

基于大语言生成模型的舆情情感文本自动生成方法

崔骥¹, 韩益亮¹, 陈雅丽², 朱率率¹

(1. 中国人民武装警察部队工程大学, 陕西 西安 710086; 2. 中国人民武装警察部队警官学院, 四川 成都 610000)

摘要: 社交网络中的负面舆情事件在传播过程中会产生巨大的消极影响。为快速自动生成正面舆情文本, 化解消极情绪、缓解舆论压力, 提出了一种基于大语言生成模型 (ERNIE-GEN) 的舆情情感文本生成方法。首先, 基于RoBERTa模型构建情感分类器, 对舆情文本进行情感极性识别与分类, 得到正面、中性和负面情绪的文本; 然后, 利用TextRank算法提取不同情感文本的关键主题, 刻画当前主题对应的舆情事件; 接下来, 将正面主题和情感文本输入ERNIE-GEN模型进行训练, 得到正面积积极舆情文本生成模型; 最后, 将负面舆情事件的主题词输入该模型, 即可生成对应的正面积积极舆情文本。以“某医院暴力伤医案”期间“医生”主题舆情事件为例, 通过模型生成与负面主题相对的正面积积极情感文本, 用于缓解和稀释热点舆情, 转移公众注意力。实验结果显示, 所提方法在网络舆情引导方面具有良好的效果。

关键词: 情感主题; 文本生成; 正面博文; 舆情引导

中图分类号: TP391

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2025074

Automatic generation method for public opinion sentiment text based on large language generation model

CUI Su¹, HAN Yiliang¹, CHEN Yali², ZHU Shuaishuai¹

1. Chinese People's Armed Police Force Engineering University, Xi'an 710086, China

2. Chinese People's Armed Police Force Officers College, Chengdu 610000, China

Abstract: Social media negative public opinion events can have a tremendous adverse impact during their dissemination process. To intelligently and rapidly generate positive public opinion texts for resolving negative emotions and alleviating public opinion pressure, a new method of generating sentiment texts based on the large language generation model (ERNIE-GEN) was proposed. Firstly, an emotion classifier was constructed based on RoBERTa model to identify and classify the emotion polarity of public opinion texts, and obtain positive, neutral and negative emotion texts. Then, the TextRank algorithm is used to extract the key themes of different emotional texts and depict the public opinion events corresponding to the current theme. Next, the positive theme and emotion text were put into the ERNE-GEN model for training, and the positive public opinion text generation model was obtained. Finally, the topic words of negative public opinion events were input into the model to generate positive public opinion event text. Taking the "doctor" theme public opinion event during the "violent assault on doctors incident at a certain hospital" as an example, positive public opinion event texts were generated according to the negative public opinion event theme, which were used to alleviate and dilute hot public opinion and divert public attention. The experimental results show that the proposed method has a very good effect in guiding online public opinion.

Keywords: emotional topic, text generation, positive text, public opinion guidance

收稿日期: 2025-02-01; 修回日期: 2025-07-10

通信作者: 韩益亮, hanyil@163.com

基金项目: 陕西省自然科学基金基础研究计划 (2024JC-YBMS-546)

Foundation Item: The Natural Science Basis Research Plan in Shaanxi Province of China (2024JC-YBMS-546)

0 引言

社交网络舆情引导^[1]是指运用多种网络媒体手段,对涉及特定事件、话题或个体的网络舆情进行研究和分析,并通过系列策略引导、塑造或控制其发展趋势,以实现特定目标的过程。该过程通常涵盖舆情监测、舆情分析、舆情塑造等环节。网络舆情引导已成为现代社会各类机构和组织管理工作的重要组成部分。

为营造健康的网络舆论环境,常规的舆情处理手段包括通过管控、删帖、限评等方式防止负面舆情扩散^[2],或在短时间内发布大量具有正面积极导向的文本^[3],还可以通过发布具有引导性的文本转移舆论焦点,从而改变人们对负面舆论及恶意舆情事件的认知,扭转社会舆论走向。其中,删帖与限评易于实施,而依靠人工撰写积极倾向的网评文本存在成本高、效率低、难以匹配舆情发展速度等问题。因此,研究基于智能写作技术的网评文本自动生成方法具有重要的现实意义。

在现有的舆情引导方法中,应用较广泛的是基于文本生成模型生成反制社交网络恶意信息的文本。其主流做法是结合主题模型与深度学习模型,依据目标主题生成指定舆论方向的文本^[4-5]。然而,这类方法生成的文本往往缺乏情感属性,而在实际的网络交互中,富含情感的文本通常具有更强的感染力和引导效力。为此,阮叶丽等^[6]提出将情感属性融入文本生成过程的方案。

目前,围绕舆情主题与情感属性协同建模、生成网络舆情引导所需文本的研究相对较少。为此,本文提出一种基于大语言生成模型 ERNIE-GEN 的舆情情感文本生成方法。具体而言,首先训练 RoBERTa 情感分类器,对舆情文本进行情绪分类;再利用 TextRank 算法提取不同情绪文本的主题词;接着将主题词与对应情感文本输入 ERNIE-GEN 模型进行训练,从而获得面向不同情感的舆情文本生成器。在实际应用中,可以将负面舆情主题输入正面舆情文本生成器,快速得到大量与负面主题对应的积极舆情文本,并将其投放到社交网络。

1 相关工作

舆论引导在应对重大公共事件中发挥着至关重要。系统的信息整合与发布、官方信息的披露

与传播,以及有效的舆论引导策略^[7],对缓解网络负面情绪具有显著效果。Wei 等^[8]提出一种基于经典有界置信模型提取信息质量变量与个体信任阈值的舆论动态模型,并构建了两阶段意见演化模型。其研究发现,论证越充分、态度越温和,越有利于引导舆论;此外,公众人物发布的引导性意见会对舆情事件产生持续的积极影响。You 等^[9]构建了基于易感-感染-易感(susceptible infected susceptible, SIS)传播模型的文本情绪倾向挖掘模型,用于分析影响(助长和阻止)舆情发展的因素。实验结果表明,当正面情绪舆论比例高于 0.8 时,可以阻断负面舆论传播;当负面与中性舆论比例均低于 0.2 时,舆情传播力度会逐渐减弱。

基于舆情引导在网络舆情事件中的重要作用,更多的学者投身于相关方法的研究。部分学者通过文本生成模型生成积极引导文本,将其投放到社交网络中以改变舆论走向。刘欣然等^[10]基于序列生成式对抗网络(SeqGAN)构建网络贴文自动生成模型,用于反制社交网络中的恶意信息,引导舆论向正确方向发展。王沛等^[11]提出一种基于外部知识筛选的主题文本生成模型,能够根据主题关键词生成相关文本,提高文本的可读性与感染力。胡益等^[5]基于 CoCon 模型提出一种结合主题模型与自监督学习的可控文本生成方法,在主题契合与常规文本生成方面表现出色。另有学者通过构建引导模型对引导效果进行计算与校对。Lyu 等^[12]将社交网络划分为舆论引导社区,基于对偶学习构建舆论引导模型,通过正面舆论引导者对负面舆论接受者进行反复引导,并建立反馈机制判断引导是否成功,最终实现双重引导。Song 等^[13]基于易感-感染-康复(susceptible infected recovered, SIR)传播模型构建未知接触者-感染-康复(unknown contact infected recovered, UCIR)动态模型,通过研究关键参数,发现不同状态之间的转移概率对舆情信息传播的影响,并提出相应的舆论引导对策。

表 1 将本文模型与现有的代表性模型进行了对比。与传统的抽取式生成方法相比,本文方法实现了文本的自动生成;与其他生成式方法相比,本文方法能够根据指定主题与情感需求生成

具有正面情感倾向的文本，从而更具针对性地实现舆论引导。

表1 本文模型与现有代表性模型的对比

方法	指定主题	指定情感	自动生成	生成方式
TextRank	—	—	—	抽取式
SeqGAN ^[10]	—	—	√	生成式
GAN ^[6]	—	√	√	生成式
CoCon ^[5]	√	—	√	生成式
本文模型	√	√	√	生成式

2 舆情情感文本自动生成模型

本文提出的基于舆情主题的文本生成模型包含3个核心部分，即情感分类器RoBERTa、主题提取模块TextRank和舆情文本生成模块ERNIE-GEN，其结构如图1所示。

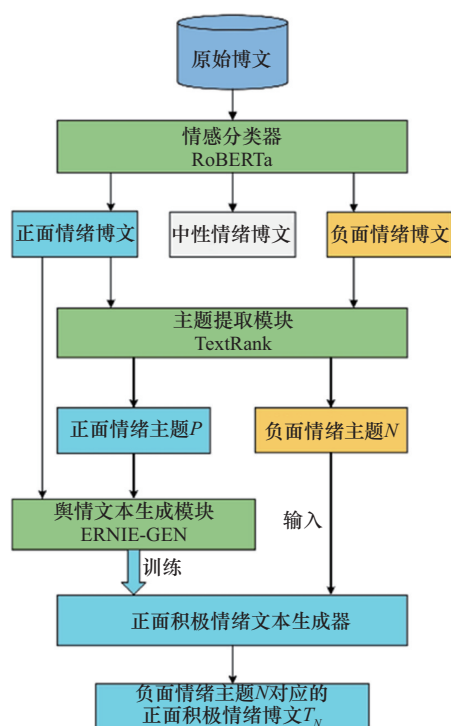


图1 模型的结构

1) 情感分类器：使用情感分类器RoBERTa^[14]对原始博文进行情感极性识别，将其分类为正面、中性与负面情绪的博文。

2) 主题提取模块：利用TextRank算法分别对已分类的正面、中性和负面情绪博文进行主题提取，得到正面情绪主题集合 P 和负面情绪主题集合 N 。

3) 舆情文本生成模块：将正面舆情主题和对应博文输入ERNIE-GEN^[15]模型进行训练，得到正面积积极情绪文本生成器。随后，将负面情绪主题 N 输入训练好的正面积积极情绪文本生成器，得到负面情绪主题 N 对应的正面积积极情绪博文 T_N 。

2.1 RoBERTa

本文采用预训练大语言模型RoBERTa进行情感分类。RoBERTa是BERT的增强版本，主要在以下方面进行了改进：扩大模型参数量、增加批处理大小、使用更丰富的训练数据；去除下一句预测（next sentence prediction, NSP）任务，将静态掩蔽策略改为动态掩蔽策略；采用字节级字节对编码（byte pair encoding, BPE）技术对预训练任务进行优化。RoBERTa的核心是Transformer结构中的多头注意力（Multi-Head Attention）机制。该机制通过映射查询（ Q ）、键（ K ）、值（ V ）向量至不同的表示空间，计算 Q 与 K 的相似度，进而对 V 进行加权聚合，从而实现对关键信息的聚焦。具体过程如下。

1) 将 Q 、 K 和 V 分别通过线性变换映射到不同的向量空间。

$$Q = W_Q x \quad K = W_K x \quad V = W_V x \quad (1)$$

其中， x 为输入向量， W_Q 、 W_K 和 W_V 为可学习的参数矩阵。

2) 通过缩放点积计算 Q 和 K 的相似度，得到注意力分数，将注意力分数通过softmax归一化后，与 V 进行加权求和得到单头输出。

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

其中， d_k 为每个头的维度。

3) 将多个注意力头（head）的输出拼接后进行线性变换，得到最终的注意力输出。

$$\text{Multi-Head Attention}(x) = W_O[\text{head}_1, \text{head}_2, \dots, \text{head}_h] \quad (3)$$

其中， head_i 为第 i 个头的输出， W_O 为输出层可学习参数矩阵。

多头注意力的输出经过残差连接与层归一化处理后被输入softmax分类层，用于构建三分类（正面、中性和负面）情感分类器。本文使用10万条情感标注数据对RoBERTa进行训练，得

到情感分类器,进而对社交网络舆情文本进行情感分类,分别得到正面、中性和负面情绪文本。

2.2 TextRank

本文采用TextRank算法对微博文本进行主题提取。TextRank是一种基于图排序的算法,被广泛应用于关键词和句子的提取。TextRank算法将文本中的语言单位(如词或句子)视为图中的节点,以语言单位间的关联关系作为边,通过迭代计算每个节点的权重,确定文本中与主题最相关的节点。这些节点可以帮助生成文本摘要或关键词列表。算法如式(4)所示。

$$WS(V_i) = (1 - d) + d \times \sum_{V_j \in \text{In}(V_i)} \frac{W_{ji}}{\sum_{V_k \in \text{Out}(V_j)} W_{jk}} WS(V_j) \quad (4)$$

其中, $WS(V_i)$ 为句子 i 的权重, d 为阻尼系数(通常取 0.85), 式(4)右侧的求和表示每个相邻句子对本句子的贡献度。在单个文档中, 可以认为所有句子都是相邻的, 无须像多文档场景那样采用滑动窗口或分块机制逐段生成和抽取特征。 W_{ji} 为句子 j 、 i 的相似度, $WS(V_j)$ 为上次迭代得到的句子 j 的权重。

针对微博文本的主题提取流程具体如下: ①将微博文本按照完整句子进行分割, 即 $S = \{S_1, S_2, \dots, S_m\}$; ②对每个句子进行分词、词性标注, 过滤停用词, 仅保留指定词性(如名词、动词、形容词等)词汇, 得到处理后的词集合 $S_i = \{t_{i,1}, t_{i,2}, \dots, t_{i,n}\}$; ③构建关键词图 $G = (V, E)$, 其中节点集 V 由步骤②得到的候选关键词组成, 边基于共现关系构建(当两个节点在长度为 5 的滑动窗口内共现时, 在对应节点间建立边); ④基于式(4)迭代更新各个节点的权重直到收敛, 将节点权重按照降序进行排列, 选取前 T 个词作为候选关键词; ⑤在原始文本中标记候选关键词, 将相邻候选词组合为多词关键词; ⑥分别对正面、负面情感文本执行上述流程, 得到两类文本对应的主题词集合。

2.3 ERNIE-GEN

在文本生成阶段, 本文采用 ERNIE-GEN 模型完成相关任务, 其整体框架如图 2 所示。需要说明的是, 原始 ERNIE-GEN 模型以预训练大语

言模型 BERT 作为基础微调框架, 而本文将这一基础框架替换为 RoBERTa。预训练大语言模型 RoBERTa 已经使用大量语料进行训练, 具备较强的语言理解与生成基础能力。为使 ERNIE-GEN 模型能够更好地执行特定文本生成任务, 本文对其进行了有监督训练。

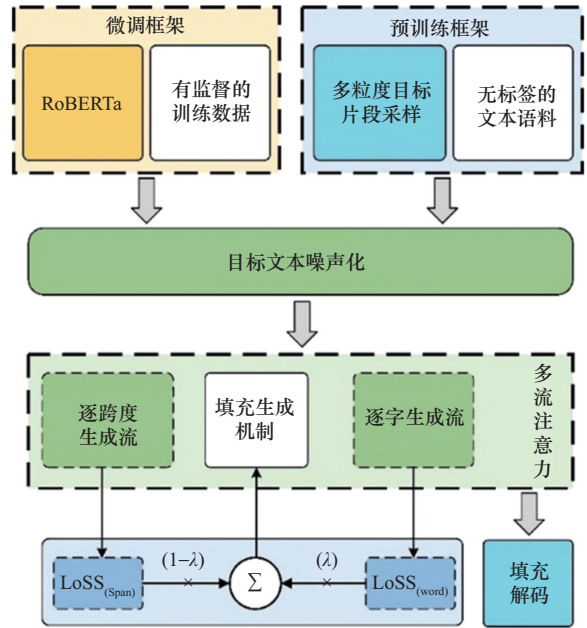


图2 ERNIE-GEN模型的整体框架

ERNIE-GEN 通过结合多粒度目标片段采样(即采样不同长度的文本片段)来构建预训练数据, 从而强化编码器与解码器之间的关联性; 同时采用填充生成机制与噪声感知生成方法, 弥合训练与推理之间的差异。在训练过程中, 模型引入多流注意力机制, 在逐字生成流(word-by-word)的基础上, 新增逐跨度(span-by-span)生成流, 使模型每次能够生成一个完整语义文本片段。

逐字生成流的形式化表示如式(5)所示。

$$\mathbf{a}_i^{(l+1)} \leftarrow \text{MH} - \text{Attn} \left(\begin{matrix} \mathbf{Q} = \mathbf{a}_i^{(l)}, \\ \mathbf{KV} = [\mathbf{S}^{(l)}, \mathbf{t}_i^{(l)}, \mathbf{a}_i^{(l)}] \end{matrix} \right) \quad (5)$$

其中, $\mathbf{Q}$ 、 $\mathbf{K}$ 、 $\mathbf{V}$ 分别表示 Multi-Head Attention 中的查询、键和值, $\mathbf{a}_i$ 表示第 l 层的人工符号序列 AW 的第 i 个向量表示, $\mathbf{a}_i^{(l)}$ 表示第 l 层的第 i 个向量表示, $\mathbf{S}^{(l)}$ 表示第 l 层源序列, $\mathbf{t}_i^{(l)}$ 表示第 l 层带噪声的目标序列。

逐跨度生成流的形式化表示如式(6)所示。

$$a_j^{(l+1)} \leftarrow \text{Multi-Head Attention} \left(\begin{matrix} Q = a_j^{(l)}, \\ KV = [S^{(l)}, t_{<b_i}^{(l)}, a_j^{(l)}] \end{matrix} \right) \quad (6)$$

其中, b_i 表示第 i 个跨度, $j \in [b_i, b_{i+1}]$, a_j 表示第 i 个跨度的 $(j - b_i)$ 向量。

由于本文所提的情感主题文本生成模型缺乏公开可用的数据集, 因此首先需要构建相应的训练数据集。原始数据不带任何标签, 本文以 TextRank 算法提取的主题词作为文本的标签。接着, 将主题词与原文进行拼接, 构成符合 ERNIE-GEN 模型输入格式的数据样本。最终, 构建出一个包含正面、中性和负面 3 种情感类别的舆情生成数据集, 其样本格式统一为“主题词+文本”。

2.4 情感文本生成建模方法

负面舆情事件通常围绕特定主题展开, 本文以这些主题词作为舆情标识, 通过生成对应主题的正面情感文本实现舆论引导。具体建模流程如下: ①使用情感分类器 RoBERTa 对舆情文本 $S = \{s_1, s_2, \dots, s_m\}$ 进行情感分类, 得到正面 $P = \{p_1, p_2, \dots, p_k\}$ 、中性 $E = \{e_1, e_2, \dots, e_i\}$ 和负面 $N = \{n_1, n_2, \dots, n_j\}$ 3 类文本; ②主题提取模块采用 TextRank 算法对分类后的情感文本 P 、 E 和 N 分别进行主题提取, 得到每条情感文本 p_k 、 e_i 和 n_j 的主题词 t_k^p 、 t_i^e 和 t_j^n ; ③将“正面主题词+正面文本”格式的数据 $t_k^p + p_k$ 输入 ERNIE-GEN 模型进行微调, 构建正面积积极情绪文本生成器 P-GEN; ④将负面舆情主题词 t_n 输入 P-GEN 文本生成器, 得到与该主题相关的正面情绪文本 p_n ; ⑤将生成的正面情绪文本投放至社交网络, 通过扩大正面舆论影响力, 实现舆情态势的引导与扭转。上述流程的形式化描述如算法 1 所示。

算法 1 基于 ERNIE-GEN 的情感文本生成建模方法

输入 微博文本数据 $S = \{s_1, s_2, \dots, s_m\}$

- 1) **for** $s = 1$ to m **do** s_m 输入 RoBERTa,
- 2) s_m prediction = (-1, 0, 1)
- 3) $n_j \leftarrow s_m(-1)$; $e_i \leftarrow s_m(0)$; $p_k \leftarrow s_m(1)$
- 4) **for** $n = 1$ to j **do** n_j 输入 TextRank,
- 5) $t_j^n \leftarrow n_j(\text{TextRank})$, $t_i^e \leftarrow e_j(\text{TextRank})$, $t_k^p \leftarrow p_j(\text{TextRank})$

end for

6) **for** $n = 1$ to k **do** $t_k^p + p_k$ 输入 ERNIE-GEN,

end for

7) P-GEN $\leftarrow t_k^p + p_k$ (ERNIE-GEN)

end for

8) **for** $n = 1$ to n **do** t_n 输入 P-GEN,

9) $p_n \leftarrow t_n$ P-GEN,

end for

输出 p_n

3 实验和结果分析

3.1 数据集和预处理

本文实验数据来源于第 26 届全国信息检索学术会议 (CCIR 2020) 评测大赛数据集^[16]。该数据集包含 2020 年 1 月 1 日至 2 月 19 日期间的 100 万条微博文本, 其中 10 万条已完成情感标注 (含正面、中性、负面 3 类标签), 其余 90 万条为无标注数据。本文采用 10 万条标注数据微调情感分类器 RoBERTa, 按 6:2:2 划分为训练集 (6 万条)、验证集 (2 万条) 与测试集 (2 万条), 分别用于模型训练、参数优化与性能评估。

微博文本中存在大量噪声, 而文本生成模型对数据质量要求较高, 因此需要对数据进行预处理。本文基于 HarvestText 工具包执行以下预处理操作: ①删除文本中的 @用户名、表情符号、电子邮箱地址等无关信息; ②删除不必要的空格与换行符; ③取消 HTML 转义标记; ④用 URL 替换文本中提到的超链接; ⑤将繁体中文转换为简体中文。

3.2 实验设置

本文的所有实验都是在百度 AI Studio 深度学习平台上进行的, 硬件配置为 4 核 CPU、Tesla V100 GPU (32 GB 显存) 及 32 GB 内存。软件环境为 Python 3.7 编程语言、PaddlePaddle 2.1.2 深度学习框架。

RoBERTa 超参数设置如下: 学习率为 5e-5、文本最大长度为 188、批处理大小为 64、训练轮次 (epoch) 为 10。

ERNIE-GEN 超参数设置如下: 学习率为 2e-5、批处理大小为 32、编码最大长度为 50、解码最大长度为 1 000、噪声概率为 0.2、最大训练步数为 60 000。

3.3 评价指标

本文从情感分类性能、文本生成质量与舆情引导效果 3 个维度设计评价指标。

1) 情感分类指标: 采用精确率 (precision)、召回率 (recall)、F1 分数与准确率 (accuracy) 评估分类性能。

2) 文本生成指标: 采用 Rouge-1 (单字重叠率) 与 Rouge-2 (双字重叠率) 评估生成文本与参考文本的语义相似度。

3) 舆情引导效果指标: 采用舆情主题算数指数 (public opinion topic arithmetic index, POI), 该指标融合爆发指数 (explosion index, EI)、情感指数 (sentiment index, SI) 和传播指数 (dissemination index, DI) 3 个维度, 能够全面反映舆情态势变化。各指数的计算式如下。

$$POI = SI + EI + DI \quad (7)$$

$$EI = \frac{\text{Relu}\left(df_{t,N}^n - \frac{1}{n-1} \sum_{i=1}^{n-1} df_{t,N}^i\right)}{1 + \frac{1}{n-1} \sum_{i=1}^{n-1} df_{t,N}^i} \quad (8)$$

$$SI = \text{Relu}\left[\frac{df_{t,N}^n}{1 + df_{t,P}^n} - \frac{\sum_{i=1}^{n-1} df_{t,N}^i}{1 + \sum_{i=1}^{n-1} df_{t,P}^i}\right] \quad (9)$$

$$DI = \ln\left(1 + \text{Relu}\left(df_{t,N}^n - \frac{1}{n-1} \sum_{i=1}^{n-1} df_{t,N}^i\right) \times \frac{1 + df_{t,N}^n}{1 + df_{t,P}^n}\right) \quad (10)$$

其中, $df_{t,N}^i$ 为第 i 天字词 t 出现的消极 (Negative) 帖子数量, $df_{t,P}^i$ 为第 i 天字词 t 出现的积极 (Positive) 帖子数量, Relu 为激活函数, 分母加 1 是为了防止公式没有意义 (避免除零错误)。

3.4 实验和结果分析

本文共进行 7 组对比实验, 组合包括: ELECTRA^[17]+ERNIE-GEN、ELECTRA-small+ERNIE-GEN、ERNIE-1.0^[18]+ERNIE-GEN、ERNIE-tiny+ERNIE-GEN、BERT^[19]+ERNIE-GEN、MacBERT^[20]+ERNIE-GEN 和 RoBERTa+ERNIE-GEN。实验时, 首先使用各预训练语言模型对微博舆情文本进行情感分类, 再基于该模型进行生成任务的训练与评估。accuracy 等情感分类指标结果是使用预训练语言模型对数据集进行情感分类所

得, 文本生成指标 Rouge-1 和 Rouge-2 是使用预训练语言模型和 ERNIE-GEN 执行生成任务所得。

情感分类实验结果见表 2, 从表中数据可以看出, ELECTRA-small 与 ERNIE-tiny 模型的分类准确率较低 (分别为 72.32% 与 72.91%), RoBERTa 模型的分类准确率最高 (76.08%)。这一结果归因于 RoBERTa 使用了更多的数据进行预训练, 拥有更多参数。这表明小规模语言模型在分类任务上的性能较大规模语言模型逊色。

表 2 情感分类实验结果

模型	F1	precision	recall	accuracy
ELECTRA-small	76.40%	74.22%	78.70%	72.32%
ELECTRA	80.42%	77.34%	83.76%	75.61%
ERNIE-tiny	76.56%	74.24%	79.02%	72.91%
ERNIE-1.0	78.68%	76.63%	80.83%	74.83%
BERT	78.30%	76.03%	78.62%	75.64%
MacBERT	80.38%	77.18%	83.86%	75.82%
RoBERTa	81.81%	77.93%	83.91%	76.08%

文本生成质量实验结果见表 3, 从表中数据可以看出, RoBERTa+ERNIE-GEN 组合的生成质量最优 (Rouge-1 与 Rouge-2 分数分别达到 22.41% 与 9.69%), 而 ELECTRA-small+ERNIE-GEN 组合表现最差 (Rouge-1 与 Rouge-2 分数分别为 10.08% 与 3.23%)。

表 3 文本生成质量实验结果

文本生成框架	预训练模型	Rouge-1	Rouge-2
	ELECTRA-small	10.08%	3.23%
	ELECTRA	14.39%	4.42%
	ERNIE-tiny	17.81%	6.56%
ERNIE-GEN	ERNIE-1.0	18.16%	6.52%
	BERT	21.12%	8.51%
	MacBERT	21.90%	9.42%
	RoBERTa	22.41%	9.69%

为验证模型的舆情引导效果, 本文以“某医院暴力伤医案”为实验场景。该事件发生后, 由于官方对真相的核查与深度报道需要一定的时间, 这与网民对即时信息的需求形成时间差。在此期间, 网络上出现大量不实猜测, 形成负面舆

情浪潮。传统的舆情处理常采用删帖、限评等方式，本文使用所提模型生成“医生”主题的正面积极情绪文本，具体见表4，通过正面舆论对冲实现引导效果。

表4 生成情绪主题文本示例

情绪倾向	主题	生成文本
正面	医生	1. 每次看到医生被伤害的新闻都忍不住落泪，今天了解**医生的一些事情后，眼泪更是止不住地流，伤心、遗憾、气愤，希望**医生早日康复，最好以后能继续正常工作！
		2. 希望所有的医疗人员都平平安安、健健康康。确实做不了什么，只能为**医生祈福。
		3. 向**医生致敬！几天来一直默默祈祷，世间善恶终有报，祝福**医生！

在仿真实验中，使用本文方法前，POI为正常数据集的算数指数。事件爆发（2022年1月20日）后，首先模拟删帖策略（次日删除与首日内容重复的负面帖）。从2022年1月22日起，每日向模拟环境中注入100条由本文模型生成的正面“医生”主题博文，并持续计算该主题的POI，结果如图3所示。

实验结果显示，2022年1月22日开始投放生成的正面文本后，POI显著下降，且后续二次舆情爆发日的POI也明显减小。这表明模型生成的正面文本能够有效稀释负面舆论浓度，缓解社会矛盾和冲突，实现舆情态势的良性扭转。

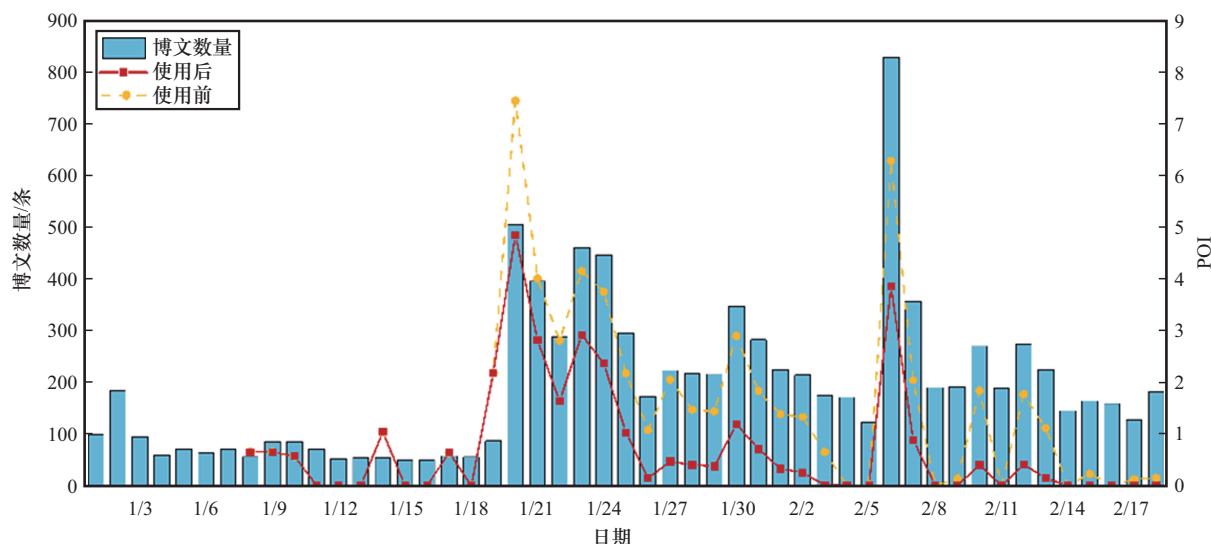


图3 “医生”舆情主题POI时间图

4 结束语

本文提出了一种基于大语言生成模型ERNIE-GEN的舆情情感文本生成方法。该方法通过RoBERTa模型实现原始博文的情感分类，利用TextRank提取不同情感文本的主题词，将情感主题和对应情感文本输入ERNIE-GEN模型进行训练，得到能够根据指定主题生成特定情感倾向文本的生成器。实验结果表明，本文所提方法在网络舆情态势引导中具有良好的效果。在下一步工作中，将训练更加智能的文本生成模型，实现聊天机器人式的问答模式，使文本生成效果更好。

参考文献：

- [1] CHEN Z T, LAN H L. Dynamics of public opinion: diverse media and audiences' choices[J]. Journal of Artificial Societies and Social Simulation, 2021, 24(2): 8.
- [2] ZHANG L J, WANG T, JIN Z L, et al. The research on social networks public opinion propagation influence models and its controllability[J]. China Communications, 2018, 15(7): 98-110.
- [3] 余晓青. 意识形态网络舆情应对机制研究[J]. 南京邮电大学学报(社会科学版), 2021, 23(5): 37-47.
- [4] YU X Q. The problems and countermeasures of coping mechanism on ideological internet public opinion[J]. Journal of Nanjing University of Posts and Telecommunications (Social Science), 2021, 23(5): 37-47.
- [5] 刘淑婷, 马颖华, 陈秀真. 面向舆情治理的信息管理机制演化博弈模型[J]. 网络与信息安全学报, 2023, 9(6): 102-115.
- [6] LIU S T, MA Y H, CHEN X Z. Evolutionary game model of infor-

- mation management mechanism for public opinion governance[J]. Chinese Journal of Network and Information Security, 2023, 9(6): 102-115.
- [5] 胡益, 刘嘉勇, 代金鞘, 等. 结合主题模型与自监督学习的可控文本生成技术研究[J]. 四川大学学报(自然科学版), 2023, 60(5): 130-138.
- HU Y, LIU J Y, DAI J Q, et al. Controllable text generation technology based on topic model and self supervised learning[J]. Journal of Sichuan University (Natural Science Edition), 2023, 60(5): 130-138.
- [6] 阮叶丽. 基于 GAN 考虑外部知识与情感属性的文本生成研究[D]. 武汉: 中南财经政法大学, 2021.
- RUAN Y L. The research of text generation with considering external knowledge and emotional attributes based on GAN[D]. Wuhan: Zhongnan University of Economics and Law, 2021.
- [7] WANG Y Q, WU R, ZENG J, et al. Research on the public opinion guidance mechanism of major public health incidents[J]. Frontiers in Psychology, 2022, 13: 872464.
- [8] WEI J, JIA Y G, TIE W Y, et al. Opinion evolution with information quality of public person and mass acceptance threshold[J]. Big Data, 2024, 12(2): 100-109.
- [9] YOU G, GAN S Q, GUO H, et al. Public opinion spread and guidance strategy under COVID-19: a SIS model analysis[J]. Axioms, 2022, 11(6): 296.
- [10] 刘欣然, 徐雅斌, 李继先. 网评贴文自动生成方法研究[J]. 数据分析与知识发现, 2023(4): 101-113.
- LIU X R, XU Y B, LI J X. Method for automatically generating online comments[J]. Data Analysis and Knowledge Discovery, 2023(4): 101-113.
- [11] 王沛, 杨频, 程芑森, 等. 基于外部知识筛选的主题文本生成技术研究[J]. 四川大学学报(自然科学版), 2024, 61(1): 75-83.
- WANG P, YANG P, CHENG P S, et al. Research on topic text generation technology based on external knowledge filtering[J]. Journal of Sichuan University (Natural Science Edition), 2024, 61(1): 75-83.
- [12] LYU B Y, DU Y J, JIANG J J, et al. Positive public opinion guidance model based on dual learning in social network[J]. Intelligent Data Analysis, 2023, 27(3): 833-853.
- [13] SONG J B, ZHU X Y. Research on public opinion guidance of converging media based on AHP and transmission dynamics[J]. Mathematical Biosciences and Engineering, 2021, 18(5): 6857-6886.
- [14] LIU Y H, OTT M, GOYAL N, et al. RoBERTa: a robustly optimized BERT pretraining approach[J]. arXiv preprint arXiv:1907.11692, 2019.
- [15] XIAO D L, ZHANG H, LI Y K, et al. ERNIE-GEN: an enhanced multi-flow pre-training and fine-tuning framework for natural language generation[C]//Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization, 2020: 3997-4003.
- [16] World Health Organization. Coronavirus disease (COVID-19) - events as they happen[EB]. 2020.
- [17] CLARK K, LUONG M T, LE Q V, et al. ELECTRA: pre-training text encoders as discriminators rather than generators[C]//Proceedings of the 8th International Conference on Learning Representations. 2020.
- [18] SUN Y, WANG S H, LI Y K, et al. ERNIE: enhanced representation through knowledge integration[J]. arXiv preprint arXiv:1904.09223, 2019.
- [19] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[J]. arXiv preprint arXiv:1810.04805, 2018.
- [20] CUI Y M, CHE W X, LIU T, et al. Revisiting pre-trained models for Chinese natural language processing[J]. arXiv preprint arXiv:2004.13922, 2020.

[作者简介]



崔骥 (1992-), 男, 山东高密人, 中国人民武装警察部队工程大学硕士生, 主要研究方向为社交网络舆情、深度学习和自然语言处理。



韩益亮 (1977-), 男, 甘肃会宁人, 博士, 中国人民武装警察部队工程大学教授、博士生导师, 主要研究方向为网络信息安全与智能化等。



陈雅丽 (1990-), 女, 湖南益阳人, 中国人民武装警察部队警官学院讲师, 主要研究方向为系统工程、风险控制和网络舆情。



朱率率 (1985-), 男, 山东高青人, 博士, 中国人民武装警察部队工程大学教授、硕士生导师, 主要研究方向为社交网络舆情、隐私保护。