

基于高斯分布引导位置相关性权重的情感分类^①



赵 振¹, 朱振方¹, 王文玲²

¹(山东交通学院 信息科学与电气工程学院 (人工智能学院), 济南 250357)

²(鲁东大学 文学院, 烟台 264025)

通信作者: 朱振方, E-mail: zhuzf_sdjtu@126.com

摘 要: 针对当前情感分类方法通常忽略不同单词之间相对位置特征, 导致模型难以学习到单词的最佳位置表示. 为了解决这一问题, 提出一种基于高斯分布引导位置相关性权重的情感分类算法. 首先, 计算每个单词与其他单词之间的位置相关性; 其次, 利用改进的高斯分布函数对位置相关性进行建模, 并将其结果与单词的特征向量相乘, 以生成单词的位置感知表示; 最后, 将算法集成到传统模型中以验证其有效性. 实验结果表明, 所提方法较传统模型获得了更高的准确率, 在域内、域外和对抗评估指标上分别提高了 2.98%、5.02% 和 10.55%, 表明其具有较好的实用价值.

关键词: 位置权重; 距离特征; 情感分类; 自然语言处理

引用格式: 赵振, 朱振方, 王文玲. 基于高斯分布引导位置相关性权重的情感分类. 计算机系统应用, 2023, 32(11): 232-239. <http://www.c-s-a.org.cn/1003-3254/9290.html>

Gaussian Distribution Guided Position Relevance Weight for Sentiment Classification

ZHAO Zhen¹, ZHU Zhen-Fang¹, WANG Wen-Ling²

¹(School of Information Science and Electrical Engineering (School of Artificial Intelligence), Shandong Jiaotong University, Jinan 250357, China)

²(School of Chinese Language and Literature, Ludong University, Yantai 264025, China)

Abstract: The current sentiment classification methods often ignore the relative positional features between different words, which makes it difficult for the model to learn the best positional representation of words. To solve this problem, a sentiment classification algorithm based on Gaussian distribution guided position relevance weight is proposed. First, the positional relevance between each word and other words is calculated. Second, the positional relevance is modeled by using an improved Gaussian distribution function, and the results are multiplied with the feature vectors of the words to generate a positional-aware representation of the words. Finally, the algorithm is integrated into the traditional model to verify its effectiveness. The experimental results show that the proposed method obtains higher accuracy than the traditional model, with improvements of 2.98%, 5.02%, and 10.55% in terms of in-domain, out-of-domain, and adversarial evaluation metrics, respectively, indicating its excellent practical value.

Key words: position weight; distance feature; sentiment classification; natural language processing (NLP)

随着互联网的迅速发展, 人们越来越热衷于在网络上发表个人观点, 因此产生了大量包含个人情感的文本数据. 这些文本数据中的情感倾向可能是正面、负面或中性的. 为了更准确地分析文本中的情感倾向,

需要采用方面级情感分析技术. 方面级情感分析是文本情感分析的一种扩展形式, 旨在分析文本中对特定方面的情感倾向. 在自然语言处理 (natural language processing, NLP) 领域中, 方面级情感分析是一个重要

① 基金项目: 山东省自然科学基金 (ZR2021MF064, ZR2021QG041)

收稿时间: 2023-04-01; 修改时间: 2023-05-06, 2023-05-23; 采用时间: 2023-06-01; csa 在线出版时间: 2023-08-21

CNKI 网络首发时间: 2023-08-22

的研究任务,其应用广泛,包括社会舆情分析、产品评价、广告营销等领域.尤其是在电商平台和在线评论系统中,它可以帮助商家了解消费者的真实需求和感受,从而优化产品和服务,提高用户体验.方面级情感分析主要包括方面提取和情感分类两个步骤.方面提取是指从文本中识别与特定方面相关的词汇或短语,例如,在餐厅评价中,“服务”“菜品”和“价格”等都可以作为方面.情感分类则是将文本的情感倾向分为正面、负面或中性.例如:“The price is reasonable although the service is poor.”,该任务需要预测“price”正面的情感类别,以及“service”负面的情感类别.方面级情感分析需要同时考虑方面和情感,因此比传统情感分析更加复杂和具有挑战性.

近年来,深度学习技术已在NLP领域取得了令人瞩目的成绩,卷积神经网络(convolutional neural network, CNN)^[1-3]和长短时记忆网络(long short-term memory, LSTM)^[4,5]均在情感分类任务上表现出良好的性能.后来,为了使模型能够记忆更多的长距离依赖信息,又有学者引入了注意力机制^[6],注意力机制是对输入权重分配的关注,分配的权重越大,表示其对应的信息越重要.例如,汪林等^[7]提出一种基于多级多尺度特征提取的CNN-BiLSTM模型,该方法利用预训练的中文词向量模型和嵌入层微调获取词级特征,并使用多尺度短语级/句子级特征表征模块获取短语/句子级特征.多尺度短语级特征表征模块采用不同卷积核尺寸的卷积网络

获取不同尺度的短语级特征,最后使用多级特征融合方法将不同级别和尺度的特征进行融合,形成多级联合特征;郭磊等^[8]提出了一种RI-GAT模型.该模型通过LSTM网络学习句子语义特征,结合位置信息生成新特征,并提取各方面词和观点词之间依存关系,从而高效地利用句法依存信息和位置信息.然而,上述神经网络仅能处理一维的序列建模任务,对于非结构化的二维数据,如树结构、图结构等,则需要使用图卷积网络(graph convolutional network, GCN)^[9,10],它可以自动地学习每个节点的特征以及节点与节点之间的关联信息,其核心思想是利用边的信息对节点信息进行聚合从而生成新的节点表示.例如,Zhu等^[11]提出一种知识引导的多粒度图卷积神经网络,该网络使用一种由概念知识指导的掩码机制来处理更多方面词,通过多粒度注意力机制,增强了方面词与情感词之间的交互.Liang等^[12]提出了一种基于SenticNet^[13]的图卷积网络,它考虑了上下文词与方面词之间的依赖关系以及意见词与方面之间的情感信息,通过整合SenticNet的情感知识来进一步增强句子的依赖图.

虽然这些神经网络模型在情感分类任务上取得了巨大的成功,但仍然存在部分缺陷.一般来说,文本中方面词相关的情感描述应该与该方面词在一维距离上非常接近.以上述句子为例,单词与方面词之间的距离如图1所示,其中深灰色区域代表方面词及其位置.

The	price	is	reasonable	although	the	service	is	poor	.
-1	0	1	2	3	4	5	6	7	8
The	price	is	reasonable	although	the	service	is	poor	.
-6	-5	-4	-3	-2	-1	0	1	2	3

图1 方面词与其他单词的相对距离例子

在图1中,“price”和“service”的观点词分别是“reasonable”和“poor”,距离各自的方面词都非常接近.为了统计这种情况在真实数据集上的分布概率,本文将一对方面词与观点词之间的相对距离除以相应句子的长度定义为方面接近度,并统计了4个不同数据集上两者的相对距离及其密度分布情况,结果如图2所示.可以看出,在所有句子中方面词与观点词的相对距离较小的占据较高比例,仅有小部分文本存在相对距离较大的情况.然而,先前的工作对于两者的相对位置信息考虑并不充分,一是赋予每个单词地相对位置权

重较为直接,没有充分利用单词的频率分布信息,二是过于弱化长文本中边界单词的位置权重,导致模型只关注方面词附近的单词,忽略了方面词与观点词之间距离较长时的情况.

针对上述问题,本文提出一种基于高斯分布引导位置相关性权重的情感分类算法(Gaussian distribution guided position relevance weight, G-POS),旨在进一步增强方面上下文的重要性,减少依赖解析过程中可能产生的数据噪声和偏差.为了实现这一目标,G-POS算法根据每个单词与方面词的位置接近程度,以相对平

滑的方式为每个单词分配位置权重. 在优化方面词与观点词距离较远时的权重赋予形式时, 该算法采用了一种特殊的高斯分布引导方法, 以确保位置权重分配的准确性和稳定性. 此外, 为了验证 G-POS 算法的有效性, 我们进行了一系列的实验, 并将本文算法与现有的情感分类算法进行比较. 实验结果表明, 结合本文算法的模型可进一步提高情感分类准确率. 这表明 G-POS 算法在处理情感分类问题时具有较高的实用性和有效性, 并有望在实际应用中发挥重要的作用.

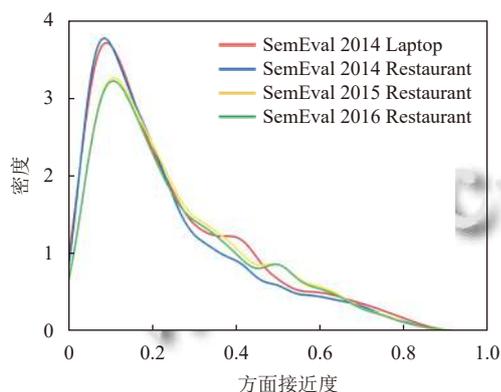


图2 不同数据集下的方面接近度分布

1 相关研究

近年来, 许多学者对文本位置权重进行了相关研究, 并取得了较好的效果, 这些研究的核心思想是越接近方面词的单词越有可能是其实际观点词. 若给定一个包含 n 个单词的句子, 方面词的起始位置为 τ , 方面词的长度为 m , 在关于文本位置权重的研究中, 有下列几种方式.

首先, Chen 等^[14] 提出 RAM 模型, 通过采用多头注意力机制来捕获被长距离分隔的情感特征, 从而使模型对无关信息具有更强的鲁棒性, 在获取单词的位置加权信息上, 作者将距离定义为单词和目标之间的单词数量, 第 i 个单词的位置权重 w_i 的计算方式如下:

$$w_i = 1 - \frac{|i - \tau|}{n} \quad (1)$$

该方法计算较为简单, 并没有考虑多个方面词对单词位置权重的影响.

之后, 有学者将上述计算方法做了进一步的改进, Li 等^[15] 提出一种新的架构 TNet, 该方法首先通过双向 RNN 获取单词嵌入表示, 然后采用 TNet 组件来生成句子中特定于目标的单词表示, 最后使用 CNN 从中

提取显著特征. 但是, 普通 CNN 可能会将方面词与不相关的观点词相关联, 为了解决这个问题, 作者提出一种邻近策略, 具体而言, 其通过计算第 i 个单词与方面词之间的位置相关性 p_i 来判断每个单词所分配的权重:

$$p_i = \begin{cases} 1 - \frac{\tau + m - i}{C}, & 0 \leq i < \tau + m \\ 1 - \frac{i - \tau}{C}, & \tau + m \leq i \leq n \end{cases} \quad (2)$$

其中, C 为预定义的常数. 该方法考虑了句子中多个方面词的情况, 基本思想是位置权重与单词到方面词的相对距离成比例的增加或减少.

后来, Zhang 等^[16] 提出 ASGCN 模型, 该方法通过在句子的依赖树上构建一个图卷积网络, 充分利用了句法信息和单词依赖, 解决了模型可能将方面词无关的上下文错误判断为方面情感线索的问题. 值得注意的是, 在将文本的隐藏状态表示送入图卷积网络之前, 该方法进行了位置权重感知计算, 第 i 个单词的位置权重 p_i 表示如下:

$$p_i = \begin{cases} 1 - \frac{\tau - i}{n}, & 0 \leq i < \tau \\ 1 - \frac{i - \tau - m + 1}{n}, & \tau + m \leq i < n \end{cases} \quad (3)$$

该方法确定了式 (2) 中常数 C 的值, 并优化了上下文与方面词之间相对距离的计算方式. 但该方法仅考虑到上下文单词的权重, 未对方面词本身的权重分配进行深入研究.

最近, Ma 等^[17] 针对先前方法存在鲁棒性较差的问题, 提出一种位置偏置权重算法 POS4ASC, 该方法提出位置偏置权重和位置偏置 dropout 两种机制, 降低了模型对情感描述词错误判断的概率, 进一步提高了模型的鲁棒性. 在位置偏置权重的计算方式上, 其第 i 个单词的位置权重 p_i 如下所示:

$$p_i = \begin{cases} 1 - \frac{\tau - i}{n - m}, & 0 \leq i < \tau \\ \frac{1}{n - m}, & \tau \leq i < \tau + m \\ 1 - \frac{i - \tau - m + 1}{n - m}, & \tau + m \leq i < n \end{cases} \quad (4)$$

该方法对式 (3) 做了进一步的改进, 为方面词本身分配了位置权重, 优化了上下文单词与方面词的相对距离在句子中的占比形式, 进一步提高了模型的性能.

本文提出的位置相关性权重算法与现有的工作既有联系又有区别. 联系在于均基于越靠近方面词的单

词越有可能是其情感描述词的思想,区别在于本文采用非线性位置权重计算方法,使单词之间的权重变化更为平滑,弥补了短文本中位置权重变化较大,长文本中边际单词位置权重分配较低的缺陷,进一步提高了模型的鲁棒能力。

2 基于高斯分布引导的位置相关性权重

2.1 高斯分布

一般情况下,若一个变量 x 服从位置参数为 μ 、尺度参数为 σ 的高斯分布,则其概率密度函数可表示为:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (5)$$

其中, μ 为数学期望,决定了分布的位置, σ 为标准差,决定了分布的幅度. 特别的,当位置参数 $\mu = 0$, 尺度参数 $\sigma = 1$ 时称为标准高斯分布. 该函数具有以下特征: (1) 函数曲线关于 $x = \mu$ 对称. (2) σ 越大, 总体分布越分散; σ 越小, 总体分布越集中.

基于上述特征,本文在高斯分布的基础上进行了如下改进.

(1) 为了使权重变化区间处于区间 $(0, 1)$, 将式 (5) 变换为:

$$f(x) = \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (6)$$

(2) 基于高斯分布函数的特征 1, 将参数 μ 取值为方面词所在位置; 基于高斯分布函数的特征 2, 将 $2\sigma^2$ 整体视为 λ 并由其控制权重分布幅度, 将式 (6) 变换为:

$$f(x) = \exp\left(-\frac{(x-\tau)^2}{\lambda}\right) \quad (7)$$

其中, τ 表示方面词的开始位置, λ 表示可变参数. 基于上述式 (7), 本文在其最大值处按方面词长度对其进行截断, 即式 (7) 的值域作为位置权重并分配给上下文单词, 对于方面词则赋予常数值权重.

2.2 算法描述

给定一个包含 n 个单词的句子, 其中方面词的长度为 m , 则该句子可表示为 $S = \{w_0, w_1, \dots, w_\gamma, w_{\gamma+1}, \dots, w_{\gamma+m-1}, \dots, w_{n-1}\}$, 其中 w_i 表示第 i 个上下文单词, θ 表示方面词的开始位置. 首先, 通过采用预先训练的语言模型, 可以将句子表示为 $V = \{d_0, d_1, \dots, d_\theta, d_{\theta+1}, \dots, d_{\theta+m-1}, \dots, d_{n-1}\}$. 其次, 使用 G-POS 算法来优化 V 并增强其表示: $E = \{h_0, h_1, \dots, h_\theta, h_{\theta+1}, \dots, h_{\theta+m-1}, \dots, h_{n-1}\}$. 最后, 将 E 合并到模型中, 即模型的下游结构都是建立在 E 而

非 V 的基础上, 以预测与不同方面词相关的情感极性.

通常, 一个方面词的情感极性由其上下文即围绕该方面的词语所决定. 因此, 可以利用相对位置信息来计算上下文单词的权重, 目的是降低那些远离方面的单词的重要性. 形式上, 本文计算第 i 个单词的位置相关性权重 $p_i \in (0, 1)$ 可表示为:

$$p_i = \begin{cases} \exp\left(-\frac{(i-\gamma)^2}{\alpha n}\right), & 0 \leq i < \gamma \\ \frac{1}{n}, & \gamma \leq i < \gamma+m \\ \exp\left(-\frac{(i-\gamma-m+1)^2}{\alpha n}\right), & \gamma+m \leq i < n \end{cases} \quad (8)$$

其中, γ 表示方面词的开始位置, m 表示方面词的长度, α 表示预定义的常数. 在 $[0, \gamma)$ 区间内, 相对距离的计算方式为上文单词位置与方面词开始位置的差值; 在 $[\gamma+m, n)$ 区间内, 相对距离的计算方式为方面词结束位置 (即 $\gamma+m-1$) 与下文单词位置的差值; 在 $[\gamma, \gamma+m)$ 区间内, 方面词始终被赋予较低的常数值权重.

图 3 显示了本文算法在实际情况下的函数图像, 曲线中每个点的纵坐标代表每个单词被赋予的位置权重, 中间部分及其两侧分别为方面词与上下文词被分配的权重.

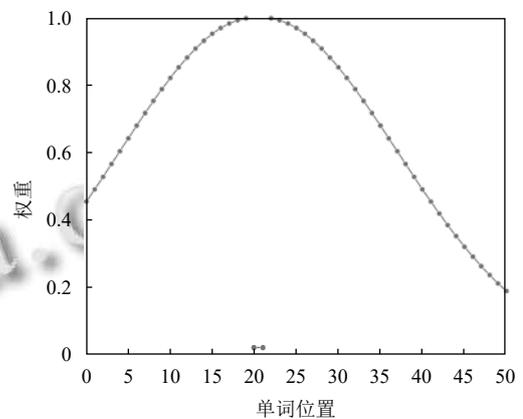


图 3 单词位置相关性权重函数图 ($\gamma=20; m=2; \alpha=10$)

本文所提算法具有以下特点: (1) 对于距离方面较远的单词, 其权重分配较为平缓. (2) 对于距离方面较近的单词, 其权重变化较大, 目的是使单词的表示更加明显, 模型更容易辨别每个单词的差异. (3) 与方面词距离相同的上下文单词具有相同的位置权重. (4) 可灵活调节 α 值以改变权重分布的幅度.

应用基于高斯分布引导位置相关性权重算法的具体过程如算法 1 所示, 其中位置权重计算函数 function 由式 (8) 所述定义.

算法 1. 基于高斯分布引导位置相关性权重算法

输入: 文本词向量矩阵 E , 方面词起止位置 (m, n) , 文本长度 t , 方面词长度 a .

输出: 综合位置权重的词向量矩阵 X .

```

1 根据  $x$  的行数和列数分别确定  $batch\ size$  和  $sequence\ length$ 
2 初始化全 0 位置权重矩阵  $weight[batch\ size \times sequence\ length]$ 
3 定义位置权重计算函数  $function$ (文本长度, 单词位置, 方面词位置)
4 FOR  $i \leftarrow 0$  to  $batch\ size$  DO
5   FOR  $j \leftarrow 0$  to  $m-1$  DO
6      $weight[i][j] \leftarrow function(t, j, m)$ 
7   END FOR
8   FOR  $j \leftarrow m$  to  $n$  DO
9      $weight[i][j] \leftarrow 1/t$ 
10  END FOR
11  FOR  $j \leftarrow n+1$  to  $t-1$  DO
12     $weight[i][j] \leftarrow function(t, j, n)$ 
13  END FOR
14 END FOR
15  $X = weight \times E$ 
16 RETURN  $X$ 

```

3 实验分析

3.1 数据集和评估标准

本文在 SemEval 2014^[18] 数据集的两个子领域: Laptop 和 Restaurant 上进行了实验, 同时使用了 Xing 等^[19] 提出的方面鲁棒性测试集 (aspect robustness test set, ARTS), 该数据集在原始 SemEval 2014 数据的基础上, 通过实施 3 种对抗策略, 分别将 Restaurant 和 Laptop 域的实例增加到 3 530 和 1 877. 由于 ARTS 只包含测试集而不包含训练集, 所以本文仅使用其进行测试. 表 1 显示了数据集的统计信息.

表 1 数据集的统计信息

Dataset	Train/Test	Positive	Negative	Neutral
Restaurant	Train	2 164	805	633
	Test	728	196	196
Laptop	Train	987	866	460
	Test	341	128	169
ARTS-Restaurant	Test	1 953	1 104	473
ARTS-Laptop	Test	883	407	587

对于实验的评估, 本文参考了文献 [17] 的评价标准, 将模型的训练及测试方式分为下列 3 种以评估算法的鲁棒性.

(1) 域内评估 (in-domain, I.D.): 模型在一个域上进行训练和测试.

(2) 域外评估 (out-of-domain, O.O.D.): 模型在一个域上训练, 另一个域上测试.

(3) 对抗评估 (adversarial, Adv.): 模型在 SemEval 数据集上训练, 在对应的 ARTS 数据集上测试.

3.2 实验设置

本文实验在 Ubuntu 20.04 系统上进行, CPU 型号为 Intel Core i9-10900X, GPU 型号为 NVIDIA Quadro RTX 6000, 深度学习框架采用 PyTorch 1.10.0, 所用编程语言为 Python 3.8.

实验参数的设置上, 本文采用 Adam 优化器. 同时, 对于非基于 BERT 的模型, 采用 300 维 GloVe^[20] 作为初始化嵌入, 学习率设置为 $1E-3$, batch size 设置为 64; 对于基于 BERT 的模型, 学习率设置为 $2E-5$, batch size 设置为 16.

3.3 基线模型

为评估算法在模型上的性能, 本文将在上述两个公开数据集上与如下相关模型进行比较.

(1) LSTM^[21]: LSTM 是一种特殊的 RNN, 它使用最后隐藏状态作为句子的最终表示, 并将结果输入到分类器中进行预测.

(2) MemNet^[22]: 该模型使用多个注意力计算层, 在推断某个方面的情感属性时明确捕获每个上下文词的重要性, 从而有助于分类.

(3) AOA^[23]: 该模型在原来 attention 的基础上再加一层 attention, 以表明每个 attention 的重要性, 从而更好地建模方面与上下文的关系.

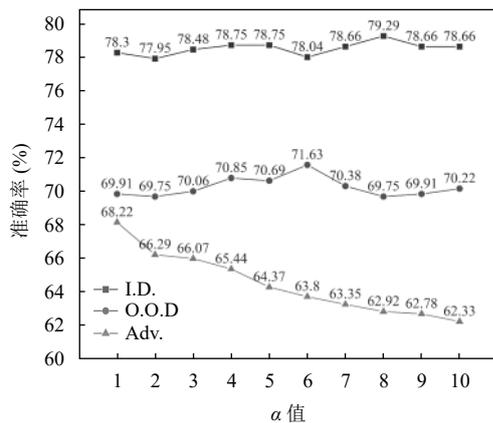
(4) RoBERTa^[24]: 该模型针对 BERT^[25] 训练不足的问题, 减小了关键超参数和训练数据大小的对下游模型的影响, 使用了更多的数据并且训练更加充分.

3.4 设置参数 α 值

在 G-POS 算法中, α 值的设置对模型的性能存在一定的影响. 为了研究不同 α 值对情感分类准确率上的影响, 本文对 G-POS 算法中的参数 α 在不同变换次数下的准确率进行对比分析, 实验在 LSTM 模型及 Restaurant 数据集上进行, 对比结果如图 4 所示.

实验对比了 α 值从 1 到 10 时的准确率情况. 从图 4 中可以看出, 对于 I.D. 和 O.O.D. 上的测试结果, 模型的准确率具有上下波动的趋势, 分别在 $\alpha=8$ 和 $\alpha=6$ 时达到了当前参数范围内的最大准确率, 对于 Adv. 上的测试结果, 模型的准确率随着 α 值的增大而减小, 说明模型的过拟合现象可能越来越明显.

依据上述方法, 本文分别在 MemNet、AOA、RoBERTa 模型上进行实验, 最终确定在各个基线模型上的 α 值, 具体参数值如表 2 所示.

图4 不同 α 值在 LSTM 模型上的准确率

3.5 结果分析

为了验证 G-POS 算法的可行性与有效性, 本文选取了相关工作中 ASGCN 和 POS4ASC 的计算位置权重方法, 分别命名为 ASGCN-pw 和 POS4ASC-pw, 并在 SemEval 2014 数据集上进行实验对比. 所有模型结果均由网络公开代码复现后所得, 实验结果如表 3 所示.

表2 超参数 α 的设置

模型	Restaurant			Laptop		
	I.D.	O.O.D.	Adv.	I.D.	O.O.D.	Adv.
LSTM	8	6	1	3	2	1
MemNet	7	1	1	9	4	1
AOA	9	4	2	9	7	2
RoBERTa	6	5	2	5	10	2

表3 本文算法与其他相关算法在基线模型上的准确率对比 (%)

模型	Restaurant			Laptop		
	I.D.	O.O.D.	Adv.	I.D.	O.O.D.	Adv.
LSTM	78.66	67.08	58.10	71.16	73.04	52.74
+ASGCN-pw	78.93 ($\uparrow 0.27$)	69.75 ($\uparrow 2.67$)	62.18 ($\uparrow 4.08$)	71.63 ($\uparrow 0.47$)	74.55 ($\uparrow 1.51$)	58.23 ($\uparrow 5.49$)
+POS4ASC-pw	78.57 ($\downarrow 0.09$)	69.91 ($\uparrow 2.83$)	62.38 ($\uparrow 4.28$)	73.20 ($\uparrow 2.04$)	74.91 ($\uparrow 1.87$)	58.39 ($\uparrow 5.65$)
+G-POS	79.29 ($\uparrow 0.63$)	71.63 ($\uparrow 4.55$)	68.22 ($\uparrow 10.12$)	74.14 ($\uparrow 2.98$)	76.25 ($\uparrow 3.21$)	63.29 ($\uparrow 10.55$)
MemNet	78.21	65.20	60.17	72.26	73.57	55.78
+ASGCN-pw	79.64 ($\uparrow 1.43$)	67.24 ($\uparrow 2.04$)	67.25 ($\uparrow 7.08$)	73.20 ($\uparrow 0.94$)	74.02 ($\uparrow 0.45$)	59.19 ($\uparrow 3.41$)
+POS4ASC-pw	79.11 ($\uparrow 0.90$)	67.55 ($\uparrow 2.35$)	66.60 ($\uparrow 6.43$)	71.94 ($\downarrow 0.32$)	73.84 ($\uparrow 0.27$)	58.71 ($\uparrow 2.93$)
+G-POS	79.82 ($\uparrow 1.61$)	69.59 ($\uparrow 4.39$)	69.43 ($\uparrow 9.26$)	74.14 ($\uparrow 1.88$)	74.55 ($\uparrow 0.98$)	62.49 ($\uparrow 6.71$)
AOA	80.54	67.71	61.50	73.20	73.84	55.25
+ASGCN-pw	79.82 ($\downarrow 0.72$)	70.53 ($\uparrow 2.82$)	67.88 ($\uparrow 6.38$)	73.51 ($\uparrow 0.31$)	75.67 ($\uparrow 1.83$)	61.11 ($\uparrow 5.86$)
+POS4ASC-pw	79.82 ($\downarrow 0.72$)	70.53 ($\uparrow 2.82$)	66.66 ($\uparrow 5.16$)	73.51 ($\uparrow 0.31$)	75.27 ($\uparrow 1.43$)	61.00 ($\uparrow 5.75$)
+G-POS	80.71 ($\uparrow 0.17$)	72.73 ($\uparrow 5.02$)	70.79 ($\uparrow 9.29$)	74.14 ($\uparrow 0.94$)	75.80 ($\uparrow 1.96$)	63.45 ($\uparrow 8.20$)
RoBERTa	88.39	82.13	82.52	83.39	85.00	77.57
+ASGCN-pw	88.75 ($\uparrow 0.36$)	83.54 ($\uparrow 1.41$)	83.74 ($\uparrow 1.22$)	84.17 ($\uparrow 0.78$)	85.80 ($\uparrow 0.80$)	78.42 ($\uparrow 0.85$)
+POS4ASC-pw	89.02 ($\uparrow 0.63$)	82.60 ($\uparrow 0.47$)	84.36 ($\uparrow 1.84$)	84.01 ($\uparrow 0.62$)	86.07 ($\uparrow 1.07$)	78.69 ($\uparrow 1.12$)
+G-POS	89.20 ($\uparrow 0.81$)	83.92 ($\uparrow 1.79$)	84.37 ($\uparrow 1.85$)	84.33 ($\uparrow 0.94$)	87.05 ($\uparrow 2.05$)	78.94 ($\uparrow 1.37$)

整体而言, 本文提出的 G-POS 算法在 4 个基线模型上取得了比其他算法更好的分类效果, 应用本文所提算法的模型在准确率上均出现了不同程度的提升.

从 G-POS 算法在 3 种评估方法上的表现来看, 首先, 在 I.D. 评估方法中, Restaurant 和 Laptop 数据集上准确率最大提升了 1.61 和 2.98 个百分点, 与其他评估方法相比提升较低, 说明对于领域内文本, 其单词的位置权重分配策略对整体模型的影响较小. 其次, 在 O.O.D. 评估方法中, 在两种数据集上的准确率较 I.D. 方法略有提升, 最高分别为 5.02 和 3.21 个百分点, 说明在跨领域文本情感分类任务中, 句法关系显得尤为重要, 而单词的位置关系可作为基本句法关系, 对模型的情感分类性能具有一定的辅助作用. 最后, 在 Adv. 的评估方法中, G-POS 算法在两种数据集上的性能提升显著, 最高分别为 10.12 和 10.55 个百分点, 说明本文所提算法具

有更好适应复杂对抗数据的能力, 对数据变化的容忍性更高, 具有较强的鲁棒性.

从 G-POS 算法在不同基线模型上的表现来看, 4 种基线模型的性能均有不同程度提升. 相比之下, 在非基于 BERT 的模型 (LSTM、MemNet、AOA) 中, 相对位置信息编码对于分类性能的影响非常显著, 表明其对于相对位置信息编码的敏感程度更强. 而对于基于 BERT 的模型 (RoBERTa), 绝对位置信息不足以建模方面与上下文之间的相对位置关系, 所以在微调预训练语言模型的过程中, 加入相对位置信息也很重要.

3.6 案例分析

为了探究位置权重算法对文本特征表示的影响, 本文从数据集上选取相对位置距离较远的样本, 对其位置权重进行可视化分析. 结果如图 5 所示, 句子中单词的背景颜色越深, 代表相应单词所获得的权重越高.

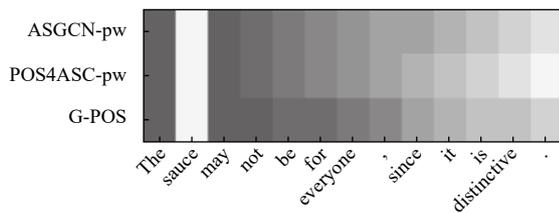


图5 位置权重在句子上的分布情况

进一步的对比分析发现,采用 ASGCN-pw 算法后,模型在处理与方面词“sauce”不相关的上下文单词时,重点关注了无关的意见表达,而其真实意见“distinctive”则被分配了较低的权重。相比之下,采用 POS4ASC-pw 算法后,方面词“sauce”附近的单词区域颜色较深,但意见表达“distinctive”较 ASGCN-pw 区域颜色变浅,说明该算法忽略了相对位置距离较远时的权重分配,而仅关注与方面较近的单词。然而,采用 G-POS 算法后,意见表达“distinctive”的区域颜色更深,说明该算法能更好地处理边界单词的权重分配,从而促使模型更有效地处理长距离单词依赖,进一步提高情感分类的准确性。

此外,为了提供定量支持,我们计算了每种算法在处理单词时分配的权重均值和标准差,结果如表4所示。

表4 算法分配的权重均值和标准差比较

指标	ASGCN-pw	POS4ASC-pw	G-POS
权重均值	0.527	0.500	0.606
权重标准差	0.289	0.292	0.302

通过上述分析和定量结果,发现每种算法在处理单词位置权重时存在差异。在权重均值方面,POS4ASC-pw 算法最低,ASGCN-pw 次之,而 G-POS 算法最高,说明在处理位置权重时,G-POS 算法更加倾向于将较高的权重分配给重要的单词,因此算法能够更准确地识别并关注与任务相关的关键信息,从而更有效地进行文本特征表示。在权重标准差方面,较小的标准差表示权重值的离散程度越小,分布越集中,而较大的权重标准差表明算法在不同单词之间分配权重时存在较大的差异。然而,在方面级情感分类任务中,较小的权重标准差是不理想的,合理的权重分布应具备较大离散度^[26,27]。因此,ASGCN-pw 和 POS4ASC-pw 算法的灵活性和变化性较差,而 G-POS 算法则能够根据文本的特定上下文和语义关系灵活调整权重的分配。

总体而言,ASGCN-pw 算法对单词的位置权重分配较高,但存在分布集中、差异化较小的问题。POS4ASC-pw 算法权重分布离散但均值较小,对重要单词的关注程度不足。G-POS 算法则对重要单词分配较高权重,且

更加注重细微的位置差异,从而更好地捕捉单词之间的重要关联和信息流动。

4 结论与展望

针对方面级文本情感分类任务,本文提出一种基于高斯分布引导位置相关性权重的情感分类算法 G-POS。该算法充分利用距离方面词越近的上下文对于情感分类越重要的特点,以渐进式的高斯分布权重分配策略,逐个增大或减小不同单词的权重差异。实验结果表明,所提算法有效地提高了现有文本情感分类算法的准确率。结合不同领域的交叉验证结果,进一步证明结合本算法的模型具有较强的鲁棒性。此外,本文还对实验结果进行了细致的分析,深入探讨了不同参数对分类准确率的影响。

尽管本文在情感分类任务上取得了不错的效果,但仍存在不足。一是所提出的算法缺乏一个良好的架构,仅依靠位置权重算法并不能很好地完成方面级情感分类任务;二是位置权重算法与文本的句法结构高度相关,对于无明显句法结构的文本,并不能发挥出最佳性能。因此,未来的研究工作将进一步考虑上述问题的解决方式。

参考文献

- 1 Kim Y. Convolutional neural networks for sentence classification. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. Doha: ACL, 2014. 1746–1751.
- 2 Kalchbrenner N, Grefenstette E, Blunsom P. A convolutional neural network for modelling sentences. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. Baltimore: ACL, 2014. 655–665.
- 3 Lei T, Barzilay R, Jaakkola T. Molding CNNs for text: Non-linear, non-consecutive convolutions. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Lisbon: ACL, 2015. 1565–1575.
- 4 Zhu XD, Sobhani P, Guo HY. Long short-term memory over recursive structures. Proceedings of the 32nd International Conference on International Conference on Machine Learning. Lille: JMLR.org, 2015. 1604–1612.
- 5 Tai KS, Socher R, Manning CD. Improved semantic representations from tree-structured long short-term memory networks. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language

- Processing. Beijing: ACL, 2015. 1556–1566.
- 6 Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 6000–6010.
 - 7 汪林, 蒙祖强, 杨丽娜. 基于多级多尺度特征提取的 CNN-BiLSTM 模型的中文情感分析. 计算机科学, 2023, 50(5): 248–254. [doi: 10.11896/jsjcx.220400069]
 - 8 郭磊, 贾真, 李天瑞. 面向方面级情感分析的交互式关系图注意力网络. 计算机应用, 1–9. <http://kns.cnki.net/kcms/detail/51.1307.tp.20230506.1134.004.html>. (2023-05-09).
 - 9 Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. Proceedings of the 5th International Conference on Learning Representations. Toulon: ICLR, 2017. 1–14.
 - 10 Yao L, Mao CS, Luo Y. Graph convolutional networks for text classification. Proceedings of the 33rd AAAI Conference on Artificial Intelligence. Honolulu: AAAI, 2019. 7370–7377.
 - 11 Zhu ZF, Zhang DY, Li L, *et al.* Knowledge-guided multi-granularity GCN for ABSA. Information Processing & Management, 2023, 60(2): 103223.
 - 12 Liang B, Su H, Gui L, *et al.* Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks. Knowledge-based Systems, 2022, 235: 107643. [doi: 10.1016/j.knosys.2021.107643]
 - 13 Cambria E, Liu Q, Decherchi S, *et al.* SenticNet 7: A commonsense-based neurosymbolic AI framework for explainable sentiment analysis. Proceedings of the 13th Language Resources and Evaluation Conference. Marseille: LREC, 2022. 3829–3839.
 - 14 Chen P, Sun ZQ, Bing LD, *et al.* Recurrent attention network on memory for aspect sentiment analysis. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Copenhagen: ACL, 2017. 452–461.
 - 15 Li X, Bing LD, Lam W, *et al.* Transformation networks for target-oriented sentiment classification. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Melbourne: ACL, 2018. 946–956.
 - 16 Zhang C, Li QC, Song DW. Aspect-based sentiment classification with aspect-specific graph convolutional networks. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. Hong Kong: ACL, 2019. 4568–4578.
 - 17 Ma F, Zhang C, Song DW. Exploiting position bias for robust aspect sentiment classification. Proceedings of the 2021 Findings of the Association for Computational Linguistics. ACL, 2021. 1352–1358.
 - 18 Pontiki M, Galanis D, Pavlopoulos J, *et al.* SemEval-2014 Task 4: Aspect based sentiment analysis. Proceedings of the 8th International Workshop on Semantic Evaluation. Dublin: ACL, 2014. 27–35.
 - 19 Xing XY, Jin ZJ, Jin D, *et al.* Tasty burgers, soggy fries: Probing aspect robustness in aspect-based sentiment analysis. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing. ACL, 2020. 3594–3605.
 - 20 Pennington J, Socher R, Manning C. GloVe: Global vectors for word representation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. Doha: ACL, 2014. 1532–1543.
 - 21 Tang DY, Qin B, Feng XC, *et al.* Effective LSTMs for target-dependent sentiment classification. Proceedings of the 26th International Conference on Computational Linguistics. Osaka: The COLING 2016 Organizing Committee, 2016. 3298–3307.
 - 22 Tang DY, Qin B, Liu T. Aspect level sentiment classification with deep memory network. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin: ACL, 2016. 214–224.
 - 23 Huang BX, Ou YL, Carley KM. Aspect level sentiment classification with attention-over-attention neural networks. Proceedings of 11th International Conference on Social Computing, Behavioral-cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation. Washington: Springer, 2018. 197–206.
 - 24 Dai JQ, Yan H, Sun TX, *et al.* Does syntax matter? A strong baseline for aspect-based sentiment analysis with RoBERTa. Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. ACL, 2021. 1816–1829.
 - 25 Devlin J, Chang MW, Lee K, *et al.* BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics. Minneapolis: ACL, 2019. 4171–4186.
 - 26 Brauwere G, Frasincar F. A survey on aspect-based sentiment classification. ACM Computing Surveys, 2022, 55(4): 65.
 - 27 Nazir A, Rao Y, Wu LW, *et al.* Issues and challenges of aspect-based sentiment analysis: A comprehensive survey. IEEE Transactions on Affective Computing, 2022, 13(2): 845–863. [doi: 10.1109/TAFFC.2020.2970399]

(校对责编: 孙君艳)