

“智慧政务”中的文本挖掘

摘要

近年来，由于各类问政平台的完善，可想而知，各类社情民意相关的文本数据量不断攀升，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。因此，利用自然语言处理和文本挖掘的方法处理问政平台中的群众留言分类，热点问题挖掘，答复质量评价等问题对于智慧政务的建设具有重大意义。

对于问题一，我们首先用 pandas 库对 Excel 中所需信息进行快速提取，再用正则表达式来进行特殊除噪，并用 jieba 分词进行中文分词，加之导入停用表，使文本遍历之后去除停用词得到标准化的文本。其后，我们用基于 TF-IDF 算法的 TfidfVectorizer 将文本数据转化成特征向量来构建词向量，将文本中的词语转换为词频矩阵，使得语义相近的词在词向量空间里聚集在一起，这对后续的文本分类操作提供了便利。接着我们把标签和语料以 2: 8 的比例随机划分成训练集和测试集，采用 SVM 分类器对其训练和测试，构建出模型。最后我们用 classification_report 和 Confusion Matrix 得出模型评估后不同分类指标的值，说明了我们模型的可靠性与稳定性。

对于问题二，我们同样得到标准化文本，将二维数组通过 doc2bow 稀疏向量，形成语料库。再用 Lsi 模型算法，词典 token2id 的特征数，计算稀疏矩阵相似度。通过 doc2bow 计算测试数据的稀疏向量，求得测试数据与样本数据的相似度，然后进行迭代执行。将求得的相似问题进行排序，再通过热度指标来进行热度排行，最终得到热点问题表。

对于问题三，我们采用在服务质量测量与评价领域较为经典的由 Parasuraman 等^[1]提出的 SERVQUAL 理论模型，用该模型从有形性，可靠性，响应性，保证性，移情性这五个互不相容的层面对答复的质量进行科学性的评价，并且我们定义了五个指标的二级指标以提高评价的准确率。在此基础上，我们建立物元可拓模型，并用 Matlab 实现层次分析法的算法获得各个指标的权重，接着用关联度使得不相容问题能够量化得出的质量评价结果。最后，我们将数据带入模型中，得出答复质量总体较好这一评价。

最后，我们还对模型的准确性和有效性进行了评估，并且总结了模型的优缺点以及对未来的工作进行了展望。

关键词：jieba 词向量 SVM 分类 doc2bow Lsi 模型 物元可拓模型

Text mining in "Smart Government Affairs"

In recent years, due to the improvement of various political platforms, it is conceivable that the volume of text data related to various social conditions and public opinions has been increasing, it brings a great challenge to the work of related departments which used to rely mainly on manual to divide the message and arrange the hot spots. Therefore, the methods of natural language processing and text mining are of great significance to the construction of intelligent government, such as the classification of people's message, the mining of hot issues and the evaluation of response quality.

For the first question, we first use the pandas library to extract the information needed in the Excel , then use the regular expression to remove the noise, and use the Jieba participle to do Chinese participle, and import the Stop List, normalizes text by erasing stop words after traversing the text. Then, we use TfidfVectorizer based on TF-IDF algorithm to transform text data into feature vectors to construct word vectors, and transform words in text into Word Frequency Matrix. Then we divide the tag and Corpus into training set and test set randomly at the ratio of 2:8, and use SVM classifier to train and test them and build the model. Finally, we use classification and fusion Matrix to get the values of different classification indexes after the model evaluation, which shows the reliability and stability of our model.

For problem two, we also get standardized text, and pass the two-dimensional array through the doc2bow sparse vector to form a corpus. Then use the Lsi model algorithm, the feature number of the dictionary token2id, and calculate the similarity of the sparse matrix. The sparse vector of the test data is calculated by doc2bow, the similarity between the test data and the sample data is obtained, and then iterative execution is performed. Sort the obtained similar problems, and then through the heat index to rank the heat, and finally get the hot issues table.

For problem three, we adopt the SERVQUAL theoretical model proposed by Parasuraman et al. [1], which is relatively classic in the field of service quality measurement and evaluation. The model is used from the five aspects of tangibility, reliability, responsiveness, guarantee, and empathy. At a mutually inconsistent level, the quality of responses is scientifically evaluated, and we have defined two secondary indicators of five indicators to improve the accuracy of the evaluation. On this basis, we establish the matter-element extension model, use Matlab to implement the analytic hierarchy process algorithm to obtain the weight of each index, and then use the correlation degree to make the incompatible problem can be quantified by the quality evaluation results. Ultimately, we bring the data into the model and arrive at the evaluation that the quality of the response is generally superior.

Finally, we also evaluate the accuracy and effectiveness of the model, and summarize the advantages and disadvantages of the model and the outlook for future work.

Key words: jieba doc2bow Word vector SVM classification Lsi model
Matter-element extension model

目录

一. 问题重述.....	4
1.1 背景.....	4
1.2 挖掘目的.....	4
二. 模型分析.....	5
2.1 总体流程图.....	5
2.2 问题分析.....	5
三. 模型建立与求解.....	6
3.1 问题一的模型建立与求解.....	6
3.1.1 问题一流程图.....	6
3.1.2 解决问题.....	6
3.2 问题二的模型建立与求解：.....	8
3.2.1 总体流程图.....	8
3.2.2 解决方案.....	8
3.3 问题三的模型建立与求解.....	10
3.3.1 流程图.....	10
3.3.2 SERVQUAL 理论模型.....	10
3.3.3 指标分析.....	11
3.3.4 物元可拓模型.....	12
3.3.5 模型的结果分析与意见提出.....	17
四. 模型的评估.....	20
五. 模型的优缺点.....	21
5.1 模型的优点.....	21
5.2 模型的缺点.....	22
六. 模型的展望.....	22
七. 参考文献.....	23

一. 问题重述

1.1 背景

随着互联网技术的逐年发展，信息化技术已经遍布生活的各个角落，在政务方面亦然。“互联网+政务服务”是深化放管服改革的关键之举。^{【2】}

加之，近年以来随着微信、微博、市长信箱、阳光热线等网络问政平台逐步成为政府了解民意、汇聚民智、凝聚民气的重要渠道，各类社情民意相关的文本数据量不断攀升，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。同时，随着大数据、云计算、人工智能等技术的发展，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，它是智慧城市建设过程中的重要构成部分和发展基础，为其他领域的智慧应用提供重要支撑，更是电子政务发展的高级阶段。^{【3】}显然，它已经成为政府部门转变管理职能、整合业务流程和转变社会服务方式的重要途径。

1.2 挖掘目的

本次建模目标是利用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，利用自然语言处理和文本挖掘，达到以下三个目标：

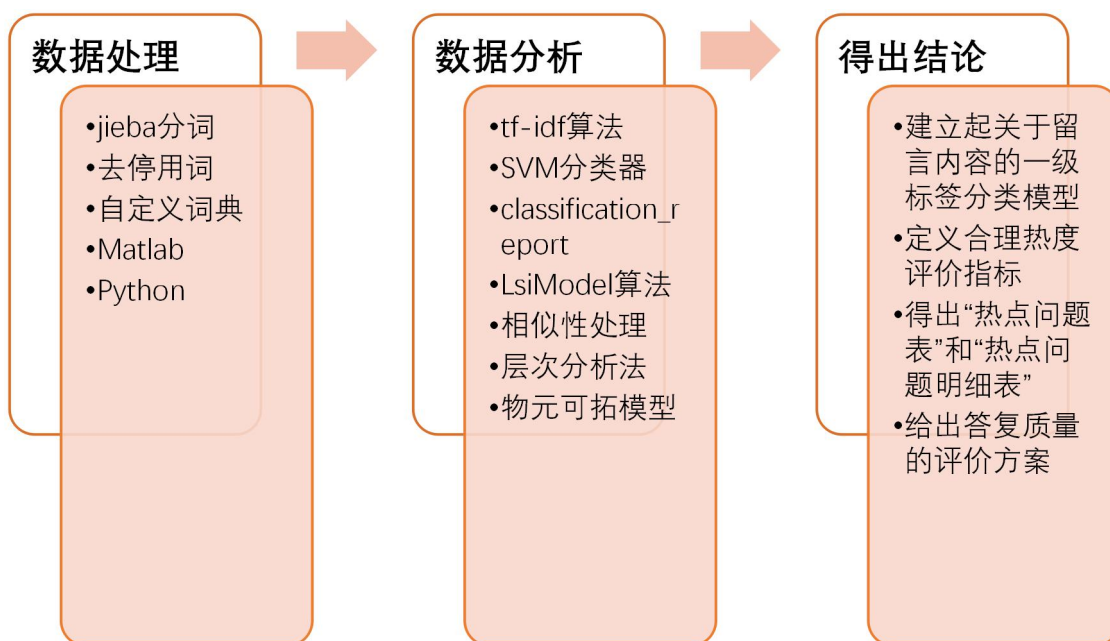
1. 利用文本分词和文本分类的方法对非结构化的数据进行测试，并建立起关于留言内容的一级标签分类模型，以便后续将群众留言分派至相应的职能部门处理。接着我们对模型进行评价，得到模型的查准率和查全率，以便我们对模型进行修改

2. 根据留言数据，建立热度统计模型，定义合理的热度评价指标，归纳总结出热度排名前五的“热点问题表”与“热点问题留言明细表”，以提升相关部门对热点问题的回复速度。

3. 根据相关部门对留言的答复意见，对答复意见的质量给出一套评价方案，为答复意见提供真实可靠的建议，并且尽量对现有数据进行评价。

二. 模型分析

2.1 总体流程图



2.2 问题分析

(1) 问题一的分析

为了对建立关于留言内容的一级标签分类模型，我们需要将非结构化数据转化为结构化数据，然后对其进行分类，建立模型，最后对建立的模型进行评估。

(2) 问题二的分析

寻找热点问题首先要将所有的留言问题筛选分类，将他们归纳为一个不同的类别，然后将他们进行热度排行，归纳出热点问题。

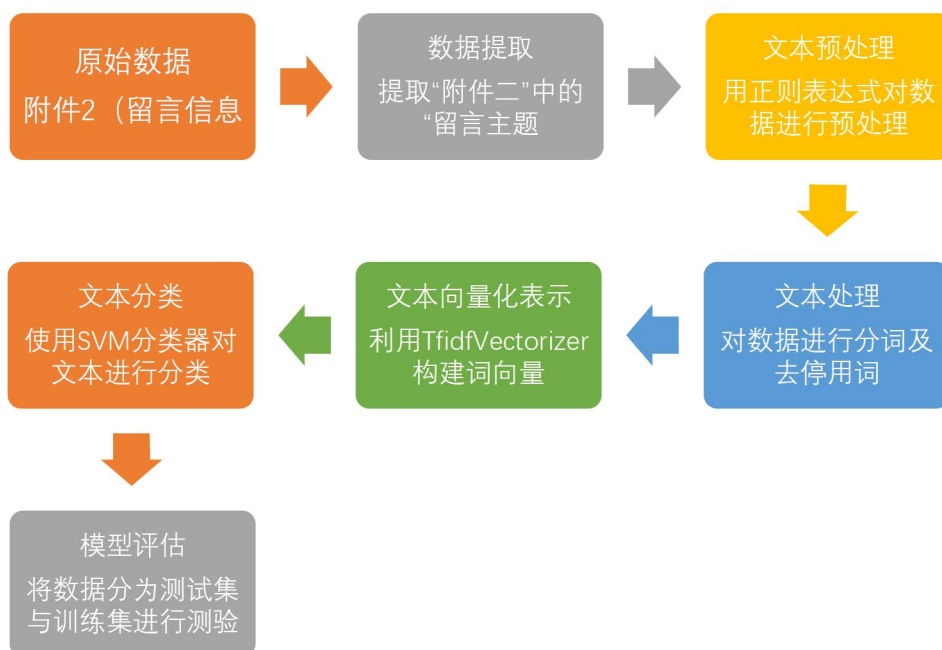
(3) 问题三的分析

为了对答复意见的质量给出一套评价方案，我们需要提出一系列的科学系评价指标，并且我们需要对这些指标的权重进行计算，同时我们还需要将答复意见的质量的等级进行定量化，从而得出过呢国家科学可信的评价结果。

三. 模型建立与求解

3.1 问题一的模型建立与求解

3.1.1 问题一流程图



3.1.2 解决问题

（1）数据抽取

我们用 pandas 读取 Excel 文件，并赋值为“data”，接着我们对文件的列名称进行赋值，依次名为“留言编号”，“留言用户”，“留言主题”，“留言时间”，“留言详情”和“一级标签”，方便我们接下来读取我们所需要的列数据。

（2）文本预处理

我们通过正则表达式对“data”中的“留言主题”去除掉无关字符，以便我们进行后面的操作。

（3）文本处理

文本处理共需要两个过程，一是进行分词，二是去除停用词。预处理中的语料库都是没有分词的原始语料（即连续的句子，而后面的工作需要我们把文本分为一个个单词），我们采用 jieba 分词对这些文本进行分词，只有这样才能在基于单词的基础上，对文档进行结构化表示。接着我们去停用词，我们将文本中出现频率很高，但实际意义又不大的词（这一类主要包括了语气助词、副词、介词、连词等，通常自身并无明确意义的词语）汇总成一份停用词表，在分词后把结果中包含的停用词剔除。这样能够让所优化的关键词更集中、更突出，提高搜索性能。

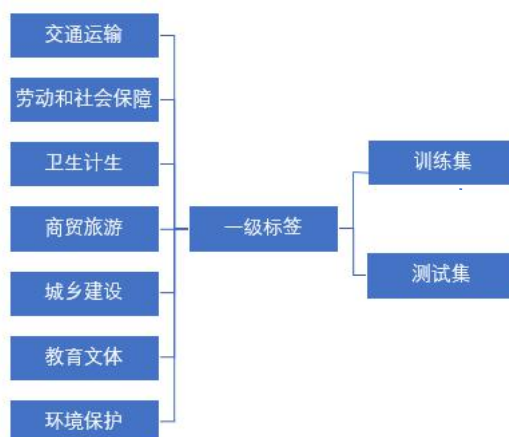
（4）文本向量化表示

我们为了将非结构化数据转为结构化数据，可以达到语义相近的词在词向量空间里聚集在一起，这对后续的文本分类操作提供了便利。据我们研究，将文本数据转化成特征向量的过程比较常用的有两种函数，一种是 CountVectorizer 函数，另一种是基于 TF-IDF 算法的 TfidfVectorizer 函数。但因为文本条目越多，TfidfVectorizer 的效果会越显著。故本文在构建词向量上采用基于 TF-IDF 算法的 TfidfVectorizer 函数。

（5）文本分类

据我们研究，文本分类器常用的有三种，分别是 NB 分类器，KNN 分类器和 SVM 分类器。经过测试，我们发现使用 SVM 分类器得出的查全率和查准率会比另外两种高，故我们采用 SVM 分类器来建立模型。我们读取列数据“一级标签”得到“交通运输”，“劳动和社会保障”，“卫生计生”，“商贸旅游”，“城乡建设”，“教育文体”，“环境保护”七大类，然后把对应的标签和语料以 2:8 的比例随机划分成训练集和测试集，放到分类器中进行训练和测试。

图 1: SVM 分类器训练的对象来源



3.2 问题二的模型建立与求解:

3.2.1 总体流程图



3.2.2 解决方案

(1) 数据的读取及文本处理

我们用 pandas 获取 excel 中的信息，提取名为“附件 3”的 xlsx 文件中的 C 列“留言主题”，然后用 jieba 分词来进行处理。

首先进行自定义词典的编辑，我们将许多的地名与必要的词语输入进去使分词时更加有规律，然后加入停用词将与关键词无关紧要的停用词（语气助词、副词、介词、连词和通常无意义的词语）去掉，从而更准确的将文本主题进行分类。

(2) 处理 excel

在这里我要新增一个名为“sheet1”的 sheet，表头有三个，分别为

“热点问题名称”，“热点问题数量”和“热点问题链接”，其中“热点问题链接”带超链接，可以直接导航到热点问题的明细数据。

(3) 相似性处理

在这里我们是用 gensim 中的 corpora 模板，然后将分词形成后的二维数组成词典，再将二维数组通过 doc2bow，形成语料库。之后就利用 Lsimodel 模型算法，将语料库计算出 Tf-idf 值。最后获取词典的 token2id 的特征数计算系数矩阵相似度，建立一个索引并生成一个 excel 表格。

(4) 热度排行

这个问题上，我们将用两步来完成，分别为“初步热度排行”和“进阶热度排行”。

①初步热度排行：我们认为一个问题的热度体现首先要从他的提出来体现。这里我们采用 excel 对分类后的问题进行数量上的排行，初步选定发布数量上最多的几个问题作为热度问题并进入下一步的进阶热度排行。

②进阶热度排行：为了排行这些热点问题，我们定义了一个话题热度的指标 hotdegree，该指标的公式如下：

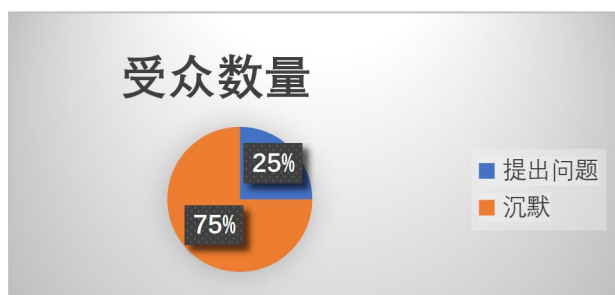
$$hot\ degree = \alpha * \text{话题留言总数} + \beta * (\text{点赞量} - \text{反对量}) + \gamma * \text{单位时间内留言数量}$$

其中： $\alpha + \beta + \gamma = 1$

根据调查可知通常会将问题进行反应的人占经历同样问题的人群（简称受众）的 1/4。^{【4】}

则我们可以把每个进行点赞的人可以代表一位受众，则发出问题的可以代表三位受众。

图 2：受众数量分析



单位时间设置为一个月，则可以计算出一个月内的受众数量，所以 $\alpha: \beta: \gamma=1:3:1$ 。计算出每个问题的热度指标，并进行排序，则可得到最终的五个热点问题。

(5) 结果分析

通过对热点问题的进阶筛选，我们最终整理出了这五个热点问题及其热度指标，见下表：

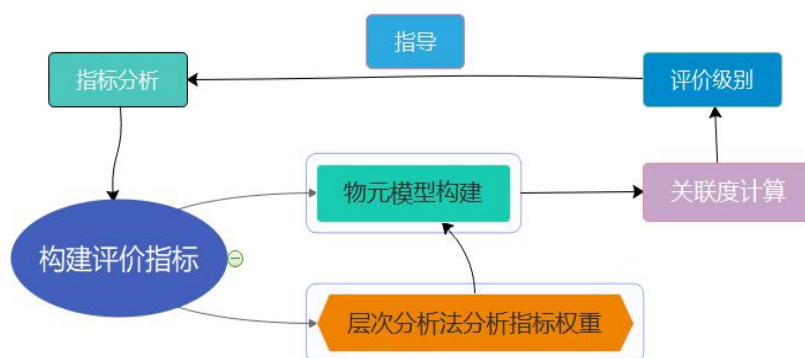
表 1：热点问题排列表

热点问题	发 布 量	点 赞 量 - 反 对 量	时间范围 (天)	单 位 时 间 发 布 量 (月)	热度指标
小区强制捆绑销售车位	40	22	35	121.71	50.74
搅拌站噪音扰民	37	47	73	64.93	44.59
对地铁的建议	26	69	231	19.09	33.22
对人才购房补贴的疑问	25	24	264	11.25	22.05
茶场村拆迁问题	10	0	308	2.92	6.60

再由热度指标大小决定热点问题的排行。

3.3 问题三的模型建立与求解

3.3.1 流程图



3.3.2 SERVQUAL 理论模型

相关研究表明，SERVQUAL 模型适合于信息系统服务质量的测量与评价。^[5]鉴于 SERVQUAL 模型在服务质量评价上的系统性优势，我们就在前人研究的基础上引入该模型构建智慧政务中答复质量的评价指标体系。Parasuraman 等^[6]我们在总结国内外专家的研究后，将该模型进行简化和归纳，最终确定了有形

性、可靠性、响应性、保证性和移情性 5 个评价维度并对这五个指标进行二级指标划分来构建评价体系，以进行更加进行科学性的评价，如表 2 所示。

表 2：智慧政务中对群众留言的答复质量评价体系

一级指标	二级指标	主要测评度
有形性(C ₁)	平台稳定性(C ₁₁)	平台运行稳定
	界面吸引性(C ₁₂)	界面设计美观，有吸引力
	网页导航清晰性(C ₁₃)	服务功能导航清晰，结构合理
可靠性(C ₂)	隐私安全性(C ₂₁)	平台具有匿名等功能
	回答信息可靠性(C ₂₂)	答复较为精准
响应性(C ₃)	回复时效性(C ₃₁)	咨询结构反馈及时
	检索便捷性(C ₃₂)	平台提供集成检索、分类检索
保证性(C ₄)	信息权威性(C ₄₁)	平台有大量权威专家指导
	平台声誉度(C ₄₂)	平台有良好的社会声誉
移情性(C ₅)	回复易读性(C ₅₁)	反馈内容表达方式合理、易于理解
	服务个性化(C ₅₂)	针对用户个性化需要推送服务
	情感关怀度(C ₅₃)	注重用户情感体验与心理疏导

3.3.3 指标分析

(1) 有形性

有形性指用户直接感知到的关于产品或服务涉及到的器材、设备与装饰等基本情况。而在智慧政务的答复建议评价过程中，我们考虑到电子政务平台的运用，并结合李宇佳对信息服务有形性的界定^[7]，我们就考虑平台稳定性，界面吸引性，网页导航清晰性这三个二级指标。

(2) 可靠性

可靠性关注服务提供者准确履行承诺的能力。我们考虑到是在互联网上进行的反馈留言，并根据 Van Iwaarden 等^[8]构建的网站信息服务评价指标我们考虑隐私安全性以及回答信息可靠性这两个二级指标。

(3) 响应性

响应性指服务提供者帮助顾客的能力及迅速提高服务水平的意愿。我们参考 VanIwaarden 与李立睿等^[9]的相关研究,考虑了回复时效性,检索便捷性这两个二级指标。

(4) 保证性

保证性指服务提供者所具有的知识、礼节以及表达出自信与可信的能力。我们参考 Rezaei 等^[10]的界定,认为服务提供者是否值得信赖和解决问题的知识储备是衡量保证性的重要指标,我们考虑信息权威性和平台声誉度。

(5) 移情性

移情性指服务者关心用户并为顾客提供个性化服务的能力。参考左文明^[11]等设计的网约车服务质量移情性指标,我们考虑了回复易读性、服务个性化和情感关怀度这三个二级指标。

3.3.4 物元可拓模型

由于我们在建立的智慧政务答复质量评价指标体系到的是有形性、可靠性、响应性、保证性和移情性等 5 个互不相容的层面,因此可以运用物元可拓模型来表征与评价答复意见的评价模型,拓展了关联函数的值域范围,且将评价指标由固定值转变为区间值,使得评价对象相对于评价集合的隶属度计算更具客观性、准确性。^{【12】}

在物元模型中,物元是描述事物名称(Name)、特征(Character)及其量值(Value)三个基本元素的简称,在形式上可记为 $R=(N, c, v)$. 在答复意见的评价中,我们认为答复意见评价等级为“物 N”、具体评价指标为“特征 c”,及 N 关于 c 的特征值 v,共同构成答复意见评价分析的一维物元。如果分析或评价对象 N 有 n 个特征向量 c_1, c_2, \dots, c_n 和相应的特征值 v_1, v_2, \dots, v_n ,可表示为如下公式:

$$R = \begin{bmatrix} N & c_1 & v_1 \\ & c_2 & v_2 \\ & \vdots & \vdots \\ & c_n & v_n \end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{bmatrix} \quad (1)$$

式中：R 为 n 维答复意见评价分析物元，记为 $R=(N, c, v)$

我们将构建物元可拓模型的步骤分为一下：

(1) 答复质量评价分析的经典域、节域确定

物元模型的评价对象在不同等级上的特征值往往有一定的取值范围，称为经典域。我们假定答复的质量等级可以划分为 m 个等级，构建的评价指标共有 n 个，则经典域物元模型如公式(1)所示。

答复意见评价分析的经典域物元矩阵可以表示为：

$$R_{oj} = (N_{oj}, c_i, v_{oji}) = \begin{bmatrix} N_{oj} & c_1 & v_{oj1} \\ & c_2 & v_{oj2} \\ & \vdots & \vdots \\ & c_n & v_{ojn} \end{bmatrix} = \begin{bmatrix} N_{oj} & c_1 & (a_{oj1}, b_{oj1}) \\ & c_2 & (a_{oj2}, b_{oj2}) \\ & \vdots & \vdots \\ & c_n & (a_{ojn}, b_{ojn}) \end{bmatrix} \quad (2)$$

其中： R_{oj} 为答复意见评价分析的经典域； N_{oj} 为所划分的第 j 个评价等级（ $j=1, 2, \dots, n$ ）； c_i 为第 i 个评价指标（ $i=1, 2, \dots, n$ ）； (a_{ojn}, b_{ojn}) 为 c_i 对应的评价指标的取值范围，即经典域。

而与其相应答复重量评价分析的节域物元矩阵可以表示为：

$$R_p = (N_p, c_i, v_{pi}) = \begin{bmatrix} N_p & c_1 & (a_{p1}, b_{p1}) \\ & c_2 & (a_{p2}, b_{p2}) \\ & \vdots & \vdots \\ & c_n & (a_{pn}, b_{pn}) \end{bmatrix} \quad (3)$$

式中： R_p 为答复意见评价分析的节域； N_p 为答复意见评价分析中的全部评价等级； $v_{pi}=(a_{pn}, b_{pn})$ ，为 N_p 对应的评价指标的取值范围，即为全体评价等级的节域。

(2) 待评物元模型确定

假设待评价的答复意见的质量等级为 N_x ，则待评价对象的物元可以表示如

下:

$$R_i = \begin{bmatrix} N_x & c_1 & v_1 \\ & c_2 & v_2 \\ & \vdots & \vdots \\ & c_n & v_n \end{bmatrix} \quad (4)$$

其中, c_i 表示已经确定的答复意见评价评价指标, v_i 表示待评价的答复意见在各个指标上的实际测量值。

(3) 计算待评物元的可拓关联函数

在可拓理论中, 关联函数用来表征元素具有某种性质的程度。基于物元可拓模型研究智慧政务答复意见质量的过程中, 需要建立不同的质量等级, 并确定各个等级下相应指标的取值范围, 而关联函数的计算能够帮助我们确定待评价对象与各个等级的接近程度。

令有界区间 $V_0 = [a, b]$ 的模定义为: $|V_0| = |b - a|$ 。某点 v 到 $[a, b]$ 的距离为:

$$\rho(v, V_0) = \left| v - \frac{1}{2}(a + b) \right| - \frac{1}{2}(b - a) \quad (5)$$

从上式可以分别得到待评物元与经典域和节域的接近程度, 进而可以得到答复意见评价指标关联度如下。可拓集合的关联函数可用代数式来表达, 就使得解决不相容问题能够定量化。关联度用于刻画待评事物各指标关于各评价等级 j 的归属程度:

$$K_j(v_i) = \begin{cases} \frac{\rho(v, V_0)}{|V_0|}, v \ni V_0 \\ \frac{\rho(v, V_0)}{\rho(v, V_p) - \rho(v, V_0)}, v \notin V_0 \end{cases} \quad (6)$$

其中 $K_j(v_i)$ 表示智慧政务答复意见质量的第 i 个评价指标与第 j 个评价等级之间的关联度。

按照关联度最大识别原则, 初步评定待评价物元所属的评价等级 j 。若满足下式, 我们就称评价指标 v_i 属于安全等级 j 。

$$K_j(v_i) = \max K_j(v_i) \quad (7)$$

(4) 用层次分析法确定权重

为有效避免主观赋权带来的偏差,我们采用层次分析法来确定各个指标的权重系数 W_i , 该分析法主要利用所有指标的内在层级关系对评价对象的本质影响, 建立层次结构模型^[13]把决策思维的过程数学化、矩阵化。其流程可表示如下:



由于是对群众留言的答复, 我们认为还是要从群众的角度出发来调查答复的质量, 于是我们在对身边不同年龄段的群众进行采访调查, 向有关专业人员进行咨询, 对所有指标进行两两比较判断的基础上, 用成对比较法和 1~9 尺度对所有判断结果取均值后构造成判定矩阵如下:

$$A_1 = \begin{bmatrix} 1 & 4 & 2 \\ 1/4 & 1 & 1/3 \\ 1/2 & 3 & 1 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 1 & 1/2 \\ 2 & 1 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 1 & 3 \\ 1/3 & 1 \end{bmatrix}$$

$$A_4 = \begin{bmatrix} 1 & 2 \\ 1/2 & 1 \end{bmatrix}$$

$$A_5 = \begin{bmatrix} 1 & 1/2 & 3 \\ 2 & 1 & 5 \\ 1/3 & 1/5 & 1 \end{bmatrix}$$

对判定矩阵特征向量归一化后可得到对应权重:

$$\omega_1 = \begin{bmatrix} 0.5584 \\ 0.1220 \\ 0.3196 \end{bmatrix} \quad \omega_2 = \begin{bmatrix} 0.3333 \\ 0.6667 \end{bmatrix} \quad \omega_3 = \begin{bmatrix} 0.7500 \\ 0.2500 \end{bmatrix} \quad \omega_4 = \begin{bmatrix} 0.5528 \\ 0.4472 \end{bmatrix} \quad \omega_5 = \begin{bmatrix} 0.3090 \\ 0.5816 \\ 0.1095 \end{bmatrix}$$

由于逻辑判断是有一定的错误率得, 所以我们需要对 AHP 的判断矩阵进行一致性检验, 在确认其逻辑上的合理性后, 才能保证分析结果的准确性。

首先, 我们要算出计算判断矩阵 A 的最大特征根 λ_{\max} :

$$\lambda_{\max} = \sum_{i=1}^n \frac{(AW)_i}{nW_i} \quad (8)$$

$$A \bullet W = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \bullet \begin{bmatrix} W_1 \\ W_2 \\ W_3 \\ W_4 \end{bmatrix}, (A \bullet W)_i = a_{ij} \quad (9)$$

随后，我们查同阶矩阵平均一致性指标 IR 表，并结合下式我们可以算出一致性比率 CR ：

$$\begin{cases} CI = \frac{\lambda_{\max} - n}{n - 1} \\ CR = \frac{CI}{RI} \end{cases} \quad (10)$$

当 $CR < 0.1$ 时， A 具备满意一致性。当 $CR \geq 0.1$ 时， A 具备非满意一致性，应调整或舍弃。当 $CR = 0$ ， A 具备完全一致性。【14】

我们在 matlab 中对我们所定的判断矩阵运行分析后的结果如下：

表 3: 用 Matlab 对矩阵的分析结果

A_i	A_1	A_2	A_3	A_4	A_5
CI_{Ai}	0.0091	0	0	0	0.0018
CR_{Ai}	0.0158	0	0	0	0.0032

以上结果均小于 0.1，所以判定矩阵均满足一致性要求。

另外在 SERVQUAL 模型中，我们运用了五个指标，在对于智慧政务的答复意见评价分析中，我们认为这五个评价指标的影响性是相同的，即认为这五个一级指标的权重相同，随后我们也通过层次分析使我们的这一假设得到了验证。

(5) 判定答复意见的质量等级

计算待分析的对群众留言的答复 R_i 关于等级的加权关联度为：

$$K_j(R_i) = \sum_{i=1}^m w_i K_{ji} \quad (11)$$

式中： $K_j(R_i)$ 为待分析答复意见 R_i 的符合评价等级的隶属度。且 $\sum_{i=1}^m w_i = 1$

经过以上计算可以得到待评价物元与所有评价等级的综合关联度，按照关联度最大识别原则，初步评定待评价物元所属的评价等级 j_0 其计算过程如公式 (12) 所示：

$$K_{i_0}(Q) = \max K_j(Q) \quad (12)$$

对待评价物元与各评价等级的综合关联度进行标准化处理可得到标准化的综合关联度如下：

$$\overline{K_j}(R_t) = \frac{K_j(R_t)}{(\max K_j(R_t) - \min K_j(R_t))} \quad (13)$$

则可以计算待评价物元的评价级别变量特征值如下：

$$j^* = \frac{\sum_{j=1}^m j \overline{K_j}(R_t)}{\sum_{j=1}^m \overline{K_j}(R_t)} \quad (14)$$

式中： j^* 为待评答复质量 R_t 的变量特征值，通过 j^* 就可以得知 R_t 偏向哪个等级。

3.3.5 模型的结果分析与意见提出

在表 1 所确定的智慧政务中对群众留言的答复质量评价指标体系的基础上，我们结合所给数据，对答复质量等级进行划分，我们将全部评价指标分别分为 4 个等级区间，即 N_1 （差）、 N_2 （一般）、 N_3 （较好）以及 N_4 （优秀），进而确定各个指标的节域及在不同评价等级的经典域。随后，我们对不同评价指标设计问卷，对身边不同年龄段的人以及进行问卷调查并咨询专家意见，最后计算得分的平均数并归一化处理后确定答复意见评价的经典域、节域和实际取值如表 4 所示。

表 4: 智慧政务答复质量的物元模型取值区间

评价指标		评价等级				节域	实际值
一级指标	二级指标	N_1 (差)	N_2 (一般)	N_3 (较好)	N_4 (优秀)	N_{ip}	N_i
(C_1)	(C_{11})	$[0.10, 0.30)$	$[0.30, 0.50)$	$[0.50, 0.80)$	$[0.80, 1]$	$[0.10, 1]$	0.70
	(C_{12})	$[0, 0.50)$	$[0.50, 0.65)$	$[0.65, 0.90)$	$[0.90, 1]$	$[0, 1]$	0.85

	(C_{13})	$[0, 0.25)$	$[0.25, 0.50)$	$[0.50, 0.75)$	$[0.75, 1]$	$[0, 1]$	0.33
(C_2)	(C_{21})	$[0.20, 0.40)$	$[0.40, 0.60)$	$[0.60, 0.80)$	$[0.80, 1]$	$[0.20, 1]$	0.43
	(C_{22})	$[0, 0.35)$	$[0.35, 0.70)$	$[0.70, 0.90)$	$[0.90, 1]$	$[0, 1]$	0.73
(C_3)	(C_{31})	$[0.10, 0.30)$	$[0.30, 0.50)$	$[0.50, 0.80)$	$[0.80, 1]$	$[0.10, 1]$	0.60
	(C_{32})	$[0, 0.35)$	$[0.35, 0.60)$	$[0.60, 0.80)$	$[0.80, 1]$	$[0, 1]$	0.87
(C_4)	(C_{41})	$[0, 0.40)$	$[0.40, 0.60)$	$[0.60, 0.80)$	$[0.80, 1]$	$[0, 1]$	0.53
	(C_{42})	$[0, 0.30)$	$[0.30, 0.60)$	$[0.60, 0.80)$	$[0.80, 1]$	$[0, 1]$	0.63
(C_5)	(C_{51})	$[0.30, 0.40)$	$[0.40, 0.60)$	$[0.60, 0.85)$	$[0.85, 1]$	$[0.30, 1]$	0.67
	(C_{52})	$[0, 0.30)$	$[0.30, 0.65)$	$[0.65, 0.80)$	$[0.80, 1]$	$[0, 1]$	0.67
	(C_{53})	$[0, 0.40)$	$[0.40, 0.70)$	$[0.70, 0.90)$	$[0.90, 1]$	$[0, 1]$	0.80

接下来，我们依据表格数据并结合公式（5）和公式（6）求得各个二级指标关于各个评价等级的关联度，结果如表 5 所示：

表 5: 智慧政务答复质量二级评价指标与各评价等级的关联度

二级指标	$K_1(V_{ik})$	$K_2(V_{ik})$	$K_3(V_{ik})$	$K_4(V_{ik})$
C_{11}	-0.571	-0.400	-0.333	-0.250
C_{12}	-0.700	-0.571	0.500	-0.250
C_{13}	-0.195	0.320	-0.340	-0.560
C_{21}	-0.115	0.150	-0.425	-0.617
C_{22}	-0.585	-0.100	0.150	-0.386
C_{31}	-0.429	-0.200	0.333	-0.333
C_{32}	-0.800	-0.675	-0.350	0.350
C_{41}	-0.217	0.350	-0.130	-0.365

C_{42}	-0.471	-0.075	0.150	-0.315
C_{51}	-0.450	-0.175	0.280	-0.353
C_{52}	-0.529	-0.057	0.065	0.283
C_{53}	-0.667	-0.333	0.500	-0.333

随后，根据公式(11)计算一级评价指标与各个评价等级的关联度，并根据公式（12）初步评定待评价物元所属的评价等级 j_0 ，再根据公式（13）以及公式（14）得到评价级别变量特征值，更精确地表明答复质量更偏向于那个等级，结果如表 6 所示。

表 6:智慧政务答复质量一级评价指标与各评价等级的关联度

一级指标	$K_1(Q_i)$	$K_2(Q_i)$	$K_3(Q_i)$	$K_4(Q_i)$	j_0	j^*
C_1	-0.467	-0.191	0.138	-0.349	3	2.842
C_2	-0.428	-0.017	-0.042	-0.463	2	2.428
C_3	-0.522	-0.319	0.162	-0.162	3	3.126
C_4	-0.331	0.160	-0.005	-0.343	2	2.382
C_5	-0.522	-0.124	0.179	0.020	3	1.751

接着，基于以上结果，我们继续利用公式(11)得到答复质量与各评价等级的整体关联度，并同前面一样得出初步评价等级以及评价级别变量，如表 7 所示：

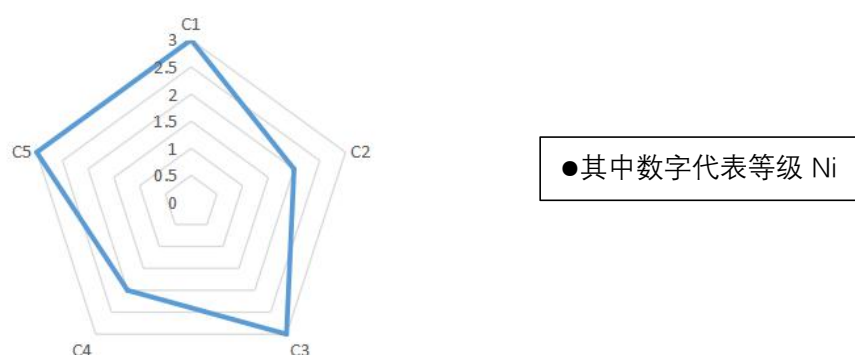
表 7:智慧政务答复质量中各评价等级的关联度

总目标	$K_1(Q)$	$K_2(Q)$	$K_3(Q)$	$K_4(Q)$	j_0	j^*
C	-0.454	-0.098	0.086	-0.223	3	3.092

最后，我们可以确定答复质量评价等级并对于答复意见给出建议。结合初步评价等级 j_0 和级别变量特征值 j^* ，得到智慧政务中对群众留言答复的质量整

体评价等级为“较好”，并且其级别变量特征值为 3.092，表示评价等级有偏向于“优秀”的倾向。同样的，我们可以根据表 5，得到各个一级指标的实际所属等级，如图 3 所示。根据我们的结果，我们可以知道响应性(C_3)和移情性(C_5)的等级为“较好”偏向“优秀”，而有形性(C_1)、可靠性(C_2)和保证性(C_4)为“较好”趋向“一般”，这说明我们在智慧政务中群众对响应性和移情性这种感知更为深刻，所以我们建议政府部门在处理社会事务的时候在更注重人文关怀的同时也不可以忽视了政策的实效性，要建立权威性，以增加群众的信任感。

图 3：智慧政务答复质量评价各个一级指标等级分布图



四. 模型的评估

我们知道常用的评估标准有：

召回率 (Recall)：

$$recall_i = \frac{a}{a+c} \times 100\%$$

准确率 (Precision)：

$$precision_i = \frac{a}{a+b} \times 100\%$$

F1-评测值 (F1-measure)：

$$F_{1i} = \frac{2 \times precision_i \times recall_i}{precision_i + recall_i}$$

这里我们采用 `classification_report` 函数和 `Confusion Matrix` 函数来对模型进行评估。`sklearn` 中的 `classification_report` 函数用于显示主要分类指标的文本报告。能够在报告中显示每个类的精确度，召回率，F1 值等信

息。 Confusion Matrix 函数所求出的矩阵对角线的值就是每一个类正确分类的文本数。

表 8: 分类结果混淆矩阵示意表

	c_1	c_2	\dots	c_p
c_1	a_{11}	e_{12}	\dots	e_{1p}
c_2	e_{21}	a_{22}	\dots	e_{2p}
\vdots	\vdots	\vdots	\vdots	\vdots
c_p	e_{p1}	e_{p2}	\dots	a_{pp}

并可以计算出以下数据:

表 9: 对模型的评估结果

	precision	recall	F1-score	support
交通运输	1.00	0.05	0.10	135
劳动和社会保障	1.00	0.05	0.10	390
卫生计生	1.00	0.05	0.10	160
商贸旅游	1.00	0.10	0.19	234
城乡建设	0.23	1.00	0.38	410
教育文体	1.00	0.07	0.13	311
环境保护	1.00	0.06	0.12	203

由上表数据可见, 数据契合度极高, 所以模型具有极佳的可行性, 可以投入使用。

五. 模型的优缺点

5.1 模型的优点

- SERVQUAL 理论模型是一个成熟的适合于信息系统服务质量的测量与评价的模型, 我们是以此为基础建立模型的;
- 我们运用层次分析法求权重, 这个方法适用于复杂问题或非单一评估标准下的决策问题;
- 用物元可拓法可以最大限度地将不相容的矛盾转化为相容关系, 从而实现全局性最佳决策目标。

- 热度排行模型大幅度的降低了人工操作，使人力在分辨热点问题上得以解脱。可以使各个热点问题的明细归纳到一个 sheet，所作出的热点问题表十分便于整理和操作，并且经过调试将相似度调节到大于 75%为最佳，使模型的效率达到最高。

5.2 模型的缺点

- 美中不足的使在输入 jieba 分词的特定词典时可能因为地点特别多而杂使得无法全部容纳与分辨，中文的同音同义词可能会出现某些不可避免的差池，但是对结果影响甚微。

- 我们的数据还存在一定的局限性，如果可以给出更多的数据，我们的模型将更精确；

- 虽然在答复质量评价模型中我们有参考国内外专家的研究成果并且有咨询专家建议，但是我们还是基于群众的主观评价，会有一定的主观误差性。

六. 模型的展望

它可以为智慧城市建设过程中提供重要的反馈机制, 显然, 它将为政府部门转变管理职能、整合业务流程和转变社会服务方式提供重要的技术基础, 早日达成官民一家亲的美好场景。

由于时间的局促性以及数据不全面性，我们的模型还有改进的方面。未来，如果条件允许的话，我们将会对 jieba 分词中的词典进行改进与完善，并且我们将会增大我们的数据量，在权重估计方面采用多种方法求权重，并用最小二乘法将其结合在一起，这将会进一步减少我们模型的主观误差并且会使我们的结果更加具有科学性。

我们相信，在大数据、云计算、人工智能等技术蓬勃发展，各个问政平台日新月异的当今社会中，我们的模型将有很大的发展空间。

七. 参考文献

- [1]Parasuraman A, Zeithaml V A, Berry L L. A Conceptual Model of Service Quality and Its Implications for Future Research[J]. Journal of Marketing, 1985, 49(4): 41-50.
- [2]刘洁. 智慧政务 APP 应用问题与对策[J]. 合作经济与技, 2020(08):169-171.
- [3]张晓海, 邓贤峰. 南京智慧政务总体架构设计思路与实现路径[J]. 电子政务, 2013(09):103-109.
- [4]王宏勇. 网络舆情热点发现与分析研究[D]. 西南交通大学, 2011
- [5]Van Dyke T P, Kappelman L A, Prybutok V R. Measuring Information Systems Service Quality: Concerns on the Use of the SERVQUAL Questionnaire[J]. MIS Quarterly, 1997, 21(2): 195-208.
- [6]Parasuraman A, Zeithaml V A, Berry L L. SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality[J]. Journal of Retailing, 1988, 64(1): 12-40.
- [7]李宇佳. 学术新媒体信息服务模式与服务质量评价研究[D]. 吉林学, 2017.
- [8]Van Iwaarden J, Van der Wiele T, Ball L, et al. Applying SERVQUAL to Web Sites: An Exploratory Study[J]. International Journal of Quality & Reliability Management, 2003, 20(8): 919-935.
- [9]李立睿, 王博雅, 邓仲华. 移动阅读应用服务质量的拓展评价研究[J]. 情报杂志, 2016, 35(7): 201-207.
- [10]Rezaei J, Kothadiya O, Tavasszy L, et al. Quality Assessment of Airline Baggage Handling Systems Using SERVQUAL and BWM[J]. Tourism Management, 2018, 66:85-93.
- [11]左文明, 朱文锋. 分享经济下基于 SERVQUAL 的网约车服务质量管理研究——以滴滴出行和优步为例[J]. 管理案例研究与评论, 2018, 11(4): 349-367.
- [12]蔡文物元模型及其应用[M]. 北京:科学技术文献出版社, 1994:1-12.
- [13]周宇. 基于模糊层次分析法的电信行业评分权重设置方法[J]. 中国招标, 2019(48):35-40.
- [14]赵蓉英, 王嵩. 基于熵权物元可拓模型的图书馆联盟绩效评价[J]. 图书情报工作, 2015, 59(12):12-18.