

刘佳伟, 陈强, 张开昱, 等. 基于 LE-BERT 的文本情感分析[J]. 智能计算机与应用, 2026, 16(3): 10-18. DOI: 10.20169/j.issn.2095-2163.24050603

基于 LE-BERT 的文本情感分析

刘佳伟¹, 陈强¹, 张开昱¹, 倪凯²

(1 上海工程技术大学 电子电气工程学院, 上海 201620; 2 上海市安全生产科学研究所, 上海 200233)

摘要: 利用情感分析技术从数以亿计的文本中提取有价值的信息, 可以把握舆论倾向, 帮助决策者做出重要决策, 显示出巨大的研究和应用价值。针对突发事件, 传统的情感分析模型难以从文本中提取相关的具体表征, 导致分类困难。为此, 本文提出了一种改进的 BERT 分类模型 LE-BERT, 采用 LDA 主题模型对文本进行分类, 并将得到的文本主题关键词与原文进行串联, 强化了文本在特定领域的表征信息。此外, 还在下游模型中引入了图像关注机制 ECA, 以改善传统 BERT 模型在下游任务中仅使用 [CLS] token 导致分类信息不足的问题, 使模型能够关注更多的全局信息。本文提出的 LE-BERT 模型在情感分类任务中的准确率为 89.92%, 宏观平均 F1 得分为 91.25%, 优于基线模型。

关键词: 情感分析; BERT; 主题模型; 注意力机制

中图分类号: TP18

文献标志码: A

文章编号: 2095-2163(2026)03-0010-09

Text sentiment analysis model based on LE-BERT

LIU Jiawei¹, CHEN Qiang¹, ZHANG Kaiyu¹, NI Kai²

(1 School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China;

2 Shanghai Institute of Work Safety Science, Shanghai 200233, China)

Abstract: The use of sentiment analysis technology to extract valuable information from hundreds of millions of texts could grasp the tendency of public opinion and help decision-makers make important decisions, and shows great research and application value. For emergencies, the conventional sentiment analysis models are difficult to extract relevant specific representations from the text. The aforementioned situation makes classification difficult. To this end, the paper proposes an improved BERT classification model LE-BERT, which uses LDA topic model to classify texts and concat keywords with original texts, strengthening the characterization information of texts in specific fields. Moreover, the paper introduces the image attention mechanism ECA in the downstream model to ameliorate the problem of insufficient classification information caused by using only [CLS] token of the traditional BERT model for downstream tasks, which enables the model to focus on more global information. The experimental results show that the introduction of LDA strengthens the representation information, leading to a positive impact on model classification. And the use of ECA in the model improves the ability of the model to extract the global information, which causes a better robustness. The proposed LE-BERT model outperforms the baseline models with an accuracy of 89.92% and a macro-averaged F1 score of 91.25% in the sentiment classification task.

Key words: sentiment analysis; BERT; topic model; attention mechanism

0 引言

根据中国互联网信息中心 (CNNIC) 发布的最新《中国互联网络发展状况统计报告》^[1] 显示, 国内互联网普及率和网民规模持续上升, 截至 2023 年 6 月分别达 76.4% 及 10.79 亿人。近年来, 全球突发

事件仍呈多发态势, 互联网不仅成为突发事件第一时间传播发布平台和公众表达相关诉求的重要途径, 也已成为有关部门及各企业知晓大众想法、回应相关质疑的重要媒介^[2]。在大数据时代背景下, 突发舆情事件打破时空限制, 爆发点多且杂, 形成聚集效应, 极易引发大规模网络舆情^[3]。为改善“事后

基金项目: 国家自然科学基金 (61673258)。

作者简介: 刘佳伟 (1999—), 女, 硕士研究生, 主要研究方向: 网络舆情分析; 张开昱 (1999—), 男, 硕士研究生, 主要研究方向: 情感分析以及视觉和 NLP 领域的其他深度学习应用; 倪凯 (1968—), 男, 博士, 高级工程师, 主要研究方向: 安全生产管理, 地理信息系统和数据挖掘。

通信作者: 陈强 (1965—), 男, 博士后, 教授, 硕士生导师, 主要研究方向: 软件工程, 地球探测与信息技术和机器学习。Email: m325121522@sues.edu.cn。

收稿日期: 2024-05-06

哈尔滨工业大学主办 ◆ 学术研究与应用

处置”的局面,化“被动”为“主动”,运用基于自然语言处理的情感分析方法在突发事件网络舆情分析领域寻找新的研究视角具有重要的现实意义。

情感分析研究一般分为2类。第一类是词典法。主要通过构建情感词典来实现情感信息的识别和分类^[4]。如邱全磊等学者^[5]针对现有词典中对符号颜文字和表情考虑不足的现状,在基础词典基础上添加了表情词典,考虑了表情对文本情感分析的影响。然而,这类方法过于依赖词典的选择和构建,领域适应性差,不能有效解决多义词问题。第二种方法基于机器学习,该方法通过朴素贝叶斯、支持向量机、KNN、条件随机场等传统机器学习方法,依据大量已标注的语料库进行模型训练,再利用训练好的模型对待分类文本进行情感分类。近年来,神经网络模型发展迅速,逐渐取代了一些基于机器学习的传统情感分析方法。深度学习通过模拟人脑的层次结构,在特征自动提取方面具有显著优势,因此被广泛应用于情感分析领域。常用的深度学习情感分类模型有基于卷积神经网络(CNN)、长短期记忆网络(LSTM)^[6]、混合神经网络等。其中,半监督学习(Semi-Supervised Learning)方法可以充分利用少量人工标注数据来训练模型,然后通过预测大量未标注数据来消除人工标注语料的过程。如Khine等学者^[7]建立了一个用于社会数据分析的半监督学习模型,该模型在情感识别方面有很好的表现。此外,在神经网络中引入注意力机制^[8]在自然语言处理中展现了巨大的前景,为情感分析带来了最先进的成果。如Lou等学者^[9]使用Bi-LSTM模型进行价格预测。该模型在LSTM的基础上增加了一个反向LSTM层,以提取文本的反向序列信息。通过神经网络提取的正向序列特征与反向序列特征相结合,使模型能够同时考虑上下文的语义信息。LSTM和门控递归神经网络^[10],在语言建模和机器翻译等序列建模和转译问题上已被公认为最先进的方法^[11]。自此以后,许多学者继续努力推动递归语言模型和编码器-解码器架构的发展^[12]。然而,LSTM的效率并不高,而且受到计算机内存的限制,一些研究通过因式分解技巧^[13]和条件计算^[14]显著提高了计算效率,同时也改善了模型性能。然而,顺序计算的基本限制仍然存在,随后出现的transformer^[15]解决了这些问题,并取得了卓越的性能。在此期间,语言模型的预训练也逐渐受到青睐,并被证明能有效改善许多自然语言处理任务^[16]。这些任务包括句子级任务,如自然语言推理^[17]和意译^[18],其目的是

通过整体分析句子来预测句子之间的关系;也包括标记级任务,如命名实体识别和问题解答,这些任务要求模型产生细粒度的输出^[19]。近年来,Bert预训练模型^[20]的提出,将语言模型的性能提升到了一个新的高度,在各种语言任务中的表现远远超过了以往的所有方法。

1 研究方法

本文情感分析方法的架构主要是建立在LDA主题模型和BERT预训练模型上,先利用LDA主题模型获得通过全体训练集总结出 n 个主题以及主题下 k 个关键词,根据设置的权重阈值对关键词进行筛选,对每个训练集样本进行主题归类并拼接超过权重阈值对应主题下的关键词;再将原评论拼接主题词后的文本输入到BERT模型中,进行词向量的融合与转换,得到包含文本主题词信息的向量。传统的BERT分类只用CLS作为分类特征来源,本文认为后面的每个token信息也具备一些有助于分类的特征,所以本文利用ECA注意力机制对后面的每个token赋予各自的权重,让网络自己学习获取有用的token进行分类,并通过一个全局平均池化层进行压缩,最后与CLS这个token进行拼接送入多层感知机(MLP)进行分类以预测情感极性。

1.1 LDA

LDA(Latent Dirichlet Allocation)^[21]是一种基于概率图模型的文本主题模型,与PLSA模型相似,同样假设文档是由潜在主题生成的,每个文档是主题的一个混合。而不同之处在于LDA引入了描述文档-主题和主题-词汇的Dirichlet先验分布,在推断中能够对参数引入先验知识,使得模型更加稳定。这有助于防止过拟合,提高模型的泛化能力。LDA模型的隐含主题拓扑结构如图1所示,该模型将文档集中的每篇文档表示为主题的混合,每个主题又表示为单词的混合。

1.2 BERT

BERT(Bidirectional Encoder Representations from Transformers)是Google于2018年提出的一种预训练语言模型。传统的语言模型通常是从左到右或从右到左单向阅读文本,而BERT通过Masked Language Model(MLM)任务,即在输入序列中随机遮蔽一些词汇然后预测这些被遮蔽词汇,使其能够同时利用上下文信息,更全面地理解语境。此外,BERT通过大规模的无标注语料进行预训练,并在特定任务上进行微调,使得其能够学到通用的语言

表示,同时适应不同的下游任务。BERT 的出现极大地推动了自然语言处理领域的发展,也催生了许多基于 BERT 的预训练模型的衍生工作。BERT 模型架构如图 2 所示。

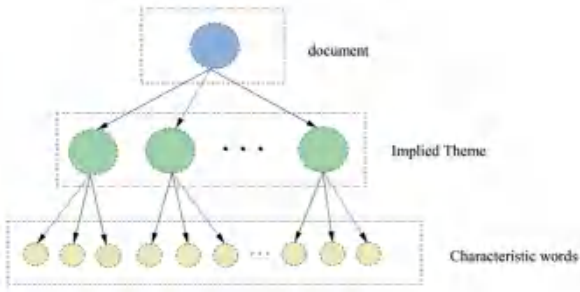


图 1 LDA 模型的隐含主题拓扑结构

Fig. 1 Implicit topic topology structure of LDA model

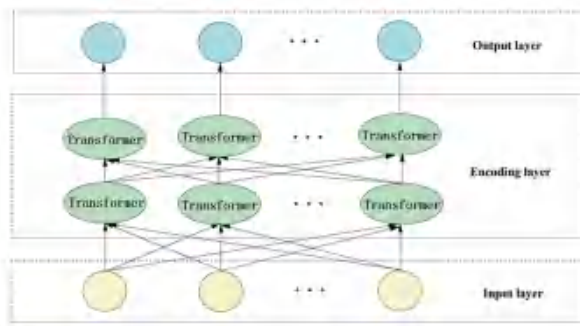


图 2 BERT 模型架构

Fig. 2 BERT model architecture

1.3 ECA

在深度学习领域,注意力机制的引入为提升模型性能提供了开发潜力。然而,传统的注意力模块,尤其是诸如 CBAM 等基于 SE (Squeeze-and-Excitation) 模块的改进方法,通常在提高性能的同时增加了模型的复杂性。解决这一问题的创新性方法之一是在 2020 年 CVPR 中提出的高效通道注意力模块 ECA (Efficient Channel Attention)^[22],其主要结构如图 3 所示。ECA 模块通过对 SE 模块的核心进行深入剖析,着眼于通道注意力中存在的问题,并提出了 2 个关键的改进。首先,通过采用全局平均池化操作,ECA 模块形成一个 $1 \times 1 \times C$ 的向量,从而解决了通道学习中可能引入降维负面效应的问题。其次,引入了自适应的卷积核大小 k ,利用一维卷积计算每个通道的权重。通过巧妙地共享权重,ECA 模块显著减少了参数量,有效地降低了模型的复杂度。

1.4 LE-BERT

在 LDA 与 BERT 融合上,如图 4 所示,本文先利用 LDA 主题模型获得通过全体训练集总结出 n 个主题以及主题下 k 个关键词,再根据设置的权重

阈值 w_0 对关键词进行筛选,然后对每个训练集样本进行主题归类并拼接对应主题下大于权重阈值 w_0 的关键词;最后将原评论拼接主题词后的文本输入到 BERT 模型中。实验表明,在本文的数据集上,相比于直接将原评论文本通过 BERT 进行情感分类,将原评论文本拼接通过 LDA 得到的相应关键词可有效提高 BERT 情感分类的结果。

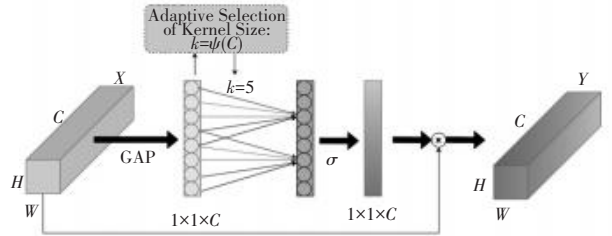


图 3 ECA 注意力机制

Fig. 3 ECA attention mechanism

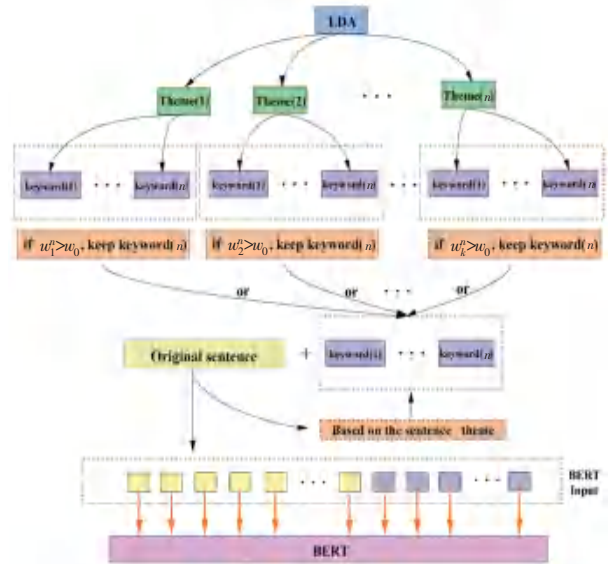


图 4 LDA+BERT 融合示意图

Fig. 4 LDA+BERT fusion diagram

在 LDA-BERT 下游模型的构造上,传统的 BERT 下游模型分类只用 [CLS] 作为分类特征来源,而本文认为后面的每个 token 信息也具备一些有助于分类的特征,所以本文创新性地从计算机视觉领域将高效通道注意力模块 ECA 引入 BERT 的下游模型。

本文利用 ECA 注意力模块对后面的每个 token 赋予各自的权重,让网络自己学习获取有用的 token 进行分类。最后通过一个全局平均池化层将除 CLS 外的 token 压缩到与 CLS 相同的维度上,并与 [CLS] 这个 token 进行拼接送入多层感知机 (MLP) 进行分类以预测情感极性,LDA-BERT 下游模型分类网络结构框架如图 5 所示。

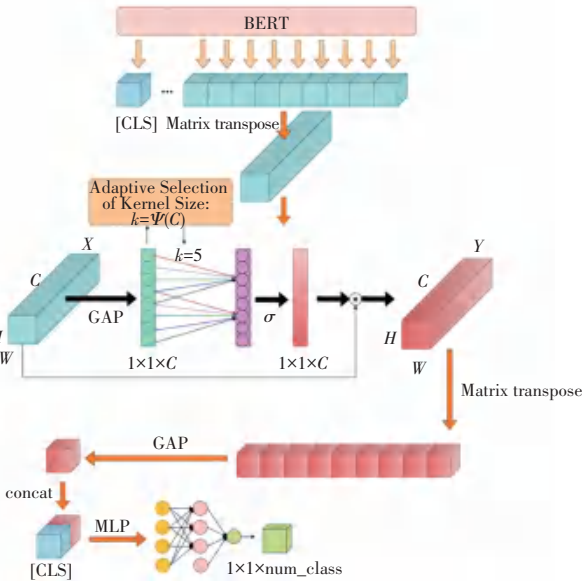


图 5 下游模型分类网络结构框架图

Fig. 5 Diagram of downstream task model classified network structure framework

LDA-BERT 下游模型前向传播过程如下:

(1) 对于 BERT 模型的输出 token 序列 $z = \{z_0, z_1, z_2, \dots, z_i\}$, 其中 z_0 为特殊符号 [CLS], 本文取其子序列 $\bar{z} = \{z_1, z_2, \dots, z_i\}$, 并将其内每个 token 进行拼接得到张量 Z , 将其转置作为 ECA 注意力机制的输入。

(2) 将输入 Z 进行全局平均池化, 之后进行卷积核大小为 k 的一维卷积运算。每个通道的权重 w 通过激活函数 σ 得到, 计算公式如下:

$$w = \sigma(C1D_k(Z^T)) \text{ where } Z = [z_1, z_2, \dots, z_i] \quad (1)$$

卷积核 k 的大小通过一个函数 $\varphi(C)$ 来进行自适应变化, 使得通道数较大的层可以更多地进行跨通道交互。推得的公式为:

$$k = \varphi(C) = \left\lfloor \frac{\log_2(c)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (2)$$

其中, $\lfloor t \rfloor_{\text{odd}}$ 表示最近的奇数 t 。

(3) 将权重 w 乘以特征 Z^T 得到最终的输出特征映射, 取转置得到 \tilde{Z} , 将其通过全局平均池化层后和 [CLS] 进行拼接, 最后送入 MLP 中, 得到分类特征 F 。对此可以表示为:

$$\tilde{Z} = (w \cdot Z^T)^T \quad (3)$$

$$F = \text{MLP}([\text{CLS}], \text{GAP}(\tilde{Z})) \quad (4)$$

2 实验

为了验证所提出的方法, 本文进行了一系列相关的实验来评估模型和算法的检测性能。本节包括

实验中一些设置的细节和使用的数据集的主要内容, 然后是实验结果中使用的评价指标。

2.1 实验环境

本文使用的是 NVIDIA GeForce RTX3090 GPU、Xeon(R) Platinum 8350C CPU、内存大小为 24 GB 的电脑来进行实验的训练部分。本文的算法框架采用了基于 Python 3.8 语言的 PyTorch 1.11.0。本文还在实验中使用了 CUDA 11.3 来调用 GPU 来进行训练加速。

2.2 数据集

在实验的训练部分中, 本文使用的是自己通过网络爬虫爬取 2022 年 12 月 2 日至 2022 年 12 月 11 日在新浪微博平台上部分典型网友评论形成的疫情期间舆情数据集。在爬虫结束后, 为了提高数据集质量, 本文进行了一系列如删除无用数据、分词、去停用词的数据预处理操作, 并人工进行微博的情感倾向二分类标注(正向为 1, 负向为 0)。最终得到 22 174 条有效数据集, 其中情感正向的有 12 881 条, 情感负向的有 9 293 条。利用 Python 得到疫情期间舆情数据集的关键词词云图如图 6 所示。

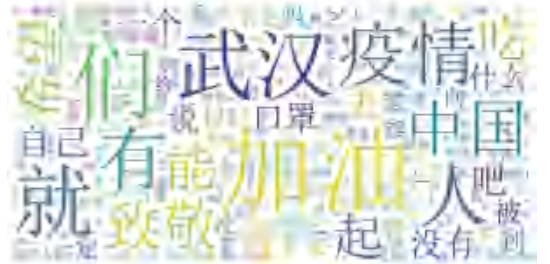


图 6 数据集关键词词云图

Fig. 6 Keyword word cloud map of the dataset

2.3 实验细节

在 LDA 上, 本文选择主题的个数为 7, 常用词个数为 100, 每个主题的关键词个数为 10。文档主题先验 Dirichlet 分布的参数为 0.1, 主题词先验 Dirichlet 分布的参数为 0.01。在 LDA 与 BERT 的拼接部分, 本文选取的关键词权重阈值为 300, 将权重大于 300 的关键词与原文进行拼接。在 BERT 上, 本文使用在 ADAM 优化算法上改进的自适应梯度方法 AdamW 来训练本文的网络, 网络采用的批量大小为 32。当本文进行对比实验及消融实验时, 本文对网络总共训练了 50 个周期。本文还设置了学习率为 0.000 01, 在 BERT 输入端, 输入的句子长度为 50, 词向量的维度为 768。

2.4 评价指标

在 LDA 模型中需要自行设置很多参数, 其中主题 (Topic) 数尤为重要。困惑度 (Perplexity) 常用来

度量一个概率分布或概率模型预测样本的好坏程度,困惑度越低,意味着更好的 LDA 模型。所以本文使用困惑度作为评价指标来确定模型的最佳主题数。困惑度计算公式为:

$$\text{Perplexity} = \exp \left\{ - \frac{\sum_{d=1}^M \log p(w_d)}{\sum_{d=1}^M N_d} \right\} \quad (5)$$

其中, $p(w_d)$ 表示测试集中每个单词出现的概率, N_d 表示 d 个文档中出现单词的总数目。

在 LDA-BERT 下游模型部分,本文使用准确率 (Accuracy)、精确率 (Precision)、召回率 (Recall) 和 $F1$ 值等方法,分析和验证了本文提出的模型检测性能。

(1) 准确率 (Accuracy)。数学定义定义为:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (6)$$

(2) 精确率 (P)。数学定义公式为:

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (7)$$

(3) 召回率 (R)。数学定义公式为:

$$R = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (8)$$

(4) $F1$ 值,是精确率和召回率的调和均值。数学定义公式为:

$$F1 = 2 \times \frac{P \times R}{P + R} \quad (9)$$

其中, TP 表示模型预测正确的正样本的数量; TN 表示模型预测正确的负样本的数量; FP 表示模型预测错误的正样本的数量; FN 表示模型预测错误的负样本的数量。

3 实验结果

本文的 LE-BERT 模型是利用 LDA 主题模型对疫情期间舆情数据集整体进行主题提取并对数据集中每个样本进行主题分类,再对原句进行主题下关键词的拼接嵌入,最终送入本文构造的引入 ECA 注意力机制的 BERT 分类模型进行情感分类。实验效果对比见表 1。从表 1 展现的数据来看,在疫情期间舆情数据集上使用本文的模型进行情感分类比单纯用 BERT 模型进行情感分类在 Accuracy, Precision, Recall 和 $F1$ 系数表现上分别提高了 4.14%, 4.39%, 2.57% 和 3.52%, 本文的模型在综合性能上有较好地提升。

表 1 LE-BERT 与 BERT 的分类效果对比

Table 1 Comparison of the classification effects of LE-BERT and BERT %

Method	Accuracy	Precision	Recall	$F1$
BERT	85.78	87.88	87.67	87.73
LE-BERT	89.92	92.27	90.24	91.25

根据式(5)计算不同主题数目下的 LDA 模型的困惑度,得到困惑度随主题数变化的曲线如图 7 所示。由图 7 可看到,在主题数为 7 时,困惑度已经到达拐点。经过验证,若主题数太多会导致 LDA 模型计算复杂度增加,且容易产生过拟合,反而影响模型分类准确度。

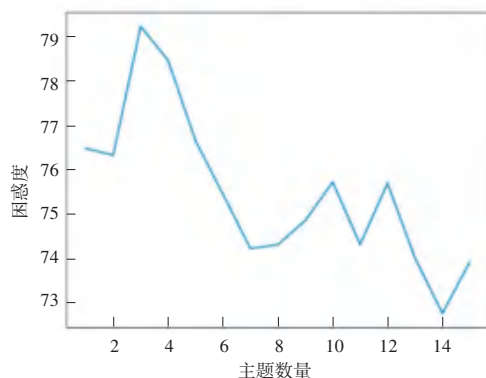


图 7 困惑度-主题数量的变化曲线

Fig. 7 The changing curve of perplexity-topic number

此后,利用 pyLDAvis 对分类结果进行可视化,得到不同主题数下数据集分类结果如图 8 所示。在运行过 LDA 主题模型之后,本文就可以知道每个主题出现的频率。将每个主题出现的频率高低用圆圈的大小来表示,同时也按顺序标号为 1 ~ n,图 8 中气泡的大小及编号即表示主题出现的频率高低。气泡相互之间的位置远近,就表达了主题之间的接近性。气泡距离采用的是 JSD 距离,气泡有重叠说明这 2 个主题里的特征词有交叉。

本文对比在 LDA 主题模型处选择不同主题数量的情况下,最终分类结果见表 2。在疫情期间舆情数据集上分析可以发现,选择不同的主题数量可以使最终分类结果的 Accuracy、Precision、Recall 和 $F1$ 值相差 2% 以上。依据困惑度随主题数变化的曲线图、分类结果的可视化及表 2 结果,本文最终确定的疫情期间舆情数据集主题数为 7。

在确定主题数为 7 后,本文对比在选取不同的关键词阈值情况下最终分类结果见表 3。由表 3 数据可知,在关键词阈值取值为 300 时模型的表现最好,故本文将关键词阈值设置为 300。

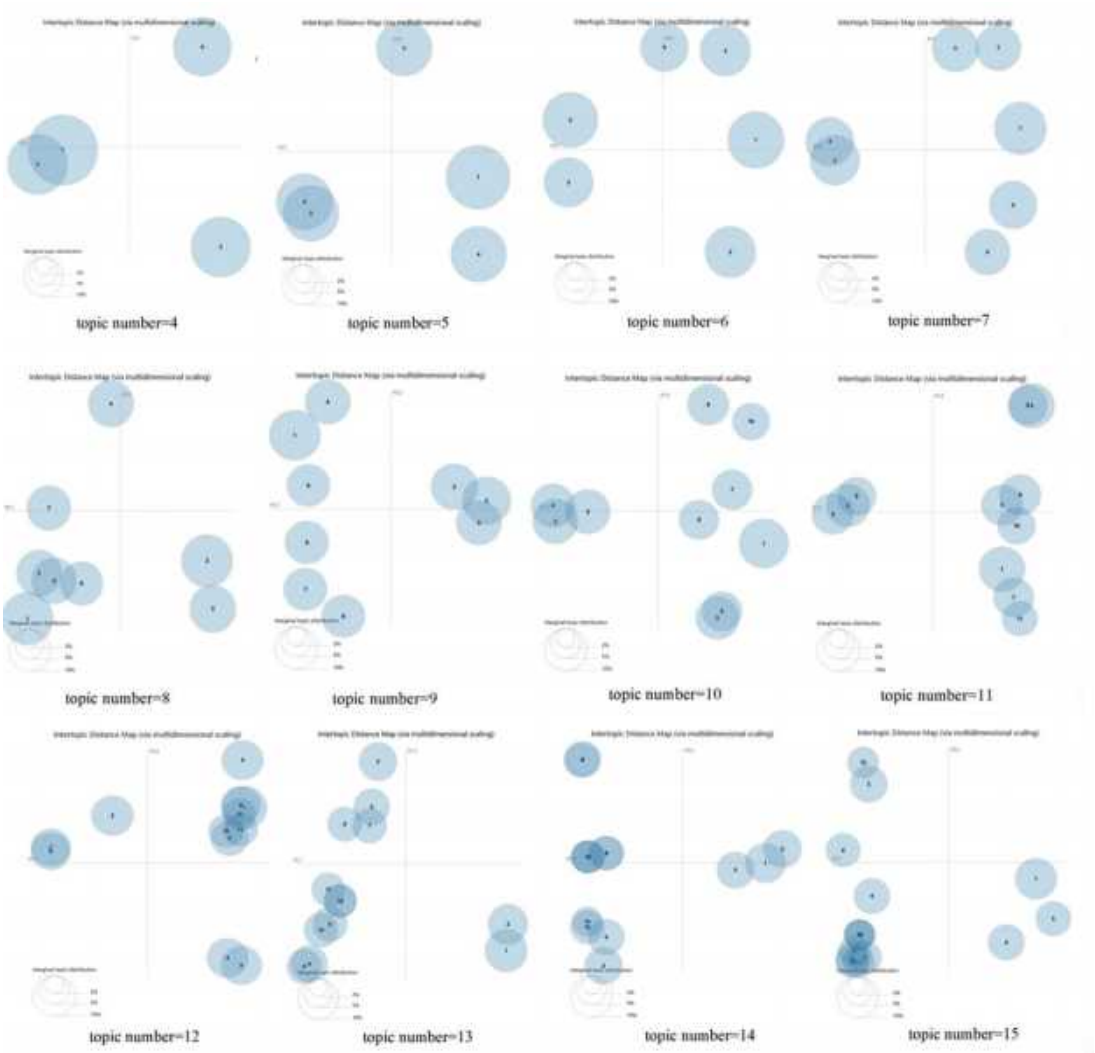


图 8 不同主题数下分类结果可视化

Fig. 8 Visualization of classification results under different number of topics

表 2 选择不同主题数量时最终分类结果的表现情况

Table 2 Performance of the final classification results when selecting different number of topics

Topic number	Confusion	Keyword number	Weight	<i>A</i> / %	<i>P</i> / %	<i>R</i> / %	<i>F1</i> / %
2	76.33	10	300	87.53	89.55	88.90	89.22
3	79.23	10	300	87.66	91.31	87.08	89.15
4	78.47	10	300	87.60	89.22	89.52	89.37
5	76.65	10	300	87.16	89.16	88.67	88.91
6	75.45	10	300	88.26	89.55	90.38	89.96
7	74.22	10	300	89.92	92.27	90.24	91.25
8	74.31	10	300	87.78	89.76	89.18	89.47
9	74.86	10	300	87.44	90.09	88.11	89.09
10	75.72	10	300	86.88	86.60	91.65	89.05
11	74.32	10	300	88.00	90.33	88.90	89.61
12	75.69	10	300	87.06	86.68	91.89	89.21
13	74.01	10	300	87.24	92.26	85.23	88.60
14	72.77	10	300	87.46	88.42	90.27	89.34
15	73.92	10	300	86.66	88.29	88.86	88.58

表3 选择不同关键词阈值时最终分类结果的表现情况

Topic number	Keyword number	Weight	A/ %	P/ %	R/ %	F1/ %
7	10	100	86.86	86.19	92.20	89.09
7	10	300	89.92	92.27	90.24	91.25
7	10	500	88.50	90.64	89.48	90.06
7	10	700	87.86	86.50	93.78	89.95
7	10	900	87.84	93.70	84.81	89.03
7	10	1 100	88.24	93.48	85.77	89.46
7	10	1 300	88.62	89.48	91.17	90.31
7	10	1 500	88.04	90.67	88.56	89.61

当选择下游模型注意力机制的时候,单纯使用 LDA-BERT 以及加入选取的 4 种注意力时的表现情况,见表 4。由表 4 可知,在疫情期间舆情数据集上可以发现,加入任意一个本文选取的注意力机制,都能使得模型性能得到较为明显的提升,而 ECA 注意力的综合性能表现是其中最好的,可使分类结果对比在下游模型中不加入注意力时的 Accuracy, Precision, Recall 和 F1 系数分别提高了 3.80%, 4.13%, 2.31% 和 3.22%。

表4 5种注意力机制加入模型时的表现情况

Table 4 Performance of five attention mechanisms when added to the model %

Attention Methods	A	P	R	F1
None	86.12	88.14	87.93	88.03
SENet	89.06	90.51	90.72	90.62
CBAM	89.60	91.82	90.18	90.99
CA	89.74	91.66	90.62	91.14
ECANet	89.92	92.27	90.24	91.25

模型进行了消融实验见表 5。由表 5 可以看到拼接由 LDA 得到的关键词使得 BERT 的分类效果在 Accuracy, Precision, Recall 和 F1 系数上分别提升 0.34%, 0.26%, 0.26% 和 0.30%; 而 ECA 注意力机制的引入极大提升了模型性能,使得 BERT 的分类效果在 Accuracy, Precision, Recall 和 F1 系数上分别提升 3.68%, 2.39%, 4.12% 和 3.29%。

表5 消融实验

Table 5 Ablation experiment

BERT	LDA	ECA	A/ %	P/ %	R/ %	F1/ %
√			85.78	87.88	87.67	87.73
√	√		86.12	88.14	87.93	88.03
√		√	89.46	90.27	91.79	91.02
√	√	√	89.92	92.27	90.24	91.25

此外,为验证模型的有效性,本节还设计了 LE-BERT 与当前主流文本情感分析模型的对比实验,使用的数据集是在文本情感分析领域模型评估中被广泛应用的电影评论数据集 MR^[23]。该数据集包含正、负情感倾向的电影评论各 5 331 条,同样以 8 : 2 的比例分割训练集和测试集。选用的对比模型有:

(1) Text-CNN^[24]。将输入文本表示为词向量序列,并通过卷积和池化操作提取特征,最终使用全连接层进行情感分类。

(2) RNN-Capsule^[25]。一种结合了循环神经网络和胶囊网络的模型,在情感分类中可以利用其对序列特征的学习和对部分-整体关系的建模能力,更好地捕获文本中的语义信息。

(3) Bi-GRU^[26]。通过 GRU 循环神经网络提取文本情感特征,结合注意力机制突显词汇重要性,将加权结果输入全连接层以输出情感分布。

(4) Bi-LSTM^[27]。结合双向 LSTM 提取上下文语义信息,同时运用注意力机制捕获文本语序与语义信息,有效处理前后词间的依赖关系,并应用于情感分类实验。

(5) BERT-Attention^[28]。首先利用 BERT 提取文本特征向量,然后运用注意力机制赋予各特征权重进行情感分类。

(6) BiLSTM-CRF^[29]。结合双向长短期记忆网络和条件随机场,通过 BiLSTM 提取文本特征,再利用 CRF 模型对序列进行标注,从而实现情感分类。

实验结果见表 6。由表 6 可见,本文提出的 LE-BERT 在公开数据集 MR 上实验得出的模型四项评价指标均优于当前主流文本情感分析对比模型,说明模型具有较好的性能和较强的鲁棒性。

表6 对比实验

Table 6 Comparative experiment %

Models	A	P	R	F1
Text-CNN ^[24]	79.19	78.84	78.83	78.84
RNN-Capsule ^[25]	83.10	82.82	82.84	82.83
Bi-GRU ^[26]	84.3	84.06	84.01	84.03
Bi-LSTM ^[27]	84.66	84.43	84.37	84.39
BERT-Attention ^[28]	85.45	85.24	85.14	85.19
BiLSTM-CRF ^[29]	85.62	85.43	85.31	85.36
LE-BERT	87.94	87.80	87.67	87.73

4 结束语

大数据时代,利用情感分析技术从数以亿计的文本信息中提取出有价值的信息并掌握舆论走向,从而帮助决策者做出重要决策具有很大的研究及应用价值。本文提出一种基于LDA和图像注意力机制ECA改进的BERT分类模型LE-BERT,能对舆情文本进行情感分析,判断其正负情感倾向从而分析舆情走势。其中主要工作包括:

(1)爬取2022年12月2日至2022年12月11日在新浪微博平台上部分典型网友评论,并进行一系列数据预处理及人工标注,形成疫情期间舆情数据集。

(2)利用LDA主题模型对数据集进行主题分类并选取主题关键词与原句进行拼接,作为BERT模型的输入,使文本具有主题信息。

(3)在下游模型中引进ECA注意力机制,一定程度上提高了模型分类的准确率。实验表明,本文提出的LE-BERT模型对比BERT原始模型在情感分类任务上的Accuracy, Precision, Recall和F1系数表现分别提高了4.14%, 4.39%, 2.57%和3.52%。

参考文献

[1] 中国互联网络信息中心. 第52次中国互联网络发展状况统计报告[R]. 北京: CNNIC, 2023.

[2] 胡稳财. 突发公共事件网络舆情影响因素分析[J]. 现代商贸工业, 2021, 42(32): 120-121.

[3] 杜宇飞, 张鹏, 王娟, 等. 我国突发事件应急管理研究可视化分析与展望[J]. 中国人民警察大学学报, 2022, 38(2): 79-85.

[4] ZARGARI H, ZAHEDI M, RAHIMI M. GINS: A global intensifier-based N-Gram sentiment dictionary[J]. Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, 2021, 40(6): 11763-11776.

[5] 邱全磊, 崔崇敏, 喻静. 基于表情和语气的情感词典用于弹幕情感分析[J]. 计算机技术与发展, 2020, 30(8): 178-182.

[6] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.

[7] KHINE W L K, AUNG N T T. Applying deep learning approach totargeted aspect-based sentiment analysis for restaurant domain [C]//Proceedings of 2019 International Conference on Advanced Information Technologies (ICAIT). Piscataway, NJ: IEEE, 2019: 206-211.

[8] SHORYU T, WANG Lei, MA Ruixin. A deep neural network approach using convolutional network and long short term memory for text sentiment classification [C]//Proceedings of 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD). Piscataway, NJ: IEEE, 2021: 763-768.

[9] LOU Jiashu, CUI Leyi, LI Ye. Bi-LSTM price prediction based on attention mechanism [J]. arXiv preprint arXiv, 2212. 03443, 2022.

[10] CHUNG J, GULCEHRE C, CHO K H, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling [J]. arXiv preprint arXiv, 1412. 3555, 2014.

[11] SUTSKEVER I, VINYALS O, LE Q V. Sequence to sequence learning with neural networks [J]. arXiv preprint arXiv, 1409. 3215, 2014.

[12] JOZEFOWICZ R, VINYALS O, SCHUSTER M, et al. Exploring the limits of language modeling [J]. arXiv preprint arXiv, 1602. 02410, 2016.

[13] KUCHAIEV O, GINSBURG B. Factorization tricks for LSTM networks [J]. arXiv preprint arXiv, 1703. 10722, 2017.

[14] SHAZEER N, MIRHOSEINI A, MAZIARZ K, et al. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer [J]. arXiv preprint arXiv, 1701. 06538, 2017.

[15] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. arXiv preprint arXiv, 1706. 03762, 2017.

[16] DAI A M, LE Q V. Semi-supervised sequence learning [C]// Proceedings of the 29th International Conference on Neural Information Processing Systems. Montreal, Canada: NeurIPS, 2015: 3079-3087.

[17] WILLIAMS A, NANGIA N, BOWMAN S. A broad-coverage challenge corpus for sentence understanding through inference [J]. arXiv preprint arXiv, 1704. 05426, 2018.

[18] DOLAN W B, BROCKETT C. Automatically constructing a corpus of sentential paraphrases [C]//Proceedings of the Third International Workshop on Paraphrasing (IWP2005). Jeju Island: Asia Federation of Natural Language Processing, 2005: 9-16.

[19] TJONG K S E F, BUCHHOLZ S. Introduction to the CoNLL-2000 shared task: Chunking [C]// Proceedings of The Workshop on Learning Language in Logic and the Conference on Computational Natural Language Learning. ACL, 2000: 127-132.

[20] DEVLIN J, CHANG Mingwei, 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, Volume 1 (Long and Short Papers). ACL, 2019: 4171 - 4186.

[21] BLEI D M, NG A, JORDAN M I. Latent dirichlet allocation [J]. The Journal of Machine Learning Research, 2003, 3: 993-1022.

[22] WANG Qilong, WU Banggu, ZHU Peifei, et al. ECA-Net: Efficient channel attention for deep Convolutional Neural Networks [C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2020: 11531-11539.

- [23] PANG Bo, LEE L. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales [C]// Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05). ACL, 2005: 115-124.
- [24] GUO B, ZHANG C, LIU J, et al. Improving text classification with weighted word embeddings via a multi-channel TextCNN model[J]. Neurocomputing, 2019, 363: 366-374.
- [25] WANG Y, SUN A, HAN J, et al. Sentiment analysis by capsules [C]//Proceedings of the 2018 World Wide Web Conference. New York: ACM, 2018: 1165-1174.
- [26] LIANG Xiao, LIU Zhiming, OUYANG Chunping. A multi-sentiment classifier based on GRU and attention mechanism[C]// Proceedings of 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS). Piscataway, NJ: IEEE, 2018: 527-530.
- [27] YAO Xianglu. Attention-based BiLSTM neural networks for sentiment classification of short texts [C]// Proceedings of the 2017 International Conference on Information Science and Cloud Computing. Guangzhou: IPEA, 2017: 110-117.
- [28] ZHANG Xuelei, SONG Xinyu, FENG Ao, et al. Multi-self-attention for aspect category detection and biomedical multilabel text classification with bert [J]. Mathematical Problems in Engineering, 2021(49):6658520.
- [29] CHEN Tao, XU Ruifeng, HE Yulan, et al. Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN[J]. Expert Systems with Applications, 2017, 72: 221-230.