制由于能够获取上下文词语与目标词语之间的重要程度,因此被引入方面级情感分析任务中。王宇欣等^[5]总结了注意力机制在情感分析中的应用。Wang 等^[6]提出了一种结合注意力机制的 LSTM 模型,当不同的方面被输入时,注意力机制能够集中在句子的不同部分,对重要部分进行情感计算。Yang 等^[7]提出了一种基于注意力机制的双向 LSTM 模型,

通过将方面目标与最显著的特征对齐,来获取其情感极性。Liu 等^[8]通过结合注意力模型,将特定方面上下文中每个单词的权重进行单独计算,从而更有效地寻找方面词对应的情感极性。申静波等^[9]将注意力机制与 LSTM 相结合,从而更有效地计算动态变化情况下的词语重要性。

随着模型网络结构的复杂,研究表明上下文的位置特征对方面词的情感极性有着重要的作用。Fan 等^[10]同时计算特定目标与上下文之间的双向交互关系,同时提出一种新的池化方法,减少在融合部分情感极性的损失。Zhai 等^[11]提出采用双层门控循环神经网络对方面词上下文语义信息进行充分提取,并采用图注意力网络对方面词信息进行更为准确的情感极性判别。杜成玉等^[12]针对方面词与上下文距离较远的情况会丢失信息,提出了一种基于方面词加权的上下文表示方法,并进行循环,最终有效提高了方面词与情感极性的匹配。该文通过两种位置特征的划分,重点研究方面词及其上下文的位置特征信息对输出结果的影响,通过 BERT 预训练模型进行词向量表示,多头注意力机制对特征进行提取,并进行池化与拼接,从而获得更为准确的特征信息,进一步提升分类准确性。

2 融合位置特征的方面级情感分类模型

2.1 模型结构

该模型主要由三部分组成:词向量层、特征层、输出层。词向量层通过两种对文本的划分方式,将方面词及其上下文进行划分,分别送入到两个 BERT 模型中,进行词向量的表示。特征层通过联合多头注意力机制,对向量矩阵分别进行特征表示。输出层将四个特征表示进行拼接和平均池化,并进行最终情感极性的分类。具体模型框架如图 1 所示。

2.2 词向量层

采用 BERT 预训练模型完成词向量的表示,能够获得更深层次的文本特征信息,该文将方面词及其上下文采用两种不同的方式进行划分,并输入到两个独立的 BERT 模型进行词向量的表示。

(1)基于方面间隔的上下文划分方法。

根据不同方面词在文本中的位置特征,将两个方面词之间的文本同时视为前一个方面词的后文及后一个方面词的前文,据此对 BERT 模型进行微调与训练。如对“In addition, the food is very good and the prices are reasonable.”,当以方面词进行间隔时,其位置关系如图 2 所示。

(a)以方面词为分隔的前文部分:“[CLS]+前文内容+[SEP]”;

(b)方面词部分:“[CLS]+方面词内容+[SEP]”;

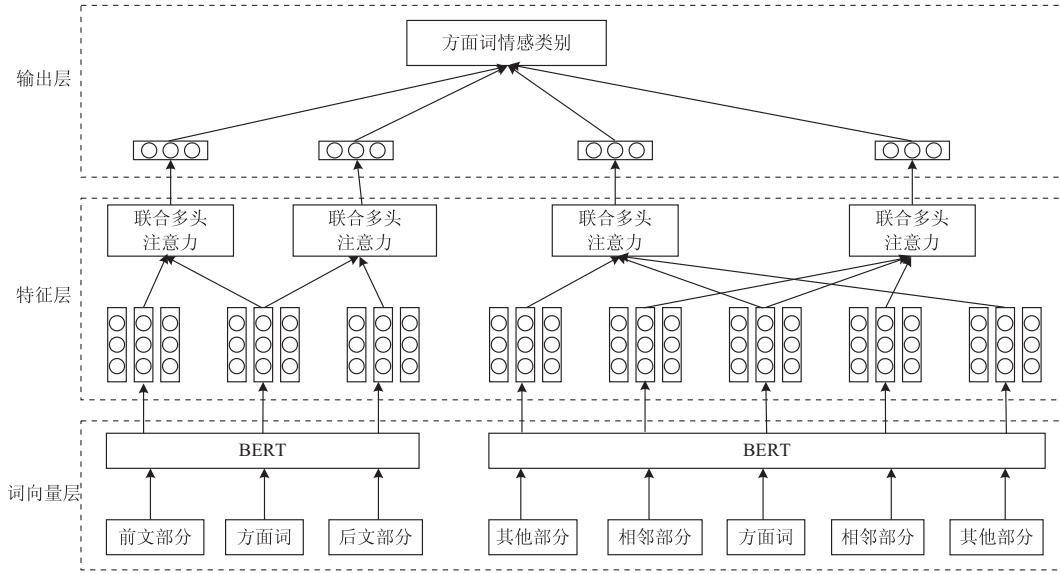


图1 融合位置特征的方面级情感分析模型结构



图2 基于方面词间隔的位置关系

(c)以方面词为分隔的后文部分:“[CLS]+后文内容+[SEP]”。

其中,[CLS]和[SEP]用来将文本进行分割,将分割后形成的部分独立输入 BERT 模型。将特定方面词及其前后文经过 BERT 网络后,得到了对应的词向量:

(a)前文部分词向量: $e_F = [e_F^1, e_F^2, \dots, e_F^f]$;

(b)方面词向量: $e_M = [e_M^1, e_M^2, \dots, e_M^m]$;

(c)后文部分词向量: $e_B = [e_B^1, e_B^2, \dots, e_B^b]$ 。

(2)基于单词距离的上下文划分方法。

根据方面词与前后文之间的位置特征,将文本分为方面词部分、相邻部分和其他部分,根据部分不同,对词向量进行权重的设置,以此对 BERT 模型进行微调与训练。

(a)其他部分:“[CLS]+其他内容+[SEP]”;

(b)相邻部分:“[CLS]+相邻内容+[SEP]”;

(c)方面词部分:“[CLS]+方面词内容+[SEP]”。

如式(1)所示,当 $d_i \leq 0$ 时,为方面词部分。当 $0 < d_i \leq \lambda$ 时,为相邻部分。当 $\lambda < d_i$ 时,为其他部分,其中, λ 为相邻部分长度。

$$d_i = |i - o_a| - \left\lfloor \frac{l_a}{2} \right\rfloor \quad (1)$$

其中, i 表示单词在句子中的位置, o_a 表示方面词的中心位置, l_a 表示方面词的长度。针对其他部分和相邻部分, d_i 表示方面词的上下文距离方面词最近的距离长度。如对“In addition, the food is very good and the prices are reasonable.”,当相邻部分长度设置为2时,其位置关系如图3所示。



图3 基于单词距离的位置关系

对于方面词部分,设置其权重系数为1。对于相邻部分的内容和其他部分的词,随着 d_i 的增大上下文词的权重系数是动态衰减的。但对相邻部分和其他部分的衰减策略会有差异,在相邻部分,设定所有词都比方面词部分的重要性小,但都比其他部分的重要性大,衰减程度小于其他部分的衰减程度,因此设定衰减系数为 α ($\beta < \alpha < 1$)。对于其他部分的词,设定所有的词的重要性都小于方面词部分及相邻部分,并会产生大幅度衰减,因此设定衰减系数 β ($0 < \beta < \alpha$)。将相同类型的部分进行拼接,并对其进行权重的设置,最终经过 BERT 网络后,得到最终对应的权重词向量:

(a)其他部分词向量: $e_o = ([e_o^1, e_o^2, \dots, e_o^o] \cdot 1)$;

(b)相邻部分词向量: $e_l = ([e_l^1, e_l^2, \dots, e_l^i] \cdot ((1 - \frac{d_i}{n})\alpha))$;

$$\frac{d_i}{n})\alpha) ;$$

(c)方面词部分词向量: $e_M = ([e_M^1, e_M^2, \dots, e_M^m] \cdot ((1 - \frac{d_i}{n})\beta))$ 。

2.3 特征层

注意力机制将一个键序列 $k = \{k_1, k_2, \dots, k_n\}$ 和一个查询序列 $q = \{q_1, q_2, \dots, q_n\}$ 映射到一个输出序列 o ,如式(2)所示:

$$o = \text{Attention}(k, q) = \text{softmax}(f_s(k, q))k \quad (2)$$

其中, f_s 为对齐函数,该函数可以学习 q_j 和 k_i 之间的语义相关性,如式(3)所示:

$$f_s(k_i, q_j) = \tanh([k_i; q_j]W_{\text{att}}) \quad (3)$$

其中, W_{att} 是一个可学习的权重,在模型的训练过程中

进行更新迭代。多头注意力机制可以在多个并子空间中学习 n 个不同的权重,将该 n 个输出拼接,并映射到指定的隐藏维度 D_{hid} ,计算方法如式(4)与式(5)所示:

$$MHA(k, q) = \text{Concat}(o^1, o^2, \dots, o^n) W_{mh} \quad (4)$$

$$o^i = \text{Attention}^i(k, q) \quad (5)$$

其中, Concat 为不同注意力的连接, $W_{mh} \in \mathbb{R}^{D_{mh} \times D_{mh}}$, $o^i = \{o_1^i, o_2^i, \dots, o_m^i\}$ 为多头注意力中的第 i 输出, $i \in [1, n]$ 。联合多头注意力的键序列和查询序列不同 $q \neq k$,可以注意到序列之间的交互关系,关注方面词的相关词,更准确地得到其情感极性。

通过给定一个方面词前文或后文词向量矩阵 e^C ,以及一个方面词向量矩阵 e^M ,联合多头注意力机制可以计算出前后文以及方面词相互关系表示 r^{inter} ,如式(6)所示:

$$r^{\text{inter}} = MHA(e^C, e^M) \quad (6)$$

其中, $r^{\text{inter}} = \{r_1^{\text{inter}}, r_2^{\text{inter}}, \dots, r_m^{\text{inter}}\}$,在特征层中,首先将前文部分的词向量 e^F 和后文部分的词向量 e^B 分别与方面词向量矩阵进行联合多头注意力,如式(7)与式(8)所示:

$$r_F^{\text{inter}} = MHA(e^F, e^M) \quad (7)$$

$$r_B^{\text{inter}} = MHA(e^B, e^M) \quad (8)$$

得到前文与实体的关系表示 r_F^{inter} 及后文与实体的关系表示 r_B^{inter} 。其次将相邻部分的词向量 e^I 与其他部分的词向量 e^O 分别与方面词向量矩阵进行联合多头注意力,如式(9)与式(10)所示:

$$r_I^{\text{inter}} = MHA(e^I, e^M) \quad (9)$$

$$r_O^{\text{inter}} = MHA(e^O, e^M) \quad (10)$$

得到前文与实体的关系表示 r_I^{inter} 及后文与实体的关系表示 r_O^{inter} 。

2.4 输出层

将特征层获得的结合前文部分的方面词表示 r_F^{inter} ,结合后文部分的方面词表示 r_B^{inter} ,结合相邻部分的方面词表示 r_I^{inter} 以及结合其他部分的方面词表示 r_O^{inter} ,将它们进行池化,并拼接形成最终的特征表示,最后使用 Softmax 函数进行结果输出,如式(11)与式(12)所示:

$$x = \text{Concat}(h_F^{\text{inter}}, h_B^{\text{inter}}, h_I^{\text{inter}}, h_O^{\text{inter}}) \quad (11)$$

$$\hat{y} = \text{softmax}(W_o^T x + b_o) \quad (12)$$

其中, x 为拼接后的特征表示, W_o^T 为权重系数矩阵, b_o 为偏置向量, \hat{y} 为预测的情感分类结果。

采用交叉熵损失函数,同时引入 L_2 正则化项,防止模型出现过拟合现象,如式(13)所示:

$$L(\theta) = - \sum_i^c \hat{y}_i \log y_i + \lambda \sum_{\theta \in \Theta} \theta^2 \quad (13)$$

其中, C 为情感类别的个数, \hat{y}_i 为预测的情感分类结果, y_i 为真实的情感分类结果, λ 为 L_2 正则化参数, Θ 为模型的参数集。

3 实验

为了验证文中模型的有效性,采用 SemEval2014 Task4^[13] 数据集和 Twitter^[14] 数据集对模型进行实验, SemEval2014 Task4 数据集包含 Laptop-14 和 Restaurant-14 两个数据集,其中 Laptop-14 是笔记本电脑和电子设备领域的评论, Restaurant-14 是餐馆和服务领域的用户评论。Twitter 数据集是经过处理的用户推文数据集。具体数据集设置如表 1 所示。

表 1 数据集设置

数据集	数据集分类	积极	中性	消极
Laptop-14	train	994	169	870
	test	341	169	128
Restaurant-14	train	2 164	637	807
	test	728	196	196
Twitter	train	1 561	3 127	1 560
	test	173	346	173

3.1 数据设置

实验中采用了 BERT Adam 优化方法,学习率为 0.001,输出层采用 softmax 进行分类,采用交叉熵作为损失函数,采用随机梯度下降算法更新权重,同时在对方面词上下文部分划分中,以 α 为 0.72, β 为 0.56 效果最好,具体详细参数如表 2 所示。

表 2 超参数设置

参数	设定值
Dropout	0.5
Learning-rate	2e-5
Batch_size	16
epoch	20
Max_seq_length	128
词嵌入维度	768
Hidden size	768
L_2 正则化权值	1e-5
α	0.72
β	0.56

3.2 评价指标

文本分类中常用的评测指标有精确率 P (Precision)、召回率 R (Recall)、F1 值 (F1-measure) 和准确率 A (Accuracy)。由于该文使用的两个数据集都比较平衡,所以使用测试集上的 F1 值和准确率 A 作为实验结果的评测指标。评测指标定义如式(14)~式(17)所示:

$$P = \frac{TP}{TP + FP} \quad (14)$$

$$R = \frac{TP}{TP + FN} \quad (15)$$

$$F1 = \frac{2PR}{P + R} \quad (16)$$

$$A = \frac{TP + TN}{TP + FN + TN + FP} \quad (17)$$

其中,TP表示实际为正向且预测为正向的数量,FP表示实际为负向但预测为正向的数量,TN表示实际为负向且预测为负向的数量,FN表示实际为正向但预测为负向的数量。TP、FP、TN、FN的分类矩阵如表3所示。

表3 分类矩阵

真实结果	预测结果	
	正向	负向
正向	实际为正向且预测为正向(TP)	实际为负向且预测为正向(FP)
负向	实际为正向且预测为负向(FN)	实际为负向且预测为负向(TN)

3.3 消融实验

本节设计了两个消融实验对模型中各个组件进行计算,首先对相邻区域长度进行优化,以此计算在数据集中性能最优的相邻区域,其次对多头注意力机制中头数的设置进行计算,以此验证注意力头数对模型性能的影响。

3.3.1 相邻区域长度

相邻区域长度是影响词向量权重的重要参数。经过实验,不同长度的相邻区域对准确率和F1值有重要的影响,如图4所示。其中当长度小于2时,在三个数据集中均呈上升趋势,当长度等于2时,准确率的效果最好,当长度大于2时,在三个数据集中呈下降趋势。这说明越是临近方面词的单词对其情感极性的判断越重要,但随着距离超过最优值,却又会引入不必要的干扰信息,最终将模型的相邻区域长度设置为2。

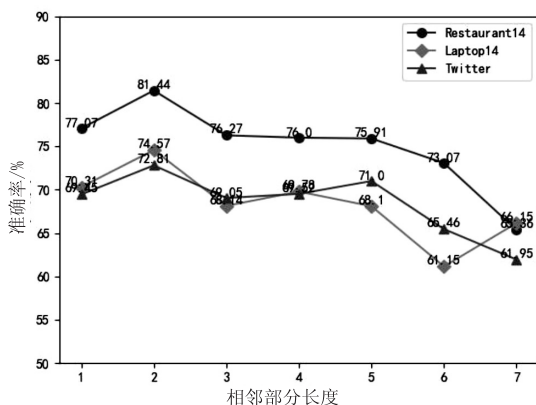


图4 相邻部分长度对准确率的影响

3.3.2 注意力机制模块数

注意力机制的头数是影响信息特征交互的重要参数。经过实验,将多头注意力机制头数分别设置为1~8,可以得到注意力机制头数的选择能够对实验结果有着重要的影响,如图5所示。当注意力头数小于5时,在三个数据集中均呈上升趋势,当长度等于5时,准确率效果最好,当长度大于5时,在三个数据集中均呈下降趋势。这说明当注意力机制头数过少时,无法有效地进行交互特征的计算,而当注意力头数过多时,则会引入大量参数,造成模型的过拟合,最终将模型注意力机制头数设置为5。

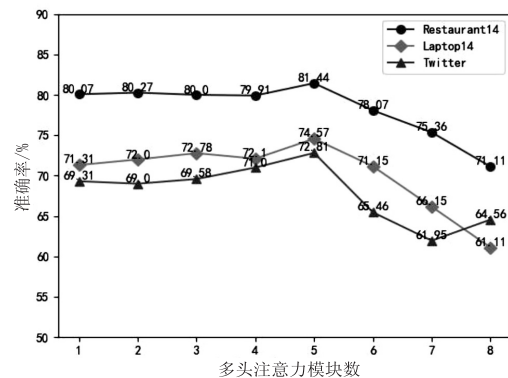


图5 多头注意力模块数对准确率的影响

3.4 对比实验

为了验证文中模型的有效性,设置了如下对比实验:

(1) Majority: 基线实验,将训练集样本中的情感极性最大的值分配给样本中的所有方面。

(2) SVM^[15]: 通过传统机器学习,从训练集样本中进行特征提取,采用支持向量机对特征进行分类。

(3) CNN^[16]: 传统的卷积神经网络模型,通过设置不同大小的卷积核,可以有效地捕获局部特征。

(4) BiLSTM^[17]: 双向长短时记忆网络,能够同时捕捉方面词前后的依赖特征。

(5) BERT^[18]: 预训练语言模型,通过双层Transformer架构捕获方面词上下文的特征信息,最后通过隐藏层获取全局分类特征信息。

表4展示了融合位置特征的方面级情感分析模型在各个数据集上与近年相关研究的效果对照,相较于对比实验在各个数据集上均有不同的提升。

4 结束语

该文提出了融合位置特征的方面级情感分析模型,该模型首先对文本进行两种区域划分,并使用了两个独立的BERT预训练模型对词向量进行编码与权重计算,其次通过多头注意力机制对方面词与上下文的特征信息进行交互,最后将特征信息进行融合输出。

与现有的模型相比,该模型能够有效地根据方面词与上下文之间位置特征的不同对信息进行计算,从而使模型能够更加准确地计算出情感极性判断结果,相对

于其他对比模型,该模型明显提高了准确率与 F1 值。在后续的工作中,将充分考虑对文本的词性或依存关系等文本特征进行计算,从而产生更好的效果。

表4 模型综合性能评估

模型	Laptop-14		Restaurant-14		Twitter	
	A / %	F1 / %	A / %	F1 / %	A / %	F1 / %
Majority	69.20	63.09	73.97	69.42	63.93	67.89
Feature-based SVM	70.49	—	80.16	—	63.40	63.30
CNN	68.33	62.29	76.43	75.23	70.97	71.88
BiLSTM	72.28	65.10	78.13	77.47	69.56	70.70
BERT	72.62	66.72	81.36	82.18	71.05	69.45
Model proposed	74.57	71.44	81.44	83.57	72.81	72.41

参考文献:

- [1] 黄萱菁,张奇,吴苑斌.文本情感倾向分析[J].中文信息学报,2011,25(6):118-126.
- [2] HAI H D, PRASAD P, MAAG A, et al. Deep learning for aspect-based sentiment analysis: a comparative review[J]. Expert Systems with Applications, 2019, 118: 272-299.
- [3] VO D T, ZHANG Y. Target-dependent twitter sentiment classification with rich automatic features[C]//Proceedings of the 24th international conference on artificial intelligence. North America: AAAI, 2015: 1347-1353.
- [4] XUE W, LI T. Aspect based sentiment analysis with gated convolutional networks[C]//Proceedings of the 56th annual meeting of the association for computational linguistics. Melbourne: ACL, 2018: 2514-2523.
- [5] 王宇欣,方浩宇,张伟,等.注意力机制在情感分析中的应用研究[J].计算机技术与发展,2022,32(4):193-199.
- [6] WANG Y, HUANG M, ZHAO L. Attention-based lstm for aspect-level sentiment classification[C]//Proceedings of the 2016 conference on empirical methods in natural language processing. Austin: EMNLP, 2016: 606-615.
- [7] YANG M, TU W, WANG J, et al. Attention based LSTM for target dependent sentiment classification[C]//Proceedings of the thirty-first AAAI conference on artificial intelligence. San Francisco: AAAI, 2017: 5013-5014.
- [8] 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. Valencia: EACL, 2017: 572-577.
- [9] 申静波,李井辉,孙丽娜.注意力机制在评论文本情感分析中的应用研究[J].计算机技术与发展,2020,30(7):169-173.
- [10] FAN Feifan, FENG Yansong, ZHAO Dongyan. Multigrained attention network for aspect-level sentiment classification[C]//Proceedings of conference on empirical methods in natural language processing. Stroudsburg: ACL, 2018: 3433-3442.
- [11] ZHAI Sheping, CHAI Yun, WANG Huizhen, et al. Aspect-level sentiment joint detection based on graph attention network[C]//Proceedings of advances in natural computation, fuzzy systems and knowledge discovery. Guiyang: Springer, 2021: 760-768.
- [12] 杜成玉,刘鹏远.基于螺旋注意力网络的方面级别情感分析模型[J].中文信息学报,2020,34(9):70-77.
- [13] CHEN P, SUN Z Q, BING L D, et al. Recurrent attention network on memory for aspect sentiment analysis[C]//Proceedings of the conference on empirical methods in natural language processing. Stroudsburg: ACL, 2017: 452-461.
- [14] DONG Li, WEI Furu, TAN Chuanqi, et al. Adaptive recursive neural network for target-dependent twitter sentiment classification[C]//Proceedings of the 52nd annual meeting of the association for computational linguistics. Baltimore: ACL, 2014: 49-54.
- [15] KIRITCHENKO S, ZHU Xiaodan, CHERRY C, et al. NRC-Canada-2014: detecting aspects and sentiment in customer reviews[C]//Proceedings of the 8th international workshop on semantic evaluation. Stroudsburg: ACL, 2014: 437-442.
- [16] KIM Y. Convolutional neural networks for sentence classification[C]//Proceedings of the 2014 conference on empirical methods in natural language processing. Doha: EMNLP, 2014: 1746-1751.
- [17] CHENG Yu-Ya, YEH Wen-Chao, CHEN Yan-Ming, et al. Using valence and arousal-infused Bi-LSTM for sentiment analysis in social media product reviews[C]//Proceedings of the 33rd conference on computational linguistics and speech processing. Taiwan, China: ROCLING, 2021: 210-217.
- [18] DEVLIN J, CHANG M W, LEE K, et al. Bert pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics; human language technologies. Minneapolis: NAACL, 2019: 4171-4186.