

# 融合双通道语义特征的情感分析研究

刘司摇,周艳玲,兰正寅,张 龔,曾张帆

(湖北大学计算机与信息工程学院,湖北武汉430062)

**摘要:** 现有大多数深度学习模型结构单一,通常会降低文本语义特征提取能力。为此,提出一种融合双通道语义特征(FDSF)的情感分析研究模型。首先,采用BERT预训练语言模型获取文本的动态特征向量表示。然后,将BiGRU-Attention通道提取的全局语义信息经注意力动态权重调整后,与CNN通道提取的局部语义信息进行特征向量融合。最后,将融合特征经过全连接层与Softmax函数,输出最终情感极性。在ChineseNLPcorpus的online\_shopping\_10\_cats、中科院谭松波学者整理的数据集上与现有主流情感分析方法进行比较实验,结果表明,FDSF模型在F1值与准确率方面均最优,证明了该模型在情感分析任务中的有效性和可行性。

**关键词:** 情感分析;深度学习;双向门循环控制单元;语义向量;双通道

DOI:10.11907/rjdk.222108

开放科学(资源服务)标识码(OSID):

中图分类号:TP391

文献标识码:A

文章编号:1672-7800(2023)009-0073-06



## Research on Sentiment Analysis Fused with Dual Channel Semantic Features

LIU Siyao, ZHOU Yanling, LAN Zhengyin, ZHANG Yan, ZENG Zhangfan

(School of Computer Science and Information Engineering, Hubei University, Wuhan 430062, China)

**Abstract:** Most existing deep learning models have a single structure, which usually reduces the ability to extract text semantic features. To this end, a sentiment analysis research model integrating dual channel semantic features (FDSF) is proposed. Firstly, the BERT pre trained language model is used to obtain the dynamic feature vector representation of the text. Then, the global semantic information extracted by the BiGRU Attention channel is adjusted by attention dynamic weights, and fused with the local semantic information extracted by the CNN channel for feature vectors. Finally, the fused features are processed through a fully connected layer and Softmax function to output the final emotional polarity. Experiment on online\_shopping\_10\_cats of ChineseNLPcorpus, and dataset compiled by scholars Tan Songbo from the Chinese Academy of Sciences, compared with existing mainstream sentiment analysis methods, the FDSF model has the best F1 value and accuracy, proving its effectiveness and feasibility in sentiment analysis tasks.

**Key Words:** sentiment analysis; deep learning; bi-directional gated recurrent unit; semantic vector; dual channel

## 0 引言

随着互联网高速发展与通信设备普及,各式各样的社交媒体平台通过网络被推送到生活中,越来越多的用户通过微博、淘宝等应用软件发表对新闻、产品的态度和评价<sup>[1]</sup>。情感分析的目的是从具有主观性的文本中提取用户情感,属于自然语言处理的基本任务之一<sup>[2]</sup>。通过情感

分析这些评论文本数据,可获取用户对商品的评价或网民对社会新闻的态度,从而精准把握用户需求,调节产品市场方向;引导社会舆论向积极方向发展,避免负面事件进一步发酵。因此,通过情感分析网络交互信息中的用户观点与情绪极具商业价值和社会意义,如何在海量数据中挖掘有效信息,利用深度学习方法更好地捕捉深层次文本语义特征便具有十分重要的研究价值<sup>[3]</sup>。

目前,文本情感分析方法主要分为:①基于情感词典

收稿日期:2022-09-22

基金项目:国家自然科学基金项目(61977021);湖北省自然科学基金项目(2021CFB503)

作者简介:刘司摇(1998-),女,湖北大学计算机与信息工程学院硕士研究生,研究方向为自然语言处理;周艳玲(1981-),女,博士,湖北大学计算机与信息工程学院副教授、硕士生导师,研究方向为自然语言处理、信号分析;兰正寅(1998-),男,湖北大学计算机与信息工程学院硕士研究生,研究方向为自然语言处理;张龔(1974-),男,博士,湖北大学计算机与信息工程学院教授、硕士生导师,研究方向为自然语言处理、网络安全;曾张帆(1983-),男,博士,湖北大学计算机与信息工程学院教授、硕士生导师,研究方向为人工智能、云计算。本文通讯作者:周艳玲。

的情感分析方法<sup>[4]</sup>,但情感词通常具有滞后性,因此通用性不足且耗时耗力;②基于传统机器学习的情感分析方法<sup>[5]</sup>,该方法通过机器学习相关技术提取文本中的情感特征,建立训练模型后预测文本情感,但对特征提取质量的依赖性较大,难以达到满意的效果;③基于深度学习的情感分析方法<sup>[6]</sup>,该方法通过模拟人脑结构,借助多层神经元自动提取、更新、优化特征,从而提升情感分析的准确度,现已取得了较好的性能<sup>[7-9]</sup>,但如何利用深度学习方法抽取更深层次的文本语义特征仍然亟待解决。

本文为了提升中文文本情感分类效果,利用各模型优势捕捉多维度语义信息,增强文本向量表征能力,提出一种融合双通道语义特征(Fused on Dual Channel Semantic Features, FDSF)的情感分析模型。首先,将BERT作为词嵌入层以获取文本词向量矩阵表示,将其输入BiGRU来提取全局文本序列特征,并引入注意力机制分配特征向量权重;然后利用CNN提取文本矩阵多个粒度下的局部特征,并将BiGRU-Attention与CNN输出特征信息进行向量融合;最后由全连接层Softmax分类器输出文本的情感倾向。

### 1 相关工作

文本向量化是情感分析研究的关键环节,通过神经网络对文本构建词向量矩阵,让机器理解文本语义,对情感极性的判断具有重要的研究意义。常用的词向量表示方法包括One-hot编码<sup>[10]</sup>、Word2vec<sup>[11]</sup>、BERT、Glove<sup>[12]</sup>等。其中,Glove词嵌入模型同时考虑了文本的局部和整体语义信息,结合了LSA与Word2vec的优点,提升了模型训练速度和分类准确度,但构建的静态词向量仍然无法解决相同文字在不同语句中的一词多义问题。

为此,Google提出采用双向Transformer编码器的语言表示模型BERT(Bidirectional Encoder Representations from Transformers)。Devlin等<sup>[13]</sup>通过BERT模型预训练的文本数据在11个自然语言处理任务上获得了较好的效果。王宇晗等<sup>[14]</sup>提出一种基于BERT的嵌入式主题模型,在主题多样性、建模时的一词多义等问题上表现优越,在大规模文本中能提取高质量、细粒度的主题词。Karimi等<sup>[15]</sup>提出一种基于BERT对抗性训练的新模型,利用对抗性训练完成情感分析中的特征提取、特征情感分类两个主要任务,研究表明在这两个任务中该方法相较于传统方法效果更优。此外,Pang等<sup>[16]</sup>基于Transformer双向编码表示(BERT),通过构造一个方面特征定位模型提出了一种有效的方面级情感分析方法。

随着情感分析研究深入,基于神经网络的深度学习模型受到了众多学者青睐,已广泛应用于自然语言处理领域中。常见的神经网络技术包括卷积神经网络(Convolutional Neural Network, CNN)<sup>[17]</sup>、循环神经网络(Recurrent Neural Network, RNN)<sup>[18]</sup>、门控循环单元(Gate Recurrent Unit,

GRU)<sup>[19]</sup>等,以上模型可从文本中提取特征并不断优化,相较于传统机器学习方法无需繁琐的人工标注就能取得更好的效果,并能提取更深层次的句子级特征信息,在下游任务情感分析中的应用效果显著。Jelodar等<sup>[20]</sup>通过长短时记忆网络(Long Short Term Memory, LSTM)分析新型冠状病毒的情感,揭示了利用公众舆论和适当的计算技术了解该问题并指导相关决策的重要性,准确度达到81.15%。

除了基于单一神经网络的深度学习算法,越来越多的研究人员考虑结合不同神经网络方法的优点组成混合神经网络的模型,并应用于情感分析。Ume等<sup>[21]</sup>提出一种结合CNN和LSTM的深度网络,在Twitter数据集上进行情感分析的结果表明,该模型优于单一机器学习分类器。Liu等<sup>[22]</sup>提出一种Bert-BiGRU-Softmax的混合模型,利用BERT模型作为输出层提取情感特征,以双向门控循环单元(BiGRU)为隐藏层计算情感权重,以注意力机制为输出层,准确率达到95.5%以上。

然而,基于单通道的混合神经网络模型往往会因融合加深网络层次,特征向量压缩损失更多文本语义信息,致使情感倾向发生偏差。基于双通道的混合神经网络模型由两个互不干扰的网络通道组成,充分利用了不同深度学习模型的优势,从多方面提取文本特征语义,在一定程度上降低了网络层次深度造成的语义信息损失。

### 2 FDSF模型

本文提出的FDSF模型结构由输入层、BERT词嵌入层、提取全局语义特征的BiGRU-Attention层、提取多粒度下局部语义特征信息的CNN层和语义向量融合输出层构成,如图1所示。

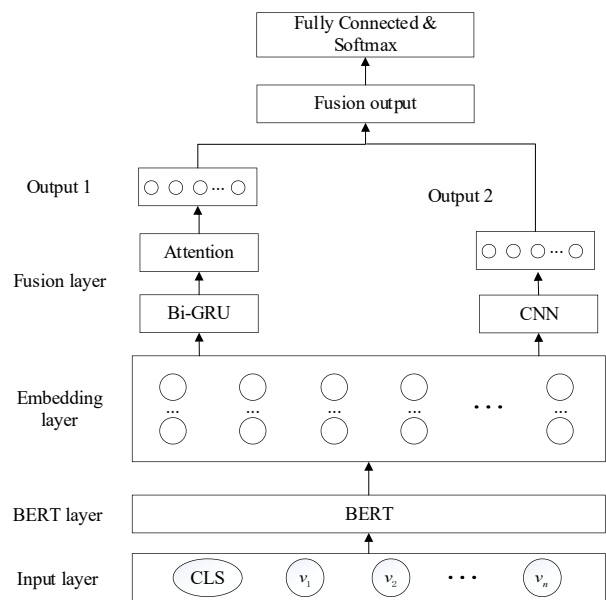


Fig. 1 FDSF model  
图1 FDSF模型

### 2.1 词嵌入层

预训练模型 BERT 以无监督方式,通过大量无标注的文本数据进行训练,得到具有文本语句情感性倾向的词向量。在不同下游任务中无需特意更改模型,只进行相对应的微调即可,相较于 Word2vec、Glove、BERT 等常用的词嵌入模型,利用双向 Transformer 编码器的多头注意力机制可捕获到更准确的语义信息,解决了相同字词在不同语境中的一词多义问题。

如图 2 所示, BERT 模型利用双向 Transformer 编码器提取文本中的语义信息。其中,  $w_1, w_2, \dots, w_n$  表示文本输

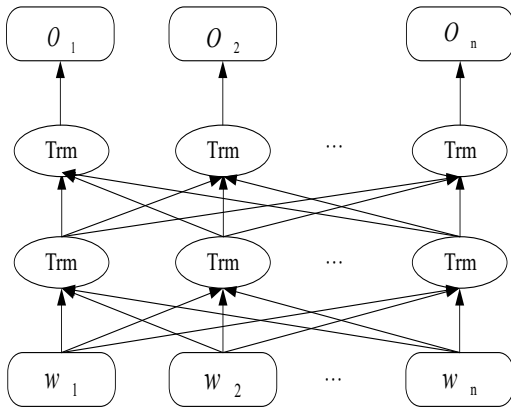


Fig. 2 BERT word embedding

图 2 BERT 词嵌入

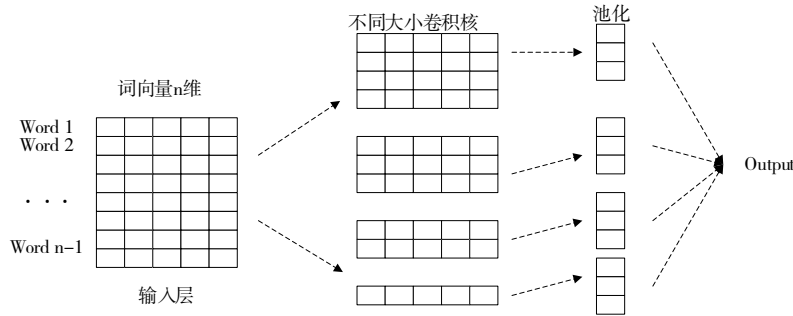


Fig. 3 CNN local feature extraction

图 3 CNN 局部特征提取

CNN 具有的局部感受野和权值共享优点,可减少模型训练参数,非常适合提取局部特征。因此,本文采用 CNN 提取局部文本情感特征信息。其中,卷积层通过卷积计算得到经过映射的特征,在池化层中深度学习模型通常采用最大池化对特征降维后得到一维向量,该一维向量可看为经过卷积层映射后得到文本数据的主要特征,最后全连接层将得到的特征进行连接。局部情感特征  $F_w$  的计算过程如式(1)所示。

$$c_i = f(W * X_{i:i+h-1} + b_c) \quad (1)$$

式中:  $b_c$  表示偏移项;  $W$  为卷积核;  $*$  为卷积运算;  $X_{i:i+h-1}$  表示从  $i$  到  $i+h-1$  个词向量;  $f(\cdot)$  为非线性激活函数。

### 2.3 BiGRU 层

RNN 将文本序列按时间顺序依次输入网络中处理,结

入,通过多层 Transformer 训练得到对应的文本输出  $O_1, O_2, \dots, O_n$ 。Transformer 编码器由多个编码器组成,每个编码器分为多头自注意力层和前馈层,自注意力层的主要作用是在捕获词向量时能考虑该单词与其他单词的上下文语义关联,有利于在深度学习任务中聚焦更重要的单词,加快训练速度。

考虑到在复杂的任务中,自注意力机制层对文本的拟合程度不够,前馈神经网络的两个线性层能加强 Transformer 的表达能力。因此,在进入解码器前,Transformer 编码器还会对输出进行残差连接和层规范化。

在预训练过程中, BERT 模型引入了两个任务,分别为遮蔽语言模型 (Masked Language Model, MLM) 和下一句预测 (Next Sentence Prediction, NSP)。其中, MLM 语言模型通过 [Mask] 掩码随机遮挡数据集中部分单词,然后让模型利用上下文预测被隐藏的单词,目的是为了更准确地捕捉单词表征; NSP 任务通过判断前后句子间的合理性来理解文本的构造。因此, BERT 预训练模型拥有优秀的语义表征能力,非常适用于情感分析任务。

### 2.2 CNN 层

CNN 是一种由输入层、卷积层、池化层和全连接层 4 部分组成的一种深度前馈神经网络,因良好的特征提取能力被广泛应用于计算机视觉等领域,近几年也有越来越多学者将其应用于情感分类,如图 3 所示。

合单词前后关联进行特征提取,然而当文本序列长度过长时会造成梯度弥散和梯度爆炸问题,导致后期模型无法有效获取前向序列文本信息。门控循环单元作为 RNN 的变体,结构中的重置门、更新门机制分别更新、重置文本序列,有效解决了 RNN 在短序列文本上的瓶颈依赖问题, GRU 结构如图 4 所示,前向传播的 GRU 网络计算公式如式(2)一式(5)所示。

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (2)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (3)$$

$$\tilde{h}_t = \tanh(W_h \cdot [r_t * h_{t-1}, x_t]) \quad (4)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (5)$$

式中:  $\sigma$  为 sigmoid 激活函数,将其函数值控制在 (0, 1) 范围内;  $W_z, W_r$  均为 GRU 网络的权值矩阵;  $h_{t-1}$  为前一时

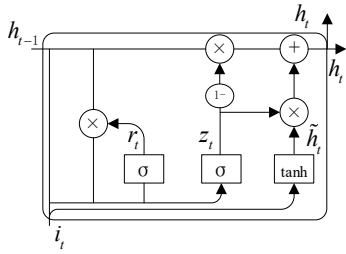


Fig. 4 GRU network model  
图4 GRU网络模型

刻的状态信息;  $h_t$  为当前隐藏状态;  $\tilde{h}_t$  为候选隐藏状态。

式(2)为更新门表达式,更新门决定了前一个时间状态信息传递到当前时间的状态信息程度,值越大说明前一时刻传递的信息越多;式(3)为重置门表达式,重置门控制丢弃前一时刻信息的程度,重置门值越小代表忽略越多。尽管GRU能有效捕捉长序列文本语义信息,但单向GRU只能从前往后读取文本数据信息,导致其只能保留前向文本特征。因此,在前后文语义关联较强的文本中,单向GRU无法较好地提取语句中的隐藏信息。

为了准确捕获文本情感倾向,FDSF模型采用双向门控单元(BiGRU)提取文本序列特征。BiGRU由两个反方向的单向GRU组成,能同时保留过去、未来文本语义特征,可结合上下文更准确地预测当前内容。如图5所示,在BiGRU网络中,设  $t$  时刻计算的前向输出为  $\vec{h}_t$ ,后向输出为  $\overleftarrow{h}_t$ ,将双向输出拼接融合得到最终输出  $h_t = [\vec{h}_t, \overleftarrow{h}_t]$ 。

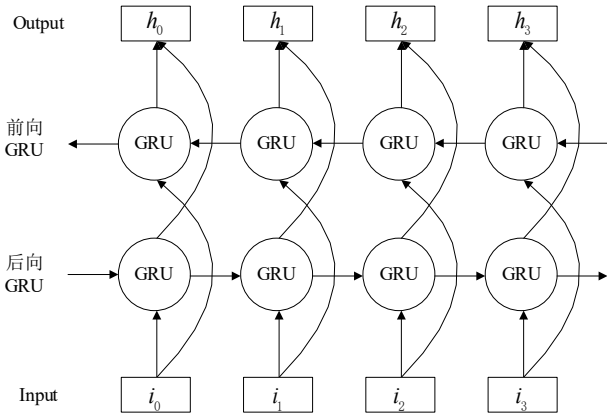


Fig. 5 Bidirectional GRU network model  
图5 双向GRU网络模型

### 2.4 Attention层

考虑到文本情感分析中每个字词对整个文本的情感倾向贡献度不同。本文为了强调关键信息在全文中的作用,通过FDSF模型将BiGRU最后时刻的隐藏层作为全局语义特征表示,并加入自注意力机制。注意力机制根据每个单词在全文中的相关程度分配权重,最终的输出即为BiGRU输出向量的加权和<sup>[23]</sup>。具体数学计算公式如式(6)所示:

$$u_t = \tanh(W_\omega h_t + b_\omega) \quad (6)$$

$$a_t = \frac{\exp(u_t)}{\sum_{i=0}^n \exp(u_i)} \quad (7)$$

$$F_c = \sum_{i=0}^n a_i h_i \quad (8)$$

式中:  $W_\omega$  为权值矩阵;  $b_\omega$  为偏置项;  $a_t$  为  $t$  时刻经过归一化处理得到的权重;  $F_c$  为经注意力机制加权后最终的全局语义特征信息。

### 2.5 融合输出层

首先将经过CNN通道获取的局部情感特征  $F_w$  和经过BiGRU-Attention通道获取的全局情感特征  $F_c$  进行拼接融合,得到最终的融合情感特征  $F = \frac{1}{2}(F_w + F_c)$ 。然后,将融合情感特征通过全连接层输出Softmax分类器运算得到最终分类预测概率。

$$p = \text{Softmax}(W_s F + b_s) \quad (9)$$

式中:  $W_s$  为全连接层的权值矩阵;  $b_s$  为偏置项。

## 3 实验与结果分析

### 3.1 实验环境

本文测试模型采用版本为Python + Pytorch1.9.0深度学习框架,实验环境为Google提供机器学习服务器Colaboratory,内置GPU为NVIDIA Tesla T4-16G。

### 3.2 实验数据集

为验证FDSF模型在中文情感分析任务上的有效性,本文使用ChinaNLPcorpus组织提供的中文情感分析数据集online\_shopping\_10\_cats和中科院谭松波学者收集的酒店评论语料数据集进行比较实验。其中,online\_shopping\_10\_cats为ChinaNLPcorpus组织公开发表的情感分析数据集,包含10种线上交易情感数据,包括正面、负面情感两种倾向,共计62773条数据(正向情感样本31727条,负向31046条),本文按照7:3的比例划分正负情感样本;酒店评论数据集包含正倾向性情感7000条,负倾向性情感3000条,本文设置训练样本9000条,测试样本1000条。

为验证模型的有效性,将本文所提方法与GRU模型<sup>[19]</sup>、BiGRU模型<sup>[24]</sup>、TextCNN模型<sup>[17]</sup>、AEN-BERT模型<sup>[25]</sup>、DPCNN模型<sup>[26]</sup>、BERT-base模型<sup>[13]</sup>这些主流深度学习分类模型进行比较。具体的,GRU模型使用单向GRU网络提取文本序列特征,将最后时刻隐藏层输入全连接层进行情感分类;BiGRU模型采用双向GRU网络结合上下文语义信息,增强模型在文本序列上的特征提取能力;TextCNN模型通过卷积层提取文本情感特征,运用最大池化提取最重要的情感特征,最后输入全连接层进行情感分类;DPCNN模型通过加深网络以增强文本情感特征提取能力,达到提升金字塔结构分类性能的目的;BERT-base模型以BERT(基于Transformer的双向编码模型,具有强大的语义表征能力)为预训练模型获取文本动态词向量后链

接全连接层,然后将其输入 Softmax 分类器中输出最终情感倾向;AEN-BERT 模型利用标签平滑方式在一定程度上解决了模糊情感文本的极性判断,通过融合注意力机制对方面词进行建模。实验参数设置如表 1 所示。

**Table 1 Experimental parameters**  
表 1 实验参数

参数描述	参数值
嵌入维度	768
卷积核大小	(1,2,3,4)
每种卷积核的数量	200
dropout	0.5
优化器	AdamW
学习率	2e-5
训练批次大小	64

### 3.3 评价标准

本文通过准确率(Accuracy)、精确率(Precision)、召回率(Recall)与 F1 值作为评价标准,如式(10)一式(13)所示,指标计算的混淆矩阵如表 2 所示。其中,TP 为预测为正的样本,FP 为预测为正的负样本,FN 为预测为负的正样本,TN 为预测为负的负样本。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (13)$$

**Table 2 Confusion matrix**

表 2 混淆矩阵

混淆矩阵		真实值	
		Positive	Negative
预测值	Positive	TP	FP
	Negative	FN	TN

### 3.4 实验结果及分析

为验证 FDSF 模型的在情感分析任务上的可靠性,避免出现偶然性结果。首先对 epoch 进行 10 次循环后采集数据测试集的准确率、精确率、召回率和 F1 值,然后求均值进行比较验证。online\_shopping\_10\_cats 数据集和酒店评论数据集上的实验结果如表 3、表 4 所示。由此可知,本

**Table 3 Experimental results of online\_shopping\_10\_cats dataset**

表 3 online\_shopping\_10\_cats 数据集上的实验结果 (%)

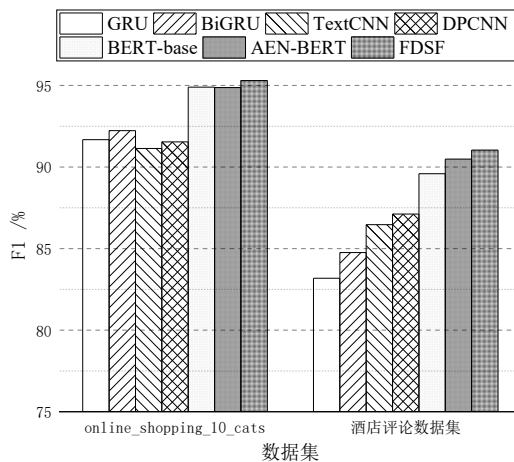
模型	准确率	精确率	召回率	F1 值
GRU	91.66	91.57	91.62	91.68
BiGRU	92.21	92.12	92.16	92.23
TextCNN	91.08	90.97	91.03	91.14
DPCNN	91.57	91.24	91.36	91.54
BERT-base	94.98	94.91	94.92	94.90
AEN-BERT	94.85	94.78	94.83	94.87
FDSF	95.28	95.21	95.26	95.30

文提出的 FDSF 模型在中文情感分类任务上相较于其它深度学习模型,在各项评级指标方面均最优。其中,各模型的 F1 值柱状图如图 6 所示。

**Table 4 Experimental results of hotel reviews dataset**

表 4 酒店评论数据集上的实验结果 (%)

模型	准确率	精确率	召回率	F1 值
GRU	84.81	84.68	83.12	83.18
BiGRU	86.38	86.14	85.03	84.76
TextCNN	87.89	86.06	85.66	86.47
DPCNN	88.63	86.49	86.01	87.12
BERT+base	91.55	89.66	89.99	89.59
AEN-BERT	91.99	90.10	90.33	90.49
FDSF	92.68	90.82	91.31	91.04



**Fig. 6 F1 value histogram analysis**

图 6 F1 值直方图分析

由图 6 可见,FDSF 模型在文本情感分析任务上性能良好,BiGRU 在两个数据集上的 F1 值相较于 GRU 分别提升 1%~3%,验证了双向门控循环机制在全局语义特征提取方面的优势。基于 BERT 的 3 类模型的各项指标均明显优于 GRU、BiGRU 和 TextCNN 模型,说明预训练模型提取的动态词向量具有更好的语义表征能力,这也是 FDSF 采用 BERT 预训练模型作为双通道词嵌入层的原因。

虽然,AEN-BERT 方法的 F1 值接近本文模型,但 FDSF 在两个数据集上相较于 AEN-BERT 方法分别提升 0.43%、0.65%,进一步说明了 FDSF 模型通过提取、融合多粒度下的局部和全局语义特征,能有效提升模型性能。

### 3.5 消融实验

本文设计消融实验,以验证 FDSF 模型中各层结构对模型的有效增益情况。其中,FDSF-CNN 为采用全局语义特征通道进行情感分析的模型;FDSF-BiGRU-Att 为原模型减去 BiGRU 结构和注意力机制后的模型,词嵌入后利用 CNN 提取多粒度下的局部语义特征进行分类;FDSF-Att 为原模型减去 BiGRU 结构后的注意力机制层模型,实验结果如表 5 所示。

**Table 5 Ablation experiment results**

表 5 消融实验结果 (%)

Model	online_shopping_10_cats	酒店评论
FDSF-CNN	94.67	90.13
FDSF-BiGRU-Att	95.01	90.67
FDSF-Att	95.12	90.89
FDSF	95.30	91.04

由表5可知,上述消融模型相较于BERT-base模型在两类数据集上的F1值均具有一定提升,说明BiGRU、CNN和注意力机制融合BERT均能效提升模型语义特征提取能力。实验发现,FDSF相较于结合门控循环单元与注意力机制进行全局语义特征提取的FDSF-CNN模型、结合CNN进行局部语义特征提取的FDSF-BiGRU-Att模型,在性能方面具有一定程度的提升,表明FDSF模型通过双通道融合特征向量的方式能提升模型性能。FDSF-Att与FDSF结果相差0.18%,说明在保持双通道结构的前提下,融入注意力机制能为模型性能带来增益。

综上所述,FDSF模型通过双通道方式融合BERT、BiGRU、Attention机制、CNN模型的优势,能显著提升模型在情感分析任务上的性能,充分提取文本语义特征信息以深入分析情感倾向。

## 4 结语

本文针对文本情感分类问题,充分结合了BERT、CNN、BiGRU模型与Attention的优势,提出融合双通道语义特征的情感特征模型。首先采用BERT预训练语言模型提取文本情感特征的动态表示,并将其分别输入两个通道进行下游分类任务;然后在CNN通道中提取局部情感特征,在BiGRU-Attention通道中提取全局情感特征;最后加权融合两个通道特征,并输入Softmax分类器得到最终结果。

在online\_shopping\_10\_cats、谭松波酒店评论数据集上,与其他深度学习分类模型进行比较实验的结果表明,FDSF模型在准确率、精确率、召回率、F1值均最优,证明了FDSF模型在文本情感分析任务中的有效性和优势。然而,考虑到通过双通道特征融合时会损失部分特征信息,未来将探究如何选择性地保留更多关键特征信息,寻求更优的融合方式获得更准确的情感倾向。

### 参考文献:

- [1] TAI K S, SOCHER R, MANNING C D. Improved semantic representations from tree-structured long short-term memory networks [DB/OL]. <https://arxiv.org/abs/1503.00075>.
- [2] ZHANG L, WANG S, LIU B. Deep learning for sentiment analysis: a survey [DB/OL]. <https://arxiv.org/abs/1801.07883>.
- [3] HUSSEIN D M E-D M. A survey on sentiment analysis challenges [J]. Journal of King Saud University Engineering Sciences, 2018, 30(4): 330-338.
- [4] ZHANG S, WEI Z, WANG Y, et al. Sentiment analysis of Chinese micro-blog text based on extended sentiment dictionary [J]. Future Generation Computer Systems, 2018, 81: 395-403.
- [5] PANG B, LEE L, VAITHYANATHAN S. Thumbs up? sentiment classification using machine learning techniques [C]// Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing, 2002:79-86.
- [6] BENGIO Y, DUCHARME R, VINCENT P. A neural probabilistic language model [J]. Advances in Neural Information Processing Systems, 2000, 3:1137-1155.
- [7] TANG L Y, XIONG C C, WANG Y, et al. Review of deep learning for short text sentiment tendency analysis [J]. Journal of Frontiers of Computer Science and Technology, 2021, 15(5):794-811.
- [8] 汤凌燕,熊聪聪,王娜,等.基于深度学习的短文本情感倾向分析综述 [J]. 计算机科学与探索,2021,15(5):794-811.
- [9] SUN X, TANG Z, ZHAO Y Y, et al. Hierarchical net-works with mixed attention for text classification [J]. Journal of Chinese Information Processing, 2021, 35(2):69-77.
- [10] 孙新,唐正,赵永妍,等.基于层次混合注意力机制的文本分类模型 [J]. 中文信息学报,2021,35(2):69-77.
- [11] HU Y L, TONG T Q, ZHANG X Y, et al. Self-attention based BGRU and CNN for sentiment analysis [J]. Computer Science, 2022, 49(1): 252-258.
- [12] 胡艳丽,童谭露,张啸宇,等.融入自注意力机制的深度学习情感分析方法 [J]. 计算机科学,2022,49(1):252-258.
- [13] BUCKMAN J, ROY A, RAFFEL C, et al. Thermometer encoding: one hot way to resist adversarial examples [C]// International Conference on Learning Representations, 2018:1-22.
- [14] LIU H X. Sentiment analysis of citations using Word2vec [DB/OL]. <https://arxiv.org/abs/1704.00177>.
- [15] PENNINGTON J, SOCHER R, MANNING C D. Glove: global vectors for word representation [C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, 2014: 1532-1543.
- [16] DEVLIN J, CHANG M W, LEE K, et al. Bert: pretraining of deep bidirectional transformers for language understanding [DB/OL]. <https://arxiv.org/abs/1810.04805>.
- [17] WANG Y H, LIN M, LI Y L, et al. Research on embedded text topic model based on BERT [J]. Computer Engineering and Applications, 2023, 59(1):169-179.
- [18] 王宇晗,林民,李艳玲,等.基于BERT的嵌入式文本主题模型研究 [J]. 计算机工程与应用,2023,59(1):169-179.
- [19] KARIMI A, ROSSI L, PRATI A. Adversarial training for aspect-based sentiment analysis with BERT [C]// 2020 25th International Conference on Pattern Recognition, 2021: 8797-8803.
- [20] PANG G, LU K, ZHU X, et al. Aspect-level sentiment analysis approach via BERT and aspect feature location model [J]. Wireless Communications and Mobile Computing, 2021(2):5534615.
- [21] KIM Y. Convolutional neural networks for sentence classification [DB/OL]. <https://arxiv.org/abs/1408.5882>.
- [22] ZHANG Y, JIANG Y, TONG Y. Study of sentiment classification for Chinese microblog based on recurrent neural network [J]. Chinese Journal of Electronics, 2016, 25(4): 601-607.
- [23] CHO K, MERRIENBOE V B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation [DB/OL]. <https://arxiv.org/abs/1406.1078>.
- [24] JELODAR H, WANG Y, ORJI R, et al. Deep sentiment classification and topic discovery on novel coronavirus or COVID-19 online discussions: NLP using LSTM recurrent neural network approach [J]. IEEE Journal of Biomedical and Health Informatics, 2020, 24(10): 2733-2742.
- [25] UMER M, ASHRAF I, MEHMOOD A, et al. Sentiment analysis of tweets using a unified convolutional neural network-long short-term memory network model [J]. Computational Intelligence, 2021, 37(1): 409-434.
- [26] LIU Y, LU J, YANG J, et al. Sentiment analysis for e-commerce product reviews by deep learning model of Bert-BiGRU-Softmax [J]. Mathematical Biosciences and Engineering, 2020, 17(6): 7819-7837.
- [27] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [DB/OL]. <https://arxiv.org/abs/1706.03762>.
- [28] DEY R, SALEM F M. Gate-variants of gated recurrent unit (GRU) neural networks [DB/OL]. <https://arxiv.org/abs/1701.05923>.
- [29] YU Q, WANG Z, JIANG K. Research on text classification based on BERT-BiGRU model [C]// Beijing:2020 3rd International Conference on Modeling, Simulation and Optimization Technologies and Applications, 2020.