

文本情感分析综述

王钦炀¹, 施水才^{1,2}, 王洪俊²

(1. 北京信息科技大学 计算机学院, 北京 100101; 2. 拓尔思信息技术股份有限公司, 北京 100096)

摘要: 随着在线平台个人意见数量激增, 情感分析变得至关重要, 其能帮助机构深入了解用户的情感倾向, 优化产品服务, 为市场决策提供有力支持, 更精准地理解社会舆论动向。综述情感分析领域的最新进展, 包括预处理技术、特征提取方法、分类技术、常用数据集等, 探讨该领域的局限性与未来研究方向, 以期对相关研究人员与从业人员提供有价值的资源。

关键词: 情感分析; 机器学习; 深度学习; 集成学习; 数据集

DOI: 10.11907/rjdk.232257

中图分类号: TP391.1

文献标识码: A

开放科学(资源服务)标识码(OSID):

文章编号: 1672-7800(2025)001-0193-10



A Review of Textual Sentiment Analysis

WANG Qinyang¹, SHI Shuicai^{1,2}, WANG Hongjun²

(1. School of Computer, Beijing Information Science and Technology University, Beijing 100101, China;
2. TRS Information Technology Co Ltd, Beijing 100096, China)

Abstract: With the surge in personal opinions on online platforms, sentiment analysis has become crucial. It can help institutions gain a deeper understanding of users' emotional tendencies, optimize product services, provide strong support for market decision-making, and more accurately understand social public opinion trends. Summarize the latest developments in the field of sentiment analysis, including preprocessing techniques, feature extraction methods, classification techniques, commonly used datasets, etc. Explore the limitations and future research directions of this field, in order to provide valuable resources for relevant researchers and practitioners.

Key Words: sentiment analysis; machine learning; deep learning; ensemble learning; datasets

0 引言

随着在线平台中个人公开表达意见与观点数量的激增, 理解这些意见背后的潜在情绪以便作出正确决策变得越来越重要。通过理解客户意见与态度背后隐藏的情绪能提高客户满意度, 增加品牌声誉, 并最终增加收入; 通过了解公众对政党、候选人与政策的看法可进行政治行情分析; 通过分析金融新闻评论与社交媒体帖子可以预测股价, 识别潜在的投资机会。由此可见, 情感分析在商业、政治、科技等领域的重要性、紧迫性日益凸显。

本文对文本情感分析领域的最新进展进行全面回顾, 将研究方法分为传统机器学习、深度学习与集成学习3

类^[1-3]。详细介绍了不同研究中的分类技术、使用的数据集以及实验结果, 同时对当前研究的局限性与该领域未来的研究方向进行了讨论。

1 文本情感分析任务简介

文本情感分析是自然语言处理的一个子领域, 专注于书面文本表达情感的自动识别。随着社交媒体数量呈指数级增长, 公众意见与情感的可用性增加, 情感分析成为理解各个领域公众情感的关键工具。情感分析框架主要包括4个模块以及一个附加的可选模块。具体如图1所示。

数据收集与标准化阶段通过网络爬虫、API(Applica-

收稿日期: 2023-12-10

扫描二维码阅读全文:



作者简介: 王钦炀(1999-), 男, 北京信息科技大学计算机学院硕士研究生, 研究方向为自然语言处理; 施水才(1966-), 男, 硕士, 北京信息科技大学计算机学院教授、硕士生导师, 研究方向为信息检索、大数据分析、人工智能; 王洪俊(1975-), 男, 博士, 拓尔思信息技术股份有限公司工程师, 研究方向为中文信息处理与信息检索。本文通讯作者: 施水才。

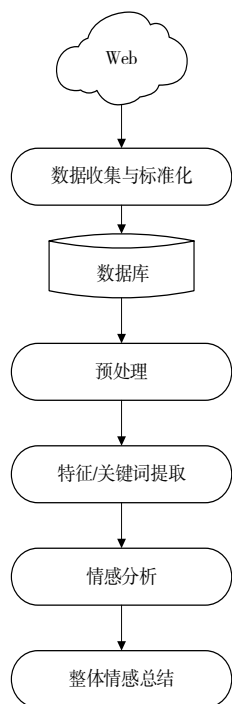


Fig. 1 Sentiment analysis framework

图1 情感分析框架

tion Programming Interface)调用与数据库查询等技术手段从各种来源收集文本数据,随后进行标准化处理,确保数据具有一致的格式与表达方式;在数据预处理阶段,原始文本数据经过分词、去除停用词、词干提取等操作转化为计算机可处理的格式,有助于减少噪音、保留关键信息,并为特征提取与模型训练作好准备;特征提取阶段通过应用TF-IDF(Term Frequency - Inverse Document Frequency)、词嵌入与N-gram模型等技术,由情感分析模块从预处理后的文本中提取关键特征或词语,这些特征将用于后续情感预测或分类任务;情感预测或分类阶段利用机器学习算法或深度学习模型,由情感分析模块对提取的特征进行预测或分类,以确定文本中包含的情感信息;整体情感总结阶段通过统计分析与可视化技术,由情感分析模块对个体情感分析结果进行综合,提供更全面的用户情感洞察。

2 情感分析数据集

情感分析数据集为情感分析模型的训练与评估提供了基础。尽管在一些情感分析工作中使用的事自建数据集,但是目前有几个公开且广泛使用的情感分析数据集,包括互联网电影数据库(IMDb)、Sentiment140、SemEval-2017任务4、Twitter US Airline Sentiment与OCEMOTION等。

2.1 IMDb

IMDb数据集包含50 000条电影评论,平均分为两个部分,每部分包含25 000条评论,一半为正面,一半为负面^[4]。数据集包括评论与情感两列,评论融合了故事情节

与个人观点,其语言复杂性增加了情感分析的难度。

2.2 Sentiment140

斯坦福大学收集的Sentiment140数据集是一份全面的客户情感数据集,包含160万个样本,均分为积极与消极两个类别^[5]。该数据集采集自推特,推文的简洁性与非正式性可能会对情感分析任务构成挑战。此外,该数据集采集自现实场景,更贴近在实际应用中可能遇到的数据类型,为研究人员提供了在实践中评估情感分析模型性能的机会。

2.3 SemEval-2017任务4

SemEval-2017任务4数据集是一个多语言数据集,包括英语与阿拉伯语,是一种综合基准数据集,包含5个子任务,涵盖了情感分析的不同方面,如消息极性分类、特定主题的消息极性分类与Tweet量化等^[6]。每个子任务都有不同数量的情感极性类别,为在不同场景下评估情感分析模型提供了全面基准。

2.4 Twitter US Airline Sentiment

Twitter US Airline Sentiment数据集由众包平台CrowdFlower于2017年收集,提供了关于美国六大航空公司的丰富客户评论。该数据集中的情绪包括积极、消极与中性3类,样本量分别为2 363、9 178与3 099。其主要挑战为类别不平衡,大多数样本属于负面情绪类,这可能会影响情感分析模型的准确性。此外,数据集中推文的简洁性与非正式性对情感分析任务提出了进一步挑战。推文通常以非正式的风格写成,篇幅简短,可能会导致上下文与重要的情感表达词被遗漏,这些挑战会使情感分析模型对推文进行错误分类。

2.5 OCEMOTION

OCEMOTION数据集的设计与组织使其成为理想的情感分析任务数据源。该数据集涵盖了丰富的情感类别,包括悲伤、快乐、厌恶、愤怒、喜欢、惊讶与恐惧,为研究者提供了多维度的情感信息^[7]。该数据集要求模型对文本中的情感进行更为详细、具体和有层次的分析,而不仅仅是简单的积极或消极分类。

表1为情感分析数据集的总结。该表提供了数据集特征的简明概述,便于研究人员比较不同数据集,并选择最适合自身需求的数据集。

3 文本情感分析方法

要完成情感分析任务,原始文本数据必须经过数据预处理、特征提取与分类3个步骤。情感分析中使用的分类器可大致分为机器学习、深度学习与集成学习3类。机器学习分类器使用数学模型预测情感;深度学习分类器利用神经网络进行情感预测;集成学习方法结合多个分类器实现情感分析。分类器的选择取决于情感分析任务的具体要求与用例。以下回顾现有情感分析研究,并重点介

Table 1 Summary of sentiment analysis datasets

表 1 情绪分析数据集总结

数据集	类别数目	非常积极	积极	中性	消极	特别消极	合计
IMDb	2	-	25 000	-	25 000	-	50 000
Sentiment140	2	-	800 000	-	800 000	-	1 600 000
SemEval-2017 4A	3	-	22 277	28 528	11 812	-	62 617
SemEval-2017 4B	2	-	17 414	-	7 735	-	25 149
SemEval-2017 4C	5	1 151	15 254	19 187	6 943	476	43 011
SemEval-2017 4D	2	-	17 414	-	7 735	-	25 149
SemEval-2017 4E	5	1 151	15 254	19 187	6 943	476	43 011
Twitter US Airline Sentiment	3	-	2 363	3 099	9 178	-	14 160
OCEMOTION	7	悲伤(12 475)、快乐(8 894)、厌恶(4 347)、愤怒(4 068)、喜欢(4 042)、惊讶(899)、恐惧(590)					35 315

绍每个研究使用的分类算法。

3.1 机器学习方法

机器学习方法首先通过预处理与删除任何无关信息的方式标准化文本数据;然后应用 TF-IDF、N-grams 等特征提取技术将文本表示为可输入机器学习分类器的数字特征。机器学习情感分析方法常用的分类器包括支持向量机、朴素贝叶斯、逻辑回归与决策树等。其中,支持向量机通过寻找最优超平面来实现数据点的二分类,这个超平面被设计为最大化两个类别的支持向量到超平面的距离,从而提高模型的泛化能力。支持向量机在处理线性可分问题时表现出色,同时也能通过核函数处理非线性问题;朴素贝叶斯是基于贝叶斯定理的分类算法,其假设特征之间相互独立,使得计算条件概率更加简化,特别适用于文本分类等领域,在处理大规模数据集与高维特征空间时具有高效性;逻辑回归作为一种二分类算法,通过使用 sigmoid 函数将线性组合输入映射到 0~1 的范围,从而得到样本属于正类别的概率,其训练过程基于最大似然估计,适用于概率估计和二分类问题;决策树是一种基于特征的分层决策方法,通过在特征空间中构建树状结构逐层作出决策,最终实现数据分类。其易于理解和解释,在处理非线性关系时具有良好的性能。图 2 为传统机器学习方法的情感分析过程。

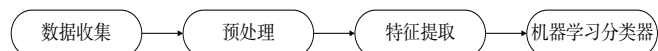


Fig. 2 Sentiment analysis process of traditional machine learning methods

图 2 传统机器学习方法情感分析过程

例如, Athindran 等^[8]利用朴素贝叶斯分析 Twitter 上的客户评论,预处理步骤包括标记化与词干提取,准确率达到 77%。Vanaja 等^[9]研究比较了两种流行的机器学习算法——朴素贝叶斯与支持向量机在亚马逊客户评论情感分析中的性能,首先对文本数据进行预处理以去除停止词,然后使用先验算法表示。结果表明,朴素贝叶斯算法的准确率为 90.42%,优于支持向量机算法的 83.42%。Iqbal 等^[10]使用朴素贝叶斯、支持向量机与最大熵 3 种机器学习方法进行情感分析,实验在 IMDb 与 Sentiment140 两个数据集上进行,结果表明,当最大熵方法结合单图与双图特征时,在 IMDb 数据集上的准确率达到 88%,在 Senti-

ment140 数据集上的准确率达到 90%。Rathi 等^[11]评估了 3 种机器学习算法(决策树、AdaBoost 与支持向量机)的情感分析性能,首先对文本进行预处理,然后使用 TF-IDF 特征表示。结果表明,支持向量机的准确率最高,达到 82%;其次是决策树,准确率为 81%;AdaBoost 的准确率最低,为 67%。Tariyal 等^[12]比较了线性判别分析、分类回归树、k 近邻、支持向量机、随机森林与 C5.0 对包含 1 150 条产品评论推文的自建数据集的情感分析性能,结果表明,分类回归树的准确率最高,达到 88.99%。Rahat 等^[13]比较了多项式朴素贝叶斯与支持向量机对包含 10 000 条 Twitter 数据集的情感分析性能。结果表明,支持向量机的准确率达到 82.48%,优于多项朴素贝叶斯模型的 76.56%。Makhmudah 等^[14]使用支持向量机对印度尼西亚与同性恋相关推文进行情感分析,取得了 99.5% 的准确率。Wongkar 等^[15]对支持向量机、朴素贝叶斯与 K 近邻 3 种机器学习方法在 2019 年印度尼西亚总统候选人相关推文上的情感分析结果进行比较,结果表明,朴素贝叶斯算法的准确率最高,为 75.58%;其次是 K 近邻算法,准确率为 73.34%;支持向量机的准确率最低,为 63.99%。Madhuri^[16]比较了 C4.5、朴素贝叶斯、支持向量机与随机森林 4 种机器学习方法对印度铁路相关推文数据集的情感分析性能。结果表明,支持向量机的准确率最高,为 91.5%;其次是随机森林,准确率为 90.5%;C4.5 的准确率为 89.5%,朴素贝叶斯的准确率为 89%。Gupta 等^[17]采用 sentiment140 数据集评估了决策树、逻辑回归、支持向量机、神经网络算法的情感分析性能,结果表明,神经网络模型的表现优于其他算法,达到 80% 的最高准确率。Prabhakar 等^[18]通过在 AdaBoost 模型中结合 boosting 与 bagging 方法设计了一种新的情感分析方法,并在美国航空公司 Twitter 数据集上进行情感分析实验, F-score 为 68%。Hourrane 等^[19]在 IMDb 与 Sentiment140 两个数据集上应用 ridge 分类方法进行情感分析。结果表明,ridge 分类器在 IMDb 数据集上的准确率为 90.54%,在 Sentiment140 数据集上的准确率为 76.84%。Alsaman^[20]使用多项朴素贝叶斯方法对阿拉伯语推文数据集进行情感分析,在将原始数据表示为 TF-IDF 特征并使用 5 倍交叉验证进行分类之前,首先使用 4-gram 进行标记,然后使用 Khoja 词干提取器进行预处理,获得了 87.5% 的准确率。Saad^[21]

使用支持向量机、逻辑回归、随机森林、XgBoost、朴素贝叶斯与决策树等6种机器学习模型对美国航空公司 Twitter 数据集进行情感分析,结果表明,支持向量机的准确率最高,为83.31%;其次是逻辑回归,准确率为81.81%。Alzyout等^[22]使用支持向量机、k近邻、朴素贝叶斯与决策树对暴力侵害妇女相关文本进行情感分析,结果表明支持向量机的准确率最高,达到78.25%。Gupta等^[23]使用TextBlob与VADER标记提取印度公民发布的新冠疫情相关推文,

构建了一个包含7 284条推文的数据集,测试了8种机器学习分类器的情感分析性能,并对每个变种执行了十折交叉验证,结果表明具有 unigram 特征的 LinearSVC 获得了最高的准确率,为84.4%。

综上所述,不同机器学习算法具有不同的准确率水平,范围为67.0%~99.5%,分类性能较好的算法为最大熵、朴素贝叶斯与支持向量机。表2为以上机器学习方法的总结。

Table 2 Summary of machine learning approaches

表2 机器学习方法总结

文献	特征提取	分类器	数据集	准确率/%
Athindran等 ^[8]		NB	自制数据集(来自 Twitter)	77
Vanaja等 ^[9]	A priori算法	NB,SVM	自制数据集(来自 Amazon)	83.42
Iqbal等 ^[10]	Unigram Bigram	NB,SVM,ME	IMDb	88
			Sentiment140	90
Rathi等 ^[11]	TF-IDF	DT,AdaBoost,SVM	Sentiment140, Polarity Dataset, University of Michigan dataset	84
				67
Tariyal等 ^[12]		分类回归树	自制数据集(来自 Twitter)	88.99
Rahat等 ^[13]		SVC	自制数据集(来自 Twitter)	82.48
		MNB		76.56
Makhmudah等 ^[14]	TF-IDF	SVM	与同性恋相关的推文	99.5
Wongkar等 ^[15]		NB	2019年印度尼西亚总统候选人相关推文	75.58
Madhuri ^[16]		SVM	印度铁路相关推文	91.5
Gupta等 ^[17]	TF-IDF	神经网络	sentiment140	80
Prabhakar等 ^[18]		AdaBoost	美国航空公司相关推文	68 (F1-score)
			IMDb	90.54
Hourrane等 ^[19]	TF-IDF	Ridge	Sentiment 140	76.84
Alsalmán ^[20]	TF-IDF	MNB	阿拉伯语推文	87.5
Saad ^[21]	Bag of Words	SVM	Twitter US Airline Sentiment	83.31
Alzyout等 ^[22]	TF-IDF	SVM	暴力侵害妇女相关文本	78.25
Gupta等 ^[23]		LinearSVC	印度新冠疫情相关推文	84.4

3.2 深度学习方法

深度学习是目前流行的情感分析方法,是多层神经网络的应用。神经网络结构如图3所示,基本构架包括输入层、隐藏层与输出层。每个神经元在输入层代表了对象的一个特征。隐藏层可以是一层或多层,其被视为一个黑盒,不直接与外界通信,既不接收外界信号,也不向外界发送信号,主要功能是对输入信息进行转换,以生成输出层可以使用的数据。输出层将隐藏层的处理结果转化为所需的最终输出。

深度学习情感分析方法首先对文本数据进行预处理;然后采用预训练的嵌入模型,如 GloVe 与 word2vec,对文本进行编码;最后将这些嵌入向量输入到深度学习模型中,用于学习文本表示或执行相应的分类任务。例如,Raza等^[24]利用具有5个隐藏层与 ReLU 激活的多层感知器(Multilayer Perceptron, MLP)对 COVID-19 相关推文进行情感分析,准确率最高达到93.73%。Harjule等^[25]在 Sentiment140 与 Twitter US Airline Sentiment 两个数据集上比较了长短期记忆网络(Long Short-Term Memory, LSTM)和支持向量机的情感分析性能。结果表明,LSTM 在 Senti-

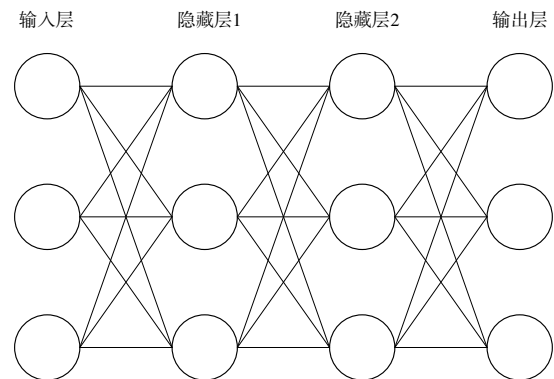


Fig. 3 Neural network structure

图3 神经网络结构

ment140数据集上达到了82%的最高准确率,而支持向量机在 Twitter US Airline Sentiment 数据集上实现了68.9%的最高准确率。Hossain等^[26]提出一种卷积神经网络(Convolutional Neural Networks, CNN)与 LSTM 结合的情感分析模型,包括 word2vec 预训练模型的嵌入层、卷积层、最大池化层、LSTM层、dropout层与分类层,在自建餐厅评论数据集上取得了75.01%的准确率。Tyagi等^[27]将 CNN 与双向长

短期记忆网络(Bidirectional Long Short-Term Memory, BiLSTM)相结合用于Sentiment140数据集的情感分析,取得了81.20%的准确率。Rhanoui等^[28]提出一种结合CNN与BiLSTM的混合模型,在包含2 003篇文章的数据集上取得了90.66%的准确率。Jang等^[29]使用Adam优化器、L2正则化、dropout、注意力机制对CNN与BiLSTM的混合模型进行改进,在IMDb数据集上的准确率达到90.26%。Janardhana等^[30]提出一种由CNN与循环神经网络(Recurrent Neural Network, RNN)组成的混合模型,通过去除停止词与标点符号对文本进行预处理,然后使用GloVe预训练模型表示文本,在电影评论数据集上达到了84%的准确率。Chowdhury等^[31]使用BiLSTM模型对Twitter US Airline Sentiment数据集进行情感分析,使用word2vec、GloVe与情感特定词嵌入来表示文本,准确率达到81.20%。Vimali等^[32]采用BiLSTM模型分析亚马逊电子商务评论的情感类型,通过数据标记、去除特殊字符进行数据清洗,然后使用word2vec预训练模型生成词嵌入,取得了90.26%的准确率。Anbukarasi等^[33]使用基于字符的深度双向长短期记忆模型DBLSTM分析泰米尔语推文的情感,采用word2vec预训练模型表示清洗后的数据作为词嵌入,准确率达到86.2%。Kumar等^[34]提出情感分析双向长短期记忆模型SAB-LSTM,该模型由196个BiLSTM单元、128个嵌入层、4个密集层与一个带有softmax激活函数的分类层组成,对由Twitter、YouTube与Facebook等不同来源收集的80 689个新闻文章样本组成的数据集进行情感分析,取得了比传统LSTM模型更好的效果。Hossen等^[35]使用RNN对从酒店预订网站收集的客户端评论进行情感分析,采用词源化、词干化以及标点和停止词去除等方法对数据进行清理,取得了86%的准确率。Younas等^[36]比较了多语言BERT(mBERT)与XLM-RoBERTa(XLM-R)两种深度学习模型在多语言社交媒体文本情感分析中的性能,使用的数据集由20 375条英语与罗马乌尔都语推文组成,结果表明,mBERT的准确率为69%,而XLM-R的准确率为71%。Dhola等^[37]比较了支持向量机、多项朴素贝叶斯、LSTM与BERT对Sentiment140数据集的情感分析性能,结果表明,BERT模型表现最好,准确率达到85.4%。Tan等^[38]设计了一种融合RoBERTa与LSTM的情感分析方法,将RoBERTa的自关注和动态掩蔽能力与LSTM捕获编码文本中远程依赖关系的能力相结合,在IMDb、Twitter US Airline Sentiment与Sentiment140数据集上分别取得92.96%、91.37%与89.70%的准确率。Kokab等^[39]提出一种基于BERT的卷积双向循环神经网络模型CBRNN,从数据中提取句子级语义与上下文特征并生成嵌入,然后使用扩展卷积提取局部和全局上下文语义特征,在Twitter US Airline Sentiment数据集上的准确率为97%,在自动驾驶汽车评论数据集上的准确率为90%,在美国总统选举评论数据集上的准确率为96%,在IMDb数据集上的准确率为93%。AlBadani等^[40]

提出一种情感转换图卷积网络ST-GCN,将情感分析建模为异构图,并使用情感图转换神经网络学习文档与单词嵌入。该模型学习了节点之间的有用连接,并确定了边缘类型与复杂关系的软选择,以学习用于情感分类的节点表示。利用拉普拉斯特征向量融合节点位置信息,使用消息传递技术学习异构图上的节点表示,并使用Transformer对具有适当位置编码的局部子结构进行聚合。该模型在真实世界的数据集上取得了最先进的结果,在SST-B数据集上的准确率达到95.43%,在Yelp2014数据集上的准确率达到72.7%,在IMDb数据集上的准确率达到94.94%。Tiwari等^[41]提出KEAHT模型,使用预训练的BERT模型在最小训练语料库上进行训练,整合了来自LDA主题建模和词汇化领域本体的显式知识,以解决情感分析中不准确的极性评分与基于效用的主题建模问题。为进一步提高分类性能,模型还整合了外部知识来源,包括情感网络图、文本长度分布、字数统计与高极性推文。在与COVID-19疫苗和印度农民抗议相关的两个基准数据集上对该模型进行评估,准确率分别达到91%和81.49%。Tefagargish等^[42]提出一种两阶段情感检测方法,第一阶段是基于句子转换器的无监督零概率学习模型,第二阶段使用集成学习方法在情感标签上训练机器学习分类器,在SemEval2017数据集上达到了87.3%的准确率。Maghsoudi等^[43]使用预训练的Transformer与Dempster-Shafer理论对300条带注释的推文进行情感分析,准确率达到84%。Kabir等^[44]使用词袋(BoW)模型、字符n-gram与词n-gram方法提取文本数据特征,使用TF-IDF与计数向量化器对特征进行向量化,在孟加拉书评数据集BANGLABOOK上的加权平均F1分数达到0.9331。Jing等^[45]提出一种用于情感分析的light-transformer模型,结合词向量表示与位置嵌入提取句子特征,与LSTM、CNN等传统方法相比,该模型分类准确率提高了0.3%~1.0%,同时大大减少了参数数量。Hu等^[46]设计了一种用于低资源非洲语言情感分析的多语言系统SACLXLMR,使用了基于词典的多语言BERT以及有监督的对抗性对比学习技术,在SemEval2023任务12上的加权F1分数达到71.0%。Karn等^[47]提出一种使用LSTM模型与连续词袋模型的情感分析方法,通过计算输入向量的最佳词大小提高分类性能,对Twitter电商评论数据集的情感分析准确性超过87%。Lyu等^[48]提出一种文本驱动方法,使用附加的用户—产品交叉上下文模块来建模用户与产品之间的关联,在IMDb、Yelp2013与Yelp-2014英语基准测试集上的准确率分别为59.0%、72.1%和72.6%。Kanwal等^[49]提出一种结合Stacked Auto-encoder与LSTM的模型,前者用于提取相关信息特征,后者基于提取的特征进行情感分析,在IMDb数据集上取得了87%的准确率。

表3为以上深度学习方法的总结。

3.3 集成学习方法

现有研究表明,通过合理集成多个模型,利用不同模

Table 3 Summary of deep learning approaches

表3 深度学习方法总结

文献	词嵌入	分类器	数据集	准确率/%
Raza 等 ^[24]	词频向量化器和 TF-IDF	MLP	COVID-19 相关推文	93.73
Harjule 等 ^[25]		LSTM	Twitter US Airline Sentiment Sentiment140	82 68.9
Hossain 和 Bhuiyan(2020) ^[26]	word2vec	CNN+LSTM	餐厅评论	75.01
Tyagi 等 ^[27]	GloVe	CNN+BiLSTM	Sentiment140	81.20
Rhanoui 等 ^[28]	doc2vec	CNN+BiLSTM	新闻文章	90.66
Jang 等 ^[29]	word2vec	Attention+CNN+BiLSTM	IMDb	90.26
Janardhan 等 ^[30]	GloVe	CNN+RNN	电影评论	84
Chowdhury 等 ^[31]	word2vec、GloVe、情感特定词嵌入	BiLSTM	Twitter US Airline Sentiment	81.2
Vimali 与 Murugan ^[32]		BiLSTM	自制数据集(来自 Amazon)	90.26
Anbukkarasi 等 ^[33]		DBLSTM	泰米尔语推文	86.2
Kumar 与 Chinnalagu ^[34]		SAB-LSTM	自制数据集	29 (POS) 50 (NEG) 21 (NEU)
Hossen 等 ^[35]		LSTM GRU	自制数据集	86 84
Younas 等 ^[36]		mBERT、 XLM-R	2018 年巴基斯坦大选相关推文	69 71
Dhola 与 Saradva ^[37]		BERT	Sentiment140	85.4
Tan 等 ^[38]		RoBERTa-LSTM	IMDb Twitter US Airline Sentiment Sentiment140	92.96 91.37 89.70
Kokab 等 ^[39]	BERT	CBRNN	Twitter US Airline Sentiment 自动驾驶汽车评论 美国总统选举评论 IMDb	97 90 96 93
AlBadani 等 ^[40]	ST-GCN	ST-GCN	SST-B IMDB	95.43 94.94
Tiwari 与 Nagpal ^[41]	BERT	KEAHT	Yelp 2014 COVID-19 疫苗相关推文 印度农民抗议相关推文	72.7 91 81.49
Tesfagerish 等 ^[42]	Zero-shot transformer	Ensemble learning	SemEval 2017	87.3
Maghsoudi 等 ^[43]	Transformer	DST	基于时间间隔分析与失眠相关的推文	84
Kabir 等 ^[44]	词频向量化器和 TF-IDF	Bangla-BERT	BANGLABOOK	93.31 (F1-score)
Jing 与 Yang(2022) ^[45]	Light-Transformer	Light-Transformer	NLPCC2014 Task2	76.40
Hu 等 ^[46]		SACLXLMR	SemEval2023	71 (F1-score)
Kam 等 ^[47]		LSTM+词袋	电商评论 IMDb	87 59.0
Lyu 等 ^[48]		文本驱动	Yelp2013 Yelp-2014	72.1 72.6
Kanwal 等 ^[49]	SAE	LSTM	IMDb	87

型的优势可进行更准确的预测。在集成学习方法中,多个模型在相同数据集上进行训练。在预测阶段,每个模型预测输入文本的情感类型,通过多数投票或加权投票决定最终预测结果。图4为集成学习方法的情感分析过程。

例如,Alrehili 等^[50]采用朴素贝叶斯、随机森林、支持向量机、bagging 与 boosting 5 种模型的集合对亚马逊客户评论数据集进行情感分析,使用词干提取、案例折叠、停止词删除等技术对数据集进行预处理。结果表明,使用 Uni-

gram 模型的多数投票集合方法达到了 89.4% 的准确率。Bian 等^[51]评估了集成逻辑回归、支持向量机、k 近邻等模型的情感分析性能,采用 TF-IDF 矢量器进行特征提取。10 倍交叉验证实验结果表明,集成模型达到了 98.99% 的准确率。Gifari 等^[52]评估了集成多项朴素贝叶斯、k 近邻、逻辑回归等机器学习方法的模型对 IMDb 数据集的情感分析性能,首先对数据文本进行标记化、停止词删除与单词提取等预处理,然后利用 TF-IDF 矢量器提取特征,获得了

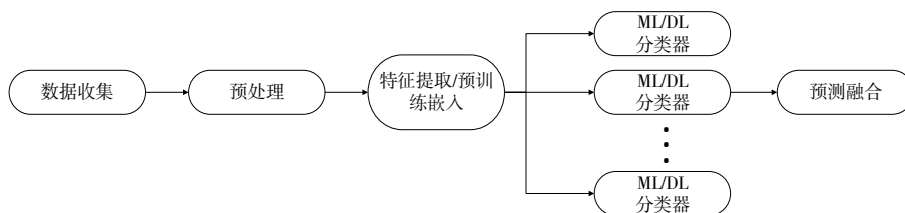


Fig. 4 Sentiment analysis process of ensemble learning method

图 4 使用集成学习方法的情感分析过程

89.40% 的准确率。Parveen 等^[53]比较了单独的伯努利朴素贝叶斯、多项朴素贝叶斯、线性支持向量分类、逻辑回归模型及其集成模型对电影评论数据集的情感分析性能,结果表明,采用多数投票方法的集成模型达到了 91% 的最高准确率。Aziz 等^[54]提出一种结合随机梯度下降、逻辑回归、朴素贝叶斯、决策树、随机森林与支持向量机的集成学习算法,分别使用简单多数投票集成与加权多数投票集成。结果表明,采用加权多数投票的集成模型在 SemEval-2017 4A、SemEval-2017 4B 与 SemEval-2017 4C 数据集上分别达到 72.95%、90.8% 与 68.89% 的准确率。Varshney 等^[55]比较了单独的朴素贝叶斯、逻辑回归、随机梯度下降与使用多数投票的集成模型的情感分析性能,均使用 TF-IDF 特征提取对数据文本进行矢量化。结果表明,集成模型在 Sentiment140 数据集上具有 80% 的正面类召回率,优于其他模型。Athar 等^[56]提出一种由逻辑回归、朴素贝叶斯、随机森林、XGBoost 与多层感知器组成的集成模型,在 IMDb 数据集上取得 89.9% 的准确率。Kamruzzaman 等^[57]比较了 6 种集成模型的情感分析性能,分别为 3 种传统集成模型:投票集成模型、bagging 集成模型、逻辑回归与随机森林的 boosting 集成模型;3 种神经网络集成模型:7 层 CNN+LSTM+注意力层、7 层 CNN+GRU、7 层 CNN+GRU+GloVe 嵌入。在在线产品评论数据集与孟加拉国达卡餐厅评论数据集上进行实验,结果表明,7 层 CNN+GRU+GloVe 嵌入模型在在线产品评论数据集上的准确率最高,为 94.19%;7 层 CNN+LSTM+注意力层模型在孟加拉国达卡餐厅评论数据集上的准确率最高,为 96.37%。Alwazrah 等^[58]评估堆叠门循环单元(SGRU)、堆叠双向门控循环单元(SBi-GRU)、来自 Transformer 的阿拉伯双向编码器表示(AraBERT)以及这些模型的集成对阿拉伯语推文数据集的情感分析性能,包括,其中 SGRU 与 SBi-GRU 模型使用预训练的阿拉伯语文本嵌入词 AraVec 来表示文本。结果表明,集成模型(SGRU+SBi-GRU+AraBERT)的准确率最高,达到 90.21%。Tan 等^[59]评估了 RoBERTa-LSTM、RoBERTa-BiLSTM 与 RoBERTa-GRU 的集成模型的情感分析性能,首先采用 RoBERTa 模型将文本序列转换为上下文嵌入,然后通过 LSTM、BiLSTM 与 GRU 模型进行编码与分类。该集成模型在 IMDb、Twitter US Airline Sentiment 与 Sentiment140 数据集上的准确率分别达到 94.9%、91.77% 与 89.81%。

集成学习是一种很有前途的情感分析方法,利用多个

模型的优势实现更好的性能,准确率为 80%~98.99%。表 4 为以上集成学习方法的总结。

3.4 情感分析方法比较

机器学习、深度学习与集成学习是情感分析领域的 3 种常用方法,具有各自的特点与适用场景。机器学习适用于小任务,多采用 TF-IDF 等基于频率的方法处理文本;深度学习在大任务上表现优越,但需要更多资源,复杂且难以解释;集成学习通过结合多个模型提高性能和鲁棒性,在处理异构数据时效果较好,但增加了计算复杂度和成本。在实践中,需要综合考虑任务的复杂性、可用的数据量、计算资源与性能需求选择合适的情感分析方法。综合比较 3 种方法的优劣势可以更好地指导实际应用中的模型选择与调优,从而适应不同的情感分析场景。表 5 为 3 种情感分析方法的优缺点总结。

4 研究现存局限性与未来展望

情感分析领域取得了重大进展,但仍然存在以下局限性:①结构不良与讽刺文本。结构不良的文本具有复杂的语法、拼写错误与不规范的语言,传统情感分析方法难以准确解释。讽刺文本往往以言外之意、反语或夸张的形式表达,需要更深层次的语境理解。传统情感分析模型难以捕捉其隐含情感,可能会导致情感误判;②粗粒度情感分析。缺乏细粒度的情感分析限制了对文本情感的深入理解,通常只能划分为几大类,如积极、消极与中性,导致对情感强度变化的不明确捕捉,忽视了同一类别下文本可能包含的细微差异,限制了对用户真实情感的准确识别。此外,过于简化的情感分类未考虑多元化的情感种类,忽略了文本语境对情感表达的影响,可能导致对情感含义的误判;③缺乏文化意识。情感分析模型在处理不同文化或地域的文本时可能因未全面考虑特定文化背景而导致情感判断误差。文化差异会影响情感表达方式,相同表达在不同文化下可能有不同情感含义。缺乏文化意识的模型可能在跨文化场景中难以理解与解释;④依赖注释数据。情感分析模型通常需要大量标注数据来训练与评估,导致其过度依赖于特定领域或标注者的主观判断,限制了在其他领域或不同群体中的泛化能力。获取大规模高质量注释数据耗费大量时间和资源,成为情感分析模型拓展到新领域或新语言的障碍;⑤词嵌入具有局限性。词嵌入在捕捉文本中词及其含义之间的复杂关系方面具有局限性,可

Table 4 Summary of ensemble learning approaches

表4 集成学习方法总结

文献	特征提取	分类器	数据集	准确率/%
Alrehili 与 Albalawi ^[50]		NB+SVM+RF+ Bagging+ Boosting	自制数据集(来自 Amazon)	89.4
Bian 等 ^[51]	TF-IDF	LR + SVM + KNN	COVID-19相关推文	98.99
Gifari 与 Lhaksmana ^[52]	TF-IDF	MNB + KNN + LR	IMDb	89.40
Parveen 等 ^[53]		MNB + BNB + LR + LSVM + NSVM	电影评论	91
Aziz 与 Dimililer(2020) ^[54]	TF-IDF	NB + LR + SGD + RF + DT + SVM	SemEval-2017 4A SemEval-2017 4B SemEval-2017 4C	72.95 90.8 68.89
Varshney 等 ^[55]	TF-IDF	LR + NB + SGD	Sentiment140	80
Athar 等 ^[56]	TF-IDF	LR + NB + XGBoost + RF +MLP	IMDb	89.9
Kamruzzaman 等 ^[57]	GloVe Attention embedding	7-Layer CNN + GRU +GloVe 7-Layer CNN + LSTM + Attention Layer	在线产品评论 孟加拉国达卡餐厅评论	94.19 96.37
Alwazrah 与 Alhumoud ^[58]	AraVec	SGRU + SBi-GRU + AraBERT	阿拉伯语推文	90.21 94.9
Tan 等 ^[59]		RoBERTa-LSTM + RoBERTa-BiLSTM + RoBERTa-GRU	IMDb Twitter US Airline Sentiment Sentiment140	91.77 89.81

Table 5 Summary of advantages and disadvantages of sentiment analysis methods

表5 情感分析方法优缺点总结

方法	优点	缺点
机器学习	1.相对简单,易于理解和实现; 2.在小规模情感分析任务中表现通常令人满意; 3.不需要大量数据进行训练	1.在复杂任务或大规模数据集上的性能可能有限; 2.对于非线性关系的建模能力相对较弱
深度学习	1.能够自动学习高级特征表示,无需手工设计特征; 2.在大规模数据和复杂任务上表现优越; 3.对于处理上下文信息和长距离依赖关系更有效	1.对于小规模数据容易过拟合; 2.训练需要更多的计算资源和时间; 3.模型结构复杂,难以解释
集成学习	1.结合多个模型进行预测,提高了整体性能和鲁棒性; 2.在处理不同类型模型时表现良好,适用于异构数据集; 3.对于数据噪声和不确定性有较好的处理能力	1.增加了模型的复杂性和计算成本 2.对于同质数据集和简单任务,可能并不总是能提升性能

能会导致模型不能准确表示文本中表达的情感。

为了应对这些挑战,未来研究可集中在以下几个方面:①细粒度情感分析。细粒度情感分析旨在更精细地划分文本情感,考虑情感的强度、种类与更具体的表达,通过深度学习架构,使模型更准确地理解用户多样化情感表达、产品评价或社交媒体评论,提高真实世界的应用效果;②情感量化。情感量化涉及基于主题的极性分布计算,通过使用主题建模技术识别文本中的底层主题,计算每个主题的情感分布可以开发更先进的模型,准确捕捉不同主题之间的情感分布;③处理歧义与讽刺文本。情感分析模型需具备上下文感知能力,以便更准确地理解歧义性语句,尤其在处理讽刺文本时能识别表面意思与实际意图的反差。通过深入挖掘多种语言特征和上下文信息,模型能更全面地理解复杂语义,适应歧义性与讽刺性文本。此外,还可使用强化学习技术训练模型处理模棱两可和讽刺文本,使其能从反馈中学习并调整预测结果;④跨语言情感

分析。当前情感分析模型主要采用英语文本进行训练,但对跨语言模型的需求日益增长。为此,研究人员可创建多语言情感分析数据集,并发展新方法捕捉不同语言之间的语义相似性,探索使用迁移学习技术在大型多语言语料库上预训练模型,再对特定语言的情感分析任务进行微调。大语言模型在情感分析领域展现出显著潜力,提供了高效的语言理解与表示学习工具;⑤社交媒体中的情感分析。社交媒体平台每天都会产生大量数据,人工处理这些数据非常困难。研究人员可以探索在社交媒体文本上训练特定领域嵌入的使用,以提高情感分析模型的准确性。还可以通过整合上下文信息和利用用户交互开发处理嘈杂或简短社交媒体文本的模型。

5 结语

文本情感分析作为自然语言处理领域的一个关键分

支,不仅深刻影响着信息检索、社交媒体分析、消费者行为研究等多个领域,而且为人类情感动态的理解提供了强有力的工具。随着深度学习、机器学习技术的不断进步,以及大数据资源的日益丰富,文本情感分析技术能够更精准、高效地识别并解析复杂多样的情感表达。虽然还面临低资源语言或特定领域数据集稀缺、微妙复杂情感状态难以区分等挑战,但相信随着持续的技术创新和理论研究,未来的情感分析模型将更加智能、精准。

参考文献:

- [1] LIGHTHART A, CATAL C, TEKINERDOGAN B. Systematic reviews in sentiment analysis: a tertiary study [J]. *Artificial Intelligence Review*, 2021(54):1-57.
- [2] DANG N C, MORENO G M N, DE L P F. Sentiment analysis based on deep learning: a comparative study[J]. *Electronics*, 2020,9(3):483.
- [3] CHAKRISWARAN P, VINCENT D R, SRINIVASAN K, et al. Emotion AI-driven sentiment analysis: a survey, future research directions, and open issues[J]. *Applied Sciences*, 2019,9(24):5462.
- [4] MAAS A, DALY R E, PHAM P T, et al. Learning word vectors for sentiment analysis[C]//*Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 2011: 142-150.
- [5] GO A, BHAYANI R, HUANG L. Twitter sentiment classification using distant supervision[J]. *CS224N Project Report*, 2009,1(12):2009.
- [6] ROSENTHAL S, FARRA N, NAKOV P. SemEval-2017 task 4: sentiment analysis in Twitter[DB/OL]. <https://arxiv.org/pdf/1912.00741>.
- [7] SOCHER R, PERELYGIN A, WU J, et al. Recursive deep models for semantic compositionality over a sentiment treebank [C]//*Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, 2013:1631-1642.
- [8] ATHINDRAN N S, MANIKANDARAJ S, KAMALESHWAR R. Comparative analysis of customer sentiments on competing brands using hybrid model approach [C]//*2018 3rd International Conference on Inventive Computation Technologies*, 2018:348-353.
- [9] VANAJA S, BELWAL M. Aspect-level sentiment analysis on e-commerce data [C]//*2018 International Conference on Inventive Research in Computing Applications*, 2018:1275-1279.
- [10] IQBAL N, CHOWDHURY A M, AHSAN T. Enhancing the performance of sentiment analysis by using different feature combinations [C]//*2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering*, 2018:1-4.
- [11] RATHI M, MALIK A, VARSHNEY D, et al. Sentiment analysis of tweets using machine learning approach [C]//*2018 Eleventh International Conference on Contemporary Computing*, 2018:1-3.
- [12] TARIYAL A, GOYAL S, TANTUBUBAY N. Sentiment analysis of tweets using various machine learning techniques [C]//*2018 International Conference on Advanced Computation and Telecommunication*, 2018:1-5.
- [13] RAHAT A M, KAHIR A, MASUM A K M. Comparison of naive Bayes and SVM algorithm based on sentiment analysis using review dataset [C]//*2019 8th International Conference System Modeling and Advancement in Research Trends*, 2019:266-270.
- [14] MAKHMUDAH U, BUKHORI S, PUTRA J A, et al. Sentiment analysis of Indonesian homosexual tweets using support vector machine method [C]//*2019 International Conference on Computer Science, Information Technology, and Electrical Engineering*, 2019:183-186.
- [15] WONGKAR M, ANGDRESEY A. Sentiment analysis using naive bayes algorithm of the data crawler: Twitter [C]//*2019 Fourth International Conference on Informatics and Computing*, 2019:1-5.
- [16] MADHURI D K. A machine learning based framework for sentiment classification: Indian railways case study [J]. *International Journal of Innovative Technology and Exploring Engineering*, 2019,8(4):441-445.
- [17] GUPTA A, SINGH A, PANDITA I, et al. Sentiment analysis of Twitter posts using machine learning algorithms [C]//*2019 6th International Conference on Computing for Sustainable Global Development*, 2019: 980-983.
- [18] PRABHAKAR E, SANTHOSH M, KRISHNAN A H, et al. Sentiment analysis of US airline Twitter data using new adaboost approach [J]. *International Journal of Engineering Research & Technology*, 2019,7(1): 1-6.
- [19] HOURLANE O, IDRISSE N. Sentiment classification on movie reviews and Twitter: an experimental study of supervised learning models [C]//*2019 1st International Conference on Smart Systems and Data Science*, 2019:1-6.
- [20] AISALMAN H. An improved approach for sentiment analysis of Arabic tweets in Twitter social media [C]//*2020 3rd International Conference on Computer Applications & Information Security*, 2020:1-4.
- [21] SAAD A I. Opinion mining on US airline Twitter data using machine learning techniques [C]//*2020 16th International Computer Engineering Conference*, 2020:59-63.
- [22] ALZYOUT M, BASHABSHEH E A L, NAJADAT H, et al. Sentiment analysis of Arabic tweets about violence against women using machine learning [C]//*2021 12th International Conference on Information and Communication Systems*, 2021:171-176.
- [23] GUPTA P, KUMAR S, SUMAN R R, et al. Sentiment analysis of lockdown in india during COVID-19: a case study on twitter [J]. *Transactions on Computational Social Systems*, 2020,8(4):992-1002.
- [24] RAZA G M, BUTT Z S, LATIF S, et al. Sentiment analysis on COVID tweets: an experimental analysis on the impact of count vectorizer and TF-IDF on sentiment predictions using deep learning models [C]//*2021 International Conference on Digital Futures and Transformative Technologies*, 2021:1-6.
- [25] HARJULE P, GURJAR A, SETH H, et al. Text classification on Twitter data [C]//*2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things*, 2020: 160-164.
- [26] HOSSAIN N, BHUIYAN M R, TUMPA Z N, et al. Sentiment analysis of restaurant reviews using combined CNN-LSTM [C]//*2020 11th International Conference on Computing, Communication and Networking Technologies*, 2020:1-5.
- [27] TYAGI V, KUMAR A, DAS S. Sentiment analysis on twitter data using deep learning approach [C]//*2020 2nd International Conference on Advances in Computing, Communication Control and Networking*, 2020: 187-190.
- [28] RHANOUI M, MIKRAM M, YOUSFI S, et al. A CNN-BiLSTM model for document-level sentiment analysis [J]. *Machine Learning and Knowledge Extraction*, 2019,1(3):832-847.
- [29] JANG B, KIM M, HARERIMANA G, et al. Bi-LSTM model to increase accuracy in text classification: combining word2vec CNN and attention mechanism [J]. *Applied Sciences*, 2020,10(17):5841.
- [30] JANARDHANA D R, VIJAY C P, SWAMY G B J, et al. Feature enhancement based text sentiment classification using deep learning model [C]//

- 2020 5th International Conference on Computing, Communication and Security, 2020: 1–6.
- [31] CHOWDHURY S, RAHMAN M L, ALI S N, et al. A RNN based parallel deep learning framework for detecting sentiment polarity from Twitter derived textual data [C]//2020 11th International Conference on Electrical and Computer Engineering, 2020: 9–12.
- [32] VIMALI J S, MURUGAN S. A text based sentiment analysis model using bi-directional lstm networks [C]//2021 6th International Conference on Communication and Electronics Systems, 2021: 1652–1658.
- [33] ANBUKKARASI S, VARADHAGANAPATHY S. Analyzing sentiment in Tamil tweets using deep neural network [C]//2020 Fourth International Conference on Computing Methodologies and Communication, 2020: 449–453.
- [34] KUMAR D A, CHINNALAGU A. Sentiment and emotion in social media COVID-19 conversations: SAB-LSTM approach [C]//2020 9th International Conference System Modeling and Advancement in Research Trends, 2020: 463–467.
- [35] HOSSEN M S, JONY A H, TABASSUM T, et al. Hotel review analysis for the prediction of business using deep learning approach [C]//2021 International Conference on Artificial Intelligence and Smart Systems, 2021: 1489–1494.
- [36] YOUNAS A, NASIM R, ALI S, et al. Sentiment analysis of code-mixed Roman Urdu-English social media text using deep learning approaches [C]//2020 IEEE 23rd International Conference on Computational Science and Engineering, 2020: 66–71.
- [37] DHOLA K, SARADVA M. A comparative evaluation of traditional machine learning and deep learning classification techniques for sentiment analysis [C]//2021 11th International Conference on Cloud Computing, Data Science & Engineering, 2021: 932–936.
- [38] TAN K L, LEE C P, ANBANANTHEN K S M, et al. RoBERTa-LSTM: a hybrid model for sentiment analysis with transformer and recurrent neural network [J]. IEEE Access, 2022, 10: 21517–21525.
- [39] KOKAB S T, ASGHAR S, NAZ S. Transformer-based deep learning models for the sentiment analysis of social media data [J]. Array, 2022, 14: 100157.
- [40] AIBADANI B, SHI R, DONG J, et al. Transformer-based graph convolutional network for sentiment analysis [J]. Applied Sciences, 2022, 12(3): 1316.
- [41] TIWARI D, NAGPAL B. KEAHT: a knowledge-enriched attention-based hybrid transformer model for social sentiment analysis [J]. New Generation Computing, 2022, 40(4): 1165–1202.
- [42] TESFAGERGISH S G, KAPOCIUTE D J, DAMASEVICIUS R. Zero-shot emotion detection for semi-supervised sentiment analysis using sentence transformers and ensemble learning [J]. Applied Sciences, 2022, 12(17): 8662.
- [43] MAGHSOUDI A, NOWAKOWSKI S, AGRAWAL R, et al. Sentiment analysis of insomnia-related tweets via a combination of transformers using dempster-shafer theory: pre - and peri - COVID-19 pandemic retrospective study [J]. Journal of Medical Internet Research, 2022, 24(12): e41517.
- [44] KABIR M, MAHFUZ O B, RAIYAN S R, et al. Banglabook: a large-scale bangla dataset for sentiment analysis from book reviews [DB/OL]. <https://arxiv.org/pdf/2305.06595>.
- [45] JING H, YANG C. Chinese text sentiment analysis based on transformer model [C]//2022 3rd International Conference on Electronic Communication and Artificial Intelligence, 2022: 185–189.
- [46] HU D, WEI L, LIU Y, et al. UCAS-IIE-NLP at SemEval-2023 task 12: enhancing generalization of multilingual BERT for low-resource sentiment analysis [DB/OL]. <https://arxiv.org/pdf/2306.01093>.
- [47] KARN A L, KARNA R K, KONDAMUDI B R, et al. Customer centric hybrid recommendation system for E-Commerce applications by integrating hybrid sentiment analysis [J]. Electronic Commerce Research, 2023, 23(1): 279–314.
- [48] LYU C, YANG L, ZHANG Y, et al. Exploiting rich textual user-product context for improving personalized sentiment analysis [DB/OL]. <https://doras.dcu.ie/29140/1/2023.findings-acl.92.pdf>.
- [49] KANWAL I, WAHID F, ALI S, et al. Sentiment analysis using hybrid model of stacked auto-encoder based feature extraction and long short term memory based classification approach [J]. IEEE Access, 2023(11): 124181–124197.
- [50] ALREHILI A, ALBALAWI K. Sentiment analysis of customer reviews using ensemble method [C]//2019 International Conference on Computer and Information Sciences, 2019: 1–6.
- [51] BIAN W S, WANG C Z, YE Z W, et al. Emotional text analysis based on ensemble learning of three different classification algorithms [C]//2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 2019: 938–941.
- [52] GIFARI M K, LHAKSMANA K M, DWIFEBRI P M. Sentiment analysis on movie review using ensemble stacking model [C]//2021 International Conference Advancement in Data Science, E-learning and Information Systems, 2021: 1–5.
- [53] PARVEEN R, SHRIVASTAVA N, TRIPATHI P. Sentiment classification of movie reviews by supervised machine learning approaches using ensemble learning & voted algorithm [C]//2nd International Conference on Data, Engineering and Applications, 2020: 1–6.
- [54] AZIZ R H H, DIMILILER N. Twitter sentiment analysis using an ensemble weighted majority vote classifier [C]//2020 International Conference on Advanced Science and Engineering, 2020: 103–109.
- [55] VARSHNEY C J, SHARMA A, YADAV D P. Sentiment analysis using ensemble classification technique [C]//IEEE Students Conference on Engineering & Systems, 2020: 1–6.
- [56] ATHAR A, ALI S, SHEERAZ M M, et al. Sentimental analysis of movie reviews using soft voting ensemble-based machine learning [C]//Eighth International Conference on Social Network Analysis, Management and Security, 2021: 1–5.
- [57] KAMRUZZAMAN M, HOSSAIN M, IMRAN M R I, et al. A comparative analysis of sentiment classification based on deep and traditional ensemble machine learning models [C]//International Conference on Science & Contemporary Technologies, 2021: 1–5.
- [58] ALWAZRAH A, ALHUMOUD S. Sentiment analysis using stacked gated recurrent unit for Arabic tweets [J]. IEEE Access, 2021, 9: 137176–137187.
- [59] TAN K L, LEE C P, LIM K M, et al. Sentiment analysis with ensemble hybrid deep learning model [J]. IEEE Access, 2022, 10: 103694–103704.

(责任编辑:尹晨茹,毛宛婷)