doi:10.3772/j.issn.2095-915x.2026.04.007

面向公众情感分析的大型语言模型有效性与智慧城市治理策略

开放科学（资源服务）标识码（OSID）

梅澎　赖先进

中共中央党校（国家行政学院）国家治理教研部（电子政务研究中心）　北京　100091

摘要：[目的 / 意义] 进入数字化智能化时代，智慧城市建设可运用的数字技术更加多样。其中，大型语言模型不仅能够处理和生成自然语言，而且能够模拟和分析智慧城市中社会公众情感、提高城市公共服务质量和效率。[方法 / 过程] 本研究以广泛使用的社交媒体平台新浪微博为数据源，采用数据挖掘技术来收集大量用户评论，并使用基于 Transformer 的双向编码器表征（Bidirectional Encoder Representations from Transformers，BERT）模型进行情感分类。通过将 BERT 模型与潜在狄利克雷分配（Latent Dirichlet Allocation，LDA）主题聚类模型和时间序列分析相结合，比较和验证了大型语言模型的情感分析功能。[结果 / 结论] 第一，BERT+LDA 模型取得了最佳性能，准确率和精密率分别为 90.49% 和 91.95%。该模型可以被选为评论文本情绪预测的主要工具。第二，在应用阶段，使用 torch.load 函数加载预训练模型并设置为评估模式，成功预测了微博评论的情绪，为智慧城市建设中政府监管和智能技术治理提供了新路径。第三，公众对大语言模型的隐私保护能力表示担忧，政府在大型语言模型和人工智能治理过程中应加强个人隐私和个人信息保护方面的监管工作。因此，基于大型语言模型，实现“用技术治理”与“对技术治理”相结合，是智慧城市建设和治理的有效策略。

关键词：大型语言模型，智慧城市，技术治理

中图分类号：G35

Effectiveness of Large Language Models for Public Sentiment Analysis and Smart City Governance Strategies

MEI Peng　LAI Xianjin

Department of National Governance (Electronic Government Research Center), Party School of the Central Committee of the Communist Party of China (National Academy of Governance), Beijing 100091, China

Abstract: [Objective/Significance] In the era of digitalization and intelligence, there is a wider variety of digital technologies available for the construction of smart cities. Among them, large language models can not only process and generate natural

基金项目　中央党校（国家行政学院）2024 年度青年科研项目“基于数字治理平台的政府公共服务成效及满意度提升机制研究”（2024QN040）。

作者简介　梅澎（1989-），博士，助理研究员，讲师，主要研究方向为电子政务、数字政府、数据要素、数字治理；赖先进（1986-），通信作者，博士，教授，主要研究方向为行政改革与国家治理现代化、科学技术与国家治理，E-mail：laixianjin2005@ccps.gov.cn。

引用格式　梅澎 , 赖先进 . 面向公众情感分析的大型语言模型有效性与智慧城市治理策略 [J]. 情报工程 , 2026, 12(4): 80-92.

language but also simulate and analyze the social public sentiment in smart cities, thereby improving the quality and efficiency of urban public services. [Methods/Process] This study uses Sina Weibo, a widely used social media platform, as the data source, employs data mining to collect a large number of user comments, and applies the Bidirectional Encoder Representations from Transformers (BERT) model for sentiment classification. By combining the BERT model with the Latent Dirichlet Allocation (LDA) topic clustering model and time-series analysis, the sentiment analysis capabilities of large language models are compared and verified. [Results/Conclusions] The results show three key findings. Firstly, the BERT+LDA model achieves optimal performance, with an accuracy of 90.49% and a precision of 91.95%, making it the preferred tool for sentiment prediction of comment texts. Secondly, during the application phase, the pre-trained model, loaded via the torch.load function and set to evaluation mode, successfully predicts the sentiment of Weibo comments, providing a new approach for government supervision and intelligent technology governance in smart city construction. Thirdly, the public expresses concerns about the privacy protections offered by large language models. Therefore, the government should strengthen its supervision over the protection of personal privacy and information in the governance of large language models and artificial intelligence. In conclusion, integrating governance through technology with governance over technology based on large language models constitutes an effective strategy for the construction and governance of smart cities.

Keywords: Large Language Model; Smart City; Technology Governance

引言

随着大数据和人工智能等现代信息技术的深度发展，人类社会逐步进入数字化智能化时代。在城市治理领域，智慧城市的概念逐渐兴起，并在全球范围内受到广泛关注。智慧城市利用物联网（Internet of Things，IoT）、大数据、云计算和人工智能（Artificial Intelligence，AI）等技术，通过高度集成和协调功能实现高效、智能和可持续的城市运营[1-3]。同样的，正是数字技术的发展，使得技术系统与社会系统相互嵌套、共同演化，技术不仅是工具，更是重塑情感互动的“物质基础”[4]。可以说，数字技术通过改变信息传递速度和互动方式，重构了情感表达的节奏和强度。而数字平台使得原本被压抑的“情感劳动”得到释放，促成情感开放[5]。网络拉近了世界各国人民之间的距离，实现了以互联网平台为载体的集体层面的情绪识别、调节与协作能力，形成情感的“制度化建构”[6]。情感作为社会互动中“意义协商”的产物，个体通过“符号解读”（如他人的表情、语气）理解情境，并据此调整自身情感表达[7]。一定程度上如果引导不慎、监管不严，会导致重大的舆情风险，引起社会发展的动荡。因此，技术、组织、社会并非孤立作用，而是以“技术中介 - 组织规训 - 社会建构”的链条共同塑造情感，这是技术、组织、社会相互作用的“社会事实”，也决定了情感预测的复杂性和必要性。

在智慧城市建设过程中，大型语言模型（Large Language Model，LLM）作为 AI 的关键分支，由于其强大的自然语言处理（Natural Language Processing，NLP）能力和广阔的应用前景，已成为智慧城市发展中不可或缺的组成部分[8]。其中，对文本情感分析为了解公众对新技术的可接受度、潜在担忧度和真实期望提供了独特的方法工具[9]。本文正是通过 LLM 的一个重要的分支技术 —— 文本情感分析技术，旨在识别、提取和量化文本数据中表达的情感

doi:10.3772/j.issn.2095-915x.2026.04.007

倾向，例如积极、消极或中立[10]，为政府决策和网信部门把握网络舆情，稳妥处理应急事件提供工具支撑。

作为LLM的突出代表，ChatGPT、DeepSeek等深度应用的智能产品，通过在大量文本数据集上进行预训练，在自然语言理解和生成方面表现出色[11]。这些模型不仅能够理解和响应复杂的查询，而且能够显著提高客户智能服务、公共安全警报和政策传播等领域的城市服务的质量和效率。借助于这些智能工具模型，政府机构可以二次开发智能客户服务系统，实现对群众咨询信息的实时应答，从而提高政府服务的响应能力和满意度。自ChatGPT发布以来，LLM技术不断进步。2023年推出的DeepSeek模型通过利用稀疏注意力机制和动态知识图谱集成，在长文本理解和特定领域的知识准确性方面取得了重大突破。与ChatGPT基于Transformer的标准架构相比，DeepSeek的三阶段预训练策略，即通用语料预训练、领域知识注入和强化对齐学习，在处理政府文件和分析政策法规方面表现出明显的优势。

然而，随着LLM越来越多地应用于智慧城市，其潜在风险和挑战也越来越明显。由于LLM是在海量数据集上训练的，其输出可能包含影响公众认知和判断的偏见或误导性信息。此外，这些模型在处理敏感信息时存在数据泄露和侵犯隐私的风险，从而威胁到个人信息安全。时至今日，LLM在各个领域得到广泛采用，同时也引发了人们对工作岗位流失和道德伦理风险的日益担忧。为确保LLM有助于智慧城市的健康发展，同时降低这些潜在风险，决策者必须制定有效的监管策略和对策。本研究试图解决的一个关键问题是，如何利用高级大语言模型来增强城市服务，同时保护数据安全。特别是研究如何在智能客户服务、公共安全预警系统和政策传播等应用中平衡技术创新和信息安全。为了解公众对LLM在智慧城市中的应用的态度和看法，本研究利用新浪微博作为数据源收集与大语言模型相关的用户评论。为了对收集到的用户评论表达的情绪进行分类，本文采用了基于Transformer的双向编码器表征模型。作为一种高级NLP模型，BERT在情感分类任务中表现出了出色的性能。通过训练BERT模型，本研究旨在实现将用户评论自动分类为积极、消极和中立情绪。为了更深入地了解公众舆论及LLM的影响，本文采用潜在狄利克雷分配（LDA）主题聚类模型与时间序列分析相结合的方法，以揭示公众情绪的时间趋势。最后，本研究提出有针对性的监管策略。研究以新浪微博社交平台数据为支撑，构建BERT+LDA情感分析模型，经实验验证形成较为完善的预测模型与评价机制；通过模型应用与实证研究反馈分析结果，评估监管策略的实际效用，以期为智慧城市治理实践提供参考。

1 研究设计

1.1 LLM在智慧城市中的应用

作为信息技术与城市基础设施深度融合的产物，智慧城市旨在通过物联网、云计算和大数据等技术实现高效、智能和可持续的城市运营。在此背景下，LLM凭借其独特的优势，已成为智慧城市发展的重要驱动力。LLM在智慧城市中的应用主要集中在四个领域：智能客户

服务与咨询、公共安全预警、政策推广与教育以及智能公共管理[12-14]。LLM 的技术实现过程包括数据预处理、模型训练、模型优化和应用程序部署。这些步骤共同构成了 LLM 在智慧城市中的应用基础。

在智能客户服务咨询领域，针对不同业务场景提供差异化公共服务时，LLM 可利用 NLP 技术处理市民的查询和投诉，给出及时准确的响应与解决方案[15-16]。这种智能化的客户服务系统不仅提高了公共服务的效率和质量，还增强了公民对政府的信任和满意度。

在公共安全预警中，LLM 的应用至关重要。通过整合和分析来自多个来源的信息，LLM 可以实时监测城市灾害风险，包括自然灾害（如地震、洪水）和人为灾害（如火灾、交通事故）。在检测到潜在风险后，LLM 可以快速生成警告消息并通过各种渠道（例如短信、社交媒体和广播）进行传播。灾后，LLM 可以通过分析灾难现场的情况、提供优化的救援资源分配建议以及为受影响人群规划疏散路线来提供实时应急响应支持。

在智能公共管理领域，以城市交通管理为例，LLM 的应用可以显著提高交通系统的运行效率和安全性。通过将 LLM 集成到智能交通信号控制系统中，城市可以实现实时交通流量监控和动态调整。LLM 可以分析历史交通数据以预测未来的交通趋势并相应地调整交通信号控制策略，从而缓解拥堵[17]。此外，利用 NLP 功能，LLM 可以处理和理解公民与交通相关的投诉和建议，为城市交通管理提供实时反馈和改进建议。同时，LLM 在交通事故预测方面也显示出相当大的潜力[18]。通过分析历史交通事故数据，LLM 可以识别与事故相关的模式和趋势，从而能够预测未来的事故。

1.2 基于LDA的注释主题聚类

LDA 模型是一种无监督主题建模方法，广泛用于从文档语料库中提取潜在主题信息。在应用 LDA 模型前，需先对文本进行分词、去停用词等预处理，并根据领域知识与困惑度指标确定主题数量[19]。建模后，LDA 为每篇文档分配主题分布、为每个主题分配词语分布。

图 1 提供了 LDA 模型的示意图。在这里，参数 α 控制文档中主题分布的稀疏性，表示文档可以包含多少个主题，而 β 控制主题中词语分布的稀疏性，确定主题中包含多少个唯一词语。θ 表示文档中 *topic* 的概率分布，Φ 表示从 *topics* 中采样的词语分布。K 是主题数，D 是训练语料库中的文档数，M 是指文档中的字数。变量 z 和 w 分别表示主题分配和相应的生成词语。在分析与智慧城市相关的评论数据时，LDA 模型可以识别评论中讨论的主要主题，例如交通管理、公共安全和环境质量等。通过对这些主题进行进一步分析，决策者可以深入了解公民对各种城市服务的满意度，并收集改进建议。需要指出的是，LDA 主题建模并非一次性输出结果即可完成分析，而是需要对同一事件进行长期监测，并结合历史数据进行前后比对，才能得出客观准确的结论。通过融合历史数据与当前数据，不仅可以明确社会信任的当前阈值，还能对其未来可能达到的峰值进行合理预测。对于平台而言，阈值的判定与峰值的预测可通过单位时间内内容发布频次、跟帖评论时间间隔及数量等指标进行测算；舆情监管

doi:10.3772/j.issn.2095-915x.2026.04.007

部门也可适时介入，对峰值出现的时间与程度进行干预和调控。然后利用这些信息来制定更有针对性和更有效的政策。

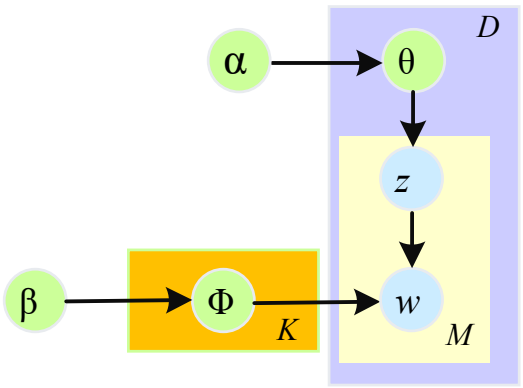

图 1　LDA 模型的示意图

1.3　基于ARIMA的情感时间序列模型

传统的情感分析方法通常会忽略时间因素，例如情感随时间变化的动态趋势。为了捕捉这些波动，本研究引入了一种基于自回归综合移动平均（ARIMA）模型的时间序列情感模型[20-21]。通过将时间序列分析与情感分析相结合，ARIMA 模型可以准确预测情感趋势。

ARIMA 模型记为 *ARIMA*(*p*,*d*,*q*)，由三个核心部分构成：自回归项（*AR*,*p*）：利用过去 *p* 个时刻的观测值预测当前值，反映序列的自相关性；差分项（*I*,*d*）：通过 *d* 阶差分将非平稳序列转化为平稳序列，消除趋势性影响；移动平均项（*MA*,*q*）：利用过去 *q* 个时刻的预测误差修正当前预测，捕捉随机冲击的滞后效应。ARIMA 建模的核心逻辑是先通过差分使序列平稳，再对平稳序列建立 ARMA 模型。模型阶数的确定通常借助自相关函数（ACF）和偏自相关函数（PACF）进行初步判断，表 2 总结了基本判定规则。

AR（*p*）分量自回归模型的数学公式表示为公式（1）：

$$Y_t=c+\phi_1 Y_{t-1}+\phi_2 Y_{t-2}+\cdots+\phi_p Y_{t-p}+\epsilon_t \quad (1)$$

Y_t 表示时间点 t 处的观测值；c 是一个常数项；$\phi_1,\phi_2,\ldots,\phi_p$ 是模型参数，用于描述当前值与过去 p 个时间点的值之间的关系；ϵ_t 表示误差项。

表 2　ARIMA 模型阶次的确定

模型	自相关函数	偏自相关函数	典型特征
AR（*p*）	拖尾（指数衰减或正弦衰减）	*p* 阶滞后后截尾	历史观测值直接影响当前值
MA（*q*）	*q* 阶滞后后截尾	拖尾（指数衰减或正弦衰减）	历史误差冲击影响当前值
ARIMA（*p*,*q*）	拖尾	拖尾	*AR* 与 *MA* 分量并存
ARIMA(*p*,*d*,*q*)	差分后 *ACF* 拖尾	差分后 *PACF* 拖尾	含 *d* 阶差分的非平稳序列建模

MA(*q*) 分量移动平均模型的数学表达式如等式（2）所示：

$$Y_t=\mu+\epsilon_t+\theta_1\epsilon_{t-1}+\theta_2\epsilon_{t-2}+\cdots+\theta_q\epsilon_{t-q} \quad (2)$$

μ 是序列的平均值；$\theta_1,\theta_2,\ldots,\theta_q$ 是模型参数，描述当前值与过去 q 个时间点的误差之间的关系。

差分是处理非平稳序列的关键手段。一阶差分通过相邻时刻观测值相减消除线性趋势，其数学表达式如公式（3）所示：

$$\nabla Y_t=Y_t-Y_{t-1} \quad (3)$$

∇Y_t 表示 Y_t 的一阶差值。

其 *d* 阶差分的数学表达式如公式（4）：

$$\nabla^d Y_t=\nabla^{d-1}Y_t-\nabla^{d-1}Y_{t-1} \quad (4)$$

令 $Y_t'=\nabla^d Y_t$ 为 *d* 阶差分后的平稳序列，则 *ARIMA*(*p*,*d*,*q*) 模型整体公式可表示为公式（5）：

$$Y_t'=c+\phi_1 Y_{t-1}'+\cdots+\phi_p Y_{t-p}'+\epsilon_t+\theta_1\epsilon_{t-1}+\cdots+\theta_q\epsilon_{t-q} \quad (5)$$

或用滞后算子 $L(LY_t=Y_{t-1})$ 紧凑表示为公式（6）：

$$\phi(L)(1-L)^d Y_t = c+\theta(L)\epsilon_t \quad (6)$$

其中，$\phi(L)=1-\phi_1(L)-\cdots-\phi_p(L)^p$，为 p 阶自回归多项式，$\theta(L)=1+\theta_1 L+\cdots+\theta_q L^q$ 为 q 阶移动平均多项式。

在情绪分析的上下文中，时间序列情绪数据被视为 Y_t，ARIMA 模型用于预测未来的情绪趋势。情感分析是时间序列情绪模型的核心部分。通过确定文本数据的情绪方向，它将文本转换为情绪分数。这些分数可以是离散的（例如，积极、消极、中性）或连续的（例如，情绪强度值）。在获得情感分数的时间序列后，下一步是时间序列建模[22-24]。

将 ARIMA 应用于情绪时间序列分析，需完成三步：一是情感评分计算，用 BERT+LDA 模型对每条评论生成情感极性数值；二是时间窗口分割，按周汇总平均情绪得分形成情绪指数；三是数据预处理，处理缺失值与异常值。完成这些步骤后，处理后的数据集就可以输入到 ARIMA 模型中了。

在建模过程中，必须确定 ARIMA 模型的参数（p,d,q）。如果时间序列是非平稳的，则必须应用差分，直到达到平稳性。在本研究中，情感指数时间序列在一阶差分后满足平稳性要求，因此 d=1。这通常是通过分析时间序列的自相关函数（ACF）和偏自相关函数（PACF）来实现的，这些函数揭示了序列中不同时间点之间的相关性。此分析确定自回归项（p）和移动平均项（q）的顺序。差分阶数（d）通常通过 Augmented Dickey-Fuller（ADF）单位根检验确定。ACF 图显示缓慢衰减，而 PACF 图在第一次滞后后急剧下降，表明 p=0 和 q=1 可能是合适的参数组合。确定 ARIMA 参数后，可以使用历史数据对模型进行训练。在训练过程中，必须选择合适的优化算法和损失函数，以最小化模型的预测误差。根据最初选择的参数（p,d,q）=（0,1,1），构建 ARIMA 模型并拟合到历史数据。使用交叉验证评估模型的性能，并对参数进行微调以进一步优化。最终选择（0,1,1）作为最佳参数组合，该组合在平衡预测准确性的同时最大限度地减少了误差。

1.4 基于BERT+LDA模型的公共舆情预测

BERT 基于 Transformer 架构，采用双向编码器建模文本，通过掩码语言建模（MLM）和下一句预测（NSP）两个预训练任务学习深层语义关系[25-27]。BERT 模型结构如图 2 所示，可简化为输入层、中间层和输出层三部分。

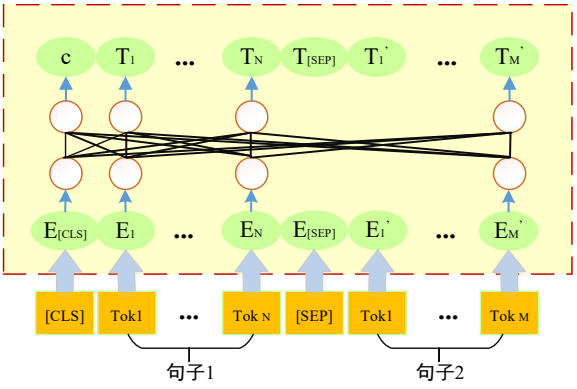

图 2 BERT 模型的结构

具体来说，BERT 采用多层 Transformer 编码器，每层包含两个主要组件：自注意力机制和前馈神经网络。自我注意机制的计算可以表示为：

$$Attention(Q,K,V)=softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (7)$$

Q、K 和 V 分别表示 *Query*、*Key* 和 *Value*，d_k 表示 *Key* 的维度。

BERT 模型的输出是每个词元的最终表示

doi:10.3772/j.issn.2095-915x.2026.04.007

形式，可用于情感分析等下游任务。对于情绪分析任务，[CLS] 标记的表示形式通常用作整个句子的表示形式：

$$Output=Transformer(Input) \quad (8)$$

Input 表示输入序列，而 *Output* 是 *Transformer* 编码器的输出，表示编码的输入序列。

对于情感分析任务，BERT 模型的输出可以连接到一个全连接层，然后用 *softmax* 层进行分类：

$$Logits=W^{T}Output+b \quad (9)$$

$$Probabilities=softmax(Logits) \quad (10)$$

W 和 *b* 分别表示全连接层的权重和偏差，*Logits* 表示激活前全连接层的输出，*Probabilities* 表示应用 *SoftMax* 激活函数后获得的概率分布。

在掩码语言模型（MLM）阶段，BERT 模型从输入文本中随机选择一些词语，并用特殊的 [MASK] 标记替换它们。该模型的目标是预测这些掩码词。在这个过程中，BERT 模型学习词语之间的上下文关系，使其能够更好地理解和生成文本。在下一个句子预测（NSP）阶段，BERT 模型将两个句子作为输入，旨在确定这些句子是否在语义上相关。这项任务有助于 BERT 模型学习句子之间的关系，提高其对句子语义和上下文的理解 [28-30]。

BERT 和 LDA 模型的组合可以构建更复杂的情感预测模型。在这个模型中，BERT 模型首先对文本数据进行编码以生成词向量。然后，LDA 模型对这些词向量进行主题建模以提取主题信息。最后，将提取的专题信息作为特征，通过机器学习算法进行情感分类。这种方法集成了深度学习和主题建模，同时考虑了文本的语义信息和主题结构，提高了情感预测的准确性。

在实际应用中，这种基于 BERT+LDA 的公众情绪预测模型可用于分析文本数据，例如社交媒体评论、新闻报道或政策反馈。通过预测公众舆情倾向，行政部门和企业可以更好地了解公众的需求和期望，从而制定更符合公共利益的政策和产品 [31-33]。

1.5 实验设计

本文选择新浪微博社交媒体平台作为数据源，通过爬虫技术，收集了大量用户评论，以更深入地了解和评估公众对 LLM 的看法。而后将 BERT+LDA 主题聚类模型与时间序列模型集成，对公众对 LLM 的关注进行全面的多角度分析。BERT+LDA 模型使用以下超参数进行训练：最大序列长度为 200，初始学习率为 2e-5，10 个训练轮次，批次大小为 256。

随后，文本定义两个关键功能。第一个函数 Weibo_LongText_data 以唯一的微博标识符（mid）作为输入，发送 HTTP 请求来检索微博文章的详情页，并从返回的 JSON 数据中提取全文内容。该功能主要用于获取长文微博的全文。第二个函数 Weibo_data 接受两个参数：keyword 和 max_page，旨在从常规的短微博帖子中收集信息。该函数遍历 HTTP 请求，获取不同页面搜索结果的 JSON 数据，提取用户昵称、发布时间、微博内容和 IP 地址信息（国家和省份）。然后将这些数据存储在列表中。

在本研究中，选择“大语言模型”“大模型”“LLM”作为搜索关键词。需要说明的是，本研究以公众对大模型技术本身的情感态度为研究对象 —— 普通网民在社交媒体中讨论相关议题时，通常以技术概念而非 " 智慧城市 " 等政策术语进行表达；对大模型的公众情感分析，

能够为智慧城市场景下大模型应用的社会接受度与治理风险评估提供重要依据。数据收集时间为 2025 年 4 月 1 日至 2025 年 11 月 23 日，涵盖 34 周，共计 52,356 条记录。删除重复数据后，保留了 49803 条唯一评论。对评论长度的统计分析显示，最短的评论仅包含 5 个字符，而最长的评论达到 6010 个字符。平均而言，每条评论的长度为 135 个字符，中位数为 68 个字符，标准差为 212 个字符。如图 3 所示，评论长度的分布表现出明显的右偏度，表明大多数评论相对较短，非常长的评论很少见。

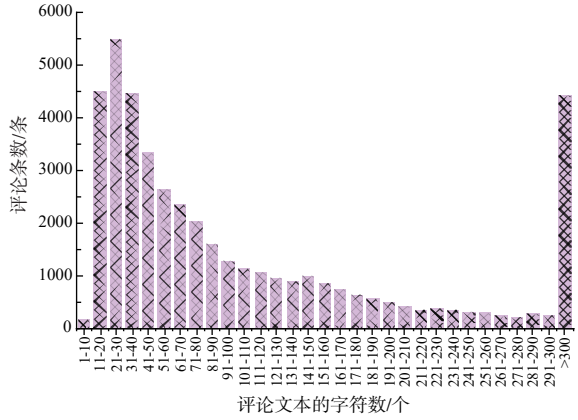

图 3 文本长度跨 Intervals 的分布

数据集划分方面，本研究采用按时间顺序划分的方式，以 2025 年 9 月 30 日为界，前期数据（2025 年 4 月 1 日至 9 月 30 日，约 35,000 条）作为训练集，后期数据（2025 年 10 月 1 日至 11 月 23 日，约 14,800 条）作为测试集。按时间划分更贴合真实应用场景，能够有效验证模型对未来未知数据的预测能力。

本研究将情感分为 3 个大类、6 个小类，采用多轮标注与交叉检验确保标注质量。统计结果显示：中性情绪评论最多（3,022 条），其次为赞扬（2,554 条）、担忧（1,145 条）、惊讶（1,012 条）、愤怒（1,023 条）和失望（842 条）。

2 研究结果与讨论

2.1 BERT+LDA 模型的性能评估

本文，对评论数据施加了 200 个字符的文本长度限制。根据此限制，模型的最大序列长度参数（Max_Seq_Length）设置为 200。初始学习率配置为 2e-5，训练过程跨越 10 个 epoch。为了在模型性能和计算效率之间实现最佳平衡，本文进行了全面、综合、系统的参数调整。最终，批量大小（Batch_Size）最终确定为 256，代表了批处理的理想大小。训练和测试数据集的准确性和损失值趋势如图 4 所示，说明了模型在不同时期的性能。

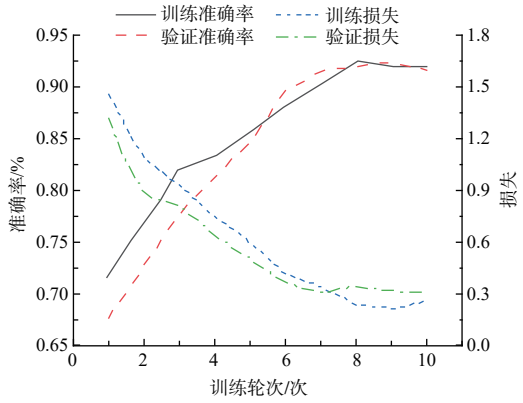

图 4 BERT + LDA 模型的精度和损耗值曲线

为了验证所提出的 BERT+LDA 模型的优势，本文对其与其他各种模型进行了全面评估，评估指标如图 5 所示。

本研究采用准确率、精确率、召回率和 F1 值四项指标评估各模型的情感分类性能，由于是三分类任务，精确率、召回率和 F1 值均采用宏平均计算。各项指标计算公式见式（11）至

doi:10.3772/j.issn.2095-915x.2026.04.007

式（14）。

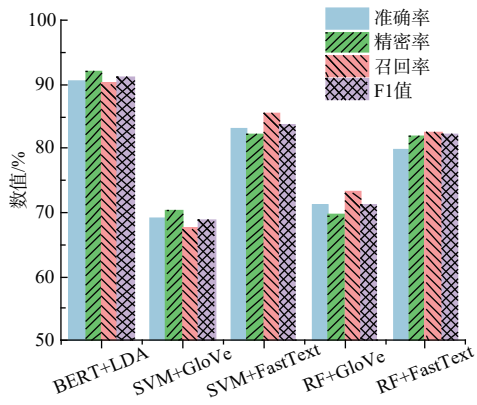

图 5 不同模型的性能比较

三分类准确率直接用“预测正确的样本数 / 总样本数”：

$$Accuracy=\frac{\sum_{i=1}^{k}TP_i}{N} \quad (11)$$

其中，TP_i 为第 i 类的真正例数，k 为类别数（本研究 k=3），N 为总样本数。

精密率表示被预测为正类的样本中真正为正类的比例，反映了模型预测结果的准确性：

$$Precision=\frac{1}{k}\sum_{i=1}^{k}\frac{TP_i}{TP_i+FP_i} \quad (12)$$

召回率表示真实正类样本中被正确预测的比例，反映了模型对正类样本的识别能力：

$$Recall=\frac{1}{k}\sum_{i=1}^{k}\frac{TP_i}{TP_i+FN_i} \quad (13)$$

F1 值是精确率和召回率的调和平均值，综合衡量模型的整体性能：

$$F1=\frac{1}{k}\sum_{i=1}^{k}2\times\frac{Precision_i\times Recall_i}{Precision_i+Recall_i} \quad (14)$$

其中，TP_i、FP_i、FN_i 分别为第 i 类的真正例、假正例和假负例数。宏平均方式将每个类别视为同等重要，不受类别数量不均衡的影响，适合本研究中各类别样本量差异不大的情形。

分析显示，BERT+LDA 模型在所有性能指标上都优于其他模型，准确率和精密率分别为 90.49% 和 91.95%。第二好的模型是 SVM+FastText，而 SVM+GloVe 模型的表现相对较弱。总体而言，在文本分类任务中，与依赖 GloVe 词向量的模型相比，使用 FastText 作为词嵌入向量的模型表现出更好的性能。在分类器选择方面，基于 SVM 的模型通常优于基于随机森林的模型。BERT+LDA 模型的出色性能进一步证实了其在 NLP 任务中的领先地位，而 FastText 和 SVM 的结合也证明了其在特定任务中的有效性。

2.2 基于时间序列的情绪预测

为从整体倾向与内部分化两个维度刻画公众情绪的动态变化，本研究构建了情绪平衡指数（Balanced Index, BI）与情绪分歧指数（Sentiment Divergence Index, SDI）。情绪平衡指数（BI）反映公众情绪的整体倾向，通过积极情绪与消极情绪的相对比例计算，取值范围为 [-1, 1]，正值表示整体情绪偏积极，负值表示偏消极，计算方式见式（15）。情绪分歧指数（SDI）反映公众内部情绪的分化程度，基于三分类情绪分布的归一化信息熵计算，取值范围为 [0, 1]，值越大表示各类情绪分布越均匀、公众意见分歧越显著，计算方式见式（16）。

$$BI=\frac{P-N}{P+N} \quad (15)$$

其中：P 为积极情绪评论数，N 为消极情绪评论数。

$$SDI=\frac{-\left(p_P\ln p_P+p_N\ln p_{N'}+p_M\ln p_M\right)}{\ln 3} \quad (16)$$

其中：p_P=P/T，p_N=N/T，p^M=P/M 分别为积极、消极、中性情绪的占比，T=P+N+M 为总评论数，分母 ln3 是三分类下的最大熵值，用于归一化。

图 6 显示了 34 周期间公众情绪指数的统计结果。一方面，积极情绪指数（BI）在坐标轴

上表现出显著波动，反映了公众情绪态度的重大变化。与此同时，情绪差异指数（SDI）在 0.7 到 1 的范围内略有波动，表明公众之间的情绪差异相对较大。在分析公众情绪时，BI 提供了一个可量化的视角，可以更准确地掌握情绪波动和趋势。另一方面，SDI 的稳定性表明，在解读公众情绪时，由于情绪波动可能显着影响整体情绪解读，因此必须考虑情绪分歧的存在。通过结合这两个指标，可以更全面地了解网民情绪的复杂性和动态变化。

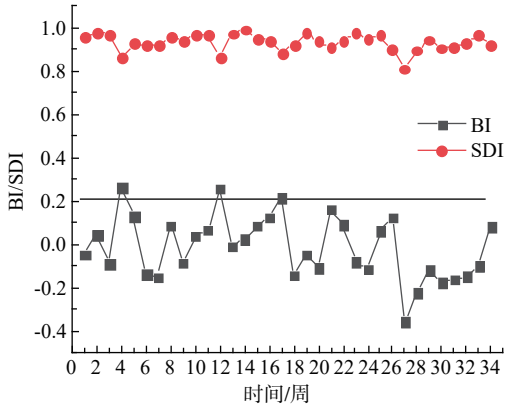

图 6　34 周公众情绪指数统计数据

对于时间序列预测，本研究采用滚动时间窗口的方式进行模型验证：以 2025 年 4 月 1 日至 9 月 30 日共 26 周的数据作为训练集，以 2025 年 10 月 1 日至 11 月 23 日共 8 周的数据作为测试集，对比模型预测值与实际值的误差，验证模型的预测有效性。在此基础上，进一步运用全量 34 周数据对模型进行重新拟合，预测未来 10 周（2025 年 11 月 24 日至 2026 年 1 月底）的公众情绪走向。

本研究采用 ARIMA(0,1,1) 模型对公众情绪进行时间序列分析，研究思路分为两个阶段：模型验证阶段与趋势预测阶段。在模型验证阶段，以 2025 年 4 月 1 日至 9 月 30 日共 26 周的情绪指数数据作为训练集拟合模型，以 2025 年 10 月 1 日至 11 月 23 日共 8 周的数据作为测试集验证预测效果。结果显示，模型在测试集上的预测值与实际值整体趋势较为吻合，表明 ARIMA(0,1,1) 模型能够较好地捕捉公众情绪的常规波动规律，具有一定的预测有效性。在趋势预测阶段，运用全部 34 周数据对 ARIMA(0,1,1) 模型进行重新拟合，对 2025 年 11 月 24 日之后未来 10 周的公众情绪走向进行了外推预测，结果如图 9 所示。预测结果表明，未来一段时间公众对大语言模型的整体情绪仍将保持上升态势，反映出社会层面对大模型技术的接受度持续提升。

需要指出的是，ARIMA 模型也存在一定局限性。一方面，ARIMA 模型本质上是线性外推，对突发事件引起的情绪突变预测能力有限。例如，2025 年 7 月 9 日 Grok 4 的发布引发了舆论的广泛关注与讨论，导致当周情绪指数出现明显波动，这类由外部冲击引起的异常变化难以被纯时序模型提前捕捉。另一方面，随着预测周期的延长，模型的不确定性会逐渐累积，预测误差通常呈扩大趋势，这也是时间序列预测模型的普遍特征。因此，在实际应用中，应结合事件监测与舆情研判对 ARIMA 预测结果进行动态修正，以提升预测的准确性与实用性。

根据 BI 指数预测，未来用户对大语言模型的情绪将呈上升趋势，这表明更多的用户可能倾向于尝试将其融入到他们的日常生活和工作中。对大语言模型的积极情绪可能由多种因素驱动，包括技术进步、用户体验的改进和社会认可的提高，所有这些都有助于大语言模型在用户中越来越受欢迎和接受。

doi:10.3772/j.issn.2095-915x.2026.04.007

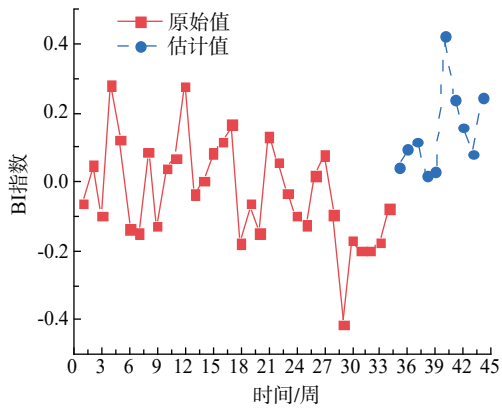

图 7　基于 ARIMA 模型的公众情绪指数预测结果

上述时间序列分析结果表明，公众对大语言模型的情绪整体呈上升趋势（BI 指数持续走高），但情绪分歧指数（SDI）始终维持在较高水平，且重大技术事件（如 Gork4 发布）会引发情绪的剧烈波动。这一发现对智慧城市治理具有重要启示：

一方面，公众对大模型接受度的整体上升，为其在智能客服咨询、公共安全预警、交通管理等智慧城市场景中的深入应用奠定了社会基础，管理者可顺势推动大模型技术在城市服务中的落地，以提升治理效率与服务体验；

另一方面，情绪的高度分歧与突发波动也提示，大模型在智慧城市中的应用推进不宜一蹴而就，需密切关注公众舆情走向，特别是在技术迭代、政策出台等关键节点做好预期引导与风险防控，避免因技术应用引发的社会情绪震荡影响城市治理的稳定性。

因此，充分发挥大模型在情感分析与趋势预测方面的技术优势，将公众情绪监测纳入智慧城市治理的常态化机制，实现技术赋能与公众接受度的动态匹配，是提升智慧城市治理水平的重要路径。

3　结论与治理启示

本研究以新浪微博公众对大语言模型的评论数据为基础，构建 BERT+LDA 情感分类模型与 ARIMA 时序预测模型，对公众情绪倾向与动态变化进行了系统分析。实证研究主要得出以下结论：

第一，BERT+LDA 模型在公众情感分类任务中表现优异，准确率和精确率分别达到 90.49% 和 91.95%，验证了深度学习与主题建模相结合在舆情分析场景中的有效性；

第二，公众对大语言模型的整体情绪呈上升趋势，但情绪分歧指数始终维持在较高水平，且重大技术事件会引发情绪的短期剧烈波动；

第三，ARIMA(0,1,1) 模型能够较好地捕捉公众情绪的常规波动规律，但对突发事件引起的情绪突变预测能力有限。

基于上述研究发现，结合智慧城市治理的现实需求，可得到以下策略启示：

（1）运用有效的大型语言模型和智能化技术，推进智慧城市建设和治理。

2025 年城市工作会议明确了新阶段我国城市发展和治理的目标，即建设创新、宜居、美丽、韧性、文明、智慧的现代化人民城市为目标。作为现代化人民城市的重要内涵，智慧是城市建设与治理的重要内容。在我国智慧城市建设过程中，积极适应大数据、人工智能时代的发展趋势，运用有效的大型语言模型和智能化技术，能够有效助推建设便捷高效的智慧城市，不断提升智慧城市建设和治理水平。按照 2025 年《中共中央、国务院关于推动城市高质量发展的意见》要求，加快打造集约统一、数据融合、

高效协同的城市数字底座，完善城市信息模型（CIM）平台。城市数字底座有利于为智慧城市建设过程中使用大型语言模型和智能化技术提供坚实的数字基础。

（2）坚持问题为导向，强化对大型语言模型使用的政府监管和技术治理。

当前，日新月异的科学技术日益渗透到社会生活、工业生产、国家治理的各方面，技术治理成为现代社会和国家面临的新型治理形态和模式。在国家治理的理论和实践场景中，技术治理既包含运用现代技术进行国家治理（用技术治理），也包含对现代技术的有效治理（对技术治理）。发挥技术对国家治理产生的正效应、避免技术对国家治理产生的负效应，加强对现代技术的监管和治理，也是技术治理不可忽视的一个重要方面。本文的情绪预测结果揭示了公众对大语言模型在智慧城市中的应用的看法的一些潜在问题和风险。比如，公众对大语言模型的隐私保护能力表示担忧。政府应加强公众个人隐私保护方面的监管和治理工作，以确保个人信息数据和公共数据的安全。在智慧城市建设和监管中，运用大型语言模型赋能城市治理还会产生新情况、新问题。城市政府和有关部门要及时跟进、科学研判，制定有效的智慧化监管措施，降低公众对大型语言模型使用的担忧和顾虑，提升政府大型语言模型的治理能力。重点要在推进国家安全体系和能力现代化的背景下，加强智慧城市的安全能力建设。在智慧城市建设过程中，推动模型算法、数据资源、基础设施、应用系统等安全能力建设，防范模型的黑箱、幻觉、算法歧视等带来的风险，完善人工智能法律法规、伦理准则等制度规范体系建设。

本文主要关注公众对大语言模型的情感倾向，对具体监管策略的制定仅进行了初步探索。在实践中，政府监管策略的制定涉及法律法规、技术标准、行业规范等多个方面，需要进一步深入和系统地研究。未来的研究在改进监管策略的制定方面，可考虑更具体的因素和变量。

参 考 文 献

[1] PENG Y F, CHEN Q Y, SHIH G. DeepSeek is openaccess and the next AI disrupter for radiology[J]. Radiology Advances, 2025, 2(1): umaf009.

[2] FU X Y, SANCHEZ T W, LI C, et al. deciphering public voices in the digital era: benchmarking ChatGPT for analyzing citizen feedback in Hamilton, New Zealand[J]. Journal of the American Planning Association, 2024, 90(4): 728-741.

[3] 姜文华 . 智慧城市背景下市政工程的技术应用 [J]. 城市建设理论研究 (电子版), 2024, 493(31): 177-179.

[4] 孟凡新 . 人工智能大模型助推制造业高质量发展：作用机理与策略选择 [J]. 改革 , 2025(12): 65-76.

[5] 田林楠 . 情感劳动的 “ 两副面孔 ”—— 从霍克希尔德到哈特与奈格里 [J]. 社会学评论 , 2022, 10(2): 88-103.

[6] 屈廖健 , 刘华聪 . 能力测评转向 : 经合组织学生社会情感能力调查项目研究 [J]. 比较教育研究 , 2020, 42(7): 90-97.

[7] 张庆熊 . 符号互动论的社会理论：以布鲁默为中心的考察 [J]. 社会科学 , 2024, 524(4): 13-23.

[8] TYAGI N, BHUSHAN B. Demystifying the role of natural language processing (NLP) in smart city applications: background, motivation, recent advances, and future research directions[J]. Wireless Personal Communications, 2023, 130(2): 857-908.

[9] CHAYNCEY S A, MCKENNA H P. Creativity and innovation in civic spaces supported by cognitive flexibility when learning with AI chatbots in smart cities[J]. Urban Science, 2024, 8(1): 16.

[10] 冯皓 . 大模型在自然语言处理中的应用方法研究 [J]. 数字通信世界 , 2024, (10): 123-125.

doi:10.3772/j.issn.2095-915x.2026.04.007

[11] CHAUNCEY S A, MCKENNA H P. Creativity and Innovation in Civic Spaces Supported By Cognitive Flexibility When Learning with Ai Chatbots in Smart Cities[J]. Urban Science, 2024, 8(1): 16.

[12] 秦雨飞 , 胡柯 , 陈国信 , 等 . 人工智能驱动下的智慧城市知识图谱研究 [J]. 信息与电脑 , 2025, 37(22): 19-22.

[13] TAN Y, SENADHEERA S, MARASINGHE R, et al. Artificial intelligence and the local government: A five-decade scientometric analysis on the evolution, state-of-the-art, and emerging trends[J]. Cities, 2024, 152: 105151.

[14] 张美菊 . 数字化转型驱动下的智慧城市管理创新研究 [J]. 中国市场 , 2025, 1237(30): 5-8.

[15] NA H, LUO Y P, EI P G. Evaluation of AI models for radiology exam preparation: DeepSeek vs. ChatGPT-3.5[J]. Medical Education Online, 2025, 30(1): 2589679.

[16] GUO D H, CHEN H X, WU R L, et al. AIGC challenges and opportunities related to public safety: A case study of ChatGPT[J]. Journal of Safety Science and Resilience, 2023, 4(4): 329-339.

[17] 徐铖铖 , 周仲昂 . 融合大语言模型的交通数据分析课程辅助教学探索 [J]. 东南大学学报 (哲学社会科学版), 2025, 27(S1): 158-163.

[18] MOHAMMAD A T, HUTHAIFA I. A, MOHAMMED E, et al. Using multimodal large language models (MLLMs) for automated detection of traffic safety-critical events[J]. Vehicles, 2024, 6(3): 1571-1590.

[19] TEMSAH A, ALHASAN K, ALTAMIMI I, et al. DeepSeek in healthcare: revealing opportunities and steering challenges of a new open-source artificial intelligence frontier[J]. Cureus, 2025, 17(2): e79221.

[20] 王雪 . 基于时间序列模型的高水平学科预测研究 [J]. 情报杂志 , 2019, 38(6): 45-49+117.

[21] 岳丽欣 , 周晓英 , 陈旖旎 . 基于 ARIMA 模型的信息构建研究主题趋势预测研究 [J]. 图书情报知识 , 2019(5): 54-63+72.

[22] AI H S, ZHOU Z Q, YAN Y. The impact of Smart city construction on labour spatial allocation: Evidence from China[J]. Applied Economics, 2024, 56(19): 2337-2356.

[23] 朱益平 , 慕钰 , 孙逸宁 . ChatGPT 大语言模型的评论情感分类预测与主题识别研究 [J]. 情报科学 , 2025, 43(5): 43-57.

[24] TUPAYACHI J, XU H W, OMITAOMU O A, et al. Towards next-generation urban decision support systems through AI-powered construction of scientific ontology using large language models—a case in optimizing intermodal freight transportation[J]. Smart Cities, 2024, 7(5): 2392-2421.

[25] LIAN Y, TANG H T, XIANG M T, et al. Public attitudes and sentiments toward ChatGPT in China: A text mining analysis based on social media[J]. Technology in Society, 2024(76): 102442.

[26] HWANG Y S, UM J S, PRADHAN B, et al. How does ChatGPT evaluate the value of spatial information in the 4th industrial revolution?[J]. Spatial Information Research, 2023, 32(2): 187-194.

[27] 杜家兵 , 王晶 , 刘胜强 , 等 . 结合多尺度特征提取和注意力机制的 NLP 数据处理模型 [J]. 兵工自动化 , 2025, 44(10): 94-101.

[28] 曹亚茹 . 认知驱动的情感支持对话生成研究 [D]. 兰州 : 西北民族大学 , 2025: 82.

[29] 张义策 . 数据和语言模型增强的细粒度情感分析研究 [D]. 哈尔滨 : 哈尔滨工业大学 , 2025: 169.

[30] 董政宇 , 张卫丰 , 许蕾 , 等 . 基于代码上下文与大语言模型的自动化合并冲突解决方法 [J]. 计算机学报 , 2025, 48(10): 2468-2486.

[31] CALZADA I. Disruptive technologies for e-diasporas: Blockchain, DAOs, data cooperatives, metaverse, and ChatGPT[J]. Futures, 2023(154): 103258.

[32] 张晓君 . 网络舆情传播行为的心理距离动态感知及导控研究 [D]. 长春 : 吉林大学 , 2024: 65.

[33] 刘伟利 . 基于事件图谱的突发事件复合风险态势感知研究 [D]. 长春 : 吉林大学 , 2025.

(责任编辑：兰天)