

## ◆大数据技术◆

融合 ChineseBERT 和双向注意力流的中文商品评论方面情感分析<sup>\*</sup>胡晓丽<sup>1</sup>,张于贤<sup>2\*\*</sup>,黄思睿<sup>3</sup>

(1. 桂林电子科技大学计算机与信息安全学院,广西桂林 541004;2. 桂林电子科技大学商学院,广西桂林 541004;3. 桂林电子科技大学,广西可信软件重点实验室,广西桂林 541004)

**摘要:**准确分类电商平台中用户评论所包含的多个方面的情感极性,能够提升购买决策的有效性。为此,提出一种融合 ChineseBERT 和双向注意力流(Bidirectional Attention Flow, BiDAF)的中文商品评论方面情感分析模型。首先,通过融合拼音与字形的 ChineseBERT 预训练语言模型获得评论文本和方面文本的词嵌入,并采用从位置编码和内存压缩注意力两个方面改进的 Transformer 来表征评论文本和方面文本的语义信息。然后,使用双向注意力流学习评论文本与方面文本的关系,找出评论文本和方面文本中关键信息所对应的词语。最后,将 Transformer 和双向注意力流的输出同时输入到多层感知机(Multilayer Perceptron, MLP)中,进行信息级联和情感极性的分类输出。测试结果表明,提出的模型在两个数据集上的准确率分别为 82.90% 和 71.08%, F1 分数分别为 82.81% 和 70.98%。

**关键词:**商品评论;方面情感分析;词嵌入模型;注意力机制;双向注意力流

中图分类号:TP391.1 文献标识码:A 文章编号:1005-9164(2023)01-0187-09

DOI:10.13656/j.cnki.gxkx.20230308.020

随着信息技术的发展,网上购买已经成为人们的主流购物方式之一。用户对网购的喜爱促进了天猫、京东等众多电商平台的发展,同时也加剧了电商平台之间的竞争。电子商务平台上的评论文本往往能比较客观真实地表达用户购买商品后的感受和看法,对其所含情感进行分析挖掘,能够有效促进电子商务平台的健康发展。对商品评论进行方面情感分析,对用

户而言,能够增强用户对新购商品的了解程度和信任感,使其更好地做出购买决策;对商户而言,能够及时有效地发掘商品的“痒点”和“痛点”,使其进一步改进商品质量、提升服务水平、开展精准营销;对平台而言,能够将商户按优劣分类,实现商户的精准化、规范化管理。

最近,针对评论文本的情感分析大多为基于卷积

收稿日期:2022-08-12

修回日期:2022-10-07

<sup>\*</sup> 国家自然科学基金项目(62267003,61967005,U18112645)和桂林市科学研究与技术开发计划项目(2020011123)资助。

## 【第一作者简介】

胡晓丽(1978-),女,高级实验师,主要从事社交网络与中文信息处理相关领域研究。

## 【\*\*通信作者】

张于贤(1963-),男,教授,主要从事管理科学与工程相关领域研究,E-mail:zyx631218@163.com。

## 【引用本文】

胡晓丽,张于贤,黄思睿.融合 ChineseBERT 和双向注意力流的中文商品评论方面情感分析[J].广西科学,2023,30(1):187-195.

HU X L,ZHANG Y X,HUANG S R. Aspect Sentiment Analysis for Chinese Commodity Reviews Based on ChineseBERT and Bidirectional Attention Flow [J]. Guangxi Sciences,2023,30(1):187-195.

神经网络(Convolutional Neural Networks, CNN)和长短期记忆网络(Long Short-Term Memory, LSTM)等深度学习的方法<sup>[1]</sup>。赵杨等<sup>[2]</sup>为了研究海淘移动电商平台的用户满意度,用支持向量机(Support Vector Machine, SVM)替换 CNN 文本分类模型中分类层的 Softmax 函数,提出 CNN-SVM 情感分类模型,增强了 CNN 模型的准确性和泛化能力。郝利栋等<sup>[3]</sup>为了增强 CNN 模型的特征抽取能力,提出多路融合的 CNN 模型,与 CNN 模型相比,其模型在商品评论情感分析任务中取得了更好的效果。王伟等<sup>[4]</sup>针对双向 LSTM (BiLSTM)模型在文本情感分类中存在的上下文语义学习不充分和训练费时的时间问题,提出一种融合注意力机制的 BiGRU 模型。除了以上使用单个 CNN 或 LSTM 模型外,一些学者还对融合多个模型的方法进行了探索<sup>[5,6]</sup>。韩虎等<sup>[7]</sup>融合 CNN 和 LSTM 两个模型,提出一种多粒度注意力层次神经网络模型,结果表明融合后得到的模型在文本情感分类任务中取得了更好的分类精确度。杜朋等<sup>[8]</sup>为了对电商平台上的商品评论进行分析,提出一种融合 Transformer 和 CNN 的模型,发现与 BiLSTM 模型相比,其模型准确率更高且训练消耗时间更短。此外,一些研究者发现通过改进神经网络中的词嵌入方式能够有效提升模型的性能<sup>[9]</sup>。库向阳等<sup>[10]</sup>提出 Sword2vect 情感向量,与 Word2Vec 相比,Sword2vect 能够更好地表示词语的上下文语义。颜礼蓉等<sup>[11]</sup>融合词性和词嵌入两种特征,解决了传统词嵌入模型存在的情景信息表征不足的问题。现有的商品评论情感分析主要以句子级的研究为主,不能细粒度地分析评论句子中各个方面的情感极性<sup>[12]</sup>。例如,在“新机到手外观变化不大,有点小失望。系统一如既往的好,性能提升较大。”这条针对手机的商品评论中,包含了“外观”和“系统”两个方面,且两个方面的情感极性不一致。在句子级的情感分析模型测试中发现,模型忽略了评论语句中对“外观”的相关消极描述,将整条评论划分在了积极的情感极性类别中。

针对现有句子级情感分析模型在处理包含多个方面,且每个方面的情感极性不一致的商品评论时,准确率较低的问题,本文提出一种融合 ChineseBERT 和双向注意力流(Bidirectional Attention

Flow, BiDAF)的中文商品评论方面情感分析模型(ChineseBERT-BiDAF)。首先采用 ChineseBERT 获得上下文动态感知的中文词向量编码;然后采用改进型 Transformer (I-Transformer)分别表征评论文本和方面文本的语义信息,同时,使用双向注意力流学习评论文本和方面文本的相互关系;最后采用多层感知机(Multilayer Perceptron, MLP)分类输出评论文本中某个方面相对应的情感极性。在 SemEval-2016 的手机和数码相机两个真实数据集上验证 ChineseBERT-BiDAF 在准确率和 F1 分数评价指标上的表现。

## 1 问题定义

中文商品评论方面情感分析任务定义如下:给定一条包含  $N$  个汉字词语的商品评论文本  $T_S = \{w_s^1, w_s^2, \dots, w_s^N\}$  和预先设定好的包含  $M$  个汉字词语的商品方面文本  $T_A = \{w_a^1, w_a^2, \dots, w_a^M\}$ , ChineseBERT-BiDAF 的目标是判断  $T_S$  在  $T_A$  上的情感极性。其中每个汉字词语可能包含一个或多个汉字,情感极性类别分为积极和消极两类。

## 2 方法

图 1 为融合 ChineseBERT 和 BiDAF 的中文商品评论方面情感分析模型,主要包括 ChineseBERT 词编码输入层、I-Transformer 语义表征层、BiDAF 层和多层感知机(Multilayer Perceptron, MLP)情感极性分类层 4 个部分。

### 2.1 ChineseBERT 词编码输入层

ChineseBERT 词编码输入层能够将评论文本和方面文本中每个汉字词语映射到一个低维( $d$ )向量空间中<sup>[13]</sup>,以便于语义表征层的 I-Transformer 学习。ChineseBERT 词编码输入层的输出分别是评论词向量序列  $V^S = \{v_s^1, v_s^2, \dots, v_s^N\} \in R^{N \times d}$  和方面词向量序列  $V^A = \{v_a^1, v_a^2, \dots, v_a^M\} \in R^{M \times d}$ 。

### 2.2 I-Transformer 语义表征层

如图 2 所示, I-Transformer 语义表征层由  $L$  层 I-Transformer 基本单元块组成。每个基本单元块主要包含位置编码、内存压缩注意力、层规范化和全连接神经网络 4 个组件。

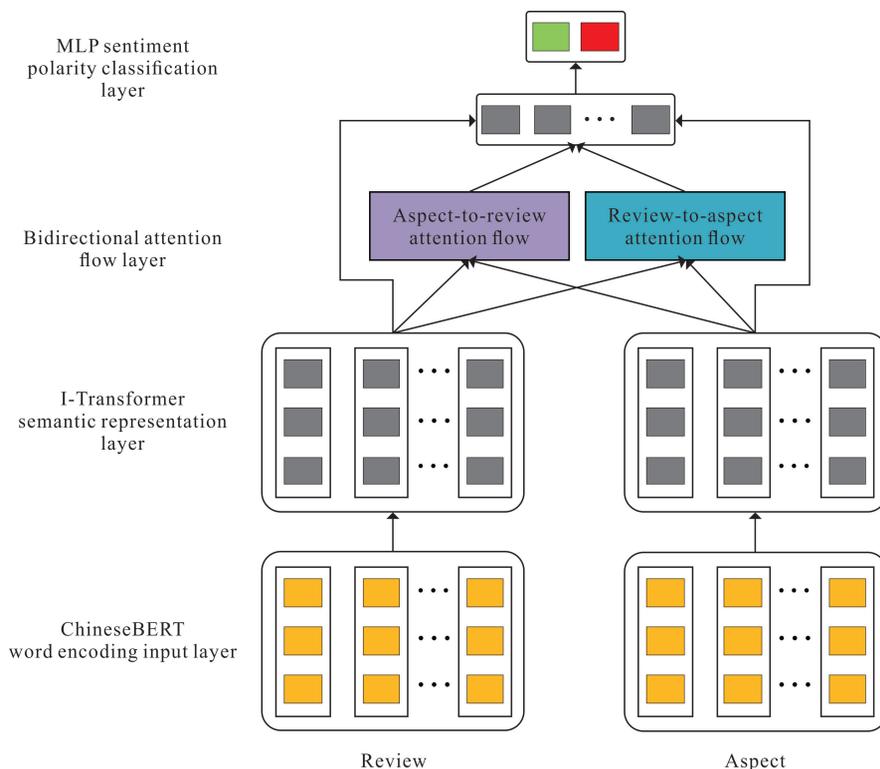


图1 模型整体架构

Fig. 1 Overall model architecture

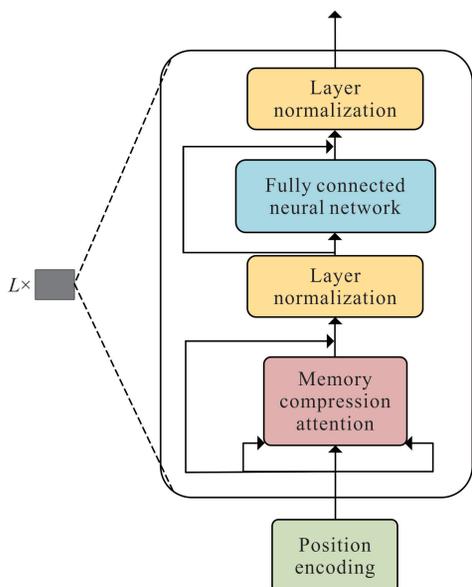


图2 I-Transformer 基本单元块结构

Fig. 2 Basic unit structure of I-Transformer

### 2.2.1 位置编码

在商品评论情感分析时,评论文本中的词语顺序对语义的表达具有重要影响,同一词语在评论文本的不同位置表达的含义可能有很大差异。然而,与循环神经网络和卷积神经网络不同,Transformer 模型采用纯注意力机制,本身不具备编码评论文本中词语绝

对位置或相对位置的能力<sup>[14]</sup>。本文采用 Complex-order 方法额外编码评论文本的位置信息<sup>[15]</sup>,解决原始 Transformer 模型中词向量与位置向量加和方式存在的位置独立性问题,能够表示评论文本中词语间的优先关系和邻接关系。Complex-order 编码方式将词编码转换为一个以位置为自变量的连续函数,如式(1)所示。

$$f(i, p) = g_i(p) \in \mathbb{R}^D, \quad (1)$$

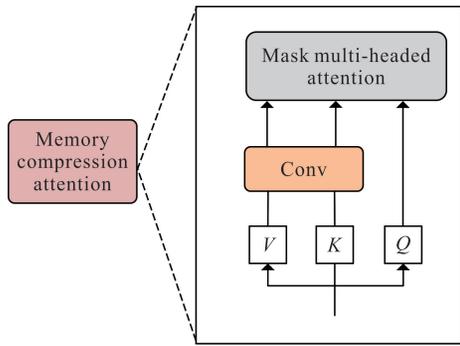
式(1)中,  $g_i(\cdot)$  表示词向量索引到连续函数的映射,  $p$  为位置索引,  $D$  为连续函数集合。

### 2.2.2 内存压缩注意力

为了降低 Transformer 模型中多头注意力机制的计算复杂度,更好地处理长序列的商品评论文本,采用如图 3 所示的内存压缩注意力替换原始的多头注意力<sup>[16]</sup>。内存压缩注意力通过限制查询(Q)和键(K)的点积,缩减多头注意力的规模和资源占用,计算公式如式(2)所示。在将令牌映射到 Q、K 和值(V)嵌入后,使用卷积步长来减少 K 和 V 的数量。

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (2)$$

式(2)中,  $d_k$  表示向量维度。

图3 内存压缩注意力结构<sup>[16]</sup>Fig. 3 Structure of memory compression attention<sup>[16]</sup>

为了增强内存压缩注意力的表征能力,使模型能够学习到评论文本中单词的差别和关系,采用了类似于Transformer模型的多头注意力机制。使用不同的参数矩阵通过式(2)获得的多个注意力进行连接,如式(3)所示。

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_n)W^0, \quad (3)$$

式(3)中,  $\text{head}_i$  表示内存压缩注意力在第  $i$  个头的输出,  $W^0$  为  $n$  个头线性输出的参数矩阵。

### 2.2.3 层规范化

为了确保模型的训练效率和稳定性,在 I-Transformer 结构中使用了如式(4)所示的层规范化操作。

$$\text{LN}(x) = \gamma \frac{x - E[x]}{\sqrt{\text{Var}[x] + \epsilon}} + \beta, \quad (4)$$

式(4)中,  $\gamma$  和  $\beta$  为模型学习所得的参数,  $E[x]$  和  $\text{Var}[x]$  分别表示均值和方差,  $\epsilon$  是避免分母为 0。

### 2.2.4 全连接神经网络

在两个层规范化之间包含一个两层激活函数,为全连接神经网络,如式(5)所示。

$$\text{FCNN}(X) = \text{GELU}(L(L(X))), \quad (5)$$

式(5)中,  $L(\cdot)$  表示线性变换,  $\text{GELU}(\cdot)$  为激活函数。

评论词向量序列和方面词向量序列经过 I-Transformer 语义表征层后,分别生成式(6)所示的评论隐藏序列  $H^S$  和式(7)所示的方面隐藏序列  $H^A$ 。

$$H^S = [h_1^s, h_2^s, \dots, h_N^s] = I\text{-Transformer}(V^S), \quad (6)$$

$$H^A = [h_1^a, h_2^a, \dots, h_M^a] = I\text{-Transformer}(V^A), \quad (7)$$

式(6)和式(7)中,  $V^S$  和  $V^A$  分别为评论词向量序列和方面词向量序列;  $h_i^s \in d^h (1 \leq i \leq N)$ ,  $h_i^a \in$

$d^h (1 \leq i \leq M)$ ,  $d^h$  为词向量维度。

### 2.3 双向注意力流层

双向注意力流从“评论到方面”“方面到评论”两个方向计算注意力,建立起评论序列和方面序列的相互关系,并融合两条序列信息<sup>[17]</sup>。为了计算这两个方向的注意力,首先通过式(8)计算方面序列和评论序列的相似性矩阵,该矩阵在计算两个方向的注意力时被共享使用。

$$S_{nm} = f(H_{:n}^S, H_{:m}^A) \in \mathbb{R}, \quad (8)$$

式(8)中,  $S_{nm}$  表示评论文本中第  $n$  个词与方面文本中第  $m$  个词的相似度,  $H_{:n}^S$  表示评论隐藏序列  $H^S$  的第  $n$  列,  $H_{:m}^A$  表示方面隐藏序列  $H^A$  的第  $m$  列,  $f(\cdot)$  为式(9)计算得到的度量相似度的标量函数。

$$f(h^s, h^a) = \omega_s^T [h^s; h^a; h^s \odot h^a], \quad (9)$$

式(9)中,  $\omega_s$  为待训练的权重矩阵,  $[\cdot; \cdot]$  表示向量间的行拼接,  $\odot$  表示向量间同位元素相乘。

#### 2.3.1 从评论到方面注意力流

从评论到方面的注意力流如图4所示,由图4可以看出,对于评论文本中的每个词,方面文本中的哪些词与其相关性最大。按照式(10)、式(11)计算评论文本中每个词的注意力分数,并将其加权求和后作为该词蕴含评论语义的方面表示( $\tilde{H}^A$ )。

$$a_n = \text{softmax}(S_{:n}) \in \mathbb{R}^M, \quad (10)$$

式(10)中,  $a_n$  为评论文本中第  $n$  个词对方面文本的注意力权重,  $\text{softmax}(\cdot)$  是归一化操作,  $S_{:n}$  表示方面文本与评论文本相似矩阵的第  $n$  行。

$$\tilde{H}_{:n}^A = \sum_m a_{nm} H_{:m}^A \in \mathbb{R}^{d^h \times M}, \quad (11)$$

式(11)中,  $\tilde{H}_{:n}^A$  表示蕴含评论语义的方面文本矩阵的第  $n$  列,将其拼接后得到  $\tilde{H}^A$ 。

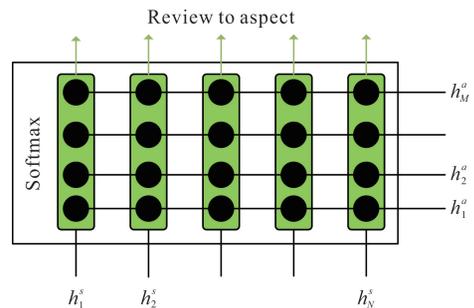


图4 从评论到方面注意力流示意图

Fig. 4 Sketch map of review-to-aspect attention flow

#### 2.3.2 从方面到评论注意力流

从方面到评论的注意力流如图5所示,由图5可以看出,对于方面文本中的每个词,评论文本中的哪

些词与其最相似。首先将相似矩阵  $S$  中每列的最大值取出后合并为一个新的向量 ( $v$ ), 如式(12)。然后通过式(13)计算评论中每个词给对应方面整体分配的相似度。最后将  $\tilde{h}^S$  复制  $N$  次形成蕴含方面语义的评论表示 ( $\tilde{H}^S$ )。

$$v = \text{softmax}(\max_{\text{col}}(S)) \in \mathbb{R}^N, \quad (12)$$

式(12)中,  $\max_{\text{col}}(S)$  表示相似性矩阵中的元素最大值所在列,  $\text{softmax}(\cdot)$  是归一化操作。

$$\tilde{h}^S = \sum_n v_n H_{:,n}^S \in \mathbb{R}^{d^h}, \quad (13)$$

式(13)中,  $H_{:,n}^S$  表示评论隐藏序列的第  $n$  列。

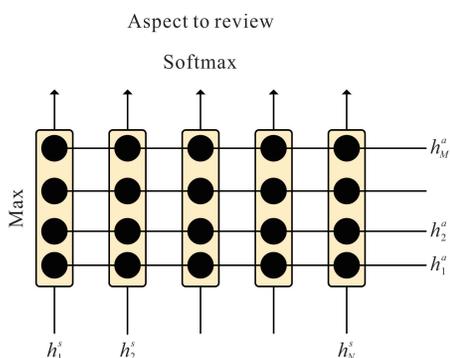


图5 从方面到评论注意力流示意图

Fig. 5 Sketch map of aspect-to-review attention flow

## 2.4 MLP 情感极性分类层

MLP 情感极性分类层的功能是级联 BiDAF 的输出<sup>[18]</sup>, 并将级联后的结果输入 Softmax 函数中求得商品评论的情感极性, 如式(14)、式(15)所示。

$$R = \text{MLP}(H^S, H^A, \tilde{H}^S, \tilde{H}^A), \quad (14)$$

$$p = \text{Softmax}(R). \quad (15)$$

模型采用交叉熵损失函数和 L2 正则化进行训练<sup>[19]</sup>, 如式(16)所示。

$$L = - \sum_{i \in D} \sum_{c \in C} y_i^c \log p_i^c + \lambda \|\theta\|^2, \quad (16)$$

式(16)中,  $p_i^c$  为模型预测的评论文本中方面对应的情感极性,  $y_i^c$  为数据集中人工标注的评论文本中方面对应的情感极性,  $C$  为评论文本中蕴含的{积极, 消极}两类情感极性,  $\theta$  为模型参数,  $\lambda$  为控制 L2 正则化对模型影响程度的参数。

## 3 实验与结果分析

### 3.1 数据集与实验环境

在 SemEval-2016 任务五 (<https://alt.qcri.org/semeval2016/task5/>) 的手机和数码相机两个中文商品评论数据集上进行实验, 实验数据集的统计信息如

表 1 所示。

表 1 数据集统计信息

Table 1 Statistical information of data set

情感极性 Emotional polarity	手机数据集 Data set on mobile phone		数码相机数据集 Data set on digital camera	
	训练数据集 Training data set	测试数据集 Test data set	训练数据集 Training data set	测试数据集 Test data set
Positive	758	310	809	344
Negative	575	219	450	137
Total	1 333	529	1 259	481

实验在谷歌 Colab 云环境进行, 模型采用 Python 编程语言, 在 Pytorch 开源深度学习框架下编程实现。ChineseBERT 预训练模型选用 base 版本 (<https://github.com/ShannonAI/ChineseBert>)。模型选用 Adam 优化函数, 学习率设定为 0.001, L2 正则化系数  $\lambda$  设定为 0.1。由于数据集规模较小, 为了避免模型训练过程中出现过拟合问题, 在模型中增加了丢弃率为 0.1 的 dropout 层。

### 3.2 评估指标和对比模型

选取准确率 (ACC) 和 F1 分数作为模型的评估指标。ACC 为模型预测正确的样本数占总样本数的比率, 如式(17)所示。

$$\text{ACC} = \frac{\text{TP} + \text{TN}}{(\text{TP} + \text{TN}) + (\text{FP} + \text{FN})}, \quad (17)$$

式(17)中, TP 表示模型预测和实际标注的情感极性均为积极, TN 表示模型预测和实际标注的情感极性均为消极, FP 表示实际标注情感为消极的样本被错误地预测为积极, FN 表示实际标注情感为积极的样本被错误地预测为消极。

F1 分数为精准率和召回率的倒数平均数, 如式(18)所示。

$$F1 = \frac{2P \times R}{P + R}, \quad (18)$$

式(18)中,  $P = \frac{\text{TP}}{\text{TP} + \text{FP}}$  为精准率,  $R = \frac{\text{TP}}{\text{TP} + \text{FN}}$  为召回率, 精准率和召回率计算公式中的 TP、TN、FP、FN 的含义与式(17)相同。

为了验证本文所提模型的准确性和可行性, 选用以下 4 种模型进行对比实验。

LSTM<sup>[20]</sup>: 将评论文本进行词向量编码后视为序列, 输入 LSTM 模型进行情感极性判断。

IAN<sup>[21]</sup>: 一种交互注意力神经网络模型。通过计算评论文本和方面文本的交互注意力, 两者相互知

晓对方文本中的重要词。首先使用 LSTM 模型获得评论文本和方面文本的隐藏层;然后使用平均池化获得方面隐藏层和评论隐藏层长度固定的向量表示,并通过交互注意力得到方面相关的评论表示和评论相关的方面表示;最后拼接评论表示和方面表示,输入 Softmax 层分类输出方面情感极性。

AOA<sup>[22]</sup>:与 IAN 模型类似,首先使用双向 LSTM 模型获得方面和评论的隐藏表示,然后通过 Attention-over-Attention 注意力模块获得方面相关的评论表示和评论相关的方面表示,最后输入 Soft-

Table 2 Comparison results of five models

模型 Model	准确率(%) ACC (%)		F1 分数(%) F1 score (%)	
	手机数据集 Data set on mobile phone	数码相机数据集 Data set on digital camera	手机数据集 Data set on mobile phone	数码相机数据集 Data set on digital camera
LSTM	81.24	81.05	69.40	69.00
IAN	81.63	81.49	69.78	69.72
AOA	81.95	81.77	70.06	69.94
BERT	82.65	82.43	70.76	70.30
This study	82.90	82.81	71.08	70.98

从表 2 可知,①本文所提模型在手机和数码相机两个数据集上的 ACC 和 F1 分数均高于对比模型,验证了通过改变词向量编码方式和引用注意力机制增强方面情感分析模型的可行性和有效性,与文献[22-23]的结论一致。②以上 5 种模型在手机数据集上的表现普遍优于在数码相机上的实验结果,可能是与数码相机数据集相比,手机数据集的数据规模更大、正负样本更加均衡,更有利于上述 5 种有监督模

Table 3 Results of bidirectional attention flow ablation experiment

模型 Model	准确率(%) ACC (%)		F1 分数(%) F1 score (%)	
	手机数据集 Data set on mobile phone	数码相机数据集 Data set on digital camera	手机数据集 Data set on mobile phone	数码相机数据集 Data set on digital camera
One-way attention flow of review-to-aspect	82.86	82.75	71.06	70.94
One-way attention flow of aspect-to-review	81.84	81.70	71.01	69.90
Bidirectional attention flow	82.90	82.81	71.08	70.98

从表 3 可知,无论是方面到评论还是评论到方面的单向注意力流,在两个数据集上的两种评价指标均低于双向注意力流,两种流向注意力共同作用的效果

max 层分类输出方面情感极性。

BERT<sup>[23]</sup>:将评论文本和方面文本按照“[CLS]评论文本[SEP]方面文本”的格式输入到 BERT 预训练模型中进行微调。

### 3.3 结果分析

#### 3.3.1 对比实验

本文所提模型与 4 种对比模型在 SemEval-2016 任务五的手机和数码相机两个数据集上的 ACC 和 F1 分数如表 2 所示。

型取得更好的实验效果。③本文所提模型和 BERT 模型比其余的 3 种模型取得更好的结果,说明“无监督预训练+有监督微调”的模型比传统的有监督神经网络更适用于中文商品评论方面情感分析任务。

#### 3.3.2 消融实验

为了验证本文提出的双向注意力流对中文商品评论方面情感分析的作用机理,针对双向注意力流进行消融实验,结果如表 3 所示。

最好,表明双向注意力流机制对模型性能的提升有帮助,使用双向注意力机制比使用单个流向的注意力机制效果更好。此外,实验结果表明评论到方面单向注

注意力流的模型效果略高于方面到评论单向注意力流的模型,说明在中文评论方面情感分析任务中,方面与评论之间相似度高的关键词对模型整体的结果影响较大。

### 3.3.3 词向量编码选择

为了探究词向量编码对模型的影响机理,同时验证选择 ChineseBERT 作为模型词向量编码的合理性,采用 Word2vec、Glove、ELMo 和 BERT 4 种词向量编码进行实验,结果如图 6 所示。

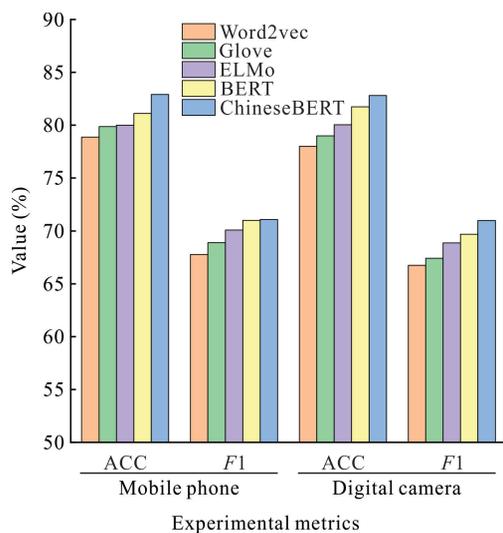


图 6 词向量编码对模型性能的影响

Fig. 6 Effect of word vector coding on model performance

从图 6 可以看出,在 5 种词向量编码中,本文所用的 ChineseBERT 在两个数据集上的 ACC 和 F1 分数均取得最高的数值。这是因为 ChineseBERT 不仅具有 BERT 词向量编码的自身优势,而且在模型预训练的中文语料中融合了汉字特有的拼音和字形信息,获得了更加准确的词向量编码。此外,实验结果表明根据上下文动态编码的 ELMo、BERT 和 ChineseBERT 3 种模型,比 Word2vec 和 Glove 两种模型在中文评论方面情感分析任务中效果更好,可能是动态编码的词向量模型能够根据上下文信息差异化表征同义词,解决了评论文本中同义词存在的一词多义问题。

### 3.3.4 模型层数选择

为了探索模型层数对模型性能的影响,本文在 [1,5] 区间内,以步长为 1 分别进行分组实验。实验时确保除模型层数不同外,其余部分均采用相同配置,实验结果如图 7 所示。

从图 7 可以看出,在手机和数码相机两个数据集上,在 [1,3] 区间内随着层数的增加,模型的 ACC 和

F1 值不断提升,说明随着层数的增加,模型对中文商品评论文本的语义理解更加正确。然而,在 [3,5] 区间内随着层数的增加,模型的性能反而有所下降,其原因可能是 ChineseBERT 是基于海量数据集通过自监督得到的预训练语言模型,本文用来实验的标注数据集与之相比规模较小,过多的层数增加了模型的复杂度,从而导致模型性能的下降。因此本文模型的层数最终设定为 3 层。

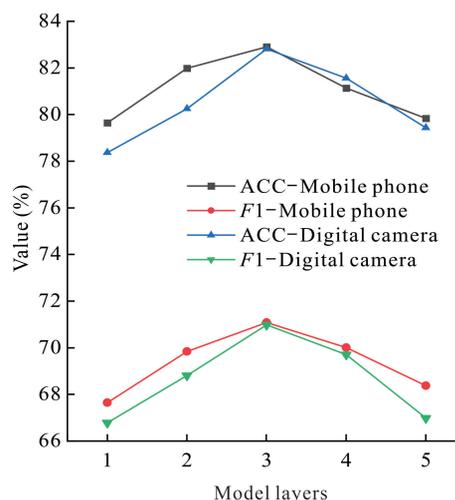


图 7 模型层数对模型性能的影响

Fig. 7 Effect of model layers on model performance

### 3.3.5 模型训练时间对比

为了评估模型的时间消耗,将本文所提模型和 4 种对比模型在相同的实验硬件环境进行训练,5 种模型在手机数据集上迭代一次所需时间如图 8 所示。

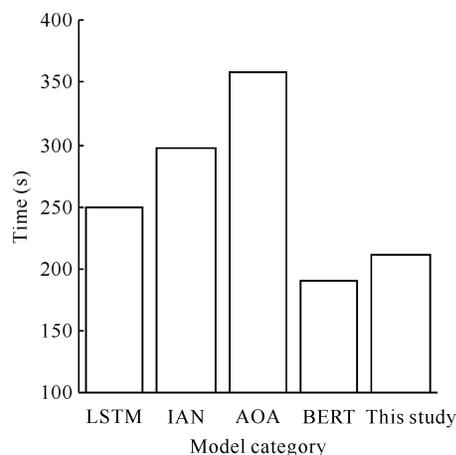


图 8 模型迭代一次所需时间

Fig. 8 Time required for model iteration

由图 8 可知,本文所提模型在保持和 BERT 几乎相同的时间消耗前提下,具有更高的 ACC 和 F1 分数。这是因为本文所提模型虽然比 BERT 模型在双向注意力流训练时需要额外花费时间,但在语义表

征层采用的 I-Transformer 与 BERT 所采用的 Transformer 相比, 时间消耗更少。同时可以观测到, 可以进行并行训练的本文所提模型和 BERT 模型的时间消耗低于不能并行训练的 LSTM、IAN 和 AOA 3 种模型。此外, 对比实验结果可以发现, LSTM、IAN 和 AOA 3 种模型的时间消耗是依次激增的, 说明模型的时间消耗与添加的注意力机制的复杂度呈正相关。这是因为 IAN 和 AOA 模型为了提升 LSTM 模型的准确性, 采取了分开训练方面文本和评论文本、增加注意力机制两种措施, 且 AOA 模型采用了比 IAN 模型结构更加复杂的注意力机制。

#### 4 结论

由于现有中文商品评论情感分析模型在处理包含多个方面, 且每个方面的情感极性不一致的评论句子时准确率较低, 本文提出一种面向中文商品评论的方面情感分析的模型 (ChineseBERT-BiDAF)。首先使用 ChineseBERT 模型分别获得方面文本和评论文本的嵌入表示, 然后采用 I-Transformer 分别学习方面和评论文本的语义信息, 通过双向注意力流学习方面与评论语句的关系, 找出方面文本与评论文本中的重要词组, 并将其与语义信息组合, 最后输入 MLP 情感极性分类层, 获得方面对应的情感极性。实验结果验证了本文所提模型的有效性, 下一步将构建中文商品评论语料库, 并基于此训练专用于中文商品评论方面情感分析的 ChineseBERT 模型, 以期进一步提升模型的性能。

#### 参考文献

- [1] 李洋, 董红斌. 基于 CNN 和 BiLSTM 网络特征融合的文本情感分析[J]. 计算机应用, 2018, 38(11): 3075-3080.
- [2] 赵杨, 李齐齐, 陈雨涵, 等. 基于在线评论情感分析的海淘 APP 用户满意度研究[J]. 数据分析与知识发现, 2018, 23(11): 19-27.
- [3] 郝利栋, 赵慧, 杨培丽. 基于多路融合卷积神经网络的网购商品情感分类[J]. 东华大学学报(自然科学版), 2018, 44(4): 555-559.
- [4] 王伟, 孙玉霞, 齐庆杰, 等. 基于 BiGRU-attention 神经网络的文本情感分类模型[J]. 计算机应用研究, 2019, 36(12): 3558-3564.
- [5] 白静, 李霏, 姬东鸿. 基于注意力的 BiLSTM-CNN 中文微博立场检测模型[J]. 计算机应用与软件, 2018, 35(3): 266-274.
- [6] YAN W, ZHOU L F, QIAN Z J, et al. Sentiment analysis of student texts using the CNN-BiGRU-AT model [J/OL]. Scientific Programming, 2021, 2021 (Pt. 9): 8405623. <https://doi.org/10.1155/2021/8405623>.
- [7] 韩虎, 刘国利, 孙天岳, 等. 多注意力层次神经网络文本情感分析[J]. 计算机工程与应用, 2020, 56(10): 100-105.
- [8] 杜朋, 卢益清, 韩长风. 基于 Transformer 模型的商品评论情感分析[J]. 中文信息学报, 2021, 35(2): 125-132.
- [9] 库向阳, 杨瑞丽, 董立红. 基于 Sword2vect 的中文在线商品评价情感分析[J]. 西安科技大学学报, 2020, 40(3): 504-511.
- [10] 库向阳, 杨瑞丽. 融合深度特征的 Sword2vect 商品在线评价情感分析[J]. 计算机应用与软件, 2020, 37(6): 212-217.
- [11] 颜礼蓉, 朱小栋, 陈曦. 基于双通道融合和 BiLSTM-attention 的评论文本情感分类算法[J]. 上海理工大学学报, 2021, 43(6): 597-605. DOI:10.13255/j.cnki.jusst.20210102001.
- [12] 冯超, 黎海辉, 赵洪雅, 等. 基于层次注意力机制和门机制的属性级别情感分析[J]. 中文信息学报, 2021, 35(10): 128-136.
- [13] SUN Z J, LI X Y, SUN X F, et al. ChineseBERT: Chinese pretraining enhanced by glyph and pinyin information [Z/OL]. (2021-06-30)[2022-08-01]. <https://doi.org/10.48550/arXiv.2106.16038>.
- [14] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [Z/OL]. (2017-06-12)[2022-08-01]. <https://doi.org/10.48550/arXiv.1706.03762>.
- [15] WANG B Y, ZHAO D H, LIOMA C, et al. Encoding word order in complex embeddings [Z/OL]. (2019-12-27)[2022-08-01]. <https://doi.org/10.48550/arXiv.1912.12333>.
- [16] LIU P J, SALEH M, POT E, et al. Generating wikipedia by summarizing long sequences [Z/OL]. (2018-01-30)[2022-08-01]. <https://doi.org/10.48550/arXiv.1801.10198>.
- [17] SEO M, KEMBHAVI A, FARHADI A, et al. Bidirectional attention flow for machine comprehension [Z/OL]. (2016-11-05)[2022-08-01]. <https://doi.org/10.48550/arXiv.1611.01603>.
- [18] TOLSTIKHIN I, HOULSBY N, KOLESNIKOV A, et al. Mlp-mixer: an all-mlp architecture for vision [Z/OL]. (2021-05-04)[2022-08-01]. <https://doi.org/10.48550/arXiv.2105.01601>.
- [19] XU Q N, ZHU L, DAI T, et al. Aspect-based sentiment classification with multi-attention network [J]. Neuro-

- computing,2020,388:135-143.
- [20] HOCHREITER S,SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9 (8): 1735-1780.
- [21] MA D,LI S J,ZHANG X D,et al. Interactive attention networks for aspect-level sentiment classification [Z/OL]. (2017-09-04)[2022-08-01]. <https://doi.org/10.48550/arXiv.1709.00893>.
- [22] HUANG B,OU Y,CARLEY K M. Aspect level sentiment classification with attention-over-attention neural networks [C]//International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation. Cham,Switzerland:Springer,2018:197-206.
- [23] DEVLIN J,CHANG M W,LEE K,et al. BERT: pre-training of deep bidirectional transformers for language understanding [Z/OL]. (2018-10-11)[2022-08-01]. <https://doi.org/10.48550/arXiv.1810.04805>.

## Aspect Sentiment Analysis for Chinese Commodity Reviews Based on ChineseBERT and Bidirectional Attention Flow

HU Xiaoli<sup>1</sup>,ZHANG Yuxian<sup>2\*\*</sup>,HUANG Sirui<sup>3</sup>

(1. School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin, Guangxi, 541004, China; 2. Business School, Guilin University of Electronic Technology, Guilin, Guangxi, 541004, China; 3. Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Guilin, Guangxi, 541004, China)

**Abstract:** Accurately classifying the sentiment polarity of various aspects contained in user reviews in E-commerce platforms can improve the effectiveness of purchase decisions. Therefore, a sentiment analysis model of Chinese product reviews based on ChineseBERT and Bidirectional Attention Flow (BiDAF) is proposed. Firstly, the word embedding of the review text and the aspect text is obtained by the ChineseBERT pre-trained language model that integrates pinyin and glyph, and the semantic information of the review text and the aspect text is represented by the Transformer improved from two aspects of position coding and memory compression attention. Then, the bidirectional attention flow is used to learn the relationship between the review text and the aspect text, and find out the words corresponding to the key information in the review text and the aspect text. Finally, the outputs of Transformer and bidirectional attention flow are simultaneously input into Multilayer Perception (MLP) for information cascade and sentiment polarity classification output. The test results show that the accuracy of the proposed model on the two data sets is 82.90% and 71.08%, respectively, and the F1 scores are 82.81% and 70.98% respectively.

**Key words:** product reviews; aspect sentiment analysis; word embedding model; attention mechanism; bidirectional attention flow

责任编辑:梁晓



微信公众号投稿更便捷

联系电话:0771-2503923

邮箱:gxkx@gxas.cn

投稿系统网址:<http://gxkx.ijournal.cn/gxkx/ch>