

基于ERNIE和融合双通道特征的文本情感分类模型

尧欢欢, 朱小栋

(上海理工大学管理学院, 上海 200093)
✉yaohhh1011@163.com; zhuxd@usst.edu.cn



摘要: 针对传统词向量模型无法获取完整的语义表达, 以及基础神经网络模型未能兼顾提取多种关联特征等问题, 提出了一种融合预训练语言模型(ERNIE)和深层金字塔神经网络结构(DPCNN)/双向门控循环单元-注意力机制(BiGRU-Attention)的双通道文本情感分类模型。基于DPCNN的左通道负责提取文本长距离依赖表示, 基于BiGRU-Attention的右通道负责提取文本时间序列特征和关键信息。此外, 均使用ERNIE模型提供动态字向量。最后, 拼接融合双通道中的信息特征以获取最终的文本表示。实验结果表明, ERNIE-DBGA模型的准确率最高达到97.05%, 优于其他对比方法, 验证该模型可以有效提升情感分类的性能。

关键词: 文本情感分类; ERNIE; 双通道; DPCNN; BiGRU; 注意力机制

中图分类号: TP391 **文献标识码:** A

Text Sentiment Classification Model based on ERNIE and Fusion of Dual-Channel Features

YAO Huanhuan, ZHU Xiaodong

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

✉yaohhh1011@163.com; zhuxd@usst.edu.cn

Abstract: In response to the problem that the traditional word vector model cannot obtain the complete semantic expression, and the basic neural network model cannot take into account the extraction of multiple associated features, this paper proposes a dual-channel sentiment classification model combining ERNIE (a pre-training language representation model enhanced by knowledge) and DPCNN (Deep Pyramid Convolutional Neural Networks) / BiGRU (Bidirectional Gated Recurrent Unit)-Attention (ERNIE-DBGA). The left channel based on DPCNN is responsible for extracting the long-distance dependency representation of the text, while the right channel based on BiGRU-Attention is responsible for extracting the text time series features and key information. In addition, the ERNIE model is used to provide dynamic word vectors. Finally, the information features in the two channels are concatenated and fused to obtain the final text representation. The experimental results show that the accuracy of ERNIE-DBGA model is up to 97.05%, which is superior to other comparison methods, verifying that the model can effectively improve the performance of sentiment classification.

Keywords: text sentiment classification; ERNIE; dual-channel; DPCNN; BiGRU; attention mechanism

1 引言(Introduction)

随着社交网络媒体和电子商务的智能化发展, 越来越多的网民倾向于在各大平台针对特定主题或者商品发表观点和意见, 由此生成海量蕴含情感色彩的文本评论数据, 这些情绪资源隐含了互联网用户的潜在诉求, 具有巨大的商业和

社会价值。通过对文本评论数据进行情感分析, 可以获得互联网用户对各种事物的意见倾向, 获取潜在的经济价值。然而, 传统采用人工分辨文本情感极性的方式耗费了大量的人力且主观性太强, 为了能够自动、高效地探究隐含在评论信息中的情感色彩, 文本情感分析技术应运而生。

文本情感分析^[1]又称文本导向分析或意见挖掘,是指对主观评论中隐含的情感色彩进行提取和分析的过程,主要方法包括情感词典规则方法、统计机器学习方法及深度学习模型。传统基于情感词典的方法依赖人工预先编制包含不同词汇和相应极性的情感词典,该方法使用简单但局限性大。传统基于机器学习的方法采用机器学习算法学习特征,性能上优于情感词典,但该方法依靠人工对文本特征进行标记,所需的人工成本高,而且人工标注的主观性也会影响分类的精度。基于深度学习的方法是机器学习算法的一个新兴分支,目前多数研究通过使用深度神经网络构建情感分类器,主流模型有卷积神经网络(CNN)、循环神经网络(RNN)和注意力机制网络等。受人工神经网络模型的启发,该方法可以自动学习文本的语义特征,弥补传统机器学习算法需人工干预特征的缺陷。然而,上述方法仅考虑了文本中的单个特征,对隐含在文本中的多角度和多关联特征的提取不充分,信息在训练过程中容易出现遗漏。

针对上述问题,本文提出了一种结合ERNIE、DPCNN和BiGRU-Attention的双通道情感分类模型。该模型利用ERNIE获取文本映射向量,一方面通过DPCNN抽取文本长距离语义依赖;另一方面采用BiGRU对整体文本序列进行时间建模,并引入注意力机制加强局部关键知识学习。实验结果表明,本文提出的融合模型在weibo_senti_100k和京东商城评论两个数据集上提高了准确率和宏平均值F1。

2 相关研究(Related research)

早期的情感分析方法分为两类:基于情感词典和基于机器学习的方法。基于情感词典的方法操作简单,但难点在于构造一个适用于各种领域的通用词典。有不少学者尝试突破这个局限,例如SIVASANKAR等^[2]融合多领域数据集术语之间的上下文关系,提出了一种改进的领域自适应增强情感词典。虽然该方法能够增强不同领域间情感分析的准确性,但是人工构建的词典仍然存在词汇覆盖度不足和分类精度不高等问题。基于传统机器学习的方法主要包括K最近邻(KNN)、朴素贝叶斯(NB)和支持向量机(SVM)等。马文等^[3]采用朴素贝叶斯算法对中文的评论分类进行研究,文本分类的准确率达到97.65%。然而,该方法依赖手动选择特征,需要采用人工干预获取输入文本的情感类别。

近年来,深度学习技术基于自动特征选择的特性,已经发展为文本情感分类领域的主流方法,包括卷积神经网络、循环神经网络和注意力机制网络等。KIM^[4]提出CNN模型用于文本语句情感分类研究并取得了不错的效果。KHAN等^[5]利用CNN-LSTM架构对罗马乌尔都语和英语方言文本进行

情感分类,在四个语料库上的大量测试表明,该混合结构模型能够提升情感分类的准确率。GAO等^[6]提出一种基于CNN模型和BiGRU网络的混合模型用于短文本方面情感分析,实验结果表明改进的CNN+BiGRU模型分类效果更好。注意力机制能够使计算机读取信息时选择性地抓住重点,已成功应用于机器视觉领域,因此许多研究者将其引入自然语言处理任务中。谢润忠等^[7]构建了双通道的注意力模型用于预测评论文本的情感极性。此外,基于CNN、RNN和注意力机制等的融合模型也被广泛应用。BHUVANESHWARI等^[8]提出一种基于BiLSTM自注意的卷积神经网络(BAC)模型,使用CNN和BiLSTM自动学习评论文本的分类特征,并应用注意机制捕捉上下文重点信息。YANG等^[9]融合了CNN和基于注意力的BiGRU模型用于文本情感分类,在约10万条当当网的真实图书数据集上进行测试,实验结果表明该模型能够显著提高文本分类的性能。李卫疆等^[10]提出了一种融合多通道特征和自注意力的BiLSTM情感分类方法用于文本情感分析,在五个基线数据集上进行实验,得到较好的分类效果。沈彬等^[11]提出一种基于知识增强语义表示和双重注意力机制的微博情感分析模型,在两个公开数据集上进行实验,得到较好的分类效果。本文融合基于知识增强语义表示的ERNIE模型和基于DPCNN、BiGRU-Attention的双通道特征对评论文本进行情感极性的预测。

3 ERNIE-DBGA模型(ERNIE-DBGA model)

本文设计的模型基于ERNIE和融合双向DPCNN/BiGRU-Attention,纵向结构分为五个层次:输入层、预训练语言模型层、特征提取层、特征融合层和输出层;横向结构分为两个数据处理通道:左侧为基于DPCNN网络的通道,用于捕捉文本长距离依赖特征,右侧为基于BiGRU网络和注意力机制的通道,用于捕获上下文时序特征和聚焦关键信息点,此外双通道均采用预训练语言模型ERNIE完成文本的向量化表示。ERNIE-DBGA模型结构如图1所示。

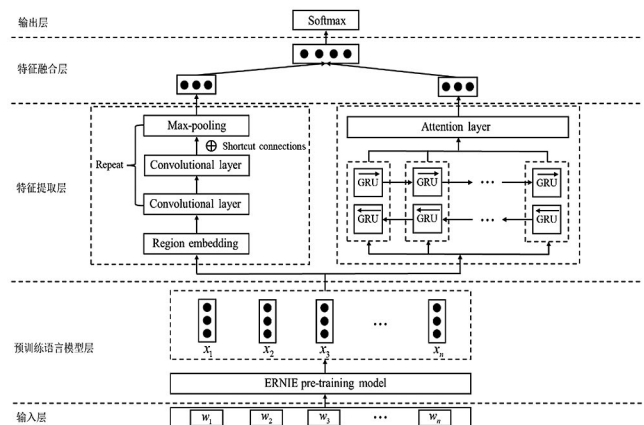


图1 ERNIE-DBGA模型结构图

Fig.1 ERNIE-DBGA model framework map

3.1 输入层

给定一串文本序列, 经预处理后得到 $w = \{w_1, w_2, w_3, \dots, w_n\}$, 其中 w_i 表示评论文本的第 i 个字, 然后输入预训练语言模型 ERNIE 层进行向量编码。

3.2 预训练语言模型层

ERNIE^[12]模型是以自编码语言模型(BERT)^[13]为基础建立的, 相较于BERT, ERNIE能够提取到更完整的语义知识信息。本文使用ERNIE作为语义表征模型, 对掩蔽策略进行了改进。不同于BERT的单粒度掩蔽, ERNIE将训练数据中的短语和实体当成一个整体单元统一掩蔽, 这样做的好处是模型在学习的时候不会遗漏任何的语料信息。ERNIE的内部网络由多层双向变压器(Transformer)的编码器(Encoder)端堆叠而成。Encoder层包含self-attention层和feed-word network层两部分。ERNIE的模型结构如图2所示。

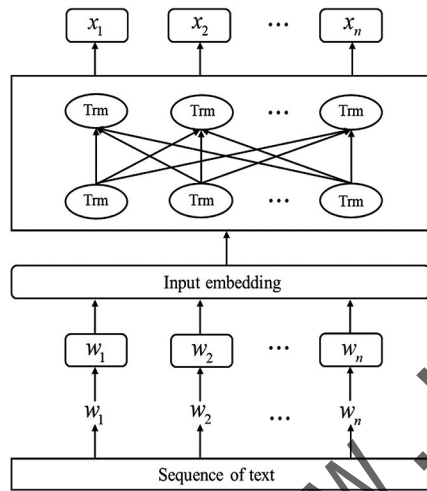


图2 ERNIE模型架构图

Fig.2 ERNIE model framework map

3.3 特征提取层

对于文本序列的信息提取, 本文构建了两个通道并行的模型结构, 减少了由于模型深度增加导致的梯度消失现象。左侧为深层金字塔神经网络结构(DPCNN)提取文本远距离依赖信息, 右侧为引入注意力机制的双向门控循环单元(BiGRUAtt)提取文本隐含的上下文关联和重要信息。

3.3.1 DPCNN模型

DPCNN由JOHNSON^[14]在自然语言处理(NLP)顶级会议中ACL 2017提出, 该模型加深了CNN的网络深度, 但不会带来太大的计算消耗。DPCNN模型结构如图3所示, 首先将输入的文本序列经过卷积操作后转化为区域嵌入层(Region embedding), 然后传入由两个固定卷积核的卷积函数构成的卷积块(block)中, 接着输入到增加了负采样的重复卷积结构中, 最后通过最大池化层(max-pooling)进行池化得到文本的

向量表示 V 。

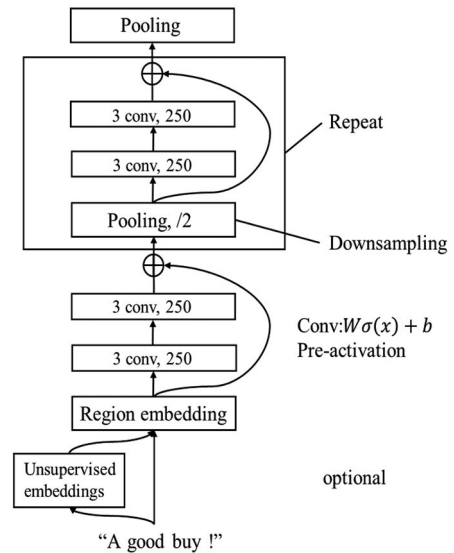


图3 DPCNN模型架构图

Fig.3 DPCNN model framework map

3.3.2 BiGRU模型

门控循环单元(GRU)是长短期记忆网络(LSTM)的一种改进模型。LSTM能够有效缓解循环单元在处理长文本时产生的长期记忆和梯度消失问题, 但模型结构复杂、参数多及训练时间长。基于此, CHO等^[15]提出了一个更简单的GRU模型。GRU将LSTM中的遗忘门和输入门整合成一个更新门, 结构更轻盈, 参数更少, 处理相同任务所需的时间成本更低。GRU的模型结构如图4所示。

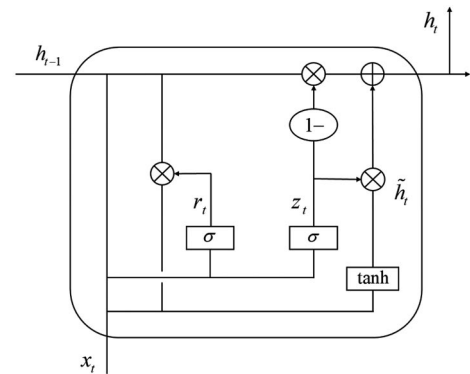


图4 GRU模型架构图

Fig.4 GRU model framework map

单向神经网络结构GRU只能处理单向语义关系, 本文使用的BiGRU模型结合了前向GRU和后向GRU, 有利于评论文本深层次特征的提取, 可以学习到更加完整的特征信息, 获得更好的分类效果。

3.3.3 注意力机制

注意力机制通过动态分配权重衡量文本中字词的重要程度, 解决了传统深度学习中将所有特征赋予相同权重的问题。本文在BiGRU模型的基础上引入注意力机制, 以期学习

更多隐藏的重要信息。

3.4 特征融合层

特征融合层通过矩阵拼接方式将左侧DPCNN模型生成的特征向量 V^1 和右侧BiGRUAtt模型生成的特征向量 V^2 融合成一个矩阵,目的是保留双通道的信息,最后生成融合全局的特征表示,计算表达式如下:

$$V^* = [V^1, V^2] \quad (1)$$

3.5 输出层

输出层采用softmax分类器进行归一化,将特征融合得到的语义信息 V^* 输入softmax层中进行概率转换,输出获得最大概率的情感极性,从而得到模型的预测结果,计算表达式如下:

$$y = \text{softmax}(WV^* + b) \quad (2)$$

其中, W 为权重系数, b 为偏置项, y 为情感分类预测的最终结果。

3.6 模型训练

模型训练是深度学习的关键步骤,本文采用交叉熵损失函数(Binary Cross-entropy)以衡量输出的真实结果与相应预测结果的吻合程度,并通过反向传播算法(BP算法)学习模型参数,计算表达式如下:

$$Loss = - \sum_{i=1}^n y_i \lg \hat{y}_i + (1 - y_i) \lg(1 - \hat{y}_i) \quad (3)$$

其中, y_i 表示输出的真实情感结果, $\hat{y}_i$ 表示输出的预测情感结果。

4 实验与分析(Experiment and analysis)

4.1 实验数据集

本文采用2个数据集进行实验。

实验数据1:采用公开数据集weibo_senti_100k,该语料包含119,988条带情感分类标签的新浪微博评论文本,其中积极评论文本为59,993条,消极评论文本为59,995条。因为社交评论数据夹杂大量噪声,这些无意义的的数据会对分类结果产生影响,所以本文使用正则表达式匹配字符的方法对微博评论文本进行数据清洗,清除URL、@人名和回复转发微博等无实际意义的数据,以提高数据集的质量。预处理后的数据集部分样例如表1所示。

表1 部分数据集样本

Tab.1 Partial data set samples

情感极性	文本内容
积极	一不留神阿姨已经把饺子包好了本来我还想看看哪个包了宝贝在里面呢嘻嘻
消极	无耻贪官挥霍的是金钱毁掉的是民族的脊梁怒怒怒无耻他们还在就是我们的噩梦

实验数据2:采用网络爬虫抓取京东商城中热销水果的评论文本,其中包括苹果、橙子、梨、奇异果/猕猴桃和葡萄/提子五个品类。为了划分情感极性便于后期的标签分类,将

评论按照星级评分爬取,判定标准为4星及以上为积极评论,2星及以下为消极评论。此外,由于电商评论具有随意性,存在少部分星级和评价内容情感严重不符的数据,所以对这部分评论进行人工打标签。此外,鉴于爬取的数据噪声大,含有大量无意义和重复的评论,因此对评论文本进行去重、过滤等数据清洗操作,最终获取积极评论和消极评论各10,000条。预处理后的数据集部分样例如表2所示。

表2 部分数据集样本

Tab.2 Partial data set samples

情感极性	文本内容
积极	橙子礼盒装的很好看感觉还不错哦个头也大
消极	这批的阿克苏苹果不是很甜也不够脆也没看到糖心一直是挺相信农夫山泉的水果的这次有点失望啊

本文将两个数据集分别随机打乱,划分为训练集、验证集和测试集,三者评论文本的数量比例为8:1:1。数据集详情如表3所示。

表3 两个数据集详情

Tab.3 Details of two data sets

数据集	训练集/条	验证集/条	测试集/条	共计/条
weibo_senti_100k	95,990	11,999	11,999	119,988
京东商城评论	16,000	2,000	2,000	20,000

4.2 实验环境

本文使用Python作为模型实现语言,Pytorch作为实验框架,实验采用的环境配置如表4所示。

表4 实验环境配置

Tab.4 Experimental environment configuration

开发环境	参数
处理器	Intel(R)Xeon(R)CPU E5-2603 v4 @1.70 GHz
内存	32.0 GB
深度学习框架	Pytorch
Python	3.7
编码格式	utf-8

4.3 实验参数

模型参数的选择和设定会对分类的精度产生直接影响。本文采用小批量梯度下降法(Mini-batch Gradient Descent)进行模型训练,采用Adam作为优化器,并采用早停法(Early Stopping)避免过拟合。由传统词向量模型搭建的神经网络模型均采用中文词向量(Chinese-Word-Vectors)提供词向量,词向量维度为300维;学习率为 $1e-3$,丢弃率为0.1;文本卷积神经网络(TextCNN)中卷积核尺寸为(2、3、4),卷积核个数为256;RNN中隐藏层特征数为256,层数为2。由预训练语言模型搭建的神经网络模型中,ERNIE采用百度的ERNIE1.0-base中文预训练模型,BERT采用谷歌的BERT-base中文预训练模型。本文模型参数设置如表5所示。

表5 模型参数设置

Tab.5 Model parameter setting

模型参数	设定值
Batch_size_data1	128
Batch_size_data2	32
Hidden_size	768
Pad_size	64
Learning_rate	5e-5
Dropout_rate	0.1
DPCNN_filter_sizes	(3,1)
DPCNN_stride	2
DPCNN_num_filters	1,536
BiGRU_num_layers	2
BiGRU_hidden_size	768

4.4 实验评价指标

为了平衡情感类别对实验结果的影响,本文使用宏平均(Macro avg)下的精确率 P (Precision)、召回率 R (Recall)和 $F1$ 值($F1$ -score)及准确率 A (Accuracy)四个标准衡量情感分类模型的性能。计算表达式如下:

$$P = \frac{TP}{TP + FP} \quad (4)$$

$$R = \frac{TP}{TP + FN} \quad (5)$$

$$F1 = \frac{2 \times P \times R}{P + R} \quad (6)$$

$$P_{\text{Macro}} = \frac{P_{\text{pos}} + P_{\text{neg}}}{2} \quad (7)$$

$$R_{\text{Macro}} = \frac{R_{\text{pos}} + R_{\text{neg}}}{2} \quad (8)$$

$$F1_{\text{Macro}} = \frac{F1_{\text{pos}} + F1_{\text{neg}}}{2} \quad (9)$$

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

4.5 实验对比结果和分析

为了全面评估ERNIE-DBG A模型的性能,本文设置了四组对比实验。第一组实验对比情感分类任务中的经典模型以验证本文模型的有效性;第二组实验比较不同词向量模型对情感极性分类的影响;第三组实验检验本文提出模型中各通道对分类结果的影响;第四组探究模型中迭代轮数对分类性能的影响。

4.5.1 情感分析模型对比实验

为了验证模型情感分类的性能,本文在两个数据集上分别设置对比实验。其中,对比模型均采用Chinese-Word-Vectors提供词向量。

TextCNN^[4]: 文本分类卷积神经网络,包括卷积层和最大池化层,通过捕捉文本中的重要信息指导情感极性预测。

LSTM^[16]: 长短时记忆网络,采用门结构和记忆单元对时

间序列建模,学习文本的上下文交互关系。

GRU^[15]: 门控循环单元,在LSTM的基础上优化了内部结构,仅保留了两个门单元但效果很好。

BiGRU^[15]: 双向门控循环单元,使用前向和后向GRU网络学习文本信息,最后将所得的向量拼接起来获得最终特征表示。

TextRCNN^[17]: 循环卷积神经网络,应用循环结构获取时序信息,并使用最大池化层压缩文本特征以提取主要信息。

BiGRUAtt^[18]: 基于注意力机制的双向门控循环单元,采用注意力机制动态分配权重使模型能提取到更深层次的隐藏特征。

DPCNN^[14]: 加深的金字塔卷积神经网络,可以获得文本的长距离依赖表示。

ERNIE-DBG A: 本文模型。

表6呈现了本文提出的ERNIE-DBG A模型与情感分类任务中经典模型的实验对比结果。从表6可以看出,ERNIE-DBG A模型在数据集1和数据集2上的各项评价指标均获得最优结果。相比TextCNN、LSTM、GRU、BiGRU、TextRCNN、DPCNN和BiGRUAtt模型,ERNIE-DBG A模型在数据集1上的宏平均 $F1$ 值和准确率 A 都相应提高了1.36%、0.99%、0.91%、0.87%、0.56%、0.82%和0.66%,在数据集2上的宏平均值 $F1$ 和准确率 A 都相应提高了2.35%、2.35%、2.20%、1.95%、1.56%、2.25%和1.70%。由此可以看出,相比当前情感分类任务中的经典模型,本文所提模型情感分类性能更佳,表明了本文模型构建的有效性。

表6 不同情感分类模型的对比结果

Tab.6 Comparison results of different sentiment classification models

语料库	模型	$P_{\text{Macro}}/\%$	$R_{\text{Macro}}/\%$	$F1_{\text{Macro}}/\%$	$A/\%$
数据集1	TextCNN	94.88	94.88	94.88	94.88
	LSTM	95.25	95.25	95.25	95.25
	GRU	95.34	95.33	95.33	95.33
	BiGRU	95.38	95.37	95.37	95.37
	TextRCNN	95.74	95.68	95.68	95.68
	DPCNN	95.44	95.42	95.42	95.42
	BiGRUAtt	95.60	95.58	95.58	95.58
	ERNIE-DBG A	96.32	96.24	96.24	96.24
数据集2	TextCNN	94.70	94.70	94.70	94.70
	LSTM	94.74	94.69	94.70	94.70
	GRU	94.89	94.84	94.85	94.85
	BiGRU	95.16	95.09	95.10	95.10
	TextRCNN	95.50	95.49	95.49	95.49
	DPCNN	94.81	94.80	94.80	94.80
	BiGRUAtt	95.39	95.34	95.35	95.35
	ERNIE-DBG A	97.08	97.04	97.05	97.05

4.5.2 词向量模型对比实验

为了探究不同词向量模型对分类性能的影响，本文在两个数据集上分别设置对比实验。

W2V-DBGA模型：采用词嵌入模型(Word2Vec)静态词向量模型作为嵌入层，然后输入双通道DBGA模型中训练。

BERT-DBGA模型：采用BERT动态预训练语言模型作为嵌入层，然后输入双通道DBGA模型中训练。

ERNIE-DBGA模型：本文模型。

基于不同词向量模型的实验对比结果如表7所示。相比静态词向量模型Word2Vec，动态预训练语言模型BERT采用基于注意力机制的双向Transformer编码器，结合上下文语境生成动态向量表示，有效解决了Word2Vec中存在的一词多义问题。从表7可以看出，BERT-DBGA模型比W2V-DBGA模型在数据集1和数据集2上的四项评价指标结果均有所提高，说明了融合文本全局知识的动态词向量模型较静态词向量模型而言，提取到的词向量蕴含更丰富的特征表示。为了进一步获取更完整的语义信息，本文采用同为动态预训练语言模型的ERNIE作为嵌入层。相比于BERT，ERNIE增加了先验语义知识单元，通过对文本中的句法和词法结构学习建模，充分吸收了序列中的知识信息，实验结果得到显著提升。

表7 不同词向量模型的对比结果

Tab.7 Comparison results of different word vector models					
语料库	模型	$P_{Macro}/\%$	$R_{Macro}/\%$	$F1_{Macro}/\%$	$A/\%$
数据集1	W2V-DBGA	95.43	95.38	95.38	95.38
	BERT-DBGA	96.05	95.97	95.97	95.97
	ERNIE-DBGA	96.32	96.24	96.24	96.24
数据集2	W2V-DBGA	94.97	94.94	94.95	94.95
	BERT-DBGA	95.93	95.89	95.90	95.90
	ERNIE-DBGA	97.08	97.04	97.05	97.05

4.5.3 各通道模型对比实验

为了检验模型中各个通道对模型性能的影响，本文在两个数据集上分别设置对比实验。

ERNIE-D模型：基于ERNIE的左侧单通道DPCNN模型。

ERNIE-BGA模型：基于ERNIE的右侧单通道BiGRUAtt模型。

ERNIE-DBGA模型：本文模型。

基于不同通道模型的实验对比结果如表8所示。从表8可以看出，相比单通道的ERNIE-D模型和ERNIE-BGA模型，本文所提出的双通道情感分类模型ERNIE-DBGA在数据集1和数据集2上的各项评价指标均得到提升。相比左通道ERNIE-D模型，ERNIE-DBGA模型在右侧引入了由

BiGRU网络和注意力机制相结合的模型，可以额外提取到文本中的时间序列特征和关键知识特征，有助于进一步挖掘更深层次的潜在语义信息，进而使分类性能得到提升。相比右通道ERNIE-BGA模型，ERNIE-DBGA模型在左侧引入了DPCNN模型，可以额外提取到文本中的远距离依赖特征，有利于捕获文本中连续丰富的知识信息，所以分类结果更佳。综上所述，基于双通道的ERNIE-DBGA模型结合了两个通道模型的优点，能够有效获取文本中多种关联特征，在一定程度上提高情感分类的性能。

表8 各通道模型的对比结果

Tab.8 Comparison results of each channel model					
语料库	模型	$P_{Macro}/\%$	$R_{Macro}/\%$	$F1_{Macro}/\%$	$A/\%$
数据集1	ERNIE-D	96.24	96.18	96.18	96.18
	ERNIE-BGA	96.22	96.19	96.19	96.19
	ERNIE-DBGA	96.32	96.24	96.24	96.24
数据集2	ERNIE-D	96.57	96.54	96.55	96.55
	ERNIE-BGA	96.62	96.59	96.60	96.60
	ERNIE-DBGA	97.08	97.04	97.05	97.05

4.5.4 迭代次数(Epoch)取值实验

迭代次数是模型训练过程中一个重要参数，它的取值将直接影响分类的准确率。本实验在保持其他参数不变的情况下改变训练轮数，通过观察模型准确率的变化确定最适合本文的迭代次数。

从图5可以看出，*epoch*取值的变化会对分类的准确率产生影响。随着*epoch*取值增加，模型准确率呈上升趋势；当*epoch*=3时，准确率达到最优值97.05%；当*epoch*取值持续增加时，模型准确率缓慢下滑后又小幅上升，最后准确率在96.5%上下小幅波动，趋于平稳。由此可以看出，合适的*epoch*取值能有效提升分类的准确性，考虑到计算的复杂度和时间成本等因素，本文将*epoch*设置为3。

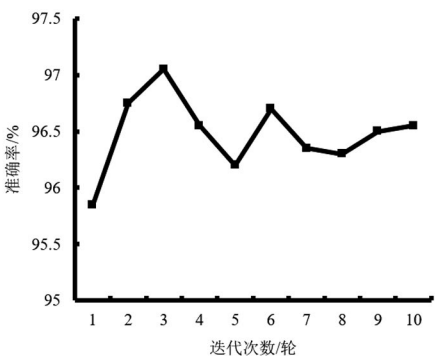


图5 不同迭代次数的对比结果

Fig.5 Comparison results of different iterations

5 结论(Conclusion)

针对文本情感分类任务中传统语言模型无法充分获取整

体情境语义信息, 以及传统深度学习模型未能有效提取文本序列中多类关联信息等问题, 提出了一种基于动态预训练语言模型和双通道融合的文本情感分类模型ERNIE-DBGA。该模型采用ERNIE作为嵌入层, 学习文本知识信息, 增强了模型获得深层语言表征的能力, 同时, 构建了基于DPCNN和BiGRUAtt的并行神经网络模型以提取和融合文本中长期语境信息、上下文状态信息和局部重点知识信息, 最后采用Softmax分类器实现文本情感分类。本文设计了情感分类任务中的经典方法对比、不同词向量工具对比、各通道模型对比和迭代次数取值四组实验, 在两个数据集weibo_senti_100k和京东电商评论上检验模型的有效性, 实验结果表明本文提出的双通道文本情感分类模型具有更好的分类性能。在未来的工作中考虑进一步研究文本情感多标签分类, 并不断优化算法模型, 以提高模型分类的准确率和泛化能力。

参考文献(References)

- [1] PANG B, LEE L. Opinion mining and sentiment analysis[J]. Foundations and Trends in Information Retrieval, 2008, 2(1-2): 1-135.
- [2] SIVASANKAR E, KRISHNAKUMARI K, BALASUBRAMANIAN P. An enhanced sentiment dictionary for domain adaptation with multi-domain dataset in Tamil language (ESD-DA)[J]. Soft Computing, 2020, 25:1-15.
- [3] 马文, 陈庚, 李昕洁, 等. 基于朴素贝叶斯算法的中文评论分类[J]. 计算机应用, 2021, 41(S2):31-35.
- [4] KIM Y. Convolutional neural networks for sentence classification[C]// MOSCHITTI A, PANG B, DAELEMANS W. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar: ACL, 2014: 1746-1751.
- [5] KHAN L, AMJAD A, AGAQ K M, et al. Deep sentiment analysis using CNN-LSTM architecture of English and Roman Urdu Text shared in social media[J]. Applied Sciences, 2022, 12(5):2694-2694.
- [6] GAO Z W, LI Z Y, LUO J Y, et al. Short text aspect-based sentiment analysis based on CNN + BiGRU[J]. Applied Sciences, 2022, 12(5):2707-2707.
- [7] 谢润忠, 李烨. 基于BERT和双通道注意力的文本情感分类模型[J]. 数据采集与处理, 2020, 35(04):642-652.
- [8] BHUVANESHWARI P, RAO A N, Robinson Y H, et al. Sentiment analysis for user reviews using Bi-LSTM self-attention based CNN model[J]. Multimedia Tools and Applications, 2022, 81(9):12405-12419.
- [9] YANG L, LI Y, WANG J, et al. Sentiment analysis for E-Commerce product reviews in Chinese based on sentiment lexicon and deep learning[J]. IEEE Access, 2020, 8: 23522-23530.
- [10] 李卫疆, 漆芳, 余正涛. 基于多通道特征和自注意力的情感分类方法[J]. 软件学报, 2021, 32(09):2783-2800.
- [11] 沈彬, 严馨, 周丽华, 等. 基于ERNIE和双重注意力机制的微博情感分析[J]. 云南大学学报(自然科学版), 2022, 44(03): 480-489.
- [12] SUN Y, WANG S H, LI Y K, et al. ERNIE: Enhanced representation through knowledge integration[EB/OL]. (2019-04-19)[2019-12-23]. <https://arxiv.org/abs/1904.09223v1>.
- [13] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding[J]. Computation and Language, 2018(10): 4171-4186.
- [14] JOHNSON R. Deep pyramid convolutional neural networks for text categorization[C]// ZHANG T. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Stroudsburg, PA: ACL, 2017:562-570.
- [15] CHO K. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]// BART V M, CAGLAR G. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar: ACL, 2014:1724-1734.
- [16] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8):1735-1780.
- [17] LAI S W. Recurrent convolutional neural networks for text classification[C]// XU L H, LIU K. Proceedings of the AAAI Conference on Artificial Intelligence, Austin, Texas, USA: ACM, 2015:2267-2273.
- [18] 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, 2018, 36(12):1-10.

作者简介:

尧欢欢(1999-), 女, 硕士生. 研究领域: 深度学习, 自然语言处理, 人工智能.

朱小栋(1981-), 男, 博士, 副教授. 研究领域: 数据挖掘与深度学习, 云计算与大数据管理, 电子商务.