

考虑心理距离的高校网络舆情情感分析研究

智路平, 赵雅欣, 赵敬华

(上海理工大学 管理学院, 上海 200093)

摘要: 为探究高校网络舆情的情感演化规律, 在传统情感分析方法基础上, 融合心理距离因素, 增强情感演化分析效果。利用潜在狄利克雷分布(LDA)主题模型进行主题识别, 提出基于预训练模型 ALBERT 的融合注意力机制的双通道特征情感分析模型(DCMAM), 对评论文本进行情感分类。结合主题识别与情感分类, 构建时间关联度、空间关联度和语义关联度 3 个评价指标, 分别从时间距离、空间距离和认知距离 3 个维度对高校网络舆情情感演化进行分析。结果表明, 多组情感分析模型对比实验中, DCMAM 模型的准确率、召回率和 F1 值分别达到 91.32%、91.55% 和 91.60%, 能够较好地识别评论文本的情感倾向。心理距离评价指标显示网络用户在不同心理距离下的舆情关注度和情感演化存在差异。

关键词: 高校网络舆情; 心理距离; LDA 主题模型; 主题识别; 情感分析

中图分类号: TP 391 **文献标志码:** A

Sentiment analysis of online public opinion in universities considering psychological distance

ZHI Luping, ZHAO Yaxin, ZHAO Jinghua

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

Abstract: In order to explore the emotional evolution law of online public opinion in universities, the traditional sentiment analysis method was enhanced by integrating the psychological distance factor to improve the effect of emotional evolution analysis. The latent Dirichlet allocation (LDA) topic model was used for topic recognition, and a dual-channel feature sentiment analysis model incorporating attention mechanism (DCMAM) based on the pre-trained model ALBERT was proposed to classify the sentiment of comment texts. By combining with topic recognition and sentiment classification, three evaluation indexes of temporal relevance, spatial relevance and semantic relevance were constructed,

收稿日期: 2024-05-12

基金项目: 国家自然科学基金青年基金资助项目(72201173); 上海市哲学社会科学规划资助项目(2024BGL001)

第一作者: 智路平(1982—), 男, 高级实验师。研究方向: 舆情情感分析。E-mail: zhi19821027@163.com

通信作者: 赵敬华(1984—), 女, 副教授。研究方向: 管理决策分析。E-mail: zhaojinhua@usst.edu.cn

引文格式: 智路平, 赵雅欣, 赵敬华. 考虑心理距离的高校网络舆情情感分析研究[J]. 上海理工大学学报, 2025, 47(3): 334-344.

Citation: ZHI Luping, ZHAO Yaxin, ZHAO Jinghua. Sentiment analysis of online public opinion in universities considering psychological distance[J]. Journal of University of Shanghai for Science and Technology, 2025, 47(3): 334-344.

and the evolution of online public opinion sentiment in universities was analyzed from three dimensions: temporal distance, spatial distance and cognitive distance. The results show that the accuracy, recall and F1 value of the DCMAM model reach 91.32%, 91.55% and 91.60%, respectively, in the comparative experiments of multiple groups of sentiment analysis models, which can better identify the sentiment tendency of the comment text. The psychological distance evaluation index shows that there are differences in the public opinion attention and emotional evolution of network users under different psychological distances.

Keywords: *online public opinion in universities; psychological distance; LDA topic model; topic recognition; sentiment analysis*

中国互联网信息中心发布的《中国青少年上网行为调查报告》显示，截至 2023 年，青少年网民规模达 2.56 亿人，在整体网民中占 41.5%，占青少年总体的 71.8%。当代大学生展现出较高的社会活跃度，在参与网络话题讨论方面表现出显著的主动性^[1]。社交媒体平台具有开放性、互动性和实时性等特点，短时间内一定数量的网络用户形成一致性看法可能会造成网络舆情的爆发^[2]。已有研究表明，相比一般网络舆情，高校网络舆情具有内容燃点低、爆发速度快、传播范围广等显著特点。各种文化、思想、观点的交融和碰撞使“真相”与“谣言”在时空发酵中相互交织，干扰大学生的思想认知并影响他们的心理健康，给高校网络舆情治理工作带来了新的挑战^[3]。

近年，“高校录取极短确认时间事件”、“天津助学金事件”等突发事件使高校频繁处于舆论的风口浪尖。然而，现实中有不少高校在应对突发网络舆情时，仍采取“亡羊补牢”式的传统治理手段，导致网络舆情热度持续攀升。大学生作为最直接的舆情受众，他们的情感变化与心理学因素密切相关，在接受舆情信息后产生的复杂心理活动可能会引发情感共鸣或导致排斥心理，进而产生喜悦、惊讶、愤怒、悲伤等情感^[4]。因此，在网络空间科学治理工作背景下，重点关注高校学生心理变化，深入分析高校网络舆情事件的发展规律至关重要。

为此，本文从心理学角度出发，以心理距离理论为核心，从时间距离、空间距离和认知距离 3 个维度对高校网络舆情事件进行情感剖析。利用潜在狄利克雷分布（latent Dirichlet allocation, LDA）主题模型识别舆情主题，构建融合注意力机制的双通道特征情感分析模型（dual-channel feature sentiment analysis model incorporating attention

mechanism, DCMAM），结合舆情主题和情感分析，深入探究不同心理距离与网络用户情感演化的关联性，并提出引导与治理策略。本文在理论层面提出时间关联度、空间关联度及语义关联度等心理距离评价指标，融合心理学因素以实现跨学科融合，同时使高校网络舆情情感分析研究的科学性得到证实。在实践层面，本文有助于高校及舆情监管部门及时把握舆情发展方向，为高校网络舆情治理提供新的研究视角。

1 相关研究

网络舆情研究已成为国内外学术界研究的热点之一，研究者们从社会学、信息学、传播学、心理学等不同学科方向对网络舆情进行了深入研究^[5]。其中，高校网络舆情因话题独特性、受众多样性和传播特殊性等特点，吸引了不少学者展开理论分析和定量研究。刘文^[6]基于符号学视角探究高校网络舆情的引爆原因及传播策略。李洋等^[7]基于社会心理学原理，提出 4 种情绪分类方式，针对如何及时疏导网络用户情绪提出建议。周欢等^[8]结合超网络理论，从社交、情感、时序等不同视角分析高校网络舆情的传播特点。当前，学术界对高校网络舆情的研究聚焦于网络传播特征与机制分析、舆情监控与引导等方面。

情感分析指通过自动分析各类评论或观点意见的内容，发现评论者对某一事件或物品的褒贬态度^[9]。目前，网络舆情情感分析方法主要有基于情感词典、基于机器学习和基于深度学习的方法。基于情感词典的方法指通过构建情感词典进行情感计算。基于机器学习的方法需先进行特征提取，然后使用支持向量机（support vector machine, SVM）^[10]、K 最近邻（K-nearest

neighbors, KNN) 算法^[11]和朴素贝叶斯 (naive Bayes, NB) 算法^[12]等机器学习算法对评论文本进行情感分类。基于深度学习的方法主要使用神经网络模型提取特征并进行情感分类, 常见的神经网络模型有卷积神经网络 (convolutional neural network, CNN)^[13]、循环神经网络 (recurrent neural network, RNN)^[14]和长短期记忆 (long short-term memory, LSTM) 网络^[15]等。2018 年, 谷歌提出预训练语言模型 BERT (bidirectional encoder representations from transformers), 在多模态情感分析研究中证实 BERT 模型较传统的深度神经网络模型表现更优异。Liu 等^[16]提出一种基于 BERT 模型的跨领域文本情感分析方法, 实现各领域文本的细粒度情感分析。Lin 等^[17]基于 Transformer 变换器的双向编码器, 利用集成学习方法识别虚假新闻并取得了较好的结果。为进一步提升模型的有效性与准确性, 研究者在 BERT 模型基础上又提出了 ALBERT (a lite BERT)^[18]、RoBERTa (robustly optimized BERT pre-training approach)^[19]和 XLNet (eXtra long net)^[20]等预训练语言模型, 并将其作为情感分析任务中的词向量嵌入工具。

心理距离一词始于 20 世纪初对美学的研究^[21]。解释水平理论认为心理距离与人们对某一事件所产生的感知风险与不确定性显著相关^[22]。建构水平理论认为心理距离是以自我为参照点, 外界客体以不同的方式远离参照点所形成的抽象距离^[23]。网络舆情研究中, 网络用户的情感表达受到舆情发生时间、地理位置和社会关系等外界客体的影响, 导致自身与舆情信息之间产生抽象距离, 符合建构水平理论所定义的心理距离。张晓君等^[24]基于心理距离探究网络用户在面对网络舆情时的传播意愿。Skalicky 等^[25]通过解释水平理论, 研究不同水平的心理距离如何影响人们对普通评论和讽刺评论的传播意愿。陈百卉^[26]从心理距离和建构水平角度考察不同心理距离和建构水平对人们情感预测强度的影响。

综上, 现有高校网络舆情的相关研究涉及传播特征、情感分析及舆情引导等方向, 研究者多将心理距离应用于舆情传播及情感强度预测方面, 但综合考虑高校网络舆情特点与受众心理距离的情感分析研究成果相对较少。针对现有研究的不足, 本文综合考虑计算资源与时间因素, 选取目前效果较为出色的 ALBERT 预训练语言模型进行词向量嵌入, 构建融合双向长短期记忆网络

(bidirectional long short-term memory, BiLSTM)、文本卷积神经网络 (text convolutional neural network, TextCNN) 及注意力机制的 DCMAM 情感分析模型。结合主题识别与情感分析, 从时间距离、空间距离、社会距离 3 个心理距离维度分析高校网络舆情的情感演变规律, 为高校网络舆情治理与监控提供参考依据。

2 研究方法

主题识别部分, 采用 LDA 主题模型进行主题聚类; 情感分析部分, 提出 DCMAM 情感分析模型, 借助 ALBERT 预训练模型进行动态词嵌入, 利用融合注意力机制的 BiLSTM-TextCNN 双通道模型充分捕捉文本上下文及局部特征, 提高文本的分类性能; 心理距离部分, 选取时间距离、空间距离和认知距离 3 个维度, 构建时间关联度、空间关联度和语义关联度作为评价指标。

2.1 LDA 主题模型

LDA 主题模型可以自动分析文档并统计文档中的词语, 根据统计信息判断当前文档包含主题的个数及各个主题的占比大小^[27]。困惑度是衡量概率模型预测性能或概率分布拟合效果的重要指标, 用于判断 LDA 主题模型的优劣, 同时调整文档的主题个数。因此, 本文首先计算出不同主题数量的困惑度数值, 然后根据困惑度曲线确定网络舆情事件的最优主题聚类数量。

2.2 DCMAM 情感分析模型

文本情感分析是指对带有感情色彩的主观性文本进行分析, 为了准确识别微博评论的情感倾向, 构建 DCMAM 模型对网络舆情相关微博评论进行情感判断。DCMAM 模型由输入层、词向量嵌入层、特征提取层、特征融合层和输出层 5 个部分构成, 结构如图 1 所示。图中, e_i 表示输入的第 i 个单词向量, $i = 1, 2, \dots, n$ 。

a. ALBERT 词嵌入模型

ALBERT 模型采用跨层参数共享的方法, 使每一层 Transformer 变换器的权重相同。ALBERT 模型采用了类循环结构, 主体结构部分只包含一层 Transformer 实体, 模型使用句子顺序预测取代 BERT 模型中的下一个句子预测, 能够学习到更加细微的语义差别和句子连贯性, ALBERT 模型结构如图 2 所示。图中: E_i 表示第 i 个输入单词; T_i 表示第 i 个输入单词对应的输出向量; Trm 表示

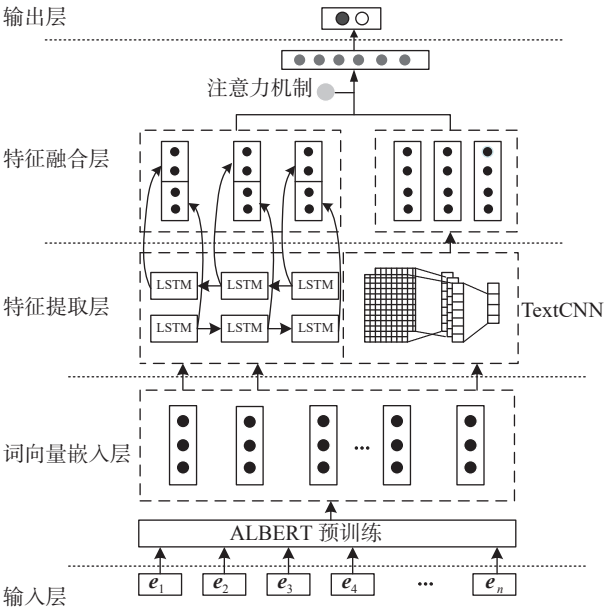


图1 DCMAM 情感分析框架

Fig.1 Sentiment analysis framework of DCMAM

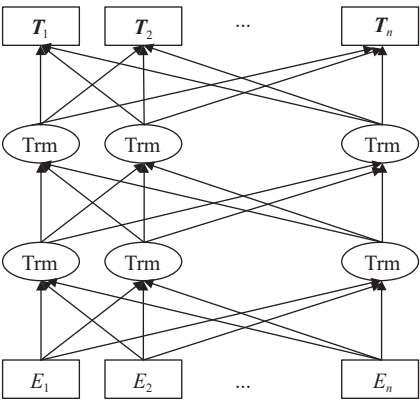


图2 ALBERT 模型结构图

Fig.2 Structure diagram of ALBERT model

经过 Transformer 变换。

b. BiLSTM 模型

BiLSTM 模型是一种基于 LSTM 模型的记忆网络，其思想是将同一个序列分别接入前向和后向两个循环神经网络中，然后将两者的输出进行拼接，使得输出的特征同时包含前后语义信息，BiLSTM 模型结构如图 3 所示。图中： x_t 、 x'_t 分别表示 t 时刻的输入、输出向量； h_t 表示 t 时刻隐藏层的输出向量。

LSTM 神经网络模型由遗忘门、输入门、输出门和记忆单元组成，相关计算公式如下：

$$f_t = \sigma(W_{f,xh}x_t + W_{f,hh}h_{t-1} + b_f) \quad (1)$$

$$i_t = \sigma(W_{i,xh}x_t + W_{i,hh}h_{t-1} + b_i) \quad (2)$$

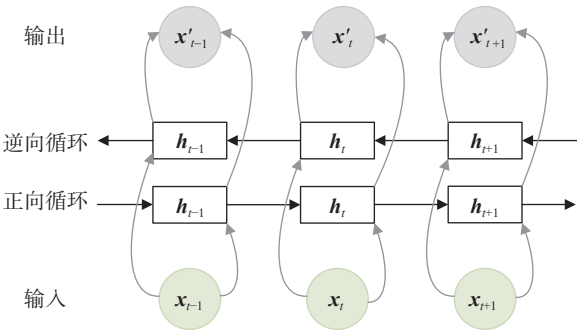


图3 BiLSTM 模型结构图

Fig.3 Structure diagram of BiLSTM model

$$o_t = \sigma(W_{o,xh}x_t + W_{o,hh}h_{t-1} + b_o) \quad (3)$$

$$u_t = \tanh(W_{u,xh}x_t + W_{u,hh}h_{t-1} + b_u) \quad (4)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot u_t \quad (5)$$

$$h_t = o_t \odot \tanh u_t \quad (6)$$

式中： σ 表示 sigmoid 激活函数； f_t 、 i_t 、 o_t 、 u_t 分别表示第 t 时刻的遗忘门、输入门、输出门、记忆单元； $W_{f,xh}$ 、 $W_{i,xh}$ 、 $W_{o,xh}$ 、 $W_{u,xh}$ 分别表示遗忘门、输入门、输出门和记忆单元对 x_t 的权重系数矩阵； $W_{f,hh}$ 、 $W_{i,hh}$ 、 $W_{o,hh}$ 、 $W_{u,hh}$ 分别表示遗忘门、输入门、输出门和记忆单元对 h_{t-1} 的权重系数矩阵； b_f 、 b_i 、 b_o 、 b_u 分别表示遗忘门、输入门、输出门和记忆单元的偏置向量。

BiLSTM 模型中，正向循环和反向循环相互独立，前向隐藏层向量 $\vec{h}_t$ 和后向隐藏层向量 $\overleftarrow{h}_t$ 经过向量拼接后得到最终的隐藏层向量 h_t ，更新过程如下：

$$\vec{h}_t = L_f(x_t, \vec{h}_{t-1}) \quad (7)$$

$$\overleftarrow{h}_t = L_b(x_t, \overleftarrow{h}_{t-1}) \quad (8)$$

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

式中： $\vec{h}_t$ 和 $\overleftarrow{h}_t$ 分别为双向 LSTM 模型在 t 时刻的前向和后向隐藏层向量； L_f 和 L_b 分别代表前向循环和后向循环。

c. TextCNN 模型

TextCNN 模型结构如图 4 所示，由输入层、卷积层、池化层和全连接层 4 个部分组成。卷积过程通常为使用尺寸不同的滤波器提取不同范围的特征。将卷积层提取的特征进行最大池化操作以保留主要特征，去除冗余信息。池化操作之后，特征被连接起来并进入全连接层，最终被映射到情感标签，并由输出层输出结果。

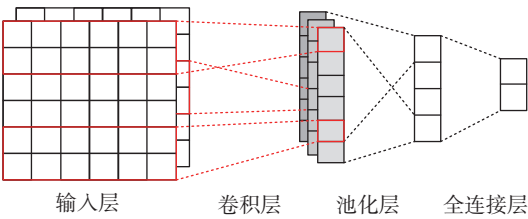


图 4 TextCNN 模型结构图

Fig.4 Structure diagram of TextCNN model

d. 特征融合层与输出层

首先，将 BiLSTM 模型与 TextCNN 模型提取的特征一起输入特征融合层进行拼接；然后，将连接后的特征向量输入到全连接层，采用注意力机制对融合的特征向量进行筛选；最后，使用 Softmax 分类器计算评论文本中包含每个情感类别的概率。计算规则如下：

$$C(f_L, f_C) = \tanh(W_1 f_L + W_2 f_C) \tag{10}$$

$$P(y|V, W_C, b_C) = \text{Softmax}(W_C C + b_C) \tag{11}$$

式中： W_1 和 W_2 为注意力机制的权重矩阵； f_L 为 BiLSTM 模型的输出特征向量； f_C 为 TextCNN 模型的输出特征向量； C 为融合 BiLSTM 模型与 TextCNN 模型的特征向量； P 表示求取积极、中立及消极情感类别的概率； y 表示不同情感类别变量； V 为所有情感类别构成的矩阵； W_C 和 b_C 为权重矩阵和偏置矩阵。

2.3 心理距离度量指标选取

Liberman 在 1998 年首次提出时间解释理论，他认为时间是一种心理距离，对个体的建构水平会产生作用，进而影响个体的认知和行为^[26]。后续研究中，心理学家陆续发现空间距离、社会距离和可能性这 3 个维度与时间距离会产生类似的效应，于是将 4 个维度进行统一，形成心理距离框架。

网络舆情研究中：时间距离为网络用户接收舆情信息的时间与舆情首次报道时间之间的距离；空间距离为网络用户所处的地理位置与舆情发生地之间的地理距离；社会距离是人们在社会结构中对其他人的社会位置和关系的感知程度^[28]；可能性指个体对某种情境、事件或行为的期望程度^[29]。心理学家认为社会距离和可能性与网络用户的认知水平有关，即受众对客观舆情信息的价值倾向评价之间的差距，可具体映射到是否产生评论、点赞、转发等行为倾向上。由认知水平所产生的个体之间在思想、经验、价值观等方面接

近或远离的程度称为认知距离。因此，本研究选择时间距离、空间距离和认知距离 3 个维度对高校网络舆情进行情感分析。

当前的网络舆情研究对心理距离度量方法的研究较少，尚未形成统一的度量标准。本文结合网络舆情特征及心理距离概念，选取心理距离各维度的度量指标，如表 1 所示。其中：时间距离度量指标为时间关联度 G ，反映不同时间节点下网络用户情感的偏移程度；空间距离的度量指标为空间关联度 Q ，反映某一时间节点下不同区域的舆情情感是否发生了偏移；认知距离的度量指标为语义关联度 R ，反映同一时间节点下不同舆情主题的情感类别波动方向及大小。

表 1 心理距离度量指标

Tab.1 Psychological distance metrics

度量维度	概念	度量指标
时间距离	评论时间与舆情最早报道时间的时间间隔	时间关联度 G
空间距离	网络用户的IP位置与舆情发生地的地理间隔	空间关联度 Q
认知距离	网络用户对舆情语义内容的认知理解差异程度	语义关联度 R

3 实验与分析

本文选取“鼠头鸭脖”数据集评估 DCMAM 情感分析模型的有效性，并与传统情感分析模型进行对比，验证 DCMAM 模型的情感分类效果。

3.1 数据采集与预处理

以“鼠头鸭脖”事件为例，通过计算机智能技术获取微博平台 2023 年 6 月 6 日至 2023 年 7 月 7 日时间段内的原创博文及评论数据，初步获取数据 18510 条。为增强数据的准确性与可靠性，对原始数据进行去重、Jieba 分词、去停用词等预处理操作，最终获得 12110 条有效数据用于构建实验语料库。

3.2 舆情生命周期划分

“鼠头鸭脖”整个事件时间跨度大，衍生话题多，为确保舆情事件生命周期的正确划分，本文使用百度热度指数辅助验证。以“鼠头鸭脖”为关键词搜索得到每日百度热度指数，如图 5 所示。图中显示，前期话题不断波动，中期呈现出峰值分布，后期趋于平稳状态，热度呈明显的阶段性变化。因此，本文根据生命周期理论将“鼠头鸭脖”事件分为起始阶段、爆发阶段、衰退阶

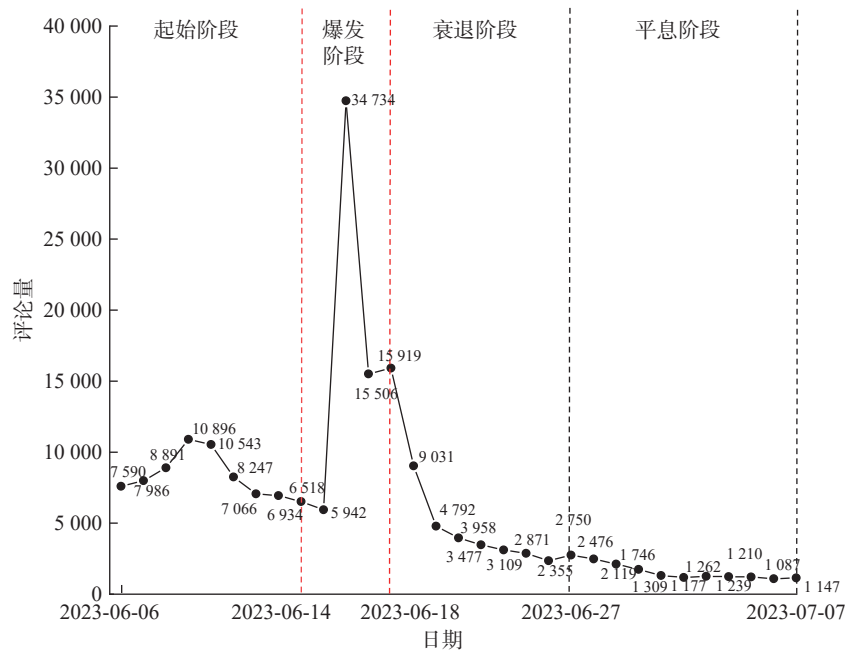


图 5 百度指数关键词搜索趋势图

Fig.5 Trend chat of Baidu index keyword search

段和平息阶段。

3.3 舆情主题识别

本文以 Sklearn 为框架，选用 Gensim 模块中的实例化 LDA 主题模型。首先，选择区间 [2, 20] 内的整数作为随机候选主题数，计算不同主题

数下的困惑度数值。综合考虑困惑度值及合理性，确定主题数为 12。其次，通过“主题-词”概率分布得出每个主题下所包含的高频词，表 2 为识别的主题内容与每个主题下出现概率最高的前 5 个关键词。

表 2 各主题关键词及所属舆情周期

Tab.2 Keywords for each topic and their corresponding public opinion cycle

主题序号	主题聚类内容	关键词	舆情周期
0	市场监督管理局确认异物为鸭脖	鼠头鸭脖 市场监督管理局 欺骗 大学生 信誓旦旦	起始阶段
1	政府公信力危机	公信力 受害者 互联网 小卖部 市场监管	爆发阶段
2	江西省教育厅介入	监管部门 安全卫生 负责人 反思 教育厅	爆发阶段
3	联合调查局公布结果	责任人 说瞎话 食品卫生 法律法规 实事求是	爆发阶段
4	绝味鸭脖股价下降	绝味鸭脖 指鹿为马 股价 指鼠为鸭 弄虚作假	起始阶段
5	江西省成立联合调查组调查	调查组 管理局 领导班子 临时工 江西人	爆发阶段
6	菁禾餐饮声明与公司无关	承包商 舆论压力 毕业证 承认错误 压不住	起始阶段
7	涉事学生澄清确实是鸭脖	老鼠头 当事人 安全事件 再次发生 避重就轻	起始阶段
8	涉事学校首次正式回应	老百姓 校领导 法定代表 学校 食品行业	起始阶段
9	食堂意见反馈群内容被曝光	学校食堂 市监局 颠倒黑白 鼠鼠我鸭 年轻人	起始阶段
10	工作人员称学生点的是猪脚饭	工作人员 猪脚饭 公共卫生 副校长 无所谓	起始阶段
11	局长“指鼠为鸭”	局长 食品安全 严重危害 沸沸扬扬 查出	起始阶段

3.4 情感倾向性分析

本文基于 Python3.10 和 Pytorch 深度学习框架构建 DCMAM 情感分析模型。DCMAM 模型使用

Adam 优化器更新网络参数以有效防止欠拟合，设置随机丢弃机制和早停机制以有效防止出现过拟合现象。为提高模型的泛化性能，采用网格搜索

技术不断调整参数，具体超参数设置如表3所示。

图6(a)为模型迭代次数不同时对模型的影响。由图可见：当迭代次数小于10时，得分呈现快速上升的趋势；当迭代次数为10时，得分达到最高；当迭代次数大于10时，得分开始缓慢下降且趋于稳定。因此，设置迭代次数为10次。

图6(b)为随机丢弃率取不同数值时对模型的影响。当随机丢弃率小于0.4时，模型性能不断提升；当随机丢弃率大于0.4时，评价指标不断下降，模型性能逐渐变低。因此，随机丢弃率设置为0.4。为探究DCMAM模型的分类效果，将DCMAM模型与BERT-TextCNN、BERT-BiLSTM、ALBERT-BiLSTM、ALBERT-TextCNN、BERT-RCNN、ALBERT-RCNN这6种模型进行对比，并

表3 DCMAM模型参数设置

Tab.3 DCMAM model parameter settings

超参数	参数值
批尺寸(batch_size)	16
迭代次数(epoch)	10
随机丢弃(dropout)	0.4
学习率(learning_rate)	1×10^{-5}
卷积核数(num_filters)	100
卷积核尺寸(filter_size)	2,3,4
权重衰退(weight_decay)	0.01
隐藏层尺寸(hidden layer size)	512

选取准确率、召回率和F1值判断模型效果，实验结果见表4。

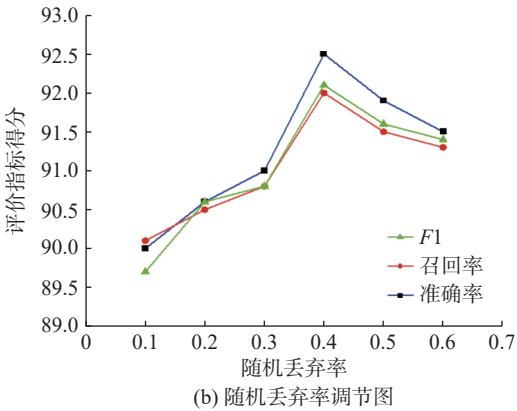
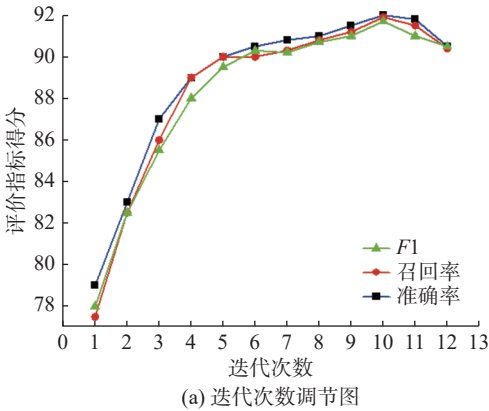


图6 DCMAM模型参数调节结果

Fig.6 Results of DCMAM model parameter adjustment

表4 实验结果对比

Tab.4 Comparison of experimental results

模型	准确率	召回率	F1值
BERT-TextCNN	90.20	90.54	90.33
BERT-BiLSTM	88.60	88.31	88.70
ALBERT-BiLSTM	88.63	88.35	88.65
ALBERT-TextCNN	90.25	90.55	90.36
BERT-RCNN	90.78	90.88	91.05
ALBERT-RCNN	90.80	90.95	90.98
DCMAM	91.32	91.55	91.60

观察实验结果，DCMAM模型的准确率、召回率和F1值分别为91.32%、91.55%、91.60%。其中，将ALBERT-TextCNN模型与BERT-TextCNN模型的实验结果进行对比，发现ALBERT-TextCNN模型在准确率、召回率和F1值上分别提

升0.05%、0.01%、0.03%。将ALBERT-RCNN模型与BERT-RCNN模型的实验结果进行对比，发现ALBERT-RCNN模型的准确率和召回率分别提升0.02%、0.07%，F1值下降0.07%。将DCMAM模型与ALBERT-RCNN模型的实验结果进行对比，发现DCMAM模型的准确率、召回率和F1值分别提升了0.52%、0.60%、0.62%。

对比实验证明DCMAM模型在情感分析方面具有一定的优势。通过该模型对评论内容的情感进行分类，将消极情感、中立情感和积极情感分别设为-1、0、1。

4 融合心理距离的情感分析

结合舆情生命周期、主题识别及情感倾向性，从时间距离、空间距离和认知距离3个维度

对“鼠头鸭脖”事件的情感演化进行分析。

4.1 基于时间距离维度的情感分析

时间关联度 G 反映了不同时间节点网络用户的情感是否发生了偏移, 其计算如下:

$$G(t, v, w) = S^k_{(t_n, v_i, w)} - S^k_{(t_0, v_i, w)} \tag{12}$$

式中: G 表示第 i 个主题的情感类别在时刻 t_n 时的情感值与在初始时刻 t_0 时的情感值之间的差值, 反映出不同情感类别受时间距离影响所产生的波动大小及方向; v_i 表示第 i 个讨论的舆情主题, $i = 0, 1, 2, \dots, 11$; w 表示网络用户 IP 显示的地理位置; $S^k_{(t_n, v_i, w)}$ 表示在时刻 t_n 、区域 w , 讨论主题 v_i 的第 k 个情感类别信息。

将各个情感类别信息输入时间关联度公式, 在保持舆情主题及地理位置不变的情况下, 计算不同时间节点之间各个情感类别的偏移程度, 部分结果如表 5 所示。在舆情起始阶段至爆发阶段, 随着时间距离增长, 消极情感占比逐渐上升, 舆情爆发至衰退阶段, 消极情感占比随时间距离的增长逐渐下降。

表 5 部分时间关联度结果

Tab.5 Results of partial temporal correlation

初始时刻 t_0	消极情感/%	中立情感/%	积极情感/%	对比时刻 t_n
2023-06-06 (8:00)	+7	-6	-1	2023-06-08 (8:00)
	+11	-7	-4	2023-06-12 (8:00)
	+12	-8	-4	2023-06-14 (8:00)
	+18	-11	-7	2023-06-18 (8:00)
	+8	-2	+6	2023-06-22 (8:00)

统计舆情生命周期各阶段不同情感类别评论占该阶段评论总量的比例, 如图 7 所示。图 7 显示, 舆情起始阶段, 消极、中立和积极情感的占比分别为 45%、35%、20%。调查发现, 该阶段产生消极情感的原因为受众对学校食堂食品卫生情况的不满及对涉事学生毕业情况的担忧。舆情爆发阶段, 时间关联度逐渐下降, 此时消极、中立和积极情感的占比分别为 70%、20%、10%。消极情感主要为对学校和基层干部的质疑、嘲讽和气愤。随着政府部门的积极回应, 舆情得到控制并开始进入衰退阶段和平息阶段, 此时网络用户的时间距离较远, 对于舆情事件的看法逐渐趋于理性。

综上, 时间距离是网络用户情感演化的重要

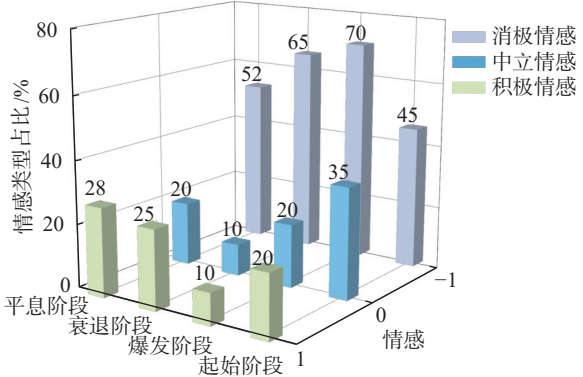


图 7 各阶段情感类型占比

Fig.7 Proportion of emotion types at each stage

影响因素, 时间距离与网络用户的情感演化并非线性相关, 而是跟随舆情生命周期呈现峰值分布。“鼠头鸭脖”舆情事件中, 网络用户的消极情感随着时间距离的增加, 在爆发期达到峰值, 直至舆情衰退阶段才开始缓慢下降。

4.2 基于空间距离维度的情感分析

空间关联度 Q 作为空间距离维度的度量指标, 反映了某一时间节点下不同区域的舆情情感是否发生了偏移, 计算如下所示:

$$Q(t, w, k) = \sum_{t-t_n}^{t-t_n+1} S^k_{(t, w_q)} - \sum_{t-t_n}^{t-t_n+1} S^k_{(t, w_m)} \tag{13}$$

式中: Q 表示在 t 时刻, 区域 w_q 相较区域 w_m 的第 k 个情感类别的波动大小及波动方向; $S^k_{(t, w_q)}$ 表示在 t 时刻、区域 w_q , 讨论主题的第 k 个情感类别信息。若 $Q > 0$, 则说明该情感类别在区域 w_q 呈上升趋势; 若 $Q < 0$, 则说明该情感类别在区域 w_q 呈下降趋势; 若 $Q = 0$, 则说明该情感类别在两个区域的变化趋势相同。

依据语义内容及用户动态筛选出学生身份的评论用户, 将他们的 IP 位置按照地理距离进行等级划分。其中: 用户 IP 位置显示在舆情发生地的省、自治区设为 3 级; 与舆情发生地相邻的省、自治区设为 2 级; 与舆情发生地不相邻的省、自治区设为 1 级。将空间等级输入空间关联度公式, 计算出不同空间等级之间的情感类别偏移大小, 部分结果如表 6 所示。

由表 6 可知, 同一时刻不同空间等级的各个情感类别占比不同, 总体 3 级区域的消极情感在所有空间等级中占比最高, 随着空间等级递减, 消极情感占比逐渐下降, 中立情感与积极情感占比逐渐上升。综上, 空间距离对网络用户的情感

表 6 部分空间关联度结果

Tab.6 Results of partial spatial relevance

时间	空间等级	消极情感/%	中立情感/%	积极情感/%
2023-06-13(10:00)	1→2	+15	-7	-8
	1→3	+18	-11	-7
2023-06-15 (10:00)	1→2	+10	-5	-5
	1→3	+15	-9	-6
2023-06-17(10:00)	1→2	+8	-4	-4
	1→3	+13	-8	-5

变化具有一定的影响作用，“鼠头鸭脖”网络舆情事件中，舆情发生地的网络用户消极情感占比比较大，消极情感随空间等级上升而逐层减弱。

4.3 基于认知距离维度的情感分析

语义关联度 R 作为认知距离维度的度量指标，反映了不同舆情主题的情感类别波动方向及大小，其计算如下所示：

$$R(t,v,k)=\sum_{i=t-t_n}^{t-t_n+1}S^k_{(t,v_i)}-\sum_{j=t-t_n}^{t-t_n+1}S^k_{(t,v_j)}\tag{14}$$

式中： R 表示在 t 时刻下，主题 v_i 与主题 v_j 在第 k 个情感类别中的波动大小及方向； $S^k_{(t,v_i)}$ 表示 t 时刻第 v_i 个舆情主题的第 k 个情感类别信息。若 $R>0$ ，则说明该情感类别在主题 v_i 中的占比呈上升趋势；若 $R<0$ ，则说明该情感类别在主题 v_i 中的占比呈下降趋势；若 $R=0$ ，则说明该情感类别在主题 v_i 与主题 v_j 中分布相同。

将各个主题及情感类别信息输入语义关联度公式，计算出各个时间节点下不同主题之间的情感类别偏移大小，按照时间结果进行排序，部分结果如表 7 所示。由表 7 可知，同一时刻不同主题之间的情感类别波动大小及方向不同，在主题 6、7、8、11 当中，主题 11 的消极情感占比最大。

为进一步探究各个主题的情感强度，本文利用评论主题、评论数量、情感强度三维数据，绘制基于认知距离的主题情感强度分布图，如图 8 所示。由图可知，网络用户由于认知水平差异，在客观舆情信息基础上衍生出不同话题，通过对语义内容进行分析发现，学生群体容易对认知距离较近的主题产生行为决策，对认知距离较远的主题产生行为决策的兴趣较低，例如对主题 0(市场监督管理局确认异物为鸭脖)、主题 7(涉事学生澄清确实是鸭脖)、主题 8(涉事学校首次正式

表 7 部分语义关联度结果

Tab.7 Results of partial semantic relevance

时间	主题	消极情感/%	中立情感/%	积极情感/%
2023-06-06(10:00)	7→6	-7	+8	+1
	8→6	-9	+6	+3
	11→6	-15	+5	+10
2023-06-07(17:00)	7→6	-8	+6	+2
	8→6	-7	+5	+2
	11→6	-12	+2	+10
2023-06-10(12:00)	7→6	-6	+4	+2
	8→6	-8	+5	+3
	11→6	-13	+9	+4

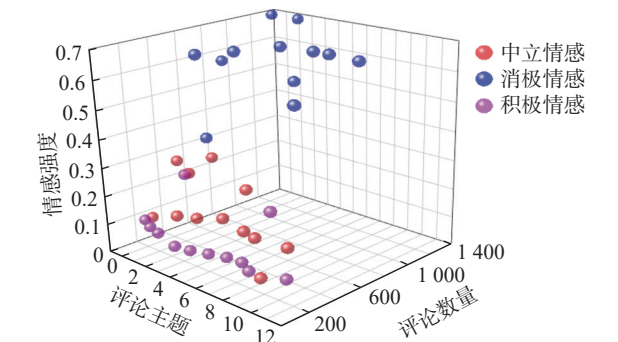


图 8 主题情感强度分布

Fig.8 Distribution of emotional intensity of topics

回应)的关注度与参与度较高，对主题 4(绝味鸭脖股价下降)、主题 10(工作人员称学生点的是猪脚饭)的关注度较低。

综上，认知距离对网络用户的行为决策及情感表达具有重要影响作用，“鼠头鸭脖”舆情事件中，学生群体更加关注与学校食堂、涉事学校、涉事学生相关的话题，社会群体则更加关注与政府官员、调查局相关的话题。各个主题的情感类别占比不同，主题 1(政府公信力危机)和主题 11(局长“指鼠为鸭”)的消极情感占比最高，高校及舆情相关部门对此类主题应及时采取有效措施进行疏导，防止产生二次舆论危机。

4.4 构建心理距离情感强度等级

综合考虑 3 个心理距离对舆情事件的情感影响，将 3 个心理距离进行分级赋值。时间距离维度方面，以舆情首次报道时间为起点，评论时间在舆情起始阶段和平息阶段则将时间距离设为 1 级，评论时间在舆情衰退阶段则将时间距离设为 2 级，评论时间在舆情爆发阶段则将时间距离设为 3 级。空间距离方面，将舆情发生地设为

3级, 与舆情发生地相邻区域设为2级, 与舆情发生地不相邻区域设为1级。认知距离方面, 与政府、公信力相关的主题设为3级, 与学校、学生相关的主题设为2级, 其他主题设为1级。将3个心理距离等级进行相加, 等级越高代表网络用户的情感越强烈, 高校及政府相关部门应及时关注并采取引导措施, 部分心理距离情感强度等级如表8所示。

表8 部分心理距离情感强度等级

Tab.8 Partial emotional intensity levels of psychological distance

时间	地区	主题	情感强度等级
2023-06-06(10:00)	江西省	7(涉事学生澄清确实是鸭脖)	6
2023-06-10(10:00)	浙江省	11(局长“指鼠为鸭”)	6
2023-06-17(10:00)	安徽省	5(江西省成立联合调查组调查)	8
2023-06-18(10:00)	广东省	3(联合调查局公布结果)	8
2023-06-25(10:00)	湖南省	2(江西省教育厅介入)	6

5 结 论

本文以高校学生持续关注的“鼠头鸭脖”事件为研究对象, 基于舆情生命周期理论和心理学因素, 利用LDA主题模型和DCMAM情感分析模型对事件各阶段评论文本进行主题识别和情感分类。结合主题与情感, 从时间距离、空间距离和认知距离对高校网络舆情进行情感演化分析并构建心理距离情感强度等级。

研究发现, 心理距离对网络用户的情感演化具有重要影响, 受众在社会从众效应的影响下在舆情初期情感表现尚不明显, 随着时间距离的延伸, 情感由爆发转至消散。空间距离越近的网络用户更容易受外界刺激而产生情感反应, 距离越远的网络用户对舆情的感知力及参与度较小, 对于舆情的传播具有压制作用。网络用户对自己认知距离近的话题更加关注与重视, 例如人们对政府及高校信任危机的关注远远超过对校园食品安全的关注。

基于此, 本文针对高校网络舆情演变周期提出相应的引导对策, 帮助高校进行有针对性的舆情引导和治理。

a. 舆情起始阶段。该阶段, 在校师生与事件的心理距离近且关注程度高, 可能会在网络上发

布带有复杂情绪的评论, 高校应密切关注在校师生情感和态度变化。以“鼠头鸭脖”事件为例, 高校可借助自然语言处理技术如BERT和DCMAM情感分析模型对获取的评论进行情感倾向性判断, 精准定位负面情绪聚集的话题。在此基础上, 依据情感判断优先响应学生关注度高、心理落差显著的议题, 建立心理距离缓冲机制, 实施分层分类的情绪疏导策略, 从而提前遏制潜在舆情风险的爆发。

b. 舆情爆发阶段。根据心理距离理论, 爆发期是舆情受众情绪最浓烈的时期, 当负面情绪占据主导地位时, 高校需进行全方位沟通: 一是确保言行一致, 避免出现前后矛盾或模糊不清的表述, 有助于提高高校的权威性和信息的可信度; 二是加强与政府、社交媒体等相关主体的合作, 通过与官方媒体的良性互动, 大力提升正面信息的传播力度, 减少谣言及虚假新闻的传播机会; 三是积极回应网络用户关注的问题, 对存在的质疑和矛盾进行归纳分析并通过官方渠道进行澄清, 针对关键问题及时作出回应。

c. 舆情衰退阶段。此阶段, 公众对舆情事件已经有了充分的了解, 在心理距离作用下情绪得到了缓和, 各种极端评论的消散有助于加速网络舆情的冷却。高校应根据舆情后果的严重性制定处置方案, 对舆情处理过程中的不当行为进行追责, 明确直接责任人和舆情处理负责人。此外, 对涉事主体进行相应的责任追究, 促使网络用户从消极情感向积极情感转变。

d. 舆情平息阶段。高校可建立舆情反馈评价机制, 对容易引发舆情危机的敏感词进行持续跟踪。此外, 做好事件的复盘工作, 及时总结经验教训, 并完善相应的监管机制, 从根本上解决问题。

本文的研究仍存在以下不足: a. 本文仅选取了微博平台的“鼠头鸭脖”舆情事件进行研究, 后续可尝试跨平台收集多类型事件数据, 尝试扩大研究范围; b. 本文在探究心理距离对舆情情感演化的影响时, 将时间距离、空间距离和认知距离作为单一变量进行研究, 后续可将心理距离各维度进行交互分析, 探究多变量因素对舆情情感演化的影响。

参考文献:

[1] 于婷. 全媒体时代高校网络舆情发展态势与应对策略

- [J]. 沧州师范学院学报, 2023, 39(4): 108–113.
- [2] 王春东, 张卉, 莫秀良, 等. 微博情感分析综述[J]. 计算机工程与科学, 2022, 44(1): 165–175.
- [3] 黄钦, 徐柳凡, 朱礼龙. 基于主题模型和情感计算的高校网络舆情演化分析[J]. 成都师范学院学报, 2023, 39(12): 39–46.
- [4] 张晓君, 黄微, 宋先智. 基于心理距离的网络舆情传播倾向性研究[J]. 情报资料工作, 2017, 38(6): 19–24.
- [5] 唐婷婷. 高校网络舆情研究综述[J]. 产业与科技论坛, 2022, 21(17): 59–60.
- [6] 刘文. 符号学视角下的高校网络舆情研究[D]. 上海: 上海理工大学, 2022.
- [7] 李洋, 李思佳, 叶琼元, 等. 面向突发事件网络舆情的社会情绪唤醒综合评价与疏导策略研究[J]. 情报资料工作, 2020, 41(6): 17–25.
- [8] 周欢, 张培颖, 王嘉仪, 等. 基于超网络的高校网络舆情态势感知及传播分析[J]. 情报理论与实践, 2023, 46(8): 131–137, 164.
- [9] 张紫琼, 叶强, 李一军. 互联网商品评论情感分析研究综述[J]. 管理科学学报, 2010, 13(6): 84–96.
- [10] 邓君, 孙绍丹, 王阮, 等. 基于 Word2Vec 和 SVM 的微博舆情情感演化分析[J]. 情报理论与实践, 2020, 43(8): 112–119.
- [11] ZHANG S C. Challenges in KNN classification[J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 34(10): 4663–4675.
- [12] LIU Y, BI J W, FAN Z P. Multi-class sentiment classification: the experimental comparisons of feature selection and machine learning algorithms[J]. Expert Systems with Applications, 2017, 80: 323–339.
- [13] ZHANG S Q, XU X Y, PANG Y W, et al. Multi-layer attention based CNN for target-dependent sentiment classification[J]. Neural Processing Letters, 2020, 51(3): 2089–2103.
- [14] BAI Q C, ZHOU J, HE L. PG-RNN: using position-gated recurrent neural networks for aspect-based sentiment classification[J]. The Journal of Supercomputing, 2022, 78(3): 4073–4094.
- [15] FENG S, WANG Y, LIU L R, et al. Attention based hierarchical LSTM network for context-aware microblog sentiment classification[J]. World Wide Web, 2019, 22(1): 59–81.
- [16] LIU N, ZHAO J H. A BERT-based aspect-level sentiment analysis algorithm for cross-domain text[J]. Computational Intelligence and Neuroscience, 2022, 2022: 8726621.
- [17] LIN S Y, KUNG Y C, LEU F Y. Predictive intelligence in harmful news identification by BERT-based ensemble learning model with text sentiment analysis[J]. Information Processing & Management, 2022, 59(2): 102872.
- [18] CHEN Y Z, KONG L Y, WANG Y, et al. Multi-grained attention representation with ALBERT for aspect-level sentiment classification[J]. IEEE Access, 2021, 9: 106703–106713.
- [19] LIAO W X, ZENG B, YIN X W, et al. An improved aspect-category sentiment analysis model for text sentiment analysis based on RoBERTa[J]. Applied Intelligence, 2021, 51(6): 3522–3533.
- [20] DANYAL M M, KHAN S S, KHAN M, et al. Proposing sentiment analysis model based on BERT and XLNet for movie reviews[J]. Multimedia Tools and Applications, 2024, 83(24): 64315–64339.
- [21] 张晓君, 黄微, 赵旭. 网络舆情用户传播行为的心理距离感知强度测度及实证研究[J]. 情报理论与实践, 2024, 47(7): 135–144.
- [22] 苏怡. 多维心理距离对在线影评有用性的影响研究[D]. 哈尔滨: 哈尔滨工业大学, 2020.
- [23] 华生旭, 吕厚超. 心理距离与建构水平的双向作用关系[J]. 心理科学, 2012, 35(6): 1519–1523.
- [24] 张晓君, 黄微, 李瑞. 心理距离视角下网络舆情传播意愿影响因素研究[J]. 图书情报工作, 2019, 63(12): 86–93.
- [25] SKALICKY S, BRUGMAN B C, DROOG E, et al. Satire from a far-away land: psychological distance and satirical news[J]. Information, Communication & Society, 2023, 26(8): 1548–1565.
- [26] 陈百卉. 心理距离对情感预测强度的影响[D]. 宁波: 宁波大学, 2014.
- [27] FARKHOD A, ABDUSALOMOV A, MAKHMUDOV F, et al. LDA-based topic modeling sentiment analysis using topic/document/sentence (TDS) model[J]. Applied Sciences, 2021, 11(23): 11091.
- [28] WON A S, SHRIRAM K, TAMIR D I. Social distance increases perceived physical distance[J]. Social Psychological and Personality Science, 2018, 9(3): 372–380.
- [29] CALDERON S, MAC GIOLLA E, ASK K, et al. Subjective likelihood and the construal level of future events: a replication study of Wakslak, Trope, Liberman, and Alony (2006)[J]. Journal of Personality and Social Psychology, 2020, 119(5): e27–e37.

(编辑: 丁红艺)