

“智慧政务”文本挖掘与应用

摘要

秉持着“群众的事无小事”的以人为本原则，近年来各大政府机构网站都陆续推出了诸如市长信箱、阳光热线、微博微信意见采集等网络问政平台。网络问政平台能够使执政者更加方便、快捷、真实地了解老百姓的意见及诉求，大大提升了群众满意度以及政府的办事效率。但是，伴随着数据量几何式增长，人工的处理手段显然已经跟不上数据增长的需求。建立基于自然语言处理技术的智慧政府系统已然是大势所趋。因此，本研究围绕互联网公开来源的群众问政留言记录以及有关部门的回复记录，运用自然语言处理技术探究其价值。主要研究成果包括以下三大部分：

一：群众留言分类（文本分类） 针对群众留言问题及问题分类总计 9210 条记录，构建了基于 BERT 衍生模型与 ERNIE 预训练模型的短文本多分类模型。尝试了诸如 EDA 文本增强、FGM 对抗训练对数据和模型进行优化，选取了分类效果最佳的 BERT、BERT-CNN、BERT-RCNN 以及 ERNIE 模型，通过五折交叉验证划分了五套训练集和测试集进行训练，并且利用投票法对五折训练之后的组内模型进行融合，组内融合后又进行了组外融合。经实验，优化后的模型在 F-score 分数上有明显的提升。

二：热点问题挖掘（文本聚类及热点问题评价） 针对群众留言文本内容总计 4326 条记录，本研究提出了双重约束多重优化的文本聚类算法：

（1）文本预处理方面，构建自定义分词表以及停用词表，并对留言主题与留言详情进行融合，并选取其中的所有动词及名词作为聚类的文本，一定程度上有效提升了待聚类文本的质量。

（2）聚类约束方面，基于规则提出留言中的市区县等作为大地点约束，基于 BERT-BiLSTM-CRF 命名实体识别模型提取出留言中的具体地点作为小地点约束，并针对不同的情况制定了一系列优化方案，最终利用 Single-Pass 短文本聚类算法进行聚类。

（3）聚类优化方面，主要解决了以下的问题：第一，NER 模型抽取地点个数可能不定，需要选取最优小地点；第二，相同小地点的表述文字可能存在些许差别，需要加一层地点聚类；第三，部分小地点并未在留言主题里面出现，需要

额外从留言详情中进行补充；第四，异词同义小地点（如：广铁集团、伊景园滨河苑）的优化处理；第五，聚类阈值方面的优化。

（4）留言问题热度评价体系方面，从内容层面、受众层面以及传播层面构建细粒度留言热度评价体系指标，并通过层次分析法确定指标权重，得出热度TOP5的留言热点问题。

三：答复意见的评价（留言答复质量评价） 针对留言答复详情总计2816条记录进行质量评价。从及时性、相关性、完整性、礼貌性、可解释性、逻辑性及专业性七个角度进行指标量化并赋予权重，并对特殊留言进行特殊处理，最终得出留言答复意见评价体系。

关键词：智慧政务；BERT模型；命名实体识别；细粒度评价

Text mining and application of "intelligent government affairs"

Abstract: Adhering to the people-oriented principle, in recent years, websites of major government agencies have successively launched online political platforms such as mayor's mailbox, sunshine hotline, weibo, WeChat opinion, etc. The network platform for political inquiry can make it more convenient, fast and true for the people to understand their opinions and demands, which greatly improves the satisfaction of the people and the efficiency of the government. However, with the geometric growth of data volume, manual processing methods are obviously unable to keep up with the growing demand of data. Therefore, the establishment of intelligent government system based on natural language processing technology is an inevitable trend. This study focused on the records of people's political messages and the replies of relevant departments from the open sources of the Internet, and explored their value by using natural language processing technology. The main research results include the following three parts:

Question 1: Classification of people's comments. A multi-classification model of short texts based on BERT derived model and ERNIE pre-training model was constructed for the total of 9210 records of message questions and problem classification. The research tried to optimize the data and models such as EDA text enhancement and FGM confrontation training, and selected BERT, BERTt-CNN, BERT-RCNN and ERNIE models with the best classification effect. Then, five sets of training sets and test sets were divided by 5 fold cross validation for training, and the in-group model after training was fused by voting method, and then the out-group model was fused.

Question 2: Hot issues mining. In this study, a text clustering algorithm with double constraints and multiple optimizations was proposed for a total of 4,326 records of the text contents of the mass message:

(1) In terms of text preprocessing, we built a custom word segmentation list and stop word list, integrated the message topic and message details, and selected all the verbs and nouns as the clustering text, which effectively improved the quality of the text to

be clustered to some extent.

(2) In terms of clustering constraints, cities and counties in the comments were proposed as large site constraints based on rules, and specific sites in the comments were extracted as small site constraints based on the bert-bilstm-crf named entity recognition model. A series of optimization schemes were developed for different situations. Finally, single-pass short text clustering algorithm was used for clustering.

(3) In terms of clustering optimization, the following problems are mainly solved: first, the number of sites extracted by NER model may be variable, so it is necessary to select the optimal small sites; Secondly, there may be some differences in the expressions of the same small sites, so it is necessary to add a layer of location clustering. Third, some small places do not appear in the message subject, need to be added from the message details; Fourth, different words are synonymous with small sites (such as: guangzhou railway group, yijingyuan binhe garden) optimization; Fifth, the optimization of clustering threshold.

(4) In terms of the heat evaluation system of message message, the index of fine-grained heat evaluation system of message message was constructed from the content level, the audience level and the communication level, and the index weight was determined by the analytic hierarchy process to get the TOP5 hot message message issues.

Question 3: A total of 2816 records were evaluated for reply details. From the perspective of timeliness, relevance, integrity, courtesy, interpretability, logicity and professionalism, the index is quantified and weighted, and special comments are processed, finally the comment response evaluation system is obtained.

Keywords: Intelligent government; BERT model; Named entity recognition; Fine-grained evaluation.

目 录

1 引言	1
1.1 问题背景.....	1
1.2 挖掘目标.....	1
1.3 挖掘流程.....	2
2 群众留言分类	4
2.1 文本分类模型构建.....	4
2.1.1 基于 BERT 的文本分类模型.....	4
2.1.2 融入卷积神经网络及周期循环神经网络.....	6
2.1.3 基于 ERNIE 的文本分类模型.....	7
2.2 尝试过的数据及模型改进方案.....	8
2.2.1 基于 EDA 的文本增强	8
2.2.2 FGM 对抗训练层优化.....	9
2.2.3 交叉验证后模型组内组外融合.....	9
2.3 实验结果与评估.....	10
2.3.1 实验评价指标.....	10
2.3.2 实验结果与比较.....	11
3 留言热点问题挖掘	12
3.1 双重约束多重优化聚类模型构建.....	12
3.2 Single-Pass 聚类算法.....	13
3.2.1 文本特征提取.....	13
3.2.2 词频-逆文本频率 TF-IDF.....	13
3.2.3 相似度计算.....	13
3.2.5 Single-Pass 算法.....	14
3.3 双重约束.....	14
3.3.1 基于规则的大地点提取.....	14
3.3.2 基于 BERT-BiLSTM-CRF 的小地点提取	14
3.4 多重优化.....	17
3.4.1 小地点优化.....	17
3.4.2 聚类优化.....	18
3.5 留言热度评价体系构建.....	19
3.5.1 内容层面.....	20
3.5.2 受众层面.....	21
3.5.3 传播层面.....	22
3.5.4 指标权重确定.....	22

3.6 热点问题结果展示.....	25
4 答复意见评价	27
4.1 答复意见评价体系构建.....	27
4.1.1 指标选择及量化.....	27
4.1.2 指标权重确定.....	30
4.1.3 特殊情况处理.....	30
4.2 效果评价.....	31
5 总结与不足	34
参考文献	35

1 引言

1.1 问题背景

大数据时代,网络逐渐成为政治生活的重要场域,直接影响着政府与公民之间的互动模式。网络问政作为我国互联网普及和应用中产生的新的政治民主发展途径,不仅拓展了我国公民政治参与的空间,成为社会各个阶层思想自由表达的新平台,而且已成为信息化时代党政领导干部沟通于民、汇聚民智的重要渠道。微信、微博、市长信箱、阳光热线等网络问政平台作为一种新的民主政治形式,可以有效搭建公民与政府的沟通桥梁,实现网络空间内的公民诉求表达与政府回应。对于推动我国社会主义民主政治发展,构建社会主义和谐社会具有重要意义。

近年来,公民通过网络表达民意和诉求的规模大幅增长,各类社情民意相关的文本数据量不断攀升。面对日益增长的网络参政,地方政府也在大力强化政府回应性建设。但当前我国网络问政制度化还处于初始阶段,网络问政发展过程中存在诸多有待改进的地方,在此过程中各种问题纷涌而至,这给以往主要依靠人工来进行留言划分和热点整理的相关部门的工作带来了极大挑战。为保障网络问政平台的平稳运行和制度化建设,相关政府需要在把握网络问政阶段性特征的基础上,从制度规范的角度出发,对网络问政存在的问题进行分析并制定相应的完善对策,以保障网络问政的常态化和长效化,同时提升市民的满意度,推动我国政治民主进程。

目前,大数据、云计算、人工智能等技术的飞速发展给网络问政平台的管理带来了新的契机,建立基于自然语言处理技术的智慧政务系统已经是社会治理创新发展的新趋势,对提升政府的管理水平和施政效率具有极大的推动作用。基于此,通过互联网公开来源的群众问政留言记录及相关部门对部分群众留言的答复意见,使用自然语言处理和文本挖掘的方法可以有效解决网络问政制度化的缺失问题,挖掘出市民真正关心的热点问题并据此制定相应的政策,还可以实现对网络问政回应的考核、监督,做到有问必答、有求必应,以保障网络问政的规范化和智慧化运行。

1.2 挖掘目标

自互联网公开来源的群众问政留言记录以及相关部门对部分群众留言的答复意见作为数据源,本文将利用自然语言处理和文本挖掘等方法逐一解决以下三大核心问题:

(1)群众留言分类:目前,网络问政平台仅依靠人工根据其经验处理问题,普遍存在工作量大、效率低,且差错率高等难题。本研究致力于通过科学的划分体系对群众留言进行精准分类,并建立关于留言内容的一级标签分类模型。通过

该步操作有利于后续将群众留言分派至相应的职能部门进行针对性和专业化的处理,既可以提高工作人员的执行效率,也能有效提升群众满意度。

(2) 留言热点问题挖掘: 热门问题是多数群众比较关注的问题, 可以理解为某一时段内群众集中反映的某一问题。如相关研究表明, 在公民的网络诉求表达中, 经济发展类议题关注度最高, 民生福利类议题次之, 然后是国土建设类、农村类和贪腐类议题^[1]。本研究的重点不仅涉及到群众关注热门问题的方向, 更重要的是将挖掘深度化和具体化——即挖掘出群众真正面临且亟需解决的实际问题。比如, 通过文本聚类算法对群众问题聚类, 并从制定合理的热度评价体系, 发现“某小区入夏以来小区楼下烧烤店深夜经营导致噪音和油烟扰民”问题的热度指数很高, 那么政府相关部门即可迅速针对该热门问题进行重点且高效的整改和管理, 真正为老百姓解决问题。因此该研究旨在及时有效地挖掘群众真正关心或困扰的热门问题, 有助于相关部门进行有针对性地管理和制度制定, 以提升服务效率, 积极推进服务型政府网络问政平台建设。

(3) 答复意见评价: 公民网络问政环境下, 政府回应能力是保障政府与公民的良性互动的重要基础。由于我国网络问政处于初级阶段, 目前平台的运行状态仍存在许多不足之处, 如回应机制不健全、缺乏监督机构、相关法律法规不完善、诉求类别不细化等。政府回应则存在回应不及时、不积极、缺乏人格化回应、信息相对封闭、对诉求的分辨能力存在缺失等问题。政府工作人员需要通过厘清相关概念及理论, 突破“瓶颈状态”, 避免陷入“回应性陷阱”的标准^[2]。因此, 本研究通过构建具有完备理论基础的答复意见评价体系, 从答复的及时性、相关性、完整性等多个角度对答复意见的质量给出一套合理的评价方案, 以监督和改善政府回应能力, 整合和提升网络问政平台公共服务水平, 最终为找寻规避政府回应失效的对策提供参考。

1.3 挖掘流程

挖掘流程将根据上述的三个挖掘目标分别进行科学制定, 以形成具有强大理论支撑的针对性挖掘体系, 以保证方案的准确性、高效性和科学性。实验的总流程图如图 1 所示。

首先, 针对群众留言分类, 也就是文本分类问题, 本研究构建了基于 BERT 与 ERNIE 预训练模型的短文本多分类模型, 并且融入 Attention 机制对文本的语义层面向量化进行优化。与此同时, 选取效果较好的 BERT、BERT-CNN、BERT-RCNN 以及 ERNIE 模型采用五折组内融合及组外融合的方法对最终的预测结果进行优化, 以保证本研究所用方法的高性能和适用性。

其次, 针对热点问题挖掘, 本研究主要从文本聚类和热点问题评价两个主要任务入手。首先, 通过构建自定义分词表以及停用词表, 并且将留言主题与留言详情进行融合以达到文本增强的效果后, 进而取得其中的所有动词及名词作为聚

类的文本。在完成文本预处理的前期工作后，本研究创新性地提出了双重约束多重优化的文本聚类算法，即通过规则提取市县区作为聚类的大地点约束，通过BERT-BiLSTM命名实体识别模型提取出具体地点作为小地点约束并进行地点优化处理，最终对聚类结果进行最终优化以得出完整的解题方案。此外，在参考相关高质量文献形成强大的理论基础后，在本研究的实际问题中，有依据地选择从内容层面、受众层面以及传播层面构建细粒度留言热度评价体系指标，并通过层次分析法确定指标权重以量化问题热度，最终取热度TOP5的留言热点问题。

最后，针对答复意见评价问题，也就是建立合理的指标体系对留言答复详情进行质量评价。在大量查阅相关文献和政府官方文件后，研究考虑到留言问题和平台服务范畴之间的误差，将问题的普适性和特殊性相结合，从及时性、相关性、完整性、礼貌性、可解释性、逻辑性及专业性七个角度进行指标体系的构建，并对特殊留言进行特殊处理，最后在指标量化时针对特殊情况赋予更为合适的权重，以得出科学性和完备性俱佳的留言答复意见评价体系。

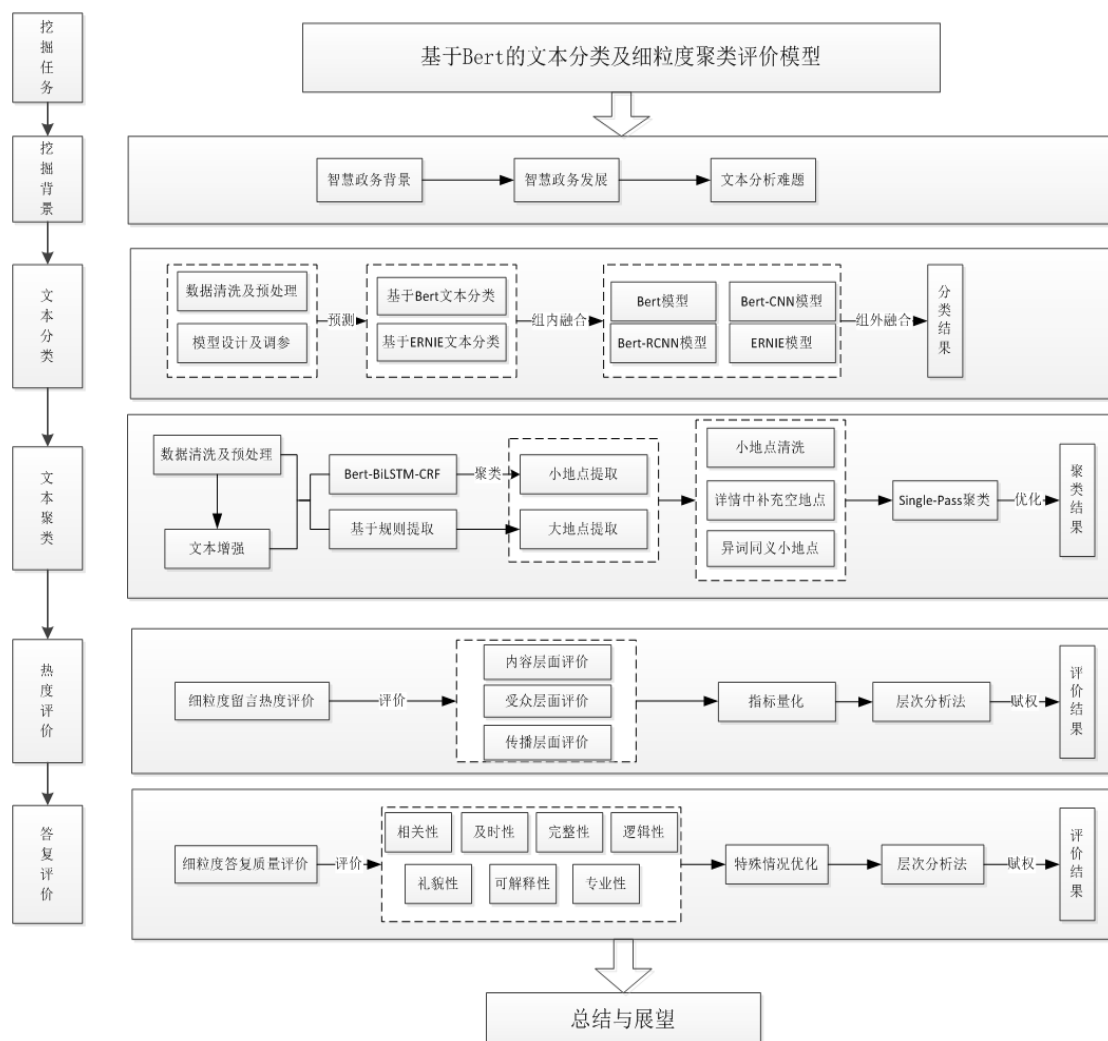


图1 挖掘总流程图

2 群众留言分类

本章主要解决的问题是留言短文本多分类问题，基于附件 1 的三级分类体系选取了 BERT、BERT-CNN、BERT-RCNN 以及 ERNIE 预训练语言模型对留言进行分类，在经过诸如 EDA 文本增强、FGM 对抗训练层优化等诸多失败尝试后，最终通过模型融合取得了较大的效果提升。旨在通过对留言进行有效分类的前提下，为相关政府问政平台的有效性、实时性留言管理提供便捷条件。

2.1 文本分类模型构建

近年来，随着深度学习的快速发展，面向自然语言处理领域的预训练技术获得了长足的进步。早期的自然语言处理领域长期使用 Word2Vec 等词向量方法对文本进行编码，这些词向量方法也可看作静态的预训练技术。然而，这种上下文无关的文本表示给其后的自然语言处理任务带来的提升非常有限，并且无法解决一词多义问题。因此，为解决此类问题，近年来 NLP 任务中的预训练语言模型如雨后春笋般喷涌而出，诸如 BERT、GPT、XLNet、ERNIE 等预训练语言模型在各种 NLP 任务中起到了决定性的提升。本研究在进行文本分类模型构建时，为确保方法的先进性和适用性，先后选取了诸如 Text-CNN、DPCNN、Text-RNN、Text-RCNN 等常规的深度学习模型，以及 ERNIE、BERT 及其改进模型等预训练语言模型作为初始方案。最终进行实验效果对比，选取了 ERNIE、BERT、BERT-CNN 以及 BERT-RCNN 作为基础模型，并通过组内组外融合对预测结果进行优化，以找到适用性和泛化性能俱佳的文本分类模型。

2.1.1 基于 BERT 的文本分类模型

由于单向语言模型对语义信息的建模能力是有限的，因此建立一个基于 Transformer 的双向预训练语言模型是一种重要的研究思路。BERT（Bidirectional Encoder Representations from Transformers）正是使用了一种特别的预训练任务来解决这个问题。BERT 通过堆叠 Transformer 子结构来构建基础模型，并且采用了多头注意力机制，可以保证文本词向量编码嵌入的过程中仍然可以一定程度上维持原有的上下文语义关系。模型结构如图 2 所示，通过 Masked-LM 这个特别的预训练方式达到了真双向语言模型的效果。

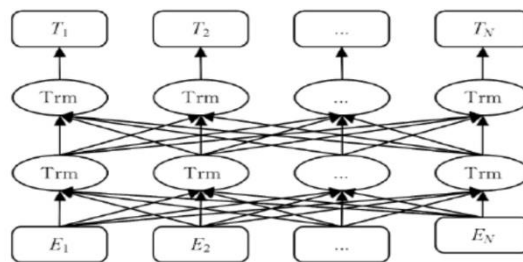


图 2 BERT 模型结构图

BERT 的内部结官网提供了两个版本，L 表示的是 transformer 的层数，H 表示输出的维度，A 表示 mutli-head attention 的个数。虽然模型的层数很大，但是由于 transformer 的 residual 模块，层数并不会引起梯度消失等问题。

$$BERT_{BASE}: L = 12, H = 768, A = 12, TotalParameters = 110M$$

$$BERT_{BASE}: L = 24, H = 1024, A = 16, TotalParameters = 340M$$

预训练模型 BERT 采用了 Transformer 编码器的模型作为语言模型，第一个阶段是语言模型的预训练，共有两个预训练任务：完形填空（Masked Language Model），它基于词语序列中其他没有遮蔽的单词提供的上下文，预测被遮蔽的词语；下一句预测（Next Sentence Prediction），在句子层面充分挖掘句子层面的语义信息，即预测第二句话是否是第一句话的下一句话；第二个阶段是基于微调的方式去解决下游任务。

当下游任务为无监督任务时，BERT 的架构无法很好地捕捉语料间的关系。故本文中采用结合注意力机制（Attention）以满足下游任务的需求。注意力机制无论是在计算机视觉领域还是在自然语言处理领域均有突出的表现。在 BERT 上采用注意力机制可以对模型中的每一层赋以不同的权重，促使模型找到对效果提升最明显的部分。BERT 的输入语料嵌入编码由词嵌入编码，段嵌入编码和位置嵌入编码三种嵌入编码求和得到。基于注意力机制的词嵌入编码由词嵌入编码，段嵌入编码和位置嵌入编码三种嵌入编码加权求和得到，如图 3 所示。在添加注意力机制的 BERT 对提升模型综合评价分数有帮助。

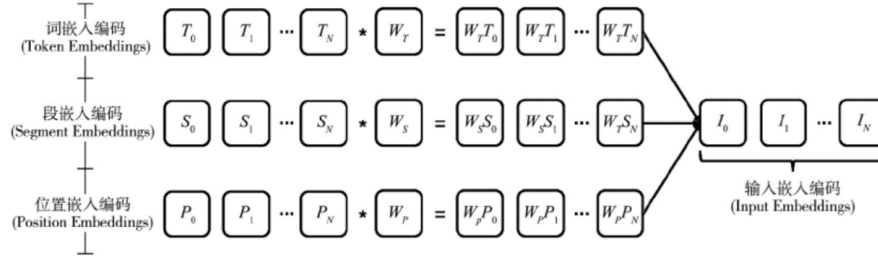


图 3 BERT 输入表征上的注意力机制

自 BERT 等预训练语言模型相继被提出，便逐渐进入了动态预训练技术的时代。BERT 击败了 11 个典型下游任务的 State-of-the-art 结果之后，已成为自然语言处理领域不可或缺的主流技术。这是自然语言处理领域预训练技术的重要里程碑，极大地推动了自然语言处理领域的发展。预训练的好处在于在特定场景使用时不需要用大量的语料来进行训练，节约时间成本，具有高效性，BERT 就是一个具有代表性的泛化能力较强的预训练模型。它开启了 NLP 领域的预训练学习，可以通过加深网络的方式增强对文本信息的挖掘能力，大大提升了模型的训练效率。相比 Word2Vec 等词嵌入模型，它可以捕获句子级别语义特征以及上下文语境。另外，它是基于无监督的语料数据进行学习的，可以减少数据搜集和人工标注的成本。这也正是本研究选择使用该模型进行文本向量化表示的诸多优点，以最终完成文本分类任务。

2.1.2 融入卷积神经网络及周期循环神经网络

BERT 预训练语言模型以其丰富的训练语料库以及融合了多层双向 Transformer 与多重注意力机制的训练结构能够有效提升词嵌入的效果。BERT 的输入语料嵌入编码由词嵌入编码,段嵌入编码和位置嵌入编码三种嵌入编码求和得到,一定程度上可以有效解决上下文语义缺失的问题。

因此,将 BERT 作为 Embedding 层来嵌入文本内容,并基于 BERT 的 Embedding 层嵌入到基础深度学习网络中,可以达到双赢的效果,既维持了网络的多样性,与此同时也能够解决普通词嵌入的弊端,能够进一步优化结果。本研究通过 BERT 获取文本向量,并接入 CNN 与 RCNN 网络中,从而建立了 BERT-CNN 和 BERT-RCNN 模型,效果均比基于 word2vec 训练文本向量所构建的 TEXT-CNN 与 TEXT-RCNN 模型的效果有所提升。

卷积神经网络 CNN 应用到文本分类任务中,可以利用多个不同 size 的 kernel 来提取句子中的关键信息,从而能够更好地捕捉局部相关性。图 4 是 TEXT-CNN 模型结构。

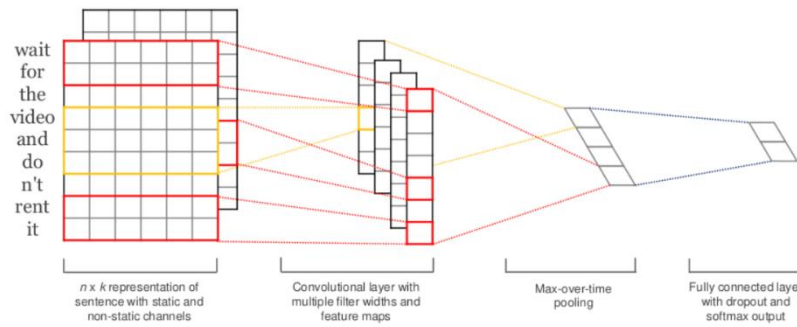


图 4 TEXT-CNN 基本网络结构

如图所示,其主要的网络结构分为 Embedding 层、Convolution 层、MaxPolling 层以及 FullConnection 全连接层和 Softmax 层来输出不同类别的概率^[3]。而 BERT-CNN 是将 Embedding 层换成了基于 BERT 的 Embedding 层。

周期循环神经网络中的双向循环结构与卷积神经网络相比能更好地减少噪声,利用最大池化层选取一句话中的最重要的特征。图 5 是 TEXT-RCNN 结构。

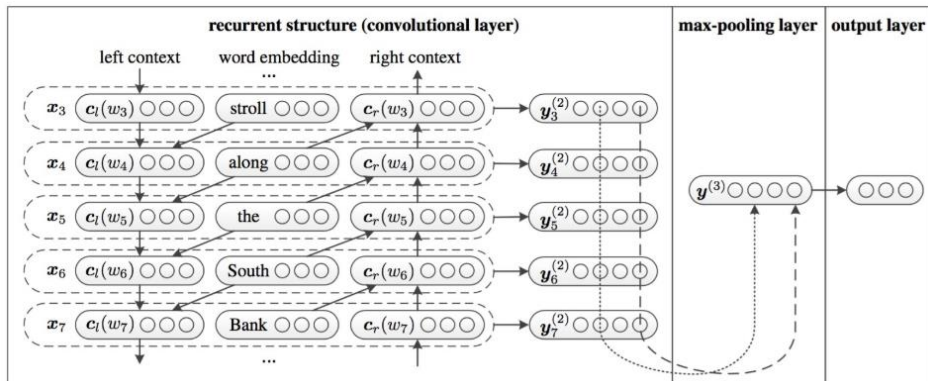


图 5 TEXT-RCNN 基本网络结构

如图所示，其主要网络结构与 TEXT-CNN 相比主要增添了双向循环结构，主要是对基于 BERT 的 Embedding 层的一种改进，经过双向 LSTM，该词的左侧的词正向输入进去得到一个词向量，该词的右侧反向输入进去得到一个词向量。再结合该词的词向量作为最终的输入结果^[4]。

2.1.3 基于 ERNIE 的文本分类模型

BERT 模型通过随机屏蔽 15% 的字或者 word 训练 Masked-LM，利用多层双向 Transformer 的建模能力，在各项 NLP 下游任务中(如 sentence pair classification task、single sentence classification task、question answering task) 都取得了很好的成绩。但 BERT 模型主要建模原始的语言内部的信号，聚焦在针对字或者英文 word 粒度的完形填空学习，没有充分利用训练数据当中词法结构、语法结构以及语义信息去学习建模，导致模型很难学出语义知识单元的完整语义表示。这个问题在中文方面尤为明显。比如“我要买华为手机”，BERT 模型将“我”、“要”、“买”、“华”、“为”、“手”、“机”每个字都统一对待进行随机 mask，这样就极可能丢失了“华为手机”这个很火爆的名词信息，属于词法信息的缺失。同时“我 + 买 + 名词”是一个非常明显的购物意图的句式，BERT 没有对此类语法结构进行专门的建模。如果预训练的语料中只有“我要买华为手机”，“我要买小米手机”，在遇到新的手机品牌出现时，它在预训练的语料中并不存在。此时没有基于词法结构以及句法结构的建模，对于新词很难给出很好的向量表示。

针对 BERT 模型的不足，百度提出基于知识增强的 ERNIE 模型，通过建模海量数据中的实体概念等先验语义知识，学习语义知识单元的完整语义表示。本节主要介绍 ERNIE 模型，它通过引入知识，使预训练模型学习到海量文本中蕴含的潜在知识，进一步提升了预训练语言模型在各个下游任务中的效果。

ERNIE 的模型结构与 BERT 基本一致，不同点在于 BERT 是对字进行随机掩码，强行掩盖把词与词、字与字之间的关系给拆散了。而 ERNIE 通过掩码词和实体概念等完整语义单元来训练 Masked-LM，从而使得模型对语义知识单元的表达更贴近真实世界。因此，相较于 BERT 基于局部词语共现学习的语义表示，ERNIE 通过对训练数据中的词法结构、语法结构和语义信息进行统一建模，增强了模型的语义表示能力，在多项任务中均取得了大幅度超越 BERT 的效果。以下对 BERT 与 ERNIE 模型的效果进行了对比说明。

表 1 BERT 与 ERNIE 模型的对比

模型	哈	尔	滨	是	黑	龙	江	省	会
BERT	哈	X	滨	是	X	龙	江	省	会
ERNIE	X	X	X	是	黑	龙	江	省	会

表 1 可见，在 BERT 模型中，通过『哈』与『滨』的局部共现，可以判断出『尔』字，然而模型没有学习与『哈尔滨』相关的知识。而 ERNIE 通过学习

词与实体的表达，使模型能够建模出『哈尔滨』与『黑龙江』的关系。ERNIE 通过全局信息去预测 mask 掉的内容，学习序列里 mask 信息里蕴含的知识，已经进行了大量的自然语言处理的分词模型，短语拼接的模型，命名实体识别的模型，能够提前把这些词或者短语给标注出来，标注出来之后再学，这样看它是基于已有的策略再去进行海量的数据上训练出一个很好的模型。ERNIE 优势在于通过学习实体概念知识，可以获得知识单元的完整语义表示，在自然语言推断、语义相似度、命名实体识别、情感分析、问答匹配任务的都取得了较好的效果。

2.2 尝试过的数据及模型改进方案

本研究在选定了基础模型之后，尝试了各种优化的方案，旨在提升模型的预测效果。如：基于分类词典与基于 EDA 的文本增强策略、基于 FGM 对抗训练网络层的优化以及五折交叉验证后的模型组内组外融合。文本增强和增加对抗训练层均没有取得较大的提升，最终选择了模型融合作为优化方案。

2.2.1 基于 EDA 的文本增强

数据增强技术在计算机视觉领域有着广泛的研究和应用，主要应用于训练数据不足的任务场景中，通过增加训练数据量能够提升模型的泛化能力。在自然语言处理任务中，数据增强技术主要包括回译（Back translation）、基于语言生成模型的文本增强以及 EDA(Easy data augmentation)技术。本文主要尝试运用了 EDA 技术来对增强题目文本数据量。

EDA（Easy data augmentation）是 19 年 Jason W 等人的研究团队总结提出的系统性文本增强策略^[5]。EDA 主要包含以下四种操作：

- （1）同义词替换 (SR)：从句子中随机选择非停止词。用随机选择的同义词替换这些单词。
- （2）随机插入 (RI)：随机的找出句中某个不属于停用词集的词，并求出其随机的同义词，将该同义词插入句子的一个随机位置。重复 n 次。
- （3）随机交换 (Random Swap, RS)：随机的选择句中两个单词并交换它们的位置。重复 n 次。
- （4）随机删除 (RD)：以概率 p 随机删除句子中每个单词。

原文本：

个人对于西医药的疑惑 5
K10县政务中心至职中路段路灯不亮 0
E8县步行街一个锦潼电位治疗的门店涉嫌诱骗老年人高额消费 3
L5县葛竹坪中心小学学生的负担很重 2
建议撤销E10县回龙镇红星小学 2
建议开发A6区县洪山村洪山古庵为旅游景点 3
关于A市C5市路绿化改造的建议 0
I2区兰溪乡羊角村村民咨询农村危房改造补贴事项 0

EDA 增强后文本：

与生俱来对于西医药的疑惑 5
K10县治政务中心至职中路段北路灯不亮 0
E8县街区一个锦潼电流治疗的门店涉嫌诱骗成年人高额消费品 3
L5县葛竹坪中心小学师生的经济负担很重 2
建议划归E10县回龙镇红星初中 2
建议开发计划A6区县石门头村石门古庵为旅游景点 3
关于A市C5市路绿化翻修的同意 0
I2区兰溪乡羊角村村民发表意见农村危房改造补贴要点 0

图 6 EDA 文本增强对比

以上是针对附件 2 留言文本分类数据进行 EDA 文本增强后的文本部分效果对比。本研究采用此种文本增强方法，双倍扩充了训练集，但是最终的预测结果准确率平均降低了千分之五左右，究其原因这是由于训练数据的数据量并未明显不足，一般训练数据量在 1000 条以下效果均有所提升。故不采用此优化方案。

2.2.2 FGM 对抗训练层优化

提到对抗训练，多数人的第一反应都是 CV 中的对抗生成网络 (GAN)，殊不知，其实对抗也可以作为一种防御机制，通过对抗训练，可以有效提高模型鲁棒性和泛化能力，基本原理：在原始输入上增加对抗扰动，得到对抗样本，再利用对抗样本进行训练，从而提高模型表现。

在 NLP 问题中，输入是离散的单词序列，一般以 one-hot vector 的形式呈现，如果直接在 raw text 上进行扰动，那么扰动的大小和方向可能都没什么意义。Goodfellow 等人^[6]在 2017 年的 ICLR 中提出了可以在连续的 embedding 上做扰动，把对抗扰动添加到嵌入层中，为了最大化对抗样本的扰动能力，使用梯度上升方式生成对抗样本。

NLP 任务中的对抗训练对 Embedding 层进行扰动主要有两种方法，主要是 FGM (Fast Gradient Method) 及 PGD (Projected Gradient Descent)。本文主要采用 FGM 在模型训练过程中对 Embedding 层进行了扰动，但是效果反而下降了很多，平均下降了 2%-3%。分析其中的原因，一般在模型中通过对抗训练对 Embedding 层进行扰动主要用于处理领域不同的数据集，本题中的数据集都属于同一个领域，且数据较为干净，没有明显的数据类型的区别，因此对抗训练扰动的作用不是很大，故不采用此优化方案。

2.2.3 交叉验证后模型组内组外融合

模型融合可以有效提升模型的预测效果，可以有效避免单个模型容易出现的过拟合现象且提高模型的泛化能力预测同时可以改进单个模型预测能力不高的缺点。对于本题而言，题目所给出的数据量不是很大，因此需要充分利用数据进行训练，这样才能够做到模型的训练效果最大化。

K 折交叉验证作为机器学习领域常用的数据集划分方法，可以避免由于训练集与测试集的划分不同从而导致的模型训练效果不佳以及随机划分测试集导致的预测结果评估有随机性的问题。不同于普通的机器学习模型，本研究所采用的深度学习模型需要同时划分训练集、验证集以及测试集，验证集主要是起到在模型的训练过程中不断评估调整网络的权值的作用。下图是使用 K-fold 划分训练集、验证集以及测试集的方法。

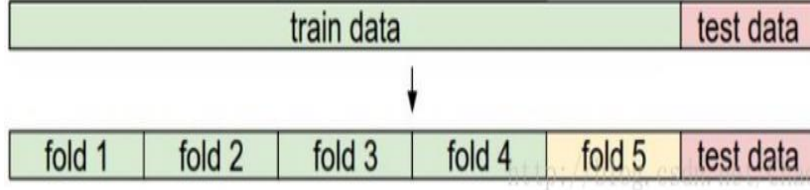


图 7 五折交叉验证切分数据集

上图表示我们拥有的数据，而后我们对数据进行了再次分割，主要是对训练集，假设将训练集分成 5 份（该数目被称为折数，5-fold 交叉验证），每次都用其中 4 份来训练模型，粉红色的那份用来验证 4 份训练出来的模型的准确率，记下准确率。然后在这 5 份中取另外 4 份做训练集，1 份做验证集，再次得到一个模型的准确率。直到所有 5 份都做过 1 次验证集，也即验证集名额循环了一圈，交叉验证的过程就结束。

不过这样划分也有缺点，就是粉红色的 test-data 这部分数据是不能参与训练的，而且 test-data 的划分也具有一定的随机性。因此本文直接采用 5 折交叉验证将数据划分为五套训练集和验证集，利用五套训练集和验证集分别训练同一模型，每个模型可以得出 5 种不同权重的模型参数。

模型融合方面采用投票法对预测结果进行融合优化，先基于选定的基础模型（BERT、BERT-CNN、BERT-RCNN、ERNIE）分别用 5 折交叉验证划分的不同的训练集和验证集进行训练，每个模型得出 5 种不同的权重，然后对同一个模型的不同权重的预测结果进行组内融合，总共得出 4 个不同模型组内融合后的结果，最后再对四种模型的组内融合预测结果再次进行组外融合，得出我们最终的预测结果。实验证明，经过五折组内融合加组外融合之后，模型预测结果的 F-score 值相比单个模型单次训练的预测结果的 F-score 值高了 2%-2.5%。

2.3 实验结果与评估

2.3.1 实验评价指标

对分类方法进行评价，两个最常见的衡量指标是准确率（给出的结果有多少是正确的）和召回率（正确的结果有多少被给出）。这两个指标通常是此消彼长的，很难兼得。很多时候使用参数来控制，通过修改参数则能得出一个准确率和召回率的曲线（ROC），该曲线与 x 和 y 轴围成的面积就是 AUC。AUC 可以综合衡量一个预测模型的好坏，该指标综合了 precision 和 recall 两个指标。但 AUC 计算相对较为麻烦，实际应用中多使用 F-score 来代替更为方便。

$$F_1 = \frac{1}{n} \sum_{i=1}^n \frac{2P_i R_i}{P_i + R_i} \quad (1)$$

其中 P_i 为第 i 类的查准率， R_i 为第 i 类的查全率。

2.3.2 实验结果与比较

本研究的实验代码均采用 python 语言进行编写，深度学习文本分类模型是基于开源框架 Pytorch 进行编写。运行环境为 Windows 10 操作系统，显卡采用的是网络算力平台的 Tesla T4 16GB 显卡。

实验结果主要分为两个方面，一方面是单个模型进行优化调参之后模型在验证集上的准确率指标。模型在验证集上的效果并不一定可以代表模型在测试集上的效果，但是也能够一定程度上体现出模型的泛化预测能力。结果见下表。

表 2 不同模型验证集上效果对比

模型	验证集上的准确率 (%)
ERNIE	92.24
Bert-RCNN	91.67
Bert-CNN	91.52
Bert	90.78
Text-RCNN	89.47
Text-CNN	89.35
Bert-DPCNN	89.05
Bert-RNN	88.71
Text-RNN	88.23
DPCNN	86.86

另一方面，为了验证模型融合之后对于预测结果的提升，与此同时防止单独切分测试集会导致测试集数据无法得到充分的利用，本研究从全部数据集里面随机抽取 20% 的数据作为测试集来验证模型融合之后的效果。但是，这样做必然会导致模型在预测时产生过拟合，导致总体结果偏高，不过，采用这样的验证方法只是为了对比模型融合前后的相对差距，真正预测的时候数值肯定会比下表的数据低且不会产生过拟合现象。对比结果如下表所示。

表 3 模型融合前后相对效果对比

F-score 模型	F 值 1	F 值 2	F 值 3	F 值 4	F 值 5	F 值均值	组内融合 F 值
Bert	96.52	96.85	95.16	95.87	96.63	96.20	98.04
Bert-CNN	96.47	95.65	94.95	95.00	94.40	95.32	97.45
Bert-RCNN	96.63	97.07	93.53	96.58	97.06	96.18	98.10
ERNIE	96.64	95.71	95.49	95.43	96.31	95.90	97.28
组外融合 F 值	98.21						

3 留言热点问题挖掘

本章重点解决的问题是：对留言热点问题进行双重约束和多重优化，基于 Single-Pass 算法对留言中的地点和留言热点问题进行聚类，并构建留言热度评价体系对热点问题进行评价，以得出有代表性的留言热点问题列表。

3.1 双重约束多重优化聚类模型构建

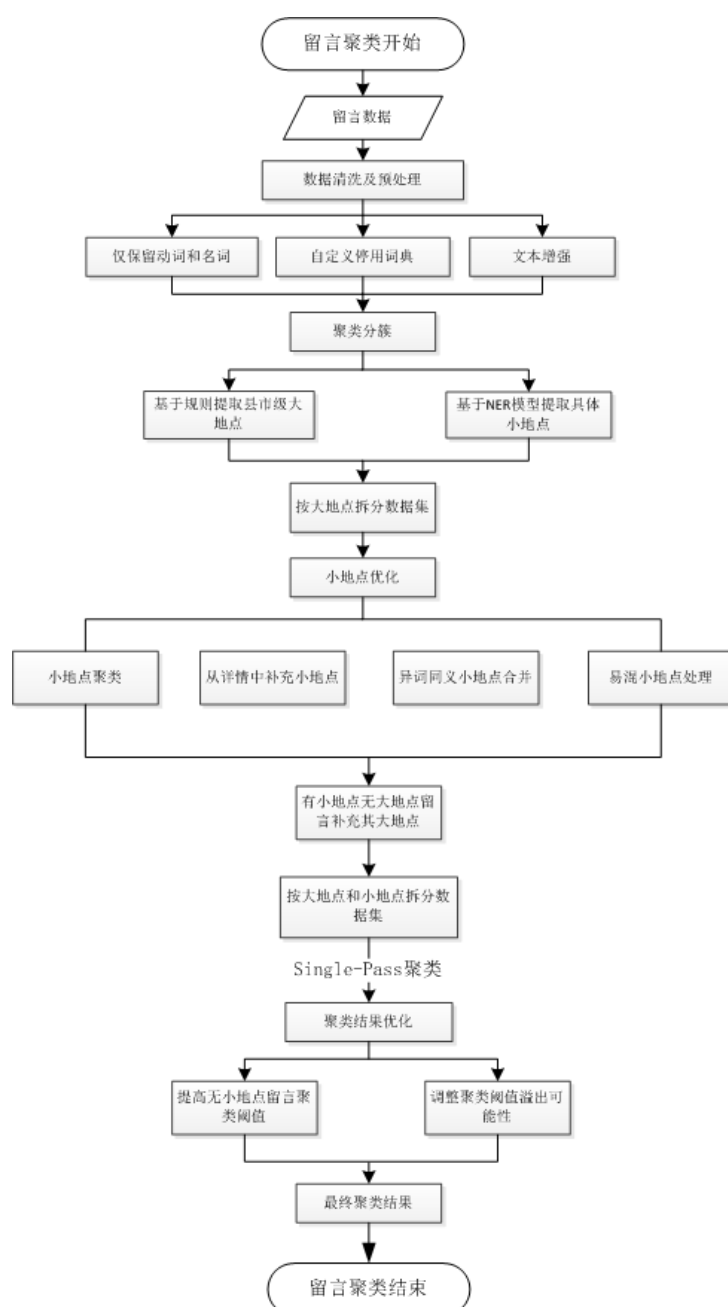


图 8 热点问题聚类算法流程图

3.2 Single-Pass 聚类算法

3.2.1 文本特征提取

文本的表征有诸多方法，如布尔模型、向量空间模型、概率模型等，其中向量空间模型是在应用中广为采用的模型，首先被用于信息检索系统。通常，文档被表示为向量，每一维均对应独立的词。每篇文档，均可以表示为规范化的特征向量：

$$d = \{(t_1, w_1), (t_2, w_2), \dots, (t_n, w_n)\} \quad (2)$$

其中， t_i 表示第 i 个特征项； w_i 表示特征项 t_i 在文本 d 中的权重，所有的文本向量构成文本集的一个特征向量。文本向量中权重值的求取最为有效的方法是使用 TF-IDF 模型，TF 称为词频，计算该词描述一篇文档内容的能力。其中，IDF 称为逆文档频率，计算该词区分文档的能力。

我们尝试了基于题目数据训练字向量、基于百科新闻训练的字向量，对比发现后者的效果更优，最终采用后者作为方案中特征向量计算的依据。

3.2.2 词频-逆文本频率 TF-IDF

TF-IDF 是词频和逆文本频率两项的乘积，用以评估一字词对于一个文件集或一个语料库中的其中一份文件的重要程度。使用在某篇文档中的原始词频是最简化的选择，如词 t 在文本内容 $content$ 中出现的次数。词频 TF 的计算公式表示为：

$$TF(t) = f(t, content) \quad (3)$$

其中， $f(t, content)$ 表示特征词 t 在文本内容 $content$ 中出现的次数。

如果一个词在很多文档中出现过，则通过这个词来区分文档的区分度越小，可以用逆文档频率 IDF 来度量，表示包含某个词的文档数目：

$$IDF(t) = \log \frac{n}{m+1} \quad (4)$$

其中， n 代表文档的数量； m 表示出现特征词的文档数量；加 1 是为了避免分母为 0。特征词的 TF-IDF 值的计算公式如下：

$$TF-IDF(t) = TF(t) \times IDF(t) \quad (5)$$

3.2.3 相似度计算

方案采用余弦相似度计算不同文本特征向量之间的相似度，余弦相似度的计算公式如下：

$$\cos \theta = \frac{\sum_{i=1}^n (x_i \times y_i)}{\sqrt{\sum_{i=1}^n x_i^2} \times \sqrt{\sum_{i=1}^n y_i^2}} \quad (6)$$

3.2.5 Single-Pass 算法

Single-Pass 算法采用增量聚类的方式将文本向量与已有话题内的报道进行比对，计算文本相似度进行匹配^[7]。若与某个话题类别匹配，则把该文本归入该话题，若该文本域所有话题类别的相似度均小于某一阈值，则将该文本表示成新的种子话题。Single-Pass 聚类算法步骤如下：1) 输入新文档 d ；2) 计算 d 与已有话题分类中每篇文档的相似度，获取与 d 相似度最大的话题并得到相似度值 T ；3) 若 T 大于阈值 θ ，则文档 d 被分类到已知的话题类别，否则作为一个新的话题类别；4) 聚类过程结束。

在小地点聚类 and 主题聚类中，分别使用了不同的阈值作为 Single-Pass 算法的参数进行聚类计算。

3.3 双重约束

3.3.1 基于规则的大地点提取

基于“不同地级市发生的事件不为同一热点问题”的假设，依据大地点（地级市）对数据集进行拆分。

观察题目所给的数据，地级行政区划为字母+“市”，县级行政区划为字母+数字+“市/县/区/州”。因此，在题目所给的数据中，大地点采用正则表达式进行匹配，如果匹配到多个，则取第一个；对于未进行脱敏处理的实际数据，则可以采用地级和县级行政区划名称字典进行匹配。取大地点函数如下：

```
1. def location(x):
2.     content=re.compile(r'[A-Z]\d*[市,县,区,州]',re.M)
3.     contents=re.findall(content,x)
4.     if contents:
5.         return contents[0]
```

由于不同留言者的表述习惯不同，属于同一地点的问题可能提取到不同级别的行政区划，例如“A市魅力之城小区”、“A5区魅力之城小区”，因此在拆分数数据集时，以大地点的首字母为依据，即：按地级市拆分数数据集。

3.3.2 基于 BERT-BiLSTM-CRF 的小地点提取

命名实体识别模型可以帮助我们有效地在序列标注的前提下，通过训练提取出我们所需要的实体，诸如：人名、地名、机构名或者时间点等。目前主流的命名

实体识别模型是生成式模型 HMM 隐马尔可夫模型与判别式模型 CRF 条件随机场模型，如图所示是 BiLSTM-CRF 命名实体识别模型的主要结构。

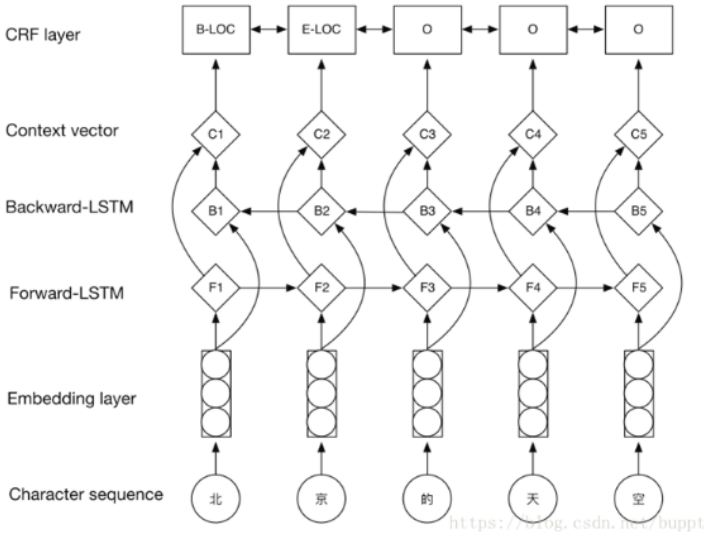


图 9 BiLSTM-CRF 命名实体识别模型结构

- (1) 最底层是基于字符级别的序列标注，将每一句话按照字符级别进行切分。
- (2) Embedding 层，Embedding 层的选择有多种，前文也有过相关的叙述，主要的方法是基于 word2vec 训练的字符向量或者词向量对需要标注的句子进行词嵌入或者字嵌入，亦或是基于预训练语言模型，如 BERT、ERNIE 等预训练语言模型进行词嵌入层的设计。
- (3) 双向的 LSTM 层，主要分为前向 LSTM 与后向 LSTM，通过双向 LSTM 层学习长距离以来信息，LSTM 通过三个门结构（输入门、遗忘门以及输出门），选择性地遗忘部分历史信息，加入部分当前的输入信息，最终整合到当前状态并产生输出状态。具体结构如下图。

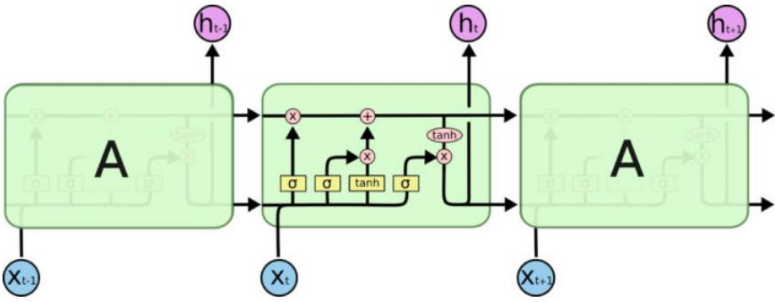


图 10 LSTM 结构图

- (4) 最后是 CRF 层，CRF 层的主要作用是让模型从训练数据中学习得到一定的约束条件，从而确保最终的预测标签序列是有效的，主要可能学习到的约束条件有。
 - ① 一句话中的第一个单词的标签应该是 B-label 或者是 O，不能是 I-label;
 - ② 同一个实体标注序列中，B-label、I-label、I-label 中的 label 应该是相同的实体标签，如 LOC、ORG 或者 TIME 等；

③ 譬如 O、I-label 这种顺序的序列标注数据是无效的，一个命名实体的第一个标签必须是 B-开头，不能以 O 或者 I-开头；

与此同时，CRF 层可以给输入序列求预测输出序列即求使目标函数最大化的最优序列，是一个动态规划问题，可以使用 Viterbi 算法解码来得到最优标签序列，其优点在于为一个位置进行标注的过程中可以利用丰富的内部及上下文特征信息。

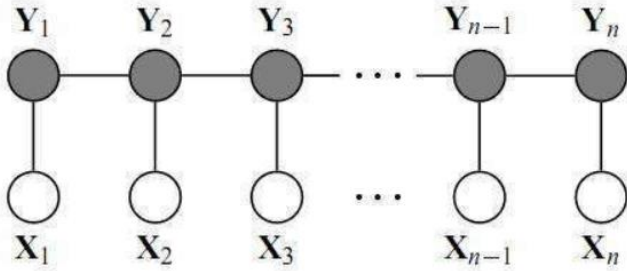


图 11 线性链条件随机场结构

本研究的命名实体识别模型，主要标记了 LOC 以及 ORG 两种实体 label，并且通过人工打标 3000 条题目中的留言主题数据，进行模型的训练。最后通过模型同时提出 LOC 以及 ORG 标签的实体，以便后续的小地点优化处理。以下是模型的训练数据格式、模型的准确率以及模型识别结果的举例展示。

表 4 BERT-BiLSTM-CRF 模型训练数据格式

数据文本	请	问	趣	步	网	络	科	技	涉	嫌	传	销	？
标签格式	O	O	B-LOC	B-LOC	B-LOC	B-LOC	B-LOC	B-LOC	O	O	O	O	O

```
LOC: precision: 95.08%; recall: 95.66%;
ORG: precision: 90.90%; recall: 92.11%;
```

图 12 LOC 及 ORG 标签精确度召回率

```
input sentence, please:请问趣步网络科技是否涉嫌传销
[{'word': '趣步网络科技', 'start': 2, 'end': 8, 'type': 'LOC'}]

input sentence, please:请对A市中山路进行改造升级
[{'word': '中山路', 'start': 4, 'end': 7, 'type': 'LOC'}]

input sentence, please:A市科技职院欺骗学生收取高额费用
[{'word': '科技职院', 'start': 2, 'end': 6, 'type': 'LOC'}]

input sentence, please:西地省三星机床公司拒不还款
[{'word': '三星机床公司', 'start': 3, 'end': 9, 'type': 'LOC'}]

input sentence, please:西地省享乐旅游公司欺诈游客
[{'word': '享乐旅游公司', 'start': 3, 'end': 9, 'type': 'LOC'}]
```

图 13 NER 小地点提取示例

3.4 多重优化

在提取出大地点约束以及小地点约束之后，还需要对已经提取出来的大小地点及聚类的结果进行优化。主要分为以下两个方面：小地点优化以及总体的聚类结果优化。

3.4.1 小地点优化

在小地点的聚类过程中，我们发现如下问题：

（一）部分留言主题不够详尽，不包含问题的小地点，而在留言详情中包含小地点，例如：

表 5 部分主题无小地点的留言

留言编号	留言主题	留言详情
195511	车位捆绑违规销售	对于 伊景园滨河苑 商品房…
241373	强行要求捆绑车位，请有关政府部门为民做主	…楼盘名称是 伊景园.滨河苑 …
204960	家里本来就困难,还要捆绑买卖车位	…内部购买 伊景园滨河苑 房子…
276016	车位属于业主所有，不应该被捆绑销售！	…定向商品房 伊景园滨河苑 …

对于这个问题，我们提出了 2 种解决办法：

（1）使用 NER 模型提取留言详情中小地点；

（2）使用其他数据中已经提取到的所有小地点，在留言详情中进行匹配，如有全字匹配的小地点，则将该小地点作为这条留言的小地点。

对比以上两种方法，由于留言详情中存在的小地点过多难以进行筛选，我们采用了第（2）种方法，补充了部分主题中没有小地点，但详情中有小地点留言的小地点。

（二）部分留言中只有小地点，没有大地点，在以地级市拆分数据集的情况下，这部分数据被单独聚为一类，影响下一步的问题聚类。

表 6 部分无大地点的留言

留言编号	大地点	小地点	留言主题
244243	na	伊景园滨河苑	关于伊景园滨河苑捆绑销售车位的投诉
285897	na	伊景园滨河苑	武广新城伊景园滨河苑违法捆绑销售车位,求解决
268920	na	伊景园滨河苑	武广新城伊景园滨河苑商品房霸王购房规定
246407	na	伊景园滨河苑	举报广铁集团在伊景园滨河苑项目非法绑定车位出售
214975	na	伊景园滨河苑	关于房伊景园滨河苑销售若干问题的投诉
251844	na	伊景园滨河苑	投诉伊景园滨河苑项目违法捆绑车位销售

对于这个问题，我们采用以下方法对由于大地点不全导致分散的小地点组进行合并：

对每一个没有大地点但有小地点（如某小区）的组，匹配存在相同小地点的组进行合并，并在合并过程中排除“xx 局”“xx 政府”等对大地点依赖性较强的小地点。

（三）不同小地点的表述方式指向同一个问题，例如“广铁集团职工住房伊景园滨河苑小区”，留言者在主题中往往只表述“广铁集团”或“伊景园滨河苑”，而这两者指向了同一问题，但在 NER 小地点提取和聚类的过程中无法归为一类。

表 7 部分不同小地点表述方式的留言

留言编号	大地点	小地点	留言主题
209571	A 市	伊景园滨河苑	伊景园滨河苑项目绑定车位出售是否合法合规
276460	A 市	伊景园滨河苑	A 市伊景园滨河苑捆绑销售车位是否合理？
255507	A 市	伊景园滨河苑	违反自由买卖的 A 市伊景园滨河苑车位捆绑销售行为
271517	A 市	广铁集团	开发商联合广铁集团捆绑车位销售
283901	A 市	广铁	A 市市政建设开发有限公司对广铁职工住宅项目操作合法吗？
280774	A 市	广铁集团	反馈广铁集团铁路职工定向商品房的一些问题

对于这个问题，我们采用以下方法进行优化：

对每一个小地点分组，匹配留言详情中出现的其他组地点，对交叉数据双向大于两条且交叉比例双向大于 0.5 的小地点分组进行合并，并在合并过程中排除“xx 局”“xx 政府”等对大地点依赖性较强的小地点。

（四）小地点中的“小区”、“社区”等较普遍的词语容易喧宾夺主，不同小区的小地点由于“小区”二字增加了相似度而被聚到一起。

对于这个问题，在聚类前删去小地点中的“小区”、“社区”等较普遍的词语后在进行聚类。

3.4.2 聚类优化

按小地点分组拆分数数据集，对每一组的留言内容采用基于余弦相似度的 Single Pass 算法进行聚类，提取热点问题。

在热点问题聚类的过程中，我们发现如下问题：

（一）留言详情中存在大量“尊敬的”、“您好”、“谢谢”等与问题本身无关的礼貌用语和虚词，容易对问题的相似度计算产生影响。

对于这个问题，我们采用了“留言主题强化”和“筛选有效词”的方法进行优化：

（1）留言主题强化：留言主题往往与内容相关性较高，我们将留言主题重复两遍加载留言详情的前面作为留言内容；

(2)筛选有效词:为了减轻礼貌用语和虚词对留言相似度计算产生的影响,我们对留言内容使用jieba模块进行分词并提取词性,仅保留词性包含“n”、“v”的词语,作为留言内容的特征。

(二) 部分留言中不包含小地点,在小地点聚类中被归为一组,而这些组的留言之间的联系较弱,容易聚出彼此不相关的问题分组。

对于这个问题,我们为没有小地点的组设定了一个相对较高的阈值,提高其聚为一类的门槛。

3.5 留言热度评价体系构建

在完成留言问题聚类后,需要对这些问题进行综合热度评价,以最终确定热点问题的排序。热度本质上是指某个问题引起的民众关注和讨论的热烈程度,如何定量地来描述留言的热度,目前在学术界还没有统一的用于热度评价的指标体系。建立适用于本研究的科学合理的群众留言热度评价体系构建需要大量的理论基础支撑。因此,笔者检索国内外相关文献并精读了大量关于热度评价体系的优质文章。

部分学者对网络舆情热度及微博热度等指标体系构建做了大量的研究。其中,曾子明等^[8]从原创微博发布量、转发量、评论量、点赞量等4个指标来描述微博舆情热度。黄怡璇等^[9]从信息源、信息人、信息环境、政府态度等一级指标共15个因素构建网络舆情热度体系。周惠子等^[10]以主题强度、主题热度、舆情状态舆情趋势作为出发点,构建了3级12指标的自媒体网络舆情热度。王晓光等^[11]通过TF-IDF词频统计和词增长速度算法对专业领域内的热点进行识别,从内容层面和传播层面两个维度设计并计算一系列指标,对识别到的热点进行热度评价与排序。吴绍忠等^[12]考虑了舆情内容、传播过程及舆情受众等关键要素,对传播规律进行了定量以及定性分析,最终从舆情、舆情传播和舆情受众三个维度共11个指标构建网络舆情相关的指标体系。谈国新等^[13]提出了网络舆情监测指标体系,并由舆情发布者指标、舆情要素指标、舆情受众指标、舆情传播指标以及区域和谐度指标五个指标群构成。此外,何跃等^[14]指出在构建评价舆情热度或微博热度的指标体系时,从用户特征、信息传播特征以及内容特征3个方面来进行,且验证了用户特征、信息传播特征以及内容特征对于热度影响的有效性和合理性。

关于热度体系的评价指标较为繁杂,研究的方向也有所差异,指标体系也在随着新闻学传播学等领域的发展得到改进。但由于每个研究都存在其特殊性,其研究数据源的获取难度和获取渠道都是不同的,因此目前没有一个标准化的指标体系可供该研究的群众留言热度评价使用。考虑到群众留言的特殊性,研究从现有指标体系中筛选提炼出能明确反映群众留言指标体系的因素,主要参考了《基于多源数据的专业领域热点探测模型研究》、《互联网络舆情预警机制研究》、《突发公共事件网络舆情监测指标体系研究》及《基于因子分析法的微博热度评价模

型》的指标体系框架^[11-14]。此外，在对该领域有深度理解后提出具有高解释性的创新指标，从而构建可广泛应用于网络问政平台上评价群众留言热度的指标体系。

首先，网络舆情或微博内容等相关的特征对热度的影响，体现在其文本内容表现的情感极性、内容充实度及相关留言的文本数量等特征上，因此在本研究中群众留言的内容层面的指标是非常重要的。其次，微博意见领袖的存在便体现了用户特征对热度的影响，同理在本研究中群众本身的影响力是不可忽视的，因此受众层面的指标是不可忽视的。而对于传播特征的热度影响力而言，舆情文本或微博的点赞率、评论率、转发率以及群众的关注度等传播特性可以最直观地反映出一条文本引起的关注热度。因此，研究最终选择内容层面、受众层面、传播层面三个维度共 12 个二级指标进行留言热度评价体系的构建。具体体系结构及计算方法如图 14 和表 8 所示。

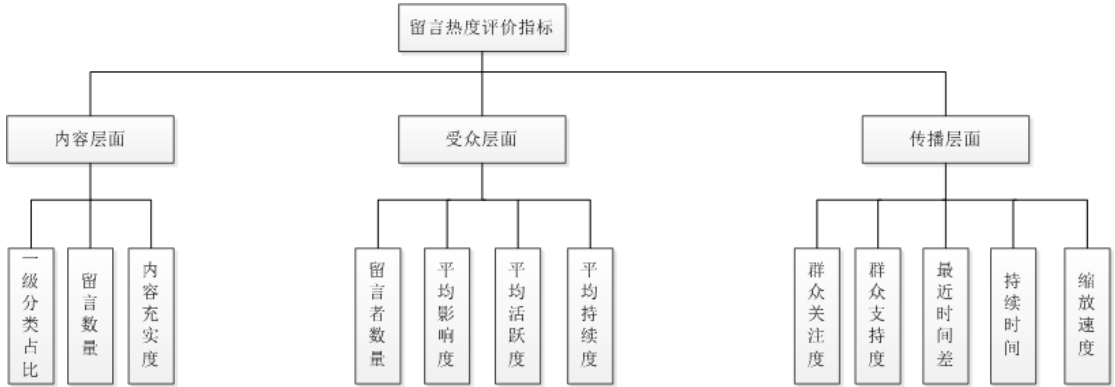


图 14 留言热度评价指标体系结构

表 8 留言热度评价指标体系表

一级指标	二级指标	计算方法
内容层面	一级分类相对数量	该类问题所属一级分类数量占所有一级分类数量之比
	留言数量	该类问题中的群众留言数量
	留言内容充实度	该类问题中所有留言的文字数量/该类问题中留言数量
受众层面	留言者数量	该类问题中留言者数量
	留言者平均影响度	该类问题中留言者的所提全部留言的点赞数/该类问题中留言者数量
	留言者平均活跃度	该类问题中留言者所提全部留言数量/该类问题中留言者数量
	留言者平均持续度	同一类留言时间间隔 1h 以上的相同留言者的留言数量/该类留言总数量
传播层面	群众关注度	该类问题中所有留言的点赞数 + 反对数
	群众支持度	该类问题中所有留言的点赞数/点赞数+反对数
	最近留言时间间隔	该类问题中时间最近的留言距离观测日期的间隔（2020.04.10）
	留言持续时间	该类问题中时间最早留言与时间最晚留言的时间间隔
	留言缩放速度	当前周期的留言条数/上一周期的留言条数

3.5.1 内容层面

首先是内容层面。生产者发布内容的数量是热度评价体系内非常重要的一大

指标^[9]，即某问题的留言数量可以较为直观地衡量问题是否热门，该指标是不容忽视且至关重要的。文本内容本身对其热度有很强的影响作用，其中，文本充实度是表征文本内容的一个有效指标^[14-15]。此外，影响微博评论的关键质量因素位居首位的就是文本信息量，并且信息内容的类型也是排名前三的^[16]。相似的是，王晓光等^[11]也使用了文档相对数量指标。在本研究中，问题所属的一级分类标签，如城乡建设、环境保护等类别，即为上述的信息内容的类型，一级分类的相对数量和文档相对数量本质上是一致的。因此，研究最终选择了一级分类相对数量、留言数量和留言内容充实度共三个指标作为内容层面的二级指标。

内容层面的指标量化过程如下：关于一级分类相对数量，量化方式为该类问题所属一级分类数量占所有一级分类的总留言数量之比，如城乡建设共 1240 条，七个一级分类的下的留言数量总和为 2816 条，那么城乡建设的比例为 44%，捆绑销售车位、餐馆油烟等热点问题都属于城乡建设的范畴，那么这些问题的一级分类相对数量的量化值都为 0.44；关于留言数量，量化方式为该类问题中的群众留言数量，如 58 车贷的相关问题留言数量为 13，那么该问题的留言数量的量化值为 13；关于留言内容充实度，取该类问题中所有留言的文字数量总和与该类问题中所有留言数量的比值，如 58 车贷的 13 条留言的总文字字数之和为 1567，那么该问题的留言内容充实度为 1567/13。

3.5.2 受众层面

其次是受众层面。吴绍忠等^[12]在受众层面使用了受众数量指标，本研究将其定义为留言者数量，可以从一定程度上反映问题热度。在这里要说明的是，留言数量和留言者数量是两个不同的指标，因为存在同一留言者多次留言相同问题的情况。此外，王青等人^[17]于 2011 年在舆情热度研究中首次提出发布者影响度指标。随后，周惠子等^[10]继续使用该指标构建网络舆情热度体系。谈国新等人^[13]使用舆情发布者影响力和活跃度指标，分别衡量发布者引起网络舆情分析者注意的程度和自身的活跃程度，用来判断具有号召力的网民，类似于“论坛领袖”、“大 V 博主”等。与其相似，本研究定义了留言者平均影响度以衡量留言者所提留言获取大众的支持度，以及留言者平均活跃度用来衡量留言者在网络问政平台上的活跃程度。最后，研究提出创新指标——留言者平均持续度，即考虑到同一个留言者反复提交相同留言的情况，该指标可反映该问题的紧急程度和群众迫切的心情，能从一定程度反映问题热度。因此，研究最终选择了留言者数量、留言者平均影响度、留言者平均活跃度和留言者平均持续度共四个指标作为受众层面的二级指标。

受众层面的指标量化过程如下：关于留言者数量，即取该类问题中所有留言者的数量作为量化值，如 58 车贷的留言者共有 12 人，则该问题的留言者数量即为 12；关于留言者平均影响度，即取该类问题中留言者所提的全部留言的点赞数/该类问题中留言者数量进行量化，如 58 车贷的 12 位留言者在该网络问政平台的所有留言一共被点赞 25 次，则留言者的平均影响度为 25/12；关于留言者平

均活跃度，取该类问题中留言者所提全部留言数量/该类问题中留言者数量，如 58 车贷的 12 位留言者在该网络问政平台的所有留言一共有 19 条，则留言者平均影响度为 19/12；关于留言者平均持续度，即计算同一类中留言时间间隔 1h 以上的相同留言者的留言数量/该类留言总数量，如聚类后发现某热点问题有群众重复留言了 2 次，且间隔时间大于 1 小时（排除系统故障或者网络卡顿等原因导致同一留言反复提交的情况），且该问题的留言总数量为 13，那么留言者平均持续度为 2/13。

3.5.3 传播层面

最后是传播层面。林文声等人^[18]在网络舆情热度评价研究中提出事件关注参与度指标，即网民对事件的参与和关注程度，本研究将其命名为群众关注度，用来衡量群众对于某留言主题的关注程度。此外，周惠子、冯江平、王青等人^[10,17,19]分别在研究中使用了网民的观点倾向度来构建网络舆情评价指标体系，本研究使用群众支持度指标，通过支持和反对态度的数量来衡量群众对某主题留言的支持倾向程度。王青等人^[19]提出创新的时效度指标，反映其生命周期中所处的阶段水平，本研究将其定为最近留言时间间隔，用来判断该主题留言的时效性。再有，何跃^[14]和林文生等^[18]在研究中分别设定出现时长和持续时间指标，以衡量热点问题出现的时间长度，因此本研究使用留言持续时间来判断每个主题的留言时间间隔。最后，王晓光等^[11]使用缩放速度指标，以判断问题的热度变化趋势，本研究也使用留言缩放速度指标，通过当前周期内的主题数量和上一时间周期内主题数量的对比来观测留言热度的变化趋势。因此，研究最终选择了群众关注度、群众支持度、最近留言时间间隔、留言持续时间和留言缩放速度共五个指标作为传播层面的二级指标。

传播层面的指标量化过程如下：群众关注度取该类问题中所有留言的点赞数与反对数之和作为量化值，如 58 车贷的所有相关留言的点赞数为 2383，反对数为 0，则群众关注度为 2383；群众支持度取该类问题中所有留言的点赞数占关注数的比例，即 58 车贷的群众支持度为 $2383 / (2380 + 0)$ ；最近留言时间间隔取该类问题中时间最近的留言距离观测日期（数据集中所有留言最近日期——2020.1.26）的间隔，如 58 车贷的最近一次留言时间为 2019.7.8，则 58 车贷的最近留言时间间隔为 202；留言持续时间取该类问题中时间最早留言与时间最晚留言的时间间隔，如 58 车贷的最早条留言时间为 2019.1.11，最后一条留言时间为 2019.7.8，则留言持续时间为 178；留言缩放速度取当前周期的留言条数与上一周期的留言条数的比值。

3.5.4 指标权重确定

为了能够更加准确评价群众留言热度，需要用权重分析方法计算评价内指标的权重。主观赋权法主要包括二项系数法、德尔菲法、层次分析法等。其中，层次分析法作为一种常用的权重分析方法，主要是对总决策目标进行分析，将其拆

分为多个不同的小目标，并设置不同的判定标准后制定不同的决策方案，从而完成对决策目标的分解；其次通过指标间的两两对比得到判断矩阵，其构造依据是分析不同的指标对于决策目标的影响程度；最后各个指标的权重由判断矩阵得出^[20]。群众留言热度评价体系的各指标权重的计算过程如下：

(1) 构建判断矩阵

群众留言热度评价体系中的指标对于留言热度的影响程度可以通过构建判断矩阵来确定，矩阵中元素标度定义见表。

表 9 判断矩阵标度定义

标度	定义
1	前者与后者同样重要
3	前者比后者稍微重要
5	前者比后者重要
7	前者比后者重要的多
9	前者比后者极其重要
2、4、6、8	相邻标度的中间值重要程度
倒数	假设参数 <i>i</i> 对参数 <i>j</i> 的标度为 a_{ij} ，则参数 <i>j</i> 对参数 <i>i</i> 的标度为 $1/a_{ij}$

(2) 计算一致性指标

判断矩阵具有良好一致性的条件是一致性指标 CI 接近与 0；反之，若 CI 的值越大，则说明矩阵的相对来说欠缺一致性， CI 的计算公式为：

$$CI = \frac{\lambda_{max} - n}{n - 1} \quad (7)$$

其中 λ_{max} 表示判断矩阵的最大特征值， n 则表示判断矩阵的阶数。

(3) 计算检验系数

通过计算检验系数 CR 的值可以分析判断矩阵的一致性，将所得的 CR 系数的值与阈值 0.1 进行比较，若小于阈值，则说明一致性检验通过。检验系数 CR 的计算公式为：

$$CR = \frac{CI}{RI} \quad (8)$$

随机一致性指标 RI 与矩阵阶数的对应关系见下表。

表 10 随机一致性指标与矩阵阶数对应关系

阶数 n	1	2	3	4	5	6
RI	0	0	0.52	0.89	1.12	1.26

(4) 模型指标权重计算

当模型指标的判断矩阵通过一致性检验后，便可利用判断矩阵来进一步确定模型各指标的权重，指标权重 w_i 计算公式为：

$$w_i = \frac{1}{n} \sum_{j=1}^n \frac{a_{ij}}{\sum_{k=1}^n a_{kj}}, i=1,2,\cdots,n \quad (9)$$

其中， a_{ij} 表示判断矩阵第*i*行第*j*列的元素，*n* 则为判断矩阵的阶数。

接下来为留言热度评价体系的指标权重确定过程。首先是一级指标权重的确定，如下表所示。经计算，模型指标所构建的判断矩阵 CR 指标符合实验标准，并通过和法求出各指标对应的权重如下：内容层面指标权重为 0.539，传播层面指标权重为 0.30，受众层面指标权重为 0.161。

表 11 留言热度评价体系一级指标判断矩阵

	内容层面	传播层面	受众层面
内容层面	1	2	3
传播层面	1/2	1	2
受众层面	1/3	1/2	1

接下来依次是内容层面、传播层面及受众层面的二级指权重确定过程。

表 12 内容层面二级指标判断矩阵

	留言数量	一级分类相对数量	留言内容充实度
留言数量	1	4	5
一级分类相对数量	1/4	1	3
留言内容充实度	1/5	1/3	1

同理，计算得知内容层面各指标对应的权重如下：留言数量指标权重为 0.665，一级分类相对数量指标权重为 0.231，留言内容充实度指标权重为 0.104。

表 13 传播层面二级指标判断矩阵

	群众关注度	群众支持度	最近留言间隔	留言缩放速度	留言持续时间
群众关注度	1	2	3	4	5
群众支持度	1/2	1	2	3	4
最近留言间隔	1/3	1/2	1	2	3
留言缩放速度	1/4	1/3	1/2	1	2
留言持续时间	1/5	1/4	1/3	1/2	1

计算得知传播层面各指标对应的权重如下：群众关注度指标权重为 0.416，群众支持度指标权重 0.262，最近留言时间间隔指标权重 0.161，留言缩放速度指标权重 0.099，留言时间间隔指标权重 0.062。

表 14 受众层面二级指标判断矩阵

	留言者数量	留言者平均影响度	留言者平均活跃度	留言者平均持续度
留言者数量	1	3	4	5
留言者平均影响度	1/3	1	2	3
留言者平均活跃度	1/4	1/2	1	2
留言者平均活跃度	1/5	1/3	1/2	1

计算得知传播层面各指标对应的权重如下：：留言者数量指标权重为 0.542，

留言者平均影响度指标权重 0.233，留言者平均活跃度指标权重 0.140，留言者平均持续度指标权重 0.085。

到此，就完成了留言热度评价体系指标权重的确定过程。留言热度评价体系指标权重一览表如下表所示。

表 15 留言热度评价体系指标权重一览表

一级指标	二级指标
内容层面(0.539)	一级分类相对数量(0.231)
	留言数量(0.665)
	留言内容充实度(0.104)
受众层面(0.161)	留言者数量(0.542)
	留言者平均影响度(0.233)
	留言者平均活跃度(0.140)
	留言者平均持续度(0.085)
传播层面(0.30)	群众关注度(0.416)
	群众支持度(0.262)
	最近留言时间间隔(0.161)
	留言持续时间(0.062)
	留言缩放速度(0.099)

3.6 热点问题结果展示

在完成热点问题聚类后，并对个别异常值进行校对调整，通过各指标的量化值及权重加权计算得到留言问题的热度排序。以下分别展示的是聚类结果和热度评价结果的 TOP10。由表可见，留言问题的热度评价结果排名和按照留言数量聚类的结果排序相似度很高，但也存在一定的差异性。在评价体系的加持下，以留言数量为热度基础，留言的内容层面、传播层面和受众层面的指标也会对热度评价结果产生影响。

此处要补充说明的是，在留言热度评价体系的量化过程中，为保证留言数量作为热度评价的绝对标准，也就是即使某主题留言中的留言者活跃性和影响力很强、且问题的持续周期等指标很好，但该问题的留言数量相较于热点问题明显偏低时，依然不能计为热点问题。因此，留言数量指标采用原值，其他指标均统一采用等比例缩放或者四分位数量化到相同值域区间，最终加权得到结果。

表 16 聚类结果的留言数量比较结果 TOP10

排名	留言数量	问题描述
1	58	丽发新城小区附近搅拌站的噪音和扬尘污染，影响居民生活和健康
2	57	伊景园滨河苑小区商品房捆绑强制销售车位，损害广铁集团职工利益
3	21	魅力之城小区底层商铺餐饮店油烟排放导致空气污染，且噪音扰民
4	13	涉嫌诈骗的 58 车贷案件经侦缓慢，案情相关信息公布拖延
5	12	辉煌国际城小区楼下商铺无专用排烟管道，属于非法营业且油烟扰民

6	11	星沙旧城改造将结束，但项目中凉塘路区域拖延进度，迟迟不动工
7	10	谷山庭院存在电梯故障等问题，且物业恶意行为损害业主合法权益
8	10	A 市涉外经济学院强制学生在固定实习实习，损害学生自由选择权
9	10	不同学校存在的问题
10	10	凯乐国际城周边道路乱停车及路灯失修问题，给市民安全带来隐患

表 17 热度评价体系比较结果 TOP10

排名	留言数量	问题描述
1	58	丽发新城小区附近搅拌站的噪音和扬尘污染，影响居民生活和健康
2	57	伊景园滨河苑小区商品房捆绑强制销售车位，损害广铁集团职工利益
3	21	魅力之城小区底层商铺餐饮店油烟排放导致空气污染，且噪音扰民
4	13	涉嫌诈骗的 58 车贷案件经侦缓慢，案情相关信息公布拖延
5	12	辉煌国际城小区楼下商铺无专用排烟管道，属于非法营业且油烟扰民
6	11	星沙旧城改造将结束，但项目中凉塘路区域拖延进度，迟迟不动工
7	10	谷山庭院存在电梯故障等问题，且物业恶意行为损害业主合法权益
8	10	凯乐国际城周边道路乱停车及路灯失修问题，给市民安全带来隐患
9	10	A 市涉外经济学院强制学生在固定实习实习，损害学生自由选择权
10	10	高铁规划线路会给绿地海外滩的居民带来严重噪音和安全隐患

观察可知，按留言数量排序的聚类结果 TOP10 与热度评价体系内的热点问题 TOP10 重合率高达 85%，存在差异的 15%均出现在 8-10 名的排序中。可见，留言数量在评价体系是占据绝对地位的，因此留言数量明显有优势的问题自然就容易归为热门问题，而留言数量很相近的问题就会因评价热度体系的计算后使其排序有所改变。因此，本研究的留言热度评价体系的设定及量化能有效地比较数量较为相近问题的热门程度。

最后展示的是赛题的提交结果，即通过某一时段内反映特定地点或特定人群问题的留言进行归类，并定义合理的热度评价指标，最终给出排名前 5 的热点问题评价结果及问题详情。可见该网络问政平台上的群众留言的热度问题分别是丽发新城小区附近搅拌站噪声污染及环境污染严重（热度指数 22.743）、广铁集团伊景园滨河苑小区商品房捆绑强制销售车位（热度指数 22.054）、魅力之城小区临街餐饮店油烟噪音扰民（热度指数 9.157）、58 车贷案件经侦缓慢且案情公布拖延（热度指数 6.721）、辉煌国际城小区楼下商铺违法开饭店（热度指数 5.970）。

表 18 热点问题表 TOP5

热度排名	问题 ID	热度指数	时间范围	地点/人群	问题描述
1	1	22.743	2019/07/03 至 2020/01/26	A 市 A2 区 丽发新城小区居民	丽发新城小区附近（百米以内）搅拌站的噪音和扬尘污染，对居民的生活和健康问题产生极大影响
2	2	22.054	2019/07/02 至 2019/09/01	A 市广铁集团 伊景园滨河苑小区居民	伊景园滨河苑小区商品房捆绑强制销售车位，不买车位就取消购房资格，损害广铁集团职工利益

3	3	9.157	2019/07/21 至 2019/12/04	A 市 A5 区劳动东路 魅力之城小区居民	魅力之城小区底层商铺的餐饮店大量油烟排放导致空气 污染严重, 且噪音扰民, 影响了居民的生活质量
4	4	6.721	2019/01/11 至 2019/07/08	A 市 A4 区 58 车贷案	涉嫌诈骗的 58 车贷案件经侦缓慢, 案情相关信息公 布拖延, 因无法获取案件进展等信息引起市民不满
5	5	5.970	2019/01/06 至 2019/05/22	A 市 A1 区 辉煌国际城居民	辉煌国际城小区楼下商铺, 商铺无专用排烟管道, 改 造为地排, 属于非法营业且油烟扰民

4 答复意见评价

针对相关部门对留言的答复意见, 研究尝试从答复的相关性、及时性、完整性、可解释性、专业性、逻辑性和礼貌性共七个角度对答复意见的质量给出一套合理的评价方案, 并通过人工打分进行结果对比及评估, 以验证答复意见评价体系的合理性。

4.1 答复意见评价体系构建

4.1.1 指标选择及量化

作为网络问政平台的重要基础, 当代政府对于群众留言的回应度和答复质量是决定政府与公众互动效果的重要指标。政府回应是了解民情、为群众解决难题并获得民众支持的重要因素, 也是政府责任的体现。互联网为我国服务型政府建设提供高新技术保障, 同时也对政府服务提出了诸多挑战。政府无法及时有效地回应人民群众日益增多的服务需求, 主要表现为政府回应动力不足、回应效率低下、回应渠道不畅、回应效果不佳等, 而问题的根源则在于政府角色错位、公仆意识不强、回应机制不健全和问责不彻底^[21]。

提升政府服务回应力的主要途径是通过加强服务理念、转变职能与作风、革新政府流程和重塑公信力等。可以考虑从主体方——政府、客体方——民众和回应机制三个方面着手, 不断提升地方政府的回应能力, 实现依法有效行政, 建设一个人民满意的服务型政府。而本研究正是从回应机制作为出发点, 通过建立合理的答复意见评价体系, 通过直观的评价打分可以有效监督政府回应的质量, 为完善政府的回应机制奠定基础, 以最终实现提升政府的服务水平的目的。

关于答复意见评价体系影响因素的选择, 需从群众的真实需求出发。

(1) 相关性

答复意见的相关性是至关重要的。政府服务人员需要做到的就是精准地捕获到群众在留言中表达的主题思想, 并从问题本身出发真正为民众解决难题。

关于相关性的量化方法, 主要计算了以下六个指标, 包括 Word2vec 词嵌入余弦、Jaccard 相似度、Euclidean 词嵌入欧氏距离、Simhash 海明距离、BOW 词袋模型、CNN 卷积网络及 GRNN 循环网络。指标的取值范围为 0-1 之间, 计算出每个指标的值后, 求其乘积作为相关性的量化值。

（2）及时性

及时性也是需要重点考虑的因素。当群众在问政平台上进行留言时，多是需要政府在一定的时间周期内进行回复，如工作人员办事效率太低则会很大程度影响到群众的满意度，也极有可能造成群众反映问题的后续不良发展，对民众的生活或者工作等方面带来不便。

关于及时性的量化方法，计算群众留言日期与政府回复日期的时间间隔，以天为时间单位，最后使用四分位数进行量化，将取值范围归为 0-1 之间。

（3）完整性

答复的完整性是值得重点考虑的。这可以体现出政府服务的规范性，设立一套完整的答复语言体系是十分有必要的，以有效提升答复的质量。

关于完整性的量化方式，将留言文本内容中是否包含开头称呼、留言获悉、结尾日期及相关的咨询电话等内容以及答复文本量的多少作为衡量标准。其中，答复文本量使用四分位数进行量化将取值范围归为 0-1 之间，其他所有指标都是根据规则判断有无，量化为 0/1。

（4）可解释性

可解释性是不可或缺的。群众反映的问题类型是具有差异的，如城乡建设、卫生计生、劳动与社会保障等很多种类，每个所属分类都会有相应的法律法规、政府文件等，群众因信息不对称极有可能获取不到所需信息，那么政府工作人员是有必要在回复过程中提供包含法律依据或相关文件中的信息，以供群众参考。

关于可解释性的量化方法，则是依据留言文本中是否出现相关法律法规法令、举例解释说明、相关解释性词以及解释的有效文本率来判断。其中，解释有效文本率指的是某答复文本中针对问题答复的有效文本字数占全部文本字数的比例，使用四分位数量化到 0-1 区间，其他指标统一根据规则判断有无，量化为 0/1。

（5）专业性

专业性高的答复能侧面反映出政府工作人员的能力，进而会在一定程度上提升政府形象。如《政务信息资源目录体系》和《国务院公文主题词表》^[22]中包含大量主题分类下的专业术语，可供量化专业化指标时使用。

关于专业性的量化方式，则是计算答复文本中所含该大类主题分类中的专业词数量。由于分类体系的差异性，为保证该指标的科学性，专业词词表是根据赛题所给的分类体系、《政务信息资源目录体系》以及《国务院公文主题词表》三个体系中融合交叉后获取所得。如赛题中的城乡建设和环境保护是区分设立的，《政务信息资源目录体系》中城乡建设和环境保护是被划分为同一类的，同时观测到赛题中所给的城乡建设的二级指标中有包含有环境保护的相关内容，因此在城乡建设主题下匹配《国务院公文主题词表》时就将环境保护主题的相关专业词也纳入其中。关于该指标的量化方式，则是计算文本中包含专业词的个数，并通过四分位数量化到 0-1 范围之内。

（6）逻辑性

逻辑性会直接影响到群众的易读性。包含逻辑词的留言答复具有更高的易读性,条理清晰、语言流畅会极大降低群众的阅读成本,使信息转化更为高效彻底。因此在留言答复中应尽可能地多使用连词、转折词、顺序词等富有逻辑的用语。

关于逻辑性的量化方法,则是根据文本中首先、然后、最后、第一、第二、第三等表示时间语序的词语以及关联词的应用。获取该类词语的数量,并通过四分位数量化到 0-1 范围之间。

(7) 礼貌性

礼貌性的答复是不可或缺的。礼貌是人类为维系社会正常生活而要求人们共同遵守的最起码的道德规范,它是人们在长期共同生活和相互交往中逐渐形成,并且以风俗、习惯和传统等方式固定下来的。礼貌性会直接影响到群众在问政平台上的体验感,因此在留言答复中应尽可能地多使用礼貌用语。

关于礼貌性的量化方法,可以通过文本开头的您好你好等敬辞、结尾是否有感谢以及祝愿等礼貌用语来确定。根据规则判断有无,量化为 0/1。

答复意见评价体系内七个一级指标的量化方法,总结如下表所示。

表 19 答复质量评价指标的量化方法

一级指标	二级指标(关键词)	量化方法	取值范围
及时性	时间间隔	四分位数量化	1/0.5/0
	Word2vec+cos	Word2vec 词嵌入余弦	0-1
相关性	Jaccrad	杰卡德相似度	0-1
	Euclidean	词嵌入欧氏距离	0-1
	Simhash	Simhash 海明距离	0-1
	BOW	词袋模型	0-1
	CNN	卷积网络	0-1
	GRNN	循环网络	0-1
完整性	开头称呼	根据规则判断有无	0/1
	留言已获悉、留言已阅等	根据规则判断有无	0/1
	答复文本量	四分位数量化	1/0.5/0
	结尾的日期	根据规则判断有无	0/1
礼貌性	电话号码	根据规则判断有无	0/1
	开头的您好、你好等敬辞	根据规则判断有无	0/1
	结尾是否有感谢	根据规则判断有无	0/1
	祝愿等礼貌用语	根据规则判断有无	0/1
可解释性	《\w+》相关法规法令	根据规则判断有无	0/1
	(\w+) 举例解释说明	根据规则判断有无	0/1
	相关解释性词(核实、调查等)	根据规则判断有无	0/1
	解释有效文本率	可解释性字数/文本量	0-1
逻辑性	逻辑词(首先 第一 ① 等)	根据规则判断有无	0-1
专业性	匹配该主题下的专业词词表	四分位数量化	0-1

4.1.2 指标权重确定

与留言热度评价体系的指标权重确定方法相似,使用层次分析法对答复意见评价体系的指标进行权重确定,最终结果见下表所示。

表 20 答复意见评价体系的指标判断矩阵

	及时性	完整性	可解释性	专业性	逻辑性	礼貌性
及时性	1	2	3	4	5	6
完整性	1/2	1	2	3	4	5
可解释性	1/3	1/2	1	2	3	4
专业性	1/4	1/3	1/2	1	2	3
逻辑性	1/5	1/4	1/3	1/2	1	2
礼貌性	1/6	1/5	1/4	1/3	1/2	1

计算可得及时性的权重为 0.379,完整性的权重为 0.248,可解释性去权重为 0.160,专业性的权重为 0.102,逻辑性的权重为 0.065,礼貌性的权重为 0.046。

综上所述,在利用层次分析法计算出内容层面、受众层面以及传播层面一级指标以及二级指标的权重之后,采用如下公式对留言热度进行加权赋分。加权公式如下:

$$\text{加权得分} = \text{相关性} * [(5 * w_1 * \text{及时性}) + (1 * w_2 * \text{完整性}) + (\frac{5}{4} * w_3 * \text{可解释性}) + (w_4 * \text{专业性}) + (5 * w_5 * \text{逻辑性}) + (\frac{5}{3} * w_6 * \text{礼貌性})] \quad (10)$$

其中, w_1 - w_6 分别代表上述所计算的及时性、完整性、可解释性、专业性、逻辑性及礼貌性的权重。与此同时,此加权得分公式主要做了以下两方面的改进:

(1)相关性指标并未直接加权而是采用直接相乘的方法,因为相关性是衡量留言答复质量较为重要的指标,改进之后,如果某一条留言的其他指标很高,但是答非所问(相关性较低),那么总体得分也会随之下降。

(2)对于二级指标数量不同的一级指标,采用统一取值范围的方法进行改进,这样可以使得所有一级指标的取值范围相同,避免了因为取值范围所引起的指标加权影响程度高低问题。

4.1.3 特殊情况处理

此处需要补充说明的是,在数据预处理过程中研究者发现在评价质量的答复体系中存有特殊情况,需要进行进一步处理。

首先,因为留言者提出问题不符合该网络问政平台的服务范畴,工作人员在回复中建议该留言者通过其他有效方法解决问题或者转交给其他部门且提供咨询电话的情况。在针对此类问题的处理中,将该问题的相关性指标提高到指标均值,该处理可以有效解决该类问题相关性极低的情况。该类问题举例如图 15 所示。

尊敬的网友：您好！您反映的问题已转交至工商部门，因反映信息不详，建议您直接拨打12315或7710659告知。

图 15 特殊情况（1）

其次，存在答复中重复了留言者问题，使得相似度过高，相关性得分极高，最终导致该问题的排名非常高。因此，在文本处理过程中，通过按句（不包含逗号）切分，以删除答复中的整句留言者问题。最后再计算文本相似度，即合理量化了相关性指标，其他指标还是以原文本为依据计算。该类问题举例如图 16 所示。

‘您好！您在信中提到的关于“泉塘街道清楚路和小塘路交汇处车流量、人流量密集，无过街天桥和地下通道，交通事故较多，此处无高杆灯一到晚上视线较差，小塘路以南至盼盼路沿线路灯较差，晚上老人、小孩过街极为不方便；希望借城市提质改造机会能在清楚路和小塘路交汇增加高杆灯，改造小塘路路灯照明，提高城市品质，增强老百姓幸福感和获得感。”的问题，我局高度重视，现回复如下：2017年我局已考虑在清楚路和小塘路安装高杆灯，但由于该路口处于高压配电网正下方，该路口无法安装高杆灯。小塘路设计之初根据道路宽度是安装的单侧路灯，根据实际需要，我局会考虑将该路段路灯增补纳入到下一年度改造计划中。感谢您对城市管理工作的关注，今后如您再次遇到城市管理相关问题，欢迎致电城管服务热线0000-00000000或12319，我局将热忱为您服务。2019年5月14日’

图 16 特殊情况（2）

4.2 效果评价

以下为答复意见评价体系的 TOP5 以及 TAIL5 的结果展示。该评价体系通过七大指标的量化能有效地将留言进行差异化排序。通过不同排序的图 17 和图 18 的对比，可以非常明显看出回复质量在各个指标上的差异性。

```
In [107]: df3["答复意见"][2811]
Out[107]: '网友：您好！留言已收悉'

In [106]: df3["答复意见"][2812]
Out[106]: '已收悉'

In [105]: df3["答复意见"][2813]
Out[105]: '网友：您好！留言已收悉'

In [104]: df3["答复意见"][2814]
Out[104]: '网友：您好！留言已收悉'

In [103]: df3["答复意见"][2815]
Out[103]: '网友：您好！留言已收悉'
```

图 17 答复意见评价体系留言 TAIL5

In [76]:	df3["答复意见"] [0]
Out [76]:	‘网友：您好！您反映“对G7县惠望溪镇马石完全小学组织学生订购牛奶问题的再质疑”的问题已收悉。我局对此高度重视，对反映的情况进行了调查核实，现将相关情况回复如下：一、关于收牛奶订购款问题经逐一到班到人核查，网友反映的学生牛奶是由牛奶销售公司委托校外人员代售，学生或家长根据自己需求自愿在该售卖人员处征订缴费。到2月28日止，马石小学共有50名学生家长自愿在代售处订购了一学期牛奶（在2月18日的基础上增加了一名学生）。该过程中马石小学学校和教师未组织、未参与学生奶的订购与收费。二、关于收取饮水费问题经调查，马石小学平时向学生提供了正常饮水。2018年上半年开学初马石完小部分家长要求孩子饮用桶装矿泉水，并愿意自费购买。其中有部分学生或家长（主要是一、二年级学生）把饮水费交给了班主任，再由班主任交到代售人员。该学期开学初发现该问题后，惠望溪镇中学已及时责令将其全部清退并对校长进行了批评教育。三、问题处理1.对于学生饮用奶等市场服务项目，惠望溪镇中学再次到马石小学重申了收费纪律和要求，严令校长及教师不得组织、参与市场服务性项目。2.对2019年春季开学马石小学涉嫌违规收费问题，县教育局已责成惠望溪镇中学免去了姚某某小学校长职务，并将该小学相关问题在全镇中小学通报处理。感谢广大网友对我县教育事业的关注和监督！2019年3月6日’
In [77]:	df3["答复意见"] [1]
Out [77]:	‘“UU0081583”您好！关于您反映的“L5县两丫坪社区一村民关于请求解决危房补的请求”这一问题，经我镇调查，现将相关问题回复如下。经调查：陈平旺，两丫坪镇两丫坪社区18组（原平安村3组）居民，系建档立卡贫困户，家庭贫困人口4人。1、关于陈平旺建档立卡贫困户的问题。陈平旺建档立卡时间为2014年，评定过程完全符合程序，在相关表格上均有其本人签字确认，且在建档立卡之后，结对帮扶责任人也按规定对其进行了结对帮扶走访，并非如他本人所述的毫不知情。2、关于陈平旺没有实施易地搬迁项目的问题。陈平旺已于2016年作为易地搬迁计划名单上报，但在搬迁选址阶段，其提供的几处选址都属于“简单位移”，不符合搬迁要求，不予通过。后社区干部也为其提供多条可行的选址建议，陈平旺都不满意不同意，遂其父亲陈异德自愿签订《退出易地搬迁退出书》，并盖有手模印，放弃享受易地搬迁政策。3、关于陈平旺危房改造项目未获补贴的问题。陈平旺于2017年10月向两丫坪社区、镇镇管理所及国土所申请危房改造建房。经村、镇两级相关部门人员现场勘查、选址放线，拟建于原宅基地前面的斜坡上，房屋出高坎1.5米处，位于水田内修建三个立柱，悬空用于作走廊支柱。2017年11月下旬，接群众举报，陈平旺私自将房屋外移占用基本农田建房。经现场勘查：陈平旺房屋现占地187平方米（建筑总面积二层共310平方米），其中占用基本农田面积为135平方米（房屋建筑占地75平方米，空坪硬化60平方米）。陈平旺占用基本农田建房违反了《中华人民共和国土地管理法》第76条和《基本农田保护条例》第17条的规定，属于非法建筑。一是根据《L5县规范农村村民建房管理办法》的通知之第九条规定：选址占用基本农田或占用农用地未办理农用地转用审批的不予批准村民建房申请。二是按照2018年危房改造项目政策要求，1人不能超过35平方米，2人不能超过45平方米，3人以上人均不能超过18平方米。所以陈平旺不能享受危房改造的资金补助，危房改造计划才未报上。4、关于陈平旺100元车费一事。2017年10月份镇政府分管国土城建工作的领导王银端、国土所所长贺军、国土所干部石辉、两丫坪社区干部吴菲、黄军长共5人应邀去看陈平旺家准备危房改造的宅基地。当一行人准备离去时，陈平旺硬要塞给吴菲300元钱说是请大家吃中饭，吴菲坚决不肯接受。两人僵持不下，陈平旺见吴菲态度坚决，只好改口说给司机100元车费。5、关于陈平旺请客两丫坪社区支部书记周元齐花费上千元一事。周元齐于2016年7月份受邀在陈平旺家吃了一顿中饭，主要咨询危房改造和易地搬迁政策，陪同人有黄军长、李模佑，当时发了一包25元的A1区胃炎，因周元齐不抽烟，离去时并未拿。其后批地基、口头下建房停工通知书周元齐都未直接参与，所以说请客周元齐花费上千元一事并不属实。2018年7月31日’
In [78]:	df3["答复意见"] [2]
Out [78]:	‘网友“UU0089835”您好！来信收悉。现回复如下：经镇综治办调查核实，信访人黄尊富，系我镇马川湖村楼梯坡组村民。2010年—2012年期间，浏酃高速公路修建经我镇马川湖村，并在该村楼梯坡组设计有一涵洞，该涵洞能安全通过小轿车、农业用车等车型。黄尊富的房屋在楼梯坡涵洞边，距离高速主线有200多米远，不符合拆迁条件。自涵洞开工建设以来，黄尊富以及其父黄刚明以楼梯坡涵洞设计位置太低，涵洞出口接线路坡度较陡，造成其本人和家人出行不便为由，多次向江背镇浏酃高速公路指挥部反映，要求解决该问题，并要求进行拆迁。江背镇党委、政府高度重视此事，从2011年至今，多次协调浏酃高速公路指挥部采取保障措施保障黄尊富一家的安全出行，并给以一定经济补偿，具体情况如下：1、在浏酃高速建设期间，积极解决黄尊富一家的出行问题并给以补偿。由浏酃高速八标段项目部承担原材料、人工等所有开支，由黄尊富之父黄刚明按照其家人出行方便的需求，现场指挥施工，对涵洞出口到其家门口的坡道泥路进行水泥硬化，有效减小了涵洞建设对黄尊富一家生产生活的的影响，已于2012年10月完工。同时该户户主黄刚明与指挥部签订了因涵洞设计较低导致出行不便的一次性补偿协议，协议明确要求黄刚明一家领取补偿款后，不得再就涵洞出口坡度较陡造成出行不便的问题提出其它要求。2012年1月20日黄刚明领取补偿款21200元。2、黄尊富否认第一次协议效力，并以自己名义签订合同再次拿到补偿。黄尊富认为2012年1月20日签订协议时他不在家，是户主黄刚明签订的，否认第一次协议的效力。2012年12月4日，指挥部再次组织马川湖村、楼梯坡组与黄尊富协商，并与其签订了2万元的一次性经济补偿协议，黄尊富承诺领取资金后，不再就此事向镇政府提出其他任何的补偿要求。3、江背镇政府积极召开信访答复会。在该次补偿后，黄尊富仍不满意，每隔一个月到江背镇政府上访，并多次向问政西地酋、县长信箱、“12345”等平台提出拆迁的无理诉求。2014年5月22日，江背镇党委、政府组织镇派出所、国土所、经经办、综治办、浏酃高速公路拆迁指挥部等部门负责人和黄尊富本人召开了有关信访答复会，明确答复黄尊富，因其房屋不在浏酃高速拆迁红线范围内，要求拆迁房屋的诉求没有相关的政策依据，也没有现实操作拆迁的可能。由于高速公路建设造成的客观影响已得到妥善解决，如果仍然不能认同，必须合理合法的到相关部门反映情况。4、2016年12月1日，江背镇综治、信访、派出所、原浏酃高速公路指挥部成员代表、黄尊富参加了最近一次的信访答复会。会上，我镇对黄尊富进行了政策宣传与解释，黄尊富不理解、不满意，并通过打电话、发短信等方式多次骚扰我镇工作人员的正常办公与休息，我镇明确告知信访人：一是由于浏酃高速建设对其家庭出行造成不便，指挥部已两次给予其家庭补助共计4万元，信访人也在协议上签字认可；二是根据信访人家庭实际情况，不符合拆迁的标准，不在拆迁范围之内；三是征地拆迁的主体是浏酃高速建设指挥部，不是江背镇人民政府；四是信访人反映诉求必须依法依规进行，不得以发短信威胁或打电话骚扰等非正常方式干扰、影响国家机关或国家工作人员的正常工、生活，在本次答复会议对其训诫的基础上，如若继续违反，将承担相应的法律责任。2016年1月1日’
In [81]:	df3["答复意见"] [3]
Out [81]:	‘网友“UU008144”您好！您的留言已收悉。经A5区教育局调查了解，现将有关情况回复如下：1.恒大城幼儿园办园的相关依据。目前，学前教育以“政府主导、社会参与、公办和民办并举”为主要发展原则。《国务院关于加强当前发展学前教育的若干意见》（国发[2010]41号）明确：“城镇小区配套幼儿园作为公共教育资源由当地政府统筹安排，举办公办幼儿园或委托办成普惠性民办幼儿园”。按政策，恒大城幼儿园可以举办公办园，也可以举办民办普惠性幼儿园。2.恒大城幼儿园移交的相关情况。经核实，恒大城幼儿园用地性质属于划拨地，按政策应该无条件移交。目前，经多方协商，恒大城开发商同意向A5区政府移交幼儿园，具体移交事项正在进行中。3.工作进展情况。2月26日，区教育局参加了由洞井街道组织召开的恒大城幼儿园的推进情况进行研讨。期间，区教育局接待了恒大城业主的多轮来电来访，向广大业主宣传相关政策，告知具体工作推进情况。3月20日，区教育局经集体研究，已形成专题报告向A5区人民政府汇报，区政府将尽快研究确定恒大城幼儿园办性质。因幼儿园移交涉及到国土、建设等多个部门的协调、对接，希望广大业主予以理解、支持。感谢您对我们工作的关心、监督与支持。2017年3月22日您好！您的留言已收悉。现将有关情况回复如下：1、恒大城幼儿园为A市鑫鑫置业有限公司开发建设的恒大城楼盘配套幼儿园，于2013年即已建成，自2012年起其开发单位与我局就该楼盘配套的小学校及幼儿园移交事宜进行了多次协商，根据2014年底A市政府出台的有关配套幼儿园移交区政府作为公办或普惠性幼儿园，但多次协商一直未能达成一致意见。2、根据国家相关政策，配套幼儿园不得改变规划性质及用途，任何单位或个人举办幼儿园均应在所属地教育主管部门办理相关许可。目前恒大城幼儿园尚未开园，我局尚未接到办理幼儿园许可的申请。3、配套幼儿园产权移交地方政府后，方可开办公办幼儿园或民办普惠性幼儿园，目前恒大城幼儿园尚未移交。如有其他疑问，请咨询A5区教育局基建规划科，联系电话0731-0000-00000000。感谢您对我们工作的关心、监督与支持。
In [82]:	df3["答复意见"] [4]
Out [82]:	‘网民朋友：你好！现就你提出的问题回复如下。一、L5县城市管理环卫执法中队情况简介2019年L5县机构改革，原L5县环境卫生管理所更名为L5县城市管理事务中心，隶属于L5县城市管理综合执法局。为加强环卫执法力度，经上级批准，成立城市管理环卫执法中队，由城管局委派正副队长，城市管理事务中心选派执法队员12名，其中包括垃圾处理费征收工作人员。二、L5县城镇生活垃圾处理费征收动向近期，我单位正在对城区酒店、宾馆进行城镇生活垃圾处理费征收工作，共派出征收人员10人，分两组同时进行，每组各安排5人。同时要求征收人员在收费过程中必须佩戴工作牌证、带齐相关政策及收费文件，实行边收费边宣传收费政策，提高市民对征收政策及标准的知晓率。三、城镇生活垃圾处理费征收相关文件《城市生活垃圾管理办法》中华人民共和国建设部令第157号、《西地省城镇生活垃圾处理收费管理办法》楚发改价服〔474〕号、《L5县城镇生活垃圾处理费征收管理办法》[政府发文]25号、《关于重新公布L5县城市生活垃圾处理收费项目和标准的通知》（派发改价服[2017]6号）等。四、L5县城市管理事务中心不存在乱收费现象L5县城市管理事务中心是隶属于L5县城市管理综合执法局的副科级公益类事业单位，是L5县城镇生活垃圾处理费征收主体，所征收的垃圾处理费全部上缴县财政，且收费全部使用财政部门印发的专用收费票据，收费标准按《关于重新公布L5县城市生活垃圾处理收费项目和标准的通知》（派发改价服[2017]6号）执行。如果你发现有人单独上门（无工作牌证）或不使用财政专用收费票据而上门收取垃圾处理费，请及时致电0745—3322571（征收办）、0745—3322454（行政办），我们将依法依纪依规严肃处理。感谢您对环卫工作的关心和支持！2019年8月1日“UU008432”您好！您反映的问题已转县城市管理和综合执法局核查、办复，感谢您的留言，祝您事事顺心。2019年7月29日’

图 18 答复意见评价体系留言 TOP5

下面，针对图 18 的 5 条留言进行详细分析。后台数据显示这五条优质留言的及时性、逻辑性和专业性基本都没有差异，且都是最优值，可见使用具有专业

性、逻辑性的流畅语言去及时回复群众是高质量答复的普遍特征。

实验数据显示五条留言答复中排名 top1 的答复是关于“马石小学组织学生订购牛奶的质疑”，从答复来看，完整性（包含您好、问题已收悉、结尾日期等字眼）、可解释性（包含调查核实等字眼且有效文本率很高）和礼貌性（感谢语句）都是不错的。但相较于这些指标来讲，该答复最值得注意的是其相关性是最处于绝对的领先优势的，明显高于其他四条留言，可见相关性是评价回复质量的重要指标。

排名 top2 的留言答复，相较于其他六个指标，其可解释性明显优于其他四条留言。观察内容可以明显发现，该条回复内容提到《中华人民共和国土地管理法》、《基本农田保护条例》、《规范农村农村建房管理办法》等条文规定，使回复具有很强的说服力和可信度。

就 top3 的留言答复来看，其七个指标都没有任何排名首位的，其可解释性（未出现法律法规等相关文件支撑）和礼貌性（未出现感谢语句）都比较低，但由于相关性排名为第二，由于相关性的权重相对较高，因此使其脱颖而出，成为优质留言的代表。

排名 top4 的留言答复，每个指标都比较好，但是没有特别突出的指标，给阅读者的直观感受是中规中矩且具有耐心的答复，除结尾缺失答复日期外没有明显漏洞，可以算是一条很好的留言答复了。

最后位居 top5 的留言答复，其完整性、礼貌性非常符合规范，在五条留言里是排名首位的，但对比发现相关性却是五条里面最低的。也就是说该留言格式规范、态度友好，但回复内容并没有完全贴合留言问题，由此也再次印证了相关性对于回复质量的决定性作用。这也与本研究设立的答复意见评价体系的初衷是完全一致的。

此外，研究者从用户角度出发，进行了答复意见评分的有效实验进行结果对比分析。即在文本中随机筛选出 500 条实验样本，通过多人对其中的留言进行无标准的打分并求其均值进行排序，以观测和本研究设立的答复意见评价体系所得结果的差距。实验结果显示两个实验的均方误差 MSE 为 0.270，由此可见该评价体系具有很强的合理性和科学性，根据实际情况进行改进后即可被应用于不同的应用场景。

5 总结与不足

“智慧政务”在文本挖掘中的应用可以有效提高政府有关部门的执政能力以及群众满意度。本题的主要任务是致力于解决针对不同留言问题的多分类文本分类任务、针对同一问题的留言文本聚类任务以及针对不同留言答复的细粒度质量评价任务。因知识贮备和时间的限制，本次研究中仍然存在以下不足亟待改进：

（1）文本分类任务中，对于各种预训练语言模型所构建的分类器，参数仍然可以进行优化。与此同时，针对某一类类别，如“商贸旅游”类别，分类的效果要明显低于其他类别，对于这种影响模型总体效果的类别，研究也尝试过通过自定义词典优化文本内容从而增加文本的区分度，但是收效甚微，对于区分度较低的类别的优化还值得继续探究。

（2）文本聚类及热点问题评价任务中，针对数据集中留言存在的一些问题进行了相应的优化，但是仅仅是针对本题中的数据集，对于不同的领域、不同的平台的数据集优化的方式也会存在不同，在以后的任务中还需要具体问题具体分析，争取设计出泛化能力更强的文本聚类算法。

（3）留言答复质量评价任务中，主要采用的指标衡量都是基于关键词匹配以及特殊情况处理优化的方案，相关性采用相似度仍然有一定不足之处，容易忽略留言答复语义层面的质量，这也是后续需要继续改进优化的地方。

参考文献

- [1] 孟天广,李锋.网络空间的政治互动:公民诉求与政府回应性——基于全国性网络问政平台的大数据分析[J].清华大学学报(哲学社会科学版),2015,30(03):17-29.
- [2] 于书鹄.网络问政平台的“回应性陷阱”[D].吉林大学,2019.
- [3] Yoon Kim. Convolutional Neural Networks for Sentence Classification[C].E MNLP,2014.
- [4] Siwei Lai,Liheng Xu,Kang Liu,Jun Zhao.Recurrent Convolutional Neural Networks for TextClassification[C].AAAI,2015.
- [5] Wei, Jason W., and Kai Zou. "Eda: Easy data augmentation techniques for boosting performance on text classification tasks." Xiv preprint arXiv:1901.11196 (2019).
- [6] Takeru Miyato, Andrew M. Dai, Ian Goodfellow. Adversarial Training Methods for Semi-Supervised Text Classification[C].ICLR,2017.
- [7] 格桑多吉,乔少杰,韩楠,张小松,杨燕,元昌安,康健.基于 Single-Pass 的网络舆情热点发现算法[J].电子科技大学学报,2015,44(04):599-604.
- [8] 曾子明,黄城莺.基于 BP 神经网络的突发传染病舆情热度趋势预测模型研究[J].现代情报,2018,38(05):37-44+52.
- [9] 黄怡璇,谢健民,秦琴,杨丽颖.影响网络舆情热度评价的主要因素识别研究[J].情报科学,2017,35(10):49-54+62.
- [10] 周惠子,张杰.基于双链量子遗传投影寻踪模型的自媒体网络舆情热度与应对等级研究[J].情报杂志,2019,38(09):126-133.
- [11] 王晓光,王宏宇,黄菡.基于多源数据的专业领域热点探测模型研究[J].图书情报工作,2019,63(14):52-61.
- [12] 吴绍忠,李淑华.互联网络舆情预警机制研究[J].中国人民公安大学学报(自然科学版),2008(03):38-42.
- [13] 谈国新,方一.突发公共事件网络舆情监测指标体系研究[J].华中师范大学学报(人文社会科学版),2010,49(03):66-70.
- [14] 何跃,蔡博驰.基于因子分析法的微博热度评价模型[J].统计与决策,2016(18):52-54.
- [15] 梁昌明,李冬强.基于新浪热门平台的微博热度评价指标体系实证研究[J].情报学报,2015,34(12):1278-1283.
- [16] 莫祖英. 微博信息内容质量评价及其对用户利用的影响分析[D].武汉大学,2014.
- [17] 王青,成颖,巢乃鹏.网络舆情监测及预警指标体系构建研究[J].图书情报工作,2011,55(08):54-57+111.
- [18] 林文声,姚一源,王志刚.食品安全事件网络舆情热度评价研究:基于 BP 神经网络的方法[J].现代管理科学,2016(09):30-32.
- [19] 冯江平,张月,赵舒贞,陈虹.网络舆情评价指标体系的构建与应用[J].云南师范大学学报(哲学社会科学版),2014,46(02):75-84.
- [20] Yuan Qilong, Shi Haibo, Zhou Xiaofeng.An optimized initialization center Kmeans clustering algorithm based on density [C].IEEE,2015.
- [21] 陈文权,余雅洁.网络环境下服务型政府建设的回应性及路径研究——以 2013 年五省(市)书记和省长集中回复网友留言为例[J].中国行政管理,2014(07):74-77.
- [22] <https://wenku.baidu.com/view/65fa7abf5ff7ba0d4a7302768e9951e79b8969e8.html>.