



基于 ERNIE 和双重注意力机制的微博情感分析

沈 彬^{1,2}, 严 馨^{1,2**}, 周丽华³, 徐广义⁴, 刘艳超⁵

(1. 昆明理工大学 信息工程与自动化学院, 云南 昆明 650500; 2. 昆明理工大学 云南省人工智能重点实验室, 云南 昆明 650500; 3. 云南大学 信息学院, 云南 昆明 650500; 4. 云南南天电子信息产业股份有限公司, 云南 昆明 650040; 5. 湖北工程学院 信息技术中心, 湖北 孝感 432000)

摘要:针对传统词向量无法在上下文中表示词的多义性,以及先验的情感资源未能在神经网络中得到充分利用等问题,提出一种基于知识增强语义表示(Enhanced Representation through Knowledge Integration, ERNIE)和双重注意力机制(Dual Attention Mechanism, DAM)的微博情感分析模型 ERNIE-DAM. 首先利用现有的情感资源构建一个包含情感词、否定词和程度副词的情感资源库;其次采用 BLSTM 网络 and 全连接网络分别对文本和文本中包含的情感信息进行编码,不同的注意力机制分别用于提取文本和情感信息中的上下文关系特征和情感特征,并且均采用 ERNIE 预训练模型获取文本的动态特征表示;最后将上下文关系特征和情感特征进行拼接融合,获取最终的特征向量表示. 实验结果表明,新模型在 COAE2014 和 weibo_senti_100k 数据集上的分类准确率分别达到了 94.50% 和 98.23%,同时也验证了将情感资源运用到神经网络中的有效性.

关键词:情感分析;知识增强语义表示;长短时记忆网络;注意力机制

中图分类号:TP391 **文献标志码:**A **文章编号:**0258-7971(2022)03-0480-10

随着微博、知乎、Twitter 等社交平台的快速发展和普及,越来越多的用户可以在这些社交平台上发表自己对某个话题的观点、态度和情感,由此产生了大量带有情感色彩的评论文本数据. 微博评论文本属于短文本的范畴,具有口语化、噪声大、网络新词多以及上下文信息有限等特点,因此给微博情感分析带来了巨大的挑战^[1]. 如何针对微博评论文本的这些特点进行情感分析,挖掘其中蕴含的有价值的信息,对于政府决策、产品营销和网络舆情监控等具有重要的意义^[2].

情感分析方法主要包括基于情感词典的方法、基于传统机器学习的方法和基于深度学习的方法. 基于情感词典的方法实质上是一种“词典+规则”的方法,该方法需要构建一个领域情感词典,并结合人工设计的一些规则来判断评论文本的情感倾向性. 例如,赵妍妍等^[3]通过构建副词、否定词、情感表情等相关词典来扩充情感词典,在微博情感分类的性能上相比基线方法提高了 1.13%. Zhang 等^[4]

使用扩展情感词典来计算微博文本情感分类的权重,相比基本情感词典提高了近 10%. Xu 等^[5]构造了一个包含基本情感词、领域情感词和多义情感词的扩展情感词典,并利用扩展情感词典和设计的评分规则进行文本情感分类. 然而,基于情感词典的方法严重依赖于情感词典的构建和人工设计的规则,人工干预大,效果较差.

基于传统机器学习的方法通过人工设计特征,并将这些特征映射为多维特征向量来训练一个分类器,性能上较基于情感词典的方法有了很大的提升,常用的分类器有支持向量机(Support Vector Machine, SVM)、朴素贝叶斯(Naive Bayes, NB)和最大熵(Maximum Entropy, ME)等. 例如,杨爽等^[6]提出一种基于 SVM 多特征融合的微博多级情感分类方法,通过融合语义、情感、词性等特征实现了情感的五级分类. Naz 等^[7]提出了一种利用 SVM 分类器和 n-gram 特征进行 Twitter 情感分析的方法. 然而,基于传统机器学习的方法依赖于情感特

收稿日期:2021-06-05; 接受日期:2021-10-05; 网络出版日期:2021-12-31

基金项目:国家自然科学基金(61462055, 61562049).

作者简介:沈 彬(1997-),男,云南人,硕士生,主要研究自然语言处理. E-mail: 1193483984@qq.com.

** 通信作者:严 馨(1969-),女,重庆人,副教授,主要研究自然语言处理. E-mail: Kg_yanxin@sina.com.

征的提取和分类器的组合选择,而且不能充分利用文本的上下文语境信息,影响了分类的准确性。

近年来,随着深度学习技术的发展,越来越多的研究者将以卷积神经网络(Convolution Neural Networks, CNN)和循环神经网络(Recurrent Neural Network, RNN)为代表的深度学习模型应用于文本情感分析领域中。Kim 等^[8]使用卷积神经网络对电影评论进行情感分类,验证了其性能优于递归神经网络。刘龙飞等^[9]使用字级别和词级别两个不同粒度的文本表示作为卷积神经网络的输入,以此来学习句子的特征信息,实验结果表明字级别的原始特征好于词级别的原始特征。Song 等^[10]提出了一种新的位置敏感 CNN,它可以通过捕获 3 种不同语言级别(单词级别、短语级别和句子级别)的位置特征来提高特征提取的能力。滕磊等^[11]提出基于胶囊网络的模型(Model based on Asp-Routing and Doc-Routing Capsule, MADC),用来预测评论文本的方面级情感。Abdi 等^[12]提出了一种基于长短时记忆网络(Long short-term memory, LSTM)的模型,该模型通过多特征融合的方法将评论情感分类的准确率提高了 5% 以上。为了充分考虑 CNN 和 RNN 的优点,杜永萍等^[13]提出一种基于 CNN-LSTM 的情感分类方法,在对含有隐含语义信息的短文本评论情感倾向性识别中取得了不错的效果。Madasu 等^[14]提出了一种顺序卷积注意递归网络(Sequential Convolutional Attentive Recurrent Network, SCARN),取得了比传统的 CNN 和 LSTM 更好的性能。由于注意力机制在视觉图像领域中的成功应用,一些研究者也将注意力机制应用到文本情感分析领域中。例如,吴小华等^[15]提出了基于字向量表示方法并结合 Self-attention 和 BiLSTM 的中文短文本情感分析算法。Lei 等^[16]提出一种多情感资源注意力模型,该模型可以从三种情感资源(情感词典、否定词、程度副词)中捕获更全面的情感信息,并在电影评论数据集(Movie Reviews, MR)和斯坦福情感树库(Stanford Sentiment Treebank, SST)上提高 1% 以上的准确率。

在深度学习中,一个关键的步骤是对文本进行向量化^[17]。Word2vec^[18]作为自然语言处理(Natural Language Processing, NLP)领域的词向量工具,通过无监督上下文无关的训练方式,将词映射为低维、稠密的分布式特征向量,使得训练好的词向量中携带着潜在的语义信息。但是 Word2vec 无法表示词

的多义性。针对这个问题,研究人员提出了各种预训练模型,通过对预训练模型微调,可以提高多项 NLP 任务的性能。曾诚等^[19]提出一种结合 ALBERT 预训练模型与卷积循环神经网络(Convolutional Recurrent Neural Network, CRNN)的弹幕文本情感分析方法。ALBERT_CRNN,实验结果均优于传统方法。2018 年, Peters 等^[20]提出了语言模型嵌入(Embeddings from Language Models, ELMo)算法,该算法采用双向 LSTM 语言模型来学习具有上下文信息的词嵌入表示,但是该方法只是将两个方向的语言模型进行拼接,并不是真正意义上的双向建模。同年, Google 的 Devlin 等^[21]提出了基于 Transformer 的双向编码器表示(Bidirectional Encoder Representation from Transformers, BERT)模型,使用具有强大特征提取能力的双向 Transformer 编码器,通过遮蔽语言模型(Masked Language Model, MLM)和下一句话预测(Next Sentence Prediction, NSP)两个无监督预测任务来获得预训练模型。2019 年, Sun 等^[22]提出了基于知识增强语义表示(Enhanced Representation through Knowledge Integration, ERNIE)模型,相较于 BERT, ERNIE 使用更大的中文训练语料,并改进了掩码策略。ERNIE 模型在训练过程中通过对字、短语和实体进行掩码,可以让模型隐式地学习到实体之间的先验语义知识和长语义依赖关系,增强了模型的语义表示能力。

尽管目前的神经网络模型在情感分析中已经取得了巨大的成功,但是情感资源在神经网络中的应用仍存在局限。文献[9, 15]虽然针对微博评论文本的相关特点验证了字级别特征更适用于微博情感分析,但是并没有解决多义性的问题。受文献[9, 16, 23-24]的启发,本文在引入情感资源和注意力机制的基础上,结合 ERNIE 预训练模型能够根据上下文语境获取字级别的动态特征表示以及具有更丰富的语义表示能力等优点,提出了一种基于 ERNIE 和双重注意力机制的微博情感分析模型。实验结果表明,新模型在 COAE2014 和 weibo_senti_100k 数据集上的分类准确率分别达到了 94.50% 和 98.23%,同时也验证了将情感资源运用到神经网络中的有效性。

1 ERNIE-DAM 情感分析模型

1.1 情感信息提取

评论文本中包含着大量的情感信息,如情感词或者修饰词与情感词的搭配组合

等,当情感词被修饰词所修饰时,有可能伴随着整句的情感极性发生变化,如极性反转、加强或者减弱等.修饰词词典一般包括连词、否定词、程度副词等.参考文献[16]构建情感资源库的方法,本文主要将情感词、程度副词、否定词加入到情感资源库中,并通过设定的提取规则从文本中提取出包含的情感信息.设定的提取规则如下:

规则 1 遍历整个文本,若当前词语是情感词,则将当前词语加入到情感信息集合中.

规则 2 遍历整个文本,若当前词语是程度副词或者否定词,且下一个词语是情感词,则将二者作为一个整体加入到情感信息集合中,并将情感词从情感信息集合中删除.

规则 3 遍历整个文本,若当前词语是程度副词,且后面紧依次紧跟否定词和情感词,则将三者

作为一个整体加入到情感信息集合中,并将否定词与情感词的整体从情感信息集合中删除;类似的,若当前词语是否定词,且后面依次紧跟程度副词和情感词,则将三者作为一个整体加入到情感信息集合中,并将程度副词与情感词的整体从情感信息集合中删除.

1.2 模型构建 本文提出的 ERNIE-DAM 模型在纵向结构上包含输入层、特征提取层、特征融合层和情感输出层 4 个部分,在横向结构上包含 2 个处理通道:基于 BLSTM 网络和注意力机制的通道负责提取评论文本的上下文关系特征,基于全连接网络和注意力机制的通道则负责提取评论文本中包含的情感信息的情感特征,并且均采用 ERNIE 预训练模型获取文本字级别的动态特征表示.模型的整体架构如图 1 所示.

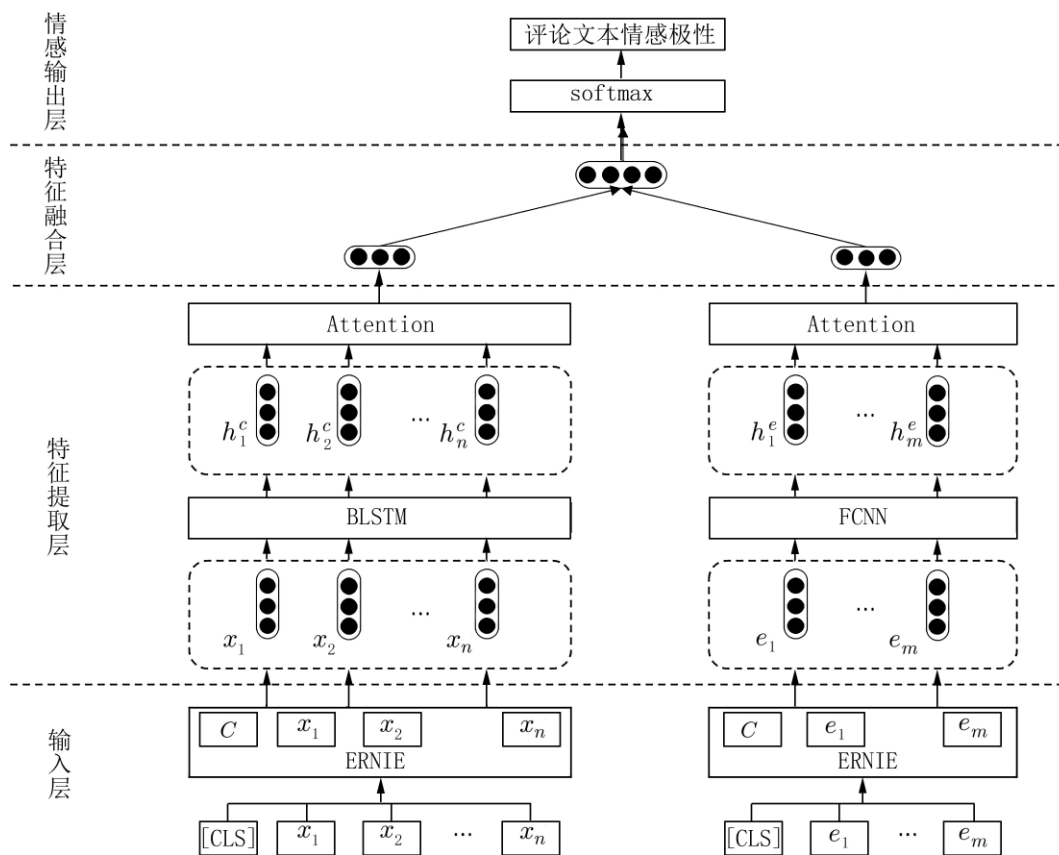


图 1 ERNIE-DAM 模型

Fig. 1 Model of ERNIE-DAM

1.2.1 输入层 输入层包括评论文本及其包含的情感信息.对于一条评论文本,根据 1.1 节中情感信息的提取规则,提取出该条评论文本中包含的情感信息.例如,评论文本“我超喜欢三星这款手机”

中包含的情感信息为“超喜欢”.将评论文本和情感信息分别按照字粒度进行切分,并在首位置拼接一个分类标志符号[CLS],即 $X = \{[CLS], x_1, x_2, \dots, x_n\}$, $E = \{[CLS], e_1, e_2, \dots, e_m\}$, 其中 x_i 和 e_i 分别表示评论

文本和情感信息中的第 i 个字. 然后输入到 ERNIE 预训练模型中, 经过 ERNIE 前向计算后得到评论文本的特征向量矩阵 X 和情感信息的特征向量矩阵 E :

$$X = x_1 \oplus x_2 \oplus \cdots \oplus x_n, \quad (1)$$

$$E = e_1 \oplus e_2 \oplus \cdots \oplus e_m, \quad (2)$$

式中, $\oplus$ 为拼接操作, n 和 m 分别表示评论文本和情感信息的长度.

1.2.2 特征提取层

1.2.2.1 上下文关系特征 在评论文本中, 不仅要考虑序列内部之间的上下文语义依赖关系, 还要关注到对于情感分类更加关键的部分, 对于重要的部分应该赋予更高的权重. 因此采用 BLSTM 网络和注意力机制相结合的方法对评论文本的上下文信息进行编码, 以获取文本的上下文关系特征.

LSTM 作为循环神经网络 RNN 的变体, 能够很好地解决 RNN 模型在训练过程中存在的梯度消失和梯度爆炸等问题, 并且具有较强的长距离语义捕捉能力, 在处理序列化文本中得到广泛使用^[24]. LSTM 神经网络的前向传播计算如下所示:

$$f_t = \sigma(W_f \cdot [x_t, h_{t-1}] + b_f), \quad (3)$$

$$i_t = \sigma(W_i \cdot [x_t, h_{t-1}] + b_i), \quad (4)$$

$$o_t = \sigma(W_o \cdot [x_t, h_{t-1}] + b_o), \quad (5)$$

$$\tilde{c}_t = \tanh(W_c \cdot [x_t, h_{t-1}] + b_c), \quad (6)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t, \quad (7)$$

$$h_t = o_t \cdot \tanh(c_t), \quad (8)$$

式中, x_t 为当前时刻的输入, h_{t-1} 为上一时刻的外部状态, W_i 、 W_f 、 W_o 和 W_c 为权重参数, b_i 、 b_f 、 b_o 和 b_c 为偏置, σ 为 sigmoid 激活函数, $\tanh$ 为双曲正切激活函数, f_t 表示遗忘门, 决定上一时刻的内部状态 c_{t-1} 中哪些信息应该被抛弃, i_t 为更新门, 决定当前时刻的候选状态 $\tilde{c}_t$ 中哪些信息应该被保留, o_t 为输出门, 决定当前时刻的内部状态 c_t 中哪些信息应该输出给外部状态 h_t .

由于 LSTM 只能捕获文本中的正向语义信息, 缺乏对逆向语义信息的捕获, 因此本文采用 BLSTM 来捕获文本中正反两个方向的长距离语义依赖关系. 在 BLSTM 中, 某一时刻 t 的输出状态 h_t^c 由正向 LSTM 的输出 $\vec{h}_t^c$ 和反向 LSTM 的输出 $\overleftarrow{h}_t^c$ 相

连接组成:

$$h_t^c = [\vec{h}_t^c, \overleftarrow{h}_t^c]. \quad (9)$$

为了提取出评论文本中更关键的信息, 将 BLSTM 模型的输出 $H^c = \{h_1^c, h_2^c, \dots, h_n^c\}$ 输入到注意力机制中, 通过注意力机制对每一时刻的输入状态 h_t^c 进行加权, 从而得到评论文本的上下文关系特征 v^c . 注意力计算方法如下所示:

$$z_t = W_{h2}(\text{ReLU}(W_{h1}h_t^c + b_{h1})) + b_{h2}, \quad (10)$$

$$\alpha_t = \frac{\exp(z_t)}{\sum_t \exp(z_t)}, \quad (11)$$

$$v^c = \sum_t \alpha_t h_t^c, \quad (12)$$

式中, W_{h1} 、 W_{h2} 和 b_{h1} 、 b_{h2} 为权重参数和偏置, α_t 为输入状态 h_t^c 的权重.

1.2.2.2 情感特征 在评论文本包含的情感信息中, 情感信息之间只是具有情感表示、情感极性的加强和减弱以及反转等功能, 不像文本序列中一样存在较强的语义依赖关系^[30]. 同时, 评论文本中可能包含多个情感词, 不同的情感词对于评论文本的情感倾向性的影响程度也是不一样的. 因此采用全连接网络和注意力机制相结合的方法对情感信息进行编码, 以获取最显著的情感特征. 首先使用两层的全连接网络对情感信息进行编码:

$$h_t^e = \text{ReLU}(W_{e2}(\text{ReLU}(W_{e1}e_t + b_{e1})) + b_{e2}), \quad (13)$$

式中, W_{e1} 、 W_{e2} 和 b_{e1} 、 b_{e2} 为全连接网络的权重参数和偏置.

然后, 将全连接网络的输出 $H^e = \{h_1^e, h_2^e, \dots, h_m^e\}$ 输入到注意力机制中, 通过注意力机制捕获情感信息中更重要的情感词, 从而提取出情感特征 v^e . 注意力计算方法如下所示:

$$u_t = \tanh(W_w h_t^e + b_w), \quad (14)$$

$$\beta_t = \frac{\exp(u_t^T u_w)}{\sum_t \exp(u_t^T u_w)}, \quad (15)$$

$$v^e = \sum_t \beta_t h_t^e, \quad (16)$$

式中, W_w 和 b_w 为权重参数和偏置, u_w 为 softmax 函数的权重参数, β_t 为输入状态 h_t^e 的权重.

1.2.3 特征融合层 特征融合层的目的是将上

下文关系特征向量和情感特征向量进行融合,从而构造新的特征向量表示.为了简化模型的复杂度和计算量,采用简单的向量拼接方式,将上下文关系特征向量 v^c 和情感特征向量 v^e 沿维度方向拼接,得到最终的特征向量表示 v^* ,如下所示:

$$v^* = [v^c, v^e]. \quad (17)$$

1.2.4 情感输出层 将特征融合层输出的特征向量 v^* 通过softmax分类器,从而得到模型最终预测该条评论文本属于哪个类别的分类结果,如下式所示:

$$\hat{y} = \text{softmax}(wv^* + b), \quad (18)$$

式中, w 为权重参数, b 为偏置, $\hat{y}$ 为该评论文本属于哪个情感类别的概率分布.

1.2.5 模型训练 在模型训练的过程中,本文使用交叉熵作为损失函数,并通过端到端的反向传播算法(back propagation)来更新网络参数,如下所示:

$$L = - \sum_{i=1}^D \sum_{j=1}^C y_i^j \lg \hat{y}_i^j, \quad (19)$$

式中, D 为训练数据集, C 为情感标签的类别数, y 为实际情感类别, $\hat{y}$ 为预测的情感类别.

2 实验

2.1 数据集 本文选取两个公开的数据集进行实验.数据集1来自COAE2014任务2中文评测数据集,共有7 000条已标注情感极性的社交评论数据,其中正向评论3 776条,负向评论3 224条,本文将数据集随机打乱后按“7:3”的比例划分为训练集和测试集.数据集2来自weibo_senti_100k数据集,共有119 988条已标注情感极性的社交评论数据,其中正向59 993条,负向评论59 995条,本文将数据集随机打乱后按“8:1:1”的比例划分为训练集、验证集和测试集.详细的数据信息如表1所示.

表1 两个数据集的相关信息

Tab. 1 The relevant information of two data sets

数据集	数据量		训练集	验证集	测试集	共计
	正向	负向				
COAE2014	3 776	3 224	4 900	—	2 100	7 000
weibo_senti_100k	59 993	59 995	95 990	11 999	11 999	119 988

表中“—”表示无数据.

情感资源库包含情感词、程度副词、否定词,如表2所示.其中,情感词来自大连理工大学情感词汇本体库、知网HowNet和台湾大学NTUSD,并将三者合并去重;程度副词和否定词来自知网HowNet.在分词过程中,本文将情感资源库作为自定义分词词典,使得文本中的情感信息能够作为一个完整的语言单元存在.

表2 情感资源库

Tab. 2 The data library of emotion resource

情感资源	数目	实例
情感词	37 795	骄傲、伟大、喜欢、自不量力
程度副词	219	稍微、有点、最、特别、愈来愈
否定词	58	不、休想、拒绝、白搭

2.2 数据预处理 评论文本中存在很多的噪声数据,影响着情感分类模型的性能,因此,在模型输入前需要对这些数据进行清洗,以提高数据的质量.

本文采用正则表达式匹配的方法对数据进行清洗,从评论文本中删除URL、@+用户名、转发微博等无意义的数,只保留中文、大小写字母和阿拉伯数字.由于话题标签hashtag中一般包含被评论的话题或者对象,于是本文对这部分数据进行了保留.

2.3 参数设置 本文使用PyTorch深度学习框架,所有实验均采用Adam^[25]优化器,在训练阶段采用“Early Stopping”法,并选择在测试集上表现最佳的结果作为最终的实验结果.文本最大长度设置为140,情感信息集合最大长度设置为40.数据集1和数据集2的batch_size分别设置为32和128.在使用传统词向量构建的神经网络模型中,Word2Vec采用Li等^[26]使用中文微博语料训练好的Chinese-Word-Vectors,词向量维度为300维,随机初始化<UNK>表示未登录词,全零初始化<PAD>表示句子补齐部分;所有模型的学习率为0.001,dropout为0.5;BLSTM网络的隐藏单元个数

为 128, 层数为 2; CNN 中滤波器的尺寸为 2、3、4, 滤波器的个数为 256; FastText 中隐藏层单元个数为 256. 在预训练模型构建的神经网络模型中, ELMo 采用哈工大的 ELMoForManyLangs, 词向量维度为 1024. BERT 采用 Google 的 BERT-Base-Chinese 预训练模型, 字向量维度为 768. ERNIE 采用百度的 ERNIE1.0 预训练模型, 字向量维度为 768. 学习率为 1×10^{-5} , dropout 为 0.5; BLSTM 网络的隐藏单元个数为 256, 层数为 2.

2.4 评测指标 文本分类中通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). 由于本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. 评测指标定义如下:

$$P = \frac{T_P}{T_P + F_P}, \quad (20)$$

$$R = \frac{T_P}{T_P + F_N}, \quad (21)$$

$$F1 = \frac{2PR}{P+R}, \quad (22)$$

$$A = \frac{T_P + T_N}{T_P + F_N + T_N + F_P}, \quad (23)$$

式中, T_P 表示实际为正向且预测为正向的数量, F_P 表示实际为负向但预测为正向的数量, T_N 表示实际为负向且预测为负向的数量, F_N 表示实际为正向但预测为负向的数量. T_P 、 F_P 、 T_N 、 F_N 的混淆矩阵如表 3 所示.

表 3 混淆矩阵
Tab. 3 Confusion matrix

预测类	实际类	
	正向	负向
正向	实际为正向且预测为正向(T_P)	实际为负向但预测为正向(F_P)
负向	实际为正向但预测为负向(F_N)	实际为负向且预测为负向(T_N)

2.5 实验结果与分析 为了验证本文提出的方法的有效性, 本文设置了 3 组实验. 第 1 组实验比较不同模型的情感分类性能, 其中 ERNIE、ERNIE-DAM 采用百度发布的中文预训练模型 ERNIE1.0 来进行文本向量化表示, 其余模型均采用 Li 等^[31]提出的 Chinese-Word-Vectors 构建词向量, 并在训

练阶段进行微调. 第 2 组实验比较不同词向量工具对本文模型情感分类性能的影响. 第 3 组实验验证不同情感信息的组合对于情感分类性能的影响.

2.5.1 分类模型对比实验 FastText: 文献 [27] 提出的文本分类模型, 其模型结构类似于 Word2Vec 中的 CBOW 模型, 通过对整个句子的词向量和 N-gram 向量取平均池化后得到句子表示.

TextCNN: 文献 [8] 提出的卷积神经网络, 通过卷积和池化操作得到句子表示.

LSTM: 文献 [28] 提出的长短时记忆网络, 取最后时刻的隐层输出作为句子表示.

BGRU/BLSTM: 文献 [28] 和文献 [29] 提出的门控循环单元和长短时记忆网络的双向变体, 拼接最后时刻正反两个方向的隐层输出作为句子表示.

RCNN: 文献 [30] 提出循环卷积神经网络, 将 BLSTM 模型各时刻的输入和输出拼接之后取最大池化得到句子表示.

BLSTM-Att: 文献 [31] 提出的基于 Attention 机制的双向长短时记忆网络, 采用注意力机制对 BLSTM 各时刻的隐层输出加权求和得到句子表示.

CNN-LSTM: 分别采用 CNN 提取局部特征和 LSTM 提取全局特征后进行拼接, 得到句子表示.

ERNIE: 取 ERNIE 模型 [CLS] 位置对应的特征向量作为句子表示.

ERNIE-DAM: 本文提出的基于 ERNIE 和双重注意力机制的情感分析模型.

不同模型的分类对比结果如表 4 所示. 从表 4 可以看出, 本文提出的 ERNIE-DAM 模型在 COAE2014 和 weibo_senti_100k 数据集上, 所有指标都取得了最好的结果. 相比 FastText、TextCNN、BLSTM、RCNN、BLSTM-Att、Word2Vec-DAM 和 ERNIE 模型, ERNIE-DAM 模型在 COAE2014 数据集上的 F1 值分别提高了 8.16%、6.94%、7.28%、6.41%、6.26%、5.28% 和 1.08%, 在 weibo_senti_100k 数据集上的 F1 值分别提高了 3.23%、1.50%、1.33%、1.03%、1.14%、1.01% 和 0.51%. 由此可以看出, 相比于其他基于传统词向量构建的神经网络模型, ERNIE、ERNIE-DAM 模型在微博评论文本情感分析任务上具有较明显的优势, 也充分说明了预训练模型能够根据文本序列中的上下文语境对同一个词的不同含义进行区分, 从而提高情感分类的性能. 而基于知识增强的 ERNIE 预训练模型由于采

用了具有强大特征提取能力的双向 Transformer 编码器,并通过对海量数据中的实体概念等先验语义知识进行建模,增强了字的语义表示能力,所以仅仅通过微调的方式就能取得比其他基于传统词向量构建的神经网络模型更好的效果。另外,将基于 Word2Vec 构建的 DAM 模型与 BLSMT-Att 模型进行比较,在两个数据集上的 F1 值分别提高了 0.98% 和 0.13%,说明将评论文本中的情感词、程度副词、否定词等情感信息提取出来,通过全连接网络和注意力机制挖掘文本中深层次的情感特征,对于文本的特征向量表示起到情感增强的作用,验证了将情感资源运用到神经网络中的有效性。

表 4 不同分类模型的对比结果

Tab. 4 Comparison results of different classification models %

模型	COAE2014		weibo_senti_100k	
	F1	A	F1	A
FastText	86.30	86.52	95.00	95.00
TextCNN	87.52	87.57	96.73	96.73
LSTM	86.28	86.49	96.12	96.12
BLSTM	87.18	87.29	96.90	96.90
BGRU	87.69	87.76	97.12	97.12
RCNN	88.05	88.10	97.20	97.20
CNN-LSTM	87.66	87.76	96.81	96.81
BLSTM-Att	88.20	88.24	97.09	97.09
Word2Vec-DAM	89.18	89.24	97.22	97.22
ERNIE	93.38	93.43	97.72	97.72
ERNIE-DAM	94.46	94.50	98.23	98.23

另外,本文还将其他基于传统词向量构建的神经网络模型进行了实验对比。BGRU 相较 BLSTM 在减少一个门函数的基础上,不仅减少了训练时间,而且效果也更佳。相比 LSTM、TextCNN 模型, CNN-LSTM 模型结合了 CNN 提取文本局部特征和 LSTM 提取文本上下文全局特征的优势,性能有进一步的提升。相比 BLSTM 模型, BLSTM-Att 模型在 BLSTM 神经网络基础上增加了注意力机制,有助于提取出文本中更重要的信息。此外, ERNIE-DAM 模型相比 ERNIE 模型在微博情感分析中具有更好的表现,证明了 DAM 模型能够充分考虑评论文本中的上下文信息以及情感信息中更显著的

情感特征,进一步提升模型的性能。

2.5.2 词向量工具对比实验 Word2Vec-DAM: 将词使用 Chinese-Word-Vectors 词向量进行初始化后,输入到 DAM 模型中进行训练。

ELMo-DAM: 将词使用 ELMoForManyLangs 预训练模型转换为词向量以后,输入到 DAM 模型中进行训练。

BERT-DAM: 采用中文预训练模型 BERT 将文本向量化表示后,输入到 DAM 模型中进行训练。

ERNIE-DAM: 采用中文预训练模型 ERNIE 将文本向量化表示后,输入到 DAM 模型中进行训练。

基于不同词向量模型分类对比结果如表 5 所示。从表 5 可以看出, ELMo、BERT、ERNIE 预训练模型相比传统的静态词向量,可以根据上下文语境对文本中同一个词的不同语义表征进行动态调整,解决了传统词向量中存在的一词多义问题,因而各项评价指标均有明显的提升。相比 ELMo 采用双向 LSTM 语言模型来提取词向量特征, BERT 和 ERNIE 采用更强大的双向 Transformer 编码器,因此进一步提高了模型的特征提取能力。ERNIE 与 BERT 相比,由于对先验语义知识建模,实验结果有明显提升。

表 5 基于不同词向量模型的对比结果

Tab. 5 Comparison results based on different word vector models %

模型	COAE2014		weibo_senti_100k	
	F1	A	F1	A
Word2Vec-DAM	89.18	89.24	97.22	97.22
ELMo-DAM	91.75	91.82	97.83	97.83
BERT-DAM	93.67	93.72	98.16	98.16
ERNIE-DAM	94.46	94.50	98.23	98.23

2.5.3 消融实验 为了验证不同情感信息的组合对于情感分类模型性能的影响,本文进行了消融实验,并选择在 COAE2014 测试集上的结果进行对比分析。当不包含情感信息时,本文模型退化为 ERNIE-BLSTM-Att 模型,在 ERNIE-BLSTM-Att 模型的基础上设置了 5 组对照实验,分别是 ERNIE-BLSTM-Att(不包含情感信息)、EBA+S(ERNIE-BLSTM-Att+情感词)、EBA+SN(ERNIE-BLSTM-Att+情感词+否定词)、EBA+SI(ERNIE-BLSTM-Att+情感词+程度副词)、ERNIE-DAM(本文模型,

ERNIE-BLSTM-Att+情感词+否定词+程度副词)。实验结果如图 2 所示。

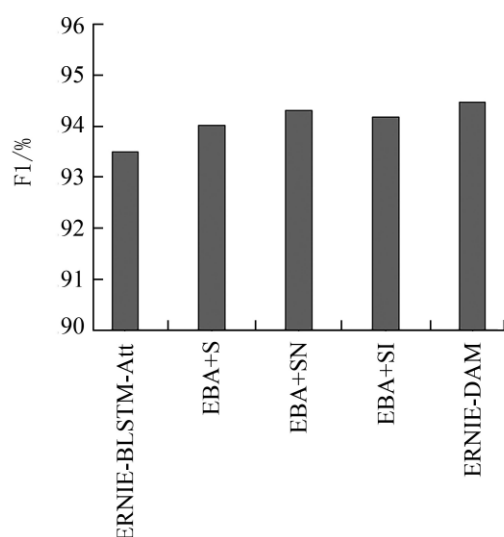


图 2 不同情感信息组合的对比结果
Fig. 2 Comparison results of different combinations of emotion information

从图 2 可以看出, 在 ERNIE-BLSTM-Att 模型的基础上单独添加情感词时, 模型的性能有了提升; 当在情感词的基础上单独添加否定词和程度副词时, 模型的性能有了进一步的提升, 但是程度副词的提升不如否定词的明显。这说明了情感词、否定词和程度副词等先验情感资源的引入可以提高情感分类模型的性能, 同时, 否定词对于文本情感极性的表现能力要强于程度副词。

3 结语

本文提出一种基于 ERNIE 和双重注意力机制的微博情感分析模型 ERNIE-DAM。该模型利用不同的注意力机制对评论文本和文本中包含的情感信息进行特征提取, 以获取文本最终的特征向量表示。通过 ERNIE 预训练模型强大的语义表示能力对文本进行动态特征表示, 解决了传统词向量中同一个词在不同上下文语境中表示不同含义的问题。同时, 将情感资源和注意力机制相结合的方法能够有效地提高模型捕获文本情感语义的能力, 进而提高文本情感分析模型的性能。实验结果表明, 本文提出的模型在 COAE2014 和 weibo_senti_100k 两个数据集上都取得了较好的分类效果。本文下一步的研究工作中, 将进一步结合微博评论文本本身的语言学特征, 融入表情符号、浅层语义信息和语义角色信息等特征, 探究其对模型性能的影响。

参考文献:

- [1] 周超, 严馨, 余正涛, 等. 融合词频特性及邻接变化数的微博新词识别[J]. 山东大学学报: 理学版, 2015, 50(3): 6-10.
Zhou Chao, Yan X, Yu Z T, et al. Weibo new word recognition combining frequency characteristic and accessor variety[J]. Journal of Shandong University: Natural Science, 2015, 50(3): 6-10.
- [2] 王婷, 杨文忠. 文本情感分析方法研究综述[J]. 计算机工程与应用, 2021, 57(12): 11-24. DOI: 10.3778/j.issn.1002-8331.2101-0022.
Wang T, Yang W Z. Review of text sentiment analysis methods[J]. Computer Engineering and Applications, 2021, 57(12): 11-24.
- [3] 赵妍妍, 秦兵, 石秋慧, 等. 大规模情感词典的构建及其在情感分类中的应用[J]. 中文信息学报, 2017, 31(2): 187-193.
Zhao Y Y, Qin B, Shi Q H, et al. Large-scale sentiment lexicon collection and its application in sentiment classification[J]. Journal of Chinese Information Processing, 2017, 31(2): 187-193.
- [4] Zhang S X, Wei Z L, 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. DOI: 10.1016/j.future.2017.09.048.
- [5] Xu G X, Yu Z H, Yao H S, et al. Chinese text sentiment analysis based on extended sentiment dictionary [J]. IEEE Access, 2019, 7: 43 749-43 762. DOI: 10.1109/ACCESS.2019.2907772.
- [6] 杨爽, 陈芬. 基于 SVM 多特征融合的微博情感多级分类研究[J]. 数据分析与知识发现, 2017, 1(2): 73-79.
Yang S, Chen F. Analysis sentiment of micro-blog posts based on support vector machine[J]. Data Analysis and Knowledge Discovery, 2017, 1(2): 73-79.
- [7] Naz S, Sharan A, Malik N. Sentiment classification on twitter data using support vector machine[C]// IEEE/WIC/ACM International Conference on Web Intelligence (WI), Santiago, Chile, 2018: 676-679.
- [8] Kim Y. Convolutional neural networks for sentence classification[C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 2014: 1 746-1 751.
- [9] 刘龙飞, 杨亮, 张绍武, 等. 基于卷积神经网络的微博情感倾向性分析[J]. 中文信息学报, 2015, 29(6): 159-165. DOI: 10.3969/j.issn.1003-0077.2015.06.021.
Liu L F, Yang L, Zhang S W, et al. Convolutional neur-

- al networks for Chinese micro-blog sentiment analysis[J]. *Journal of Chinese Information Processing*, 2015, 29(6): 159-165.
- [10] Song Y, Hu Q V, He L. P-CNN: Enhancing text matching with positional convolutional neural network[J]. *Knowledge-Based Systems*, 2019, 169: 67-79. DOI: [10.1016/j.knosys.2019.01.028](https://doi.org/10.1016/j.knosys.2019.01.028).
- [11] 滕磊, 严馨, 徐广义, 等. 使用胶囊网络的细粒度情感分析方法[J]. *小型微型计算机系统*, 2020, 41(12): 2 550-2 556. DOI: [10.3969/j.issn.1000-1220.2020.12.015](https://doi.org/10.3969/j.issn.1000-1220.2020.12.015).
- Teng L, Yan X, Xu G Y, et al. Fine-grained sentiment analysis based on capsule network[J]. *Journal of Chinese Computer Systems*, 2020, 41(12): 2 550-2 556.
- [12] Abdi A, Shamsuddin S M, Hasan S, et al. Deep learning-based sentiment classification of evaluative text based on Multi-feature fusion[J]. *Information Processing & Management*, 2019, 56(4): 1 245-1 259. DOI: [10.1016/j.ipm.2019.02.018](https://doi.org/10.1016/j.ipm.2019.02.018).
- [13] 杜永萍, 赵晓铮, 裴兵兵. 基于 CNN-LSTM 模型的短文本情感分类[J]. *北京工业大学学报*, 2019, 45(7): 662-670. DOI: [10.11936/bjtxxb2017120035](https://doi.org/10.11936/bjtxxb2017120035).
- Du Y P, Zhao X Z, Pei B B. Short text sentiment classification based on CNN-LSTM model[J]. *Journal of Beijing University of Technology*, 2019, 45(7): 662-670.
- [14] Madasu A, Rao V A. Sequential learning of convolutional features for effective text classification[J]. *arXiv*: 1909.00080, 2019.
- [15] 吴小华, 陈莉, 魏甜甜, 等. 基于 Self-Attention 和 Bi-LSTM 的中文短文本情感分析[J]. *中文信息学报*, 2019, 33(6): 100-107. DOI: [10.3969/j.issn.1003-0077.2019.06.015](https://doi.org/10.3969/j.issn.1003-0077.2019.06.015).
- Wu X H, Chen L, Wei T T, et al. Sentiment analysis of Chinese short text based of Self-Attention and Bi-LSTM[J]. *Journal of Chinese Information Processing*, 2019, 33(6): 100-107.
- [16] Lei Z Y, Yang Y J, Yang M, et al. A multi-sentiment-resource enhanced attention network for sentiment classification[C]// *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Melbourne, Australia, 2018: 758-763.
- [17] Bengio Y, Ducharme R, Vincent P, et al. A neural probabilistic language model[J]. *The Journal of Machine Learning Research*, 2003, 3(2): 1 137-1 155.
- [18] Mikolov T, Sutskever I, Chen K, et al. Distributed representations of words and phrases and their compositionality[C]// *Proceedings of the 26th International Conference on Information Processing Systems*, Lake Tahoe, California, USA, 2013: 3 111-3 119.
- [19] 曾诚, 温超东, 孙瑜敏, 等. 基于 ALBERT-CRNN 的弹幕文本情感分析[J]. *郑州大学学报: 理学版*, 2021, 53(3): 1-8. DOI: [10.13705/j.issn.1671-6841.2020359](https://doi.org/10.13705/j.issn.1671-6841.2020359).
- Zeng C, Wen C D, Sun Y M, et al. Barrage text sentiment analysis based on ALBERT-CRNN[J]. *Journal of Zhengzhou University: Natural Science*, 2021, 53(3): 1-8.
- [20] Peters M E, Neumann M, Iyyer M, et al. Deep contextualized word representations[C]// *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, New Orleans, Louisiana, 2018: 2 227-2 237.
- [21] 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, Minnesota, 2019: 4 171-4 186.
- [22] Sun Y, Wang S H, Li Y K, et al. Ernie: Enhanced representation through knowledge integration[J]. *arXiv*: 1904.09223, 2019.
- [23] 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(4): 642-652. DOI: [10.16337/j.1004-9037.2020.04.005](https://doi.org/10.16337/j.1004-9037.2020.04.005).
- Xie R Z, Li Y. Text sentiment classification model based on BERT and dual channel attention[J]. *Journal of Data Acquisition & Processing*, 2020, 35(4): 642-652.
- [24] 张仰森, 郑佳, 黄改娟, 等. 基于双重注意力模型的微博情感分析方法[J]. *清华大学学报: 自然科学版*, 2018, 58(2): 122-130. DOI: [10.16511/j.cnki.qhdxxb.2018.22.015](https://doi.org/10.16511/j.cnki.qhdxxb.2018.22.015).
- Zhang Y S, Zheng J, Huang G J, et al. Microblog sentiment analysis method based on a double attention model[J]. *Journal of Tsinghua University: Science and Technology*, 2018, 58(2): 122-130.
- [25] Kingma D P, Ba J. Adam: A method for stochastic optimization[C]// *Proceedings of the 3rd International Conference on Learning Representations*, San Diego, California, USA, 2015: 1-15.
- [26] Li S, Zhao Z, Hu R F, et al. Analogical reasoning on Chinese morphological and semantic relations[C]// *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Melbourne, Australia, 2018: 138-143.
- [27] Joulin A, Grave E, Bojanowski P, et al. Bag of tricks for efficient text classification[C]// *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, Valencia, Spain,

- 2017: 427-431.
- [28] Zhu X, Sobihani P, Guo H. Long short-term memory over recursive structures[C]// Proceedings of the 32nd International Conference on Machine Learning, Lille, France, 2015: 1 604-1 612.
- [29] Cho K, Bart V M, Caglar G, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 2014: 1 724-1 734.
- [30] Lai S W, Xu L H, Liu K, et al. Recurrent convolutional neural networks for text classification[C]// Proceedings of the AAAI Conference on Artificial Intelligence, Austin, Texas, USA, 2015, 29(1): 2 267-2 273.
- [31] Zhou P, Shi W, Tian J, et al. Attention-based bidirectional long short-term memory networks for relation classification[C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, 2016: 207-212.

Microblog sentiment analysis based on ERNIE and dual attention mechanism

SHEN Bin^{1,2}, Yan Xin^{1,2**}, ZHOU Li-hua³, XU Guang-yi⁴, LIU Yan-chao⁵

- (1. School of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, Yunnan, China;
2. Yunnan Key Laboratory of Artificial Intelligence, Kunming University of Science and Technology, Kunming 650500, Yunnan, China;
3. School of Information Science & Engineering, Yunnan University, Kunming 650500, Yunnan, China;
4. Yunnan Nantian Electronic Information Industry Co., Ltd., Kunming 650040, Yunnan, China;
5. The Information Technology Center, Hubei Engineering University, Xiaogan 432000, Hubei, China)

Abstract: Aiming at the problems that the traditional word vector could not represent the polysemy of words in the context, and the priori emotion resources could not be fully utilized in the neural network, a microblog sentiment analysis model based on ERNIE and Dual Attention Mechanism (DAM) was proposed. Firstly, we use the existing emotional resources to construct an emotional resource library including emotional words negative words and adverbs of degree. Secondly, BLSTM network and fully connected network were used to encode the text and the emotional information contained in the text respectively. Different attention mechanisms were used to extract the contextual features and emotional features of the text and the emotional information respectively, and ERNIE pretraining model was used to obtain the dynamic feature representation of the text. Finally, the contextual features and emotional features are spliced and fused to obtain the final feature vector representation. Experimental results show that the classification accuracy of the new model on COAE2014 and weibo_senti_100k data sets reaches 94.50% and 98.23% respectively, which also verifies the effectiveness of applying emotional resources into neural networks.

Key words: sentiment analysis; Enhanced Representation through Knowledge Integration (ERNIE); long short-term memory network; attention mechanism