“智慧政务”中的文本挖掘应用

摘要

本文旨在利用机器学习模型帮助政府相关部门实现自动化留言分类，挑选出其中的热点问题，并对政府的回复意见进行客观评价。

针对问题一群众留言分类问题，我们首先选取了自己所需的数据，然后对截取数据进行了预处理，以达到我们的使用要求，再建立多个模型对对数据集进行训练测试，选出每个模型的最优参数，再选出最优模型，本文对于此文采用了线性分类模型，从结果可以看出此模型的精确度、召回率和 F1 值均高达 90%以上。

针对问题二热点问题挖掘，我们选择采用 Reddit 社区的排序算法以及 TF-IDF 模型来解决，首先使用 Reddit 的排序算法得出排序前五的热点问题，再根据 TF-DIF 模型以及 jieba 和 gensim 模块，将与排名前五的热点问题有关的话题皆罗列出来，形成了热点问题留言明细表。

针对问题三答复意见的评价，本文构建了多个评价方法来评测答复意见的相关性，时效性，可理解性等。使用了 DNN，线性回归等算法来实现评价。从多个指标的评测，让评测的结果更加完整。

目录

一、问题重述

（一）问题背景

近年来，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，对提升政府的管理水平和施政效率具有极大的推动作用。

（二）要解决的问题

1.群众留言分类

在处理网络问政平台的群众留言时，工作人员首先按照一定的划分体系（参考附件 1 提 供的内容分类三级标签体系）对留言进行分类，以便后续将群众留言分派至相应的职能部门 处理。目前，大部分电子政务系统还是依靠人工根据经验处理，存在工作量大、效率低，且差错率高等问题。

2.热点问题挖掘

某一时段内群众集中反映的某一问题可称为热点问题，如“XXX 小区多位业主多次反映 入夏以来小区楼下烧烤店深夜经营导致噪音和油烟扰民”。及时发现热点问题，有助于相关 部门进行有针对性地处理，提升服务效率。请根据附件 3 将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，定义合理

的热度评价指标，并给出评价结果，按表 1 的格式给出 排名前 5 的热点问题，并保存为文件“热点问题表.xls”。按表 2 的格式给出相应热点问题 对应的留言信息，并保存为“热点问题留言明细表.xls”。

表 1-热点问题表

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	…	2019/08/18 至 2019/09/04	A 市 A5 区魅力之城小区	小区临街餐饮店油烟噪音扰民
2	2	…	2017/06/08 至 2019/11/22	A 市经济学院学生	学校强制学生去定点企业实习
…	…	…	…	…	…

表 2-热点问题留言明细表

问题 ID	留言编号	留言用户	留言主题	留言时间	留言详情	点赞数	反对数
1	360104	A012417	A 市魅力之城商铺无排烟管道，小区内到处油烟味	2019/08/18 14:44:00	A 市魅力之城小区自打交房入住后，底层商铺无排烟管道，经营餐馆导致大量油烟排入小区内，每天到凌晨还在营业……	0	0
1	360105	A120356	A5 区魅力之城小区一楼被搞成商业门面，噪音扰民严重	2019/08/26 08:33:03	我们是魅力之城小区居民，小区朝北大门两侧的楼栋下面一楼，本来应是架空层，现搞成商业门面，噪声严重扰民，有很大的油烟味往楼上窜，没办法居住……	1	0
1	360106	A235367	A 市魅力之城小区底层商铺营业到凌晨，各种噪音好痛苦	2019/08/26 01:50:38	2019 年 5 月起，小区楼下商铺越发嚣张，不仅营业到凌晨不休息，各种烧烤、喝酒的噪音严重影响了小区居民休息……	0	0

3.答复意见的评价

针对附件 4 相关部门对留言的答复意见，从答复的相关性、完整性、可解释性等角度对 答复意见的质量给出一套评价方案，并尝试实现。

二、群众留言分类

在处理网络问政平台的群众留言时，工作人员首先按照一定的划分体系对留言进行分类，以便后续将群众留言分派至相应的职能部门 处理。目前，大部分电子政务系统还是依靠人工根据经验处理，存在工作量大、效率低，且差错率高等问题。根据比赛给出的数据，建立关于留言内容的一级标签分类模型。

（一）数据的选取与预处理

本节主要详细介绍了本小节数据来源以及对原始数据进行预处理的步骤。

1.数据的选取

本节选取附件 2 给出的数据作为数据来源，根据附件 2 分类将留言分为七大类：城乡建设、劳动和社会保障、教育文体、商贸旅游、环境保护、卫生计生、交通运输。数据质量较高，对其各类标签下的数量进行了一个统计如图。

	一级标签	count
0	城乡建设	2009
1	劳动和社会保障	1969
2	教育文体	1589
3	商贸旅游	1215
4	环境保护	938
5	卫生计生	877
6	交通运输	613

根据比赛题意本节我们截取了附件 2 中的留言详情和一级标签作为本节的使用数据（如图）。

	留言详情	一级标签
0	\n\t\t\t\t\t\n\t\t\t\t\tA3区大道西行便道，未管所路口至加油站路段，...	城乡建设
1	\n\t\t\t\t\t\n\t\t\t\t\t位于书院路主干道的在水一方大厦一楼至四楼人为...	城乡建设
2	\n\t\t\t\t\t\n\t\t\t\t\t尊敬的领导：A1区苑小区位于A1区火炬路，小...	城乡建设
3	\n\t\t\t\t\t\n\t\t\t\t\tA1区A2区华庭小区高层为二次供水，楼顶水箱...	城乡建设
4	\n\t\t\t\t\t\n\t\t\t\t\tA1区A2区华庭小区高层为二次供水，楼顶水箱...	城乡建设
5	\n\t\t\t\t\t\n\t\t\t\t\t我在2015年购买了盛世耀凯小区17栋3楼，...	城乡建设
6	\n\t\t\t\t\t\n\t\t\t\t\t由于西地省地区常年阴冷潮湿的气候，加之近年气...	城乡建设
7	\n\t\t\t\t\t\n\t\t\t\t\t尊敬的胡书记：您好！家住A市A3区桐梓坡西路...	城乡建设
8	\n\t\t\t\t\t\n\t\t\t\t\t我们是梅家田社区辖区内的小区居民，我们每年都...	城乡建设

2.数据的预处理

2.1 数据清洗

利用 pandas 库中一个十分便利的 isnull()函数来判断选取数据中是否有空值，清洗结果为选取数据无空值。

```
一级标签中总共有 0 个空值.
留言详情中总共有 0 个空值.
```

2.2 一级标签数字化

使用 pandas 库中的 factorize 函数，将 Series 中的标称型数据映射称为一组数字，相同的标称型映射为相同的数字。以此达到将一级标签用数字代替的目的，结果如图。

	留言详情	一级标签	一级标签_id
2573	\n\t\t\t\t\t\n\t\t\t\t\t 青龙预制场建于M4市禾青镇杨坪村境内，因...	环境保护	1
2492	\n\t\t\t\t\t\n\t\t\t\t\t 马坪开发区下一百五十米喷漆厂毒气每天排...	环境保护	1
5057	\n \n 最近返回昔日工作过的学校墨山学校，往事历历在目。都说十多年过去了...	教育文体	3
7531	\n\t\t\t\t\t\n\t\t\t\t\t\t尊敬的领导您好！我是上海花园13栋公寓业主，...	商贸旅游	5
9151	\n \n I6市合作医疗7月份开通以来各乡镇大药房租借部分卫生室帐号疯狂刷...	卫生计生	6
1805	\n\t\t\t\t\t\n\t\t\t\t\t 泓泰国际公寓与康南加油站只有一墙之隔，...	城乡建设	0
6623	\n \n 职工退休金的发放以前分几个时间段分批发放。有的十五号，有的二十五...	劳动和社会保障	4
1277	\n\t\t\t\t\t\n\t\t\t\t\tL5县凯旋城二栋全体业主投诉！本栋楼于201...	城乡建设	0
4408	\n \n J6县一中马上就要高考报名了，J6县一中却不允许一部分后进生...	教育文体	3
7056	\n \n 我是E7县花桥村的一个普通老百姓，也年年按时交农村合作医疗保险，...	劳动和社会保障	4

2.3 留言语句清洗

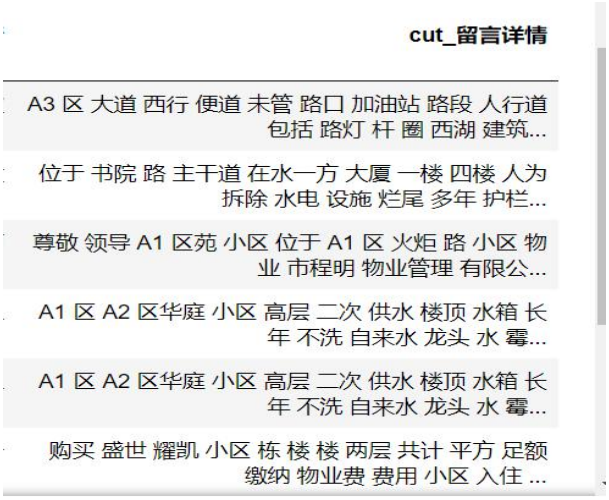
利用正则匹配表达式用户自定义一个用于删除每则留言除字母,数字，汉字以外的所有字符的函数，然后对数据集中的留言详情使用此函数达到清洗数据的目的。结果如图。可以清楚的看出达到了预定的效果。

clean_留言详情
A3区大道西行便道未管所路口至加油站路段人行道包括路灯杆被圈西湖建筑集团燕子山安置房项目施工...
位于书院路主干道的在水一方大厦一楼至四楼人为拆除水电等设施后烂尾多年用护栏围着不但占用人行道...
尊敬的领导A1区苑小区位于A1区火炬路小区物业A市程明物业管理有限公司未经小区业主同意利用业...
A1区A2区华庭小区高层为二次供水楼顶水箱长年不洗现在自来水龙头的水严重霉味大家都知道水是我...
A1区A2区华庭小区高层为二次供水楼顶水箱长年不洗现在自来水龙头的水严重霉味大家都知道水是我...
我在2015年购买了盛世耀凯小区17栋3楼4楼两层共计2千平方一直以来我们按时足额缴纳物业费...

2.4 分词并过滤停用词

使用 jieba 分词对我们清洗后的数据进行分词处理，并使用一个简单的判断

语句判断分词是否处于我们的停用词表中来剔除一些意义不大的停用词，从而提高分词的质量，更有利于后续操作的进行。结果展示如下。



cut_留言详情
A3 区 大道 西行 便道 未管 路口 加油站 路段 人行道 包括 路灯 杆 圈 西湖 建筑...
位于 书院 路 主干道 在水一方 大厦 一楼 四楼 人为 拆除 水电 设施 烂尾 多年 护栏...
尊敬 领导 A1 区苑 小区 位于 A1 区 火炬 路 小区 物业 市程明 物业管理 有限公...
A1 区 A2 区华庭 小区 高层 二次 供水 楼顶 水箱 长年 不洗 自来水 龙头 水 霉...
A1 区 A2 区华庭 小区 高层 二次 供水 楼顶 水箱 长年 不洗 自来水 龙头 水 霉...
购买 盛世 耀凯 小区 栋 楼 楼 两层 共计 平方 足额 缴纳 物业费 费用 小区 入住 ...

2.5 分词的稀疏矩阵

从 sklearn 库中导入 TfidfVectorizer，使用 fit_transform()函数先拟合数据，然后转化它将其转化为标准形式，从而将文档转换为文档 - 词矩阵。返回稀疏矩阵，以便后面训练模型时使用。结果展示如图。

```
(0, 399496)    0.15622076750032005
(0, 396738)    0.15622076750032005
(0, 399553)    0.15622076750032005
(0, 606400)    0.15622076750032005
(0, 264356)    0.14104922086949562
(0, 626866)    0.15622076750032005
(0, 89636)     0.14950725116664162
```

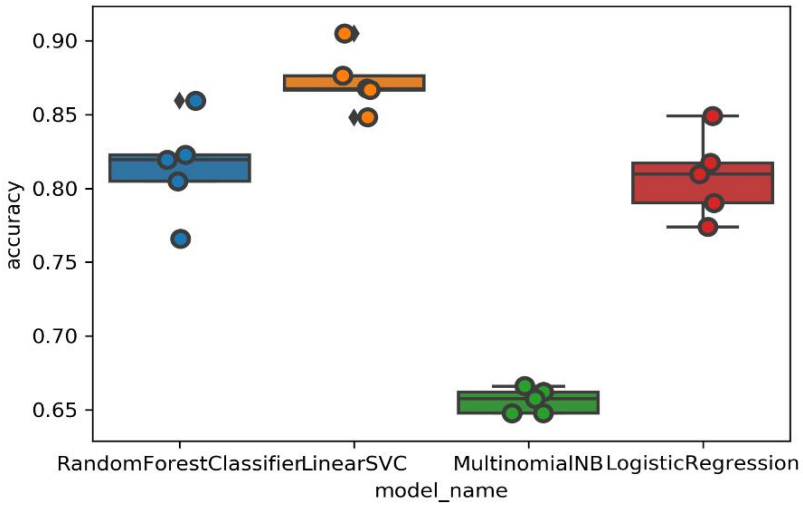
（二）训练模型

1.模型选择

首先从 sklearn.model_selection 中调用 train_test_split 函数将原始数据按照比例分割为“测试集”和“训练集”，首先预选四个模型：逻辑回归（LogisticRegression）、随机森林（RandomForestClassifier）、朴素贝叶斯（MultinomialNB）、线性分类（LinearSVC），然后使用 sklearn 的 cross_val_score 进行交叉验证来评估同一模型不同参数时的能力。最终使每个模型参数设置最优。过程展示如图。

	model_name	fold_idx	accuracy
0	RandomForestClassifier	0	0.765727
1	RandomForestClassifier	1	0.822668
2	RandomForestClassifier	2	0.819316
3	RandomForestClassifier	3	0.859316
4	RandomForestClassifier	4	0.804679
5	LinearSVC	0	0.848156
6	LinearSVC	1	0.876356
7	LinearSVC	2	0.867607
8	LinearSVC	3	0.904943
9	LinearSVC	4	0.866703
10	MultinomialNB	0	0.647505
11	MultinomialNB	1	0.665944
12	MultinomialNB	2	0.661964
13	MultinomialNB	3	0.657251
14	MultinomialNB	4	0.647443
15	LogisticRegression	0	0.773861

再使用 Seaborn 作图直观展示各个模型的预测准确度情况从而帮助我们选出最优模型。结果如图。

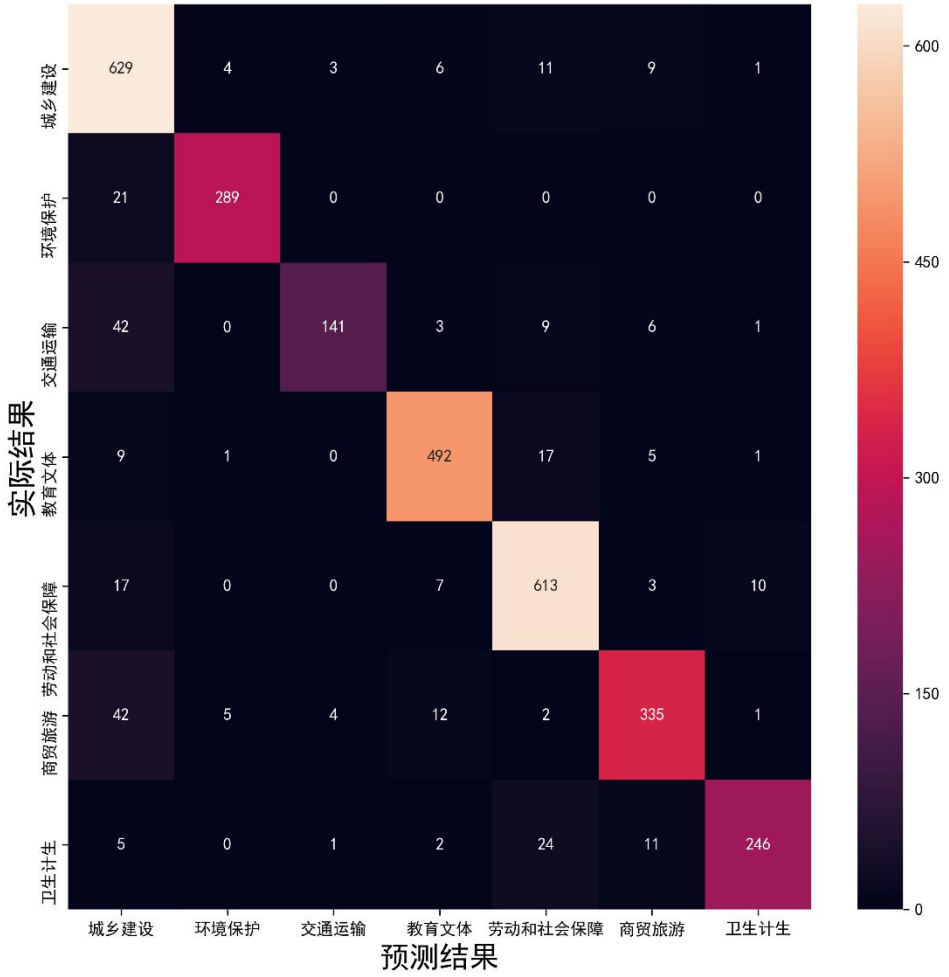


2.模型训练结果展示

我们已经通过上述步骤选出了最佳模型：线性分类（LinearSVC）模型，接下来我们便使用此模型对数据进行训练拟合，利用 train_test_split 函数划分训练集和测试集，先使用 fit()函数求得训练集的一些固有属性。再使用 predict 函数对测试集进行预测，再将实际结果和预测结果输出混淆矩阵。结果如图。

```
array([[629,   4,   3,   6,  11,   9,   1],
       [ 21, 289,   0,   0,   0,   0,   0],
       [ 42,   0, 141,   3,   9,   6,   1],
       [  9,   1,   0, 492,  17,   5,   1],
       [ 17,   0,   0,   7, 613,   3,  10],
       [ 42,   5,   4,  12,   2, 335,   1],
       [  5,   0,   1,   2,  24,  11, 246]], dtype=int64)
```

然后用 seaborn 的热度图绘制出混淆矩阵数据



最后使用 sklearn 中的 classification_report 函数显示分类指标的文本报告．在报告中显示每个类的精确度，召回率，F1 值等信息。即结果如图。

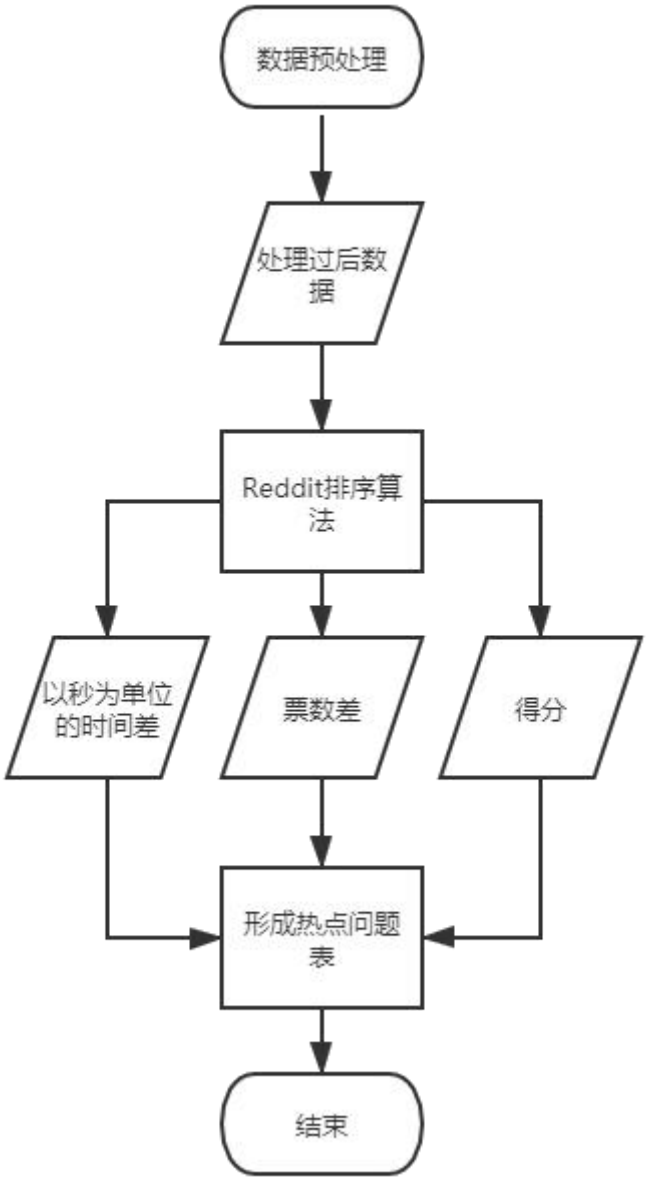
	precision	recall	f1-score	support
城乡建设	0.82	0.95	0.88	663
环境保护	0.97	0.93	0.95	310
交通运输	0.95	0.70	0.80	202
教育文体	0.94	0.94	0.94	525
劳动和社会保障	0.91	0.94	0.92	650
商贸旅游	0.91	0.84	0.87	401
卫生计生	0.95	0.85	0.90	289
avg / total	0.91	0.90	0.90	3040

（三）小结

针对问题一群众留言分类问题，我们首先选取了自己所需的数据，然后对截取数据进行了预处理，以达到我们的使用要求，再建立多个模型对对数据集进行训练测试，选出每个模型的最优参数，再选出最优模型，本文对于此文采用了线性分类模型，从结果可以看出此模型的精确度、召回率和 F1 值均高达 90%以上。

三、热点问题挖掘

在本小问中，我们小组采用了 Reddit 社区排序的算法来获得热度值，从而进行话题的排序，其优点在于一旦话题时间确定，其时间就为固定值。而且话题越新，其 t 值越大，从而使得新话题比旧话题排名靠前，从而提高新话题的热度值。解题思路如下：



（一）模型建立与求解

设时间变量 t,文本时间 A 以及一固定时间 B。（而对于此题，我们经讨论研究，决定选择文本中最早的时间作为 B 的值），将两者做差值，以秒为单位，从而得到变量 t。

$$t = A - B \tag{2-1}$$

并且我们还设差值变量为 x，文件中的点赞数为 ups,反对数 downs，将两者

做差值，从而得到变量 x，如下所示：

$$x = ups - downs \tag{2-2}$$

因为有了投票数，因此我们引入了一个符号变量 y，其含义为投票方向，表示对这个话题的总体看法如何，如果点赞数比反对数多，则 y=+1；若点赞数小于反对数，则 y=-1；此外，若点赞数等于反对数，则 y=0。+1 表示对该话题具有正面评价；-1 表示对该话题具有负面评价；0 表示对该话题不持有倾向性。其式子如下所示：

$$y = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases} \tag{2-3}$$

此外，我们还引入了一个新参数 z 来表示话题的受肯定程度，z 表示点赞数超过反对数的数量，如果点赞数少于或等于反对数，那么 z 就等于 1。其式子如下所示：

$$z = \begin{cases} |x| & |\mathrm{x}| \geq 1 \\ 1 & |\mathrm{x}| < 1 \end{cases} \tag{2-4}$$

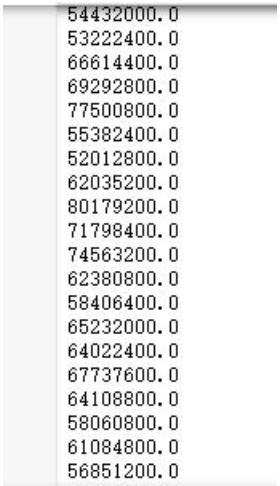
根据上面的参数，我们最终可以引入 Reddit 的最终得分的计算公式如下所示：

$$Score = log_{10} z + \frac{yt}{45000} \tag{2-5}$$

该式子可看做两部分构成，第一部分为 $\log_{10} z$，第二部分为 $\frac{yt}{45000}$，第一部分的 log 函数表示，赞成数比反对数越多，得分就越高，其中如果 z=10 那么可得 1 分；若 z=100，则可以得 2 分。明显这两者之间还差了很多，这是一个缺点，但是若一个话题十分受欢迎，那么后面 z 越大，其影响也就不怎么大了；第二部

分 t 越大，得分也就越高，这也就突出了新鲜话题占比比陈旧话题得分高的目的，它也起到了将旧话题往下拉的目的；而其中 y 的作用是用来表示正负关系的，如上所示，其取值为+1、-1、0，这也就让得到大量净点赞数的话题得到更多的分。

经过计算可得类似与如下所示的以秒为单位的时间：



此外，我们也可得到类似于如下图所示的排名前五的得分：

热度指数
148.5101
140.9382
139.4859
136.5573
131.6035

根据这些得分，我们可以将话题热度前五的话题提取出来，将之提取出来，对其留言详情进行概括，并且从中提取出相应的地点和人群。

（二）热度问题留言

由一中的模型可得出类似于题干所给的表一类型形式的文件，我们根据这个排名，就需要选出与之相关的话题，因此我们选择了 TF-IDF 模型来解决这个问题。

1.模型建立与求解

首先，我们先引入一个参数 D 用于存放我们需要处理的文本集，由此，可引发出三个不同的参数$n_{w,d}$、$\{w_d\}$、n_w，其中$n_{w,d}$表示词 w 在文本中出现的次数；$\{w_d\}$表示文本中所有词的集合；n_w表示包含文本词的所有集合。

从而我们可以引出计算词频的公式如下所示：

$$TF(w,d)=\frac{n_{w,d}}{\sum_{u\in\{w_d\}} n_{u,d}} \tag{2-6}$$

用该公式来描述词 w 在文档 d 中出现的的频率。

因为其为 TF-IDF 模型，我们在上面只给出了其 TF 公式，因此，我们又可以引入 IDF 的公式如下图所示：

$$IDF(w,d)=\log\frac{n}{n_w} \tag{2-7}$$

IDF 这个公式用于计算逆文档的频率，其中包含词 w 的文档数越多，IDF 也就越小。

综合公式 2-6、2-7，我们可得到 TF-IDF 的计算公式：

$$TF-IDF(w,d)=TF(w,d)*IDF(w,d) \tag{2-8}$$

有了 TF-IDF 的模型，我们就要着手于去找到与文本相似的文本，因此就涉及到了，文本相似度的计算。

首先，我们引入两个参数w_{d1}、w_{d2}，(需要比几份文本就引入多少个参数)，再引入参数 w 用于算两个文本的并集，如下所示：

$$w=w_{d1}\cup w_{d2}=\{w_1,w_2,....,w_l\} \tag{2-9}$$

再分别计算 w 中每个词与 d1、d2 之间的相似度，得到新的两个参数 v1 和 v2，然后使用余弦公式，计算 v1 和 v2 之间的余弦距离如下所示：

$$\cos(v1, v2) = \frac{v1 * v2}{|v1| * |v2|} \tag{2-10}$$

根据以上的模型，我们就可以先使用 jieba 模块以及 gensim 模块来进行，并且使用 TF-IDF 来训练模型，根据相似度从而将与热度排名前五的的文本罗列出来，如下图所示：

问题ID	留言编号	留言用户	留言主题	留言时间	留言详情	点赞数	反对数
1	188006	A000102948	A3区一米阳光婚纱艺术摄影是否合法纳税了?	2019/2/28 11:25:05	\n\t\t\t\t\t\n\t\t\t\t\t座落在A市A3区联丰路米兰春天G2栋320，...	0	0
1	189176	A00093880	A3区咸嘉湖西路车辆违停及占道经营乱象为何久治无效	2019/7/24 11:04:12	\n\t\t\t\t\t\n\t\t\t\t\tA3区咸嘉湖西路麓景路——西二环路段，车辆违...	0	0
1	191154	A00044102	举报A市A3区柏家塘小区私营ktv严重扰民	2019/7/13 18:19:08	\n\t\t\t\t\t\n\t\t\t\t\t尊敬的胡书记：您好！作为一名在西地省读完四年...	16	0
1	191584	A00029988	请严查A3区杜容路工程车环卫车乱停扰民问题	2019/7/10 10:44:27	\n\t\t\t\t\t\n\t\t\t\t\tA3区杜容路本是交警划定的禁停路段，但一些大...	1	0
1	191968	A00032980	A3区实验小学东西校区明明属同一个学校，为什么这么多不同?	2019/11/22 10:38:01	\n\t\t\t\t\t\n\t\t\t\t\t您好！我们看好A3区的发展，12年得知A3区...	1	0

四、答复意见的评价

在 C 题，答复意见的评价所给的数据示例如下：

留言编号	留言用户	留言主题	留言时间	留言详情	答复意见	答复时间
744	A089211	建议增加 A 小区快递柜	2019/10/18 14:44	我们是某小区居民...	网民‘A089211’你好...	2019/10/19 8:40

图 4-1 C 题第三问示例数据

在示例中，我们看到数据包含以下几个方面的内容：留言编号，留言用户，留言主题，留言时间，答复意见以及答复时间。为了对答复意见的质量建立起一套评价方案，我们从以下角度对答复意见进行评价。

（一）评价方法

1.时间层面

观察所给的数据我们发现，在留言时间和答复时间之间有一定的时间差。在对时间处理时，我们采用正则表达式对时间做了截取处理，只保留了年月日，删除了时分。处理前后的数据示例如下：

```
处理前的时间如下
0      2019/4/25 9:32:09
1     2019/4/24 16:03:40
2     2019/4/24 15:40:04
3     2019/4/24 15:07:30
4     2019/4/23 17:03:19
Name: 留言时间, dtype: object
```

图 2 处理前的时间格式

处理后的时间如下：

```
0    [2019/4/25]
1    [2019/4/24]
2    [2019/4/24]
3    [2019/4/24]
4    [2019/4/23]
Name: 留言时间, dtype: object
```

图 3 处理后的时间格式

接下来将表格中的时间由 list 格式转变为 datetime 格式，随即做时间差处理。

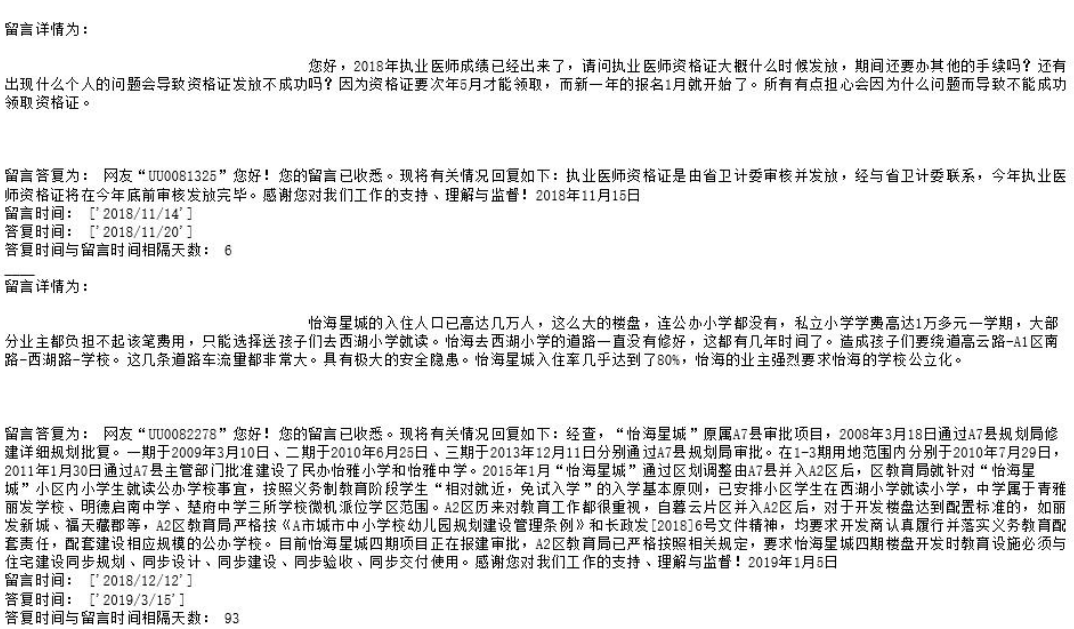


图 4 时间处理结果示例

不同的答复意见，其回复时间效率差距十分大。在上面的两天留言中，第二条在留言后的 93 后才给予回复，而第一条却在不到一个星期了就得到了回复。

2.回答和问题的相似度

按照我们的常理思维，一个好的回答，必定和其问题有一定的关联。我们想到通过计算回答和问题的相似度，来衡量答复意见和问题详情之间的相关性。

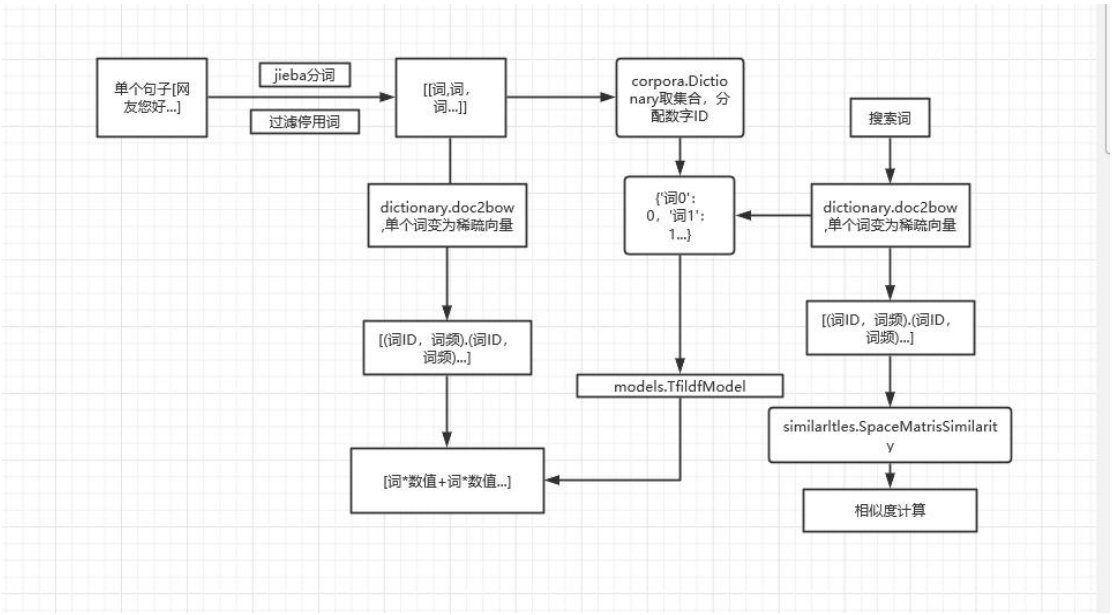


图 5 相似度处理思路

计算文本的相似度分为了以下几个步骤：

1，将文本集生成分词列表；

2，基于文本集建立词典，获得词典特征数；

3，基于词典，将分词列表集转换成稀疏向量集，同时也将搜索词转换为稀疏向量；

4，创建 TF-IDF 模型，传入语料库来训练，TF-IDF 是一种统计方法，TF 是词频(Term Frequency)，IDF 是逆文本频率指数(Inverse Document Frequency)。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。字词的重要性随着它在文件中出现的次数成正比增加，但同时会随着它在语料库中出现的频率成反比下降。；

5，用训练好的 TF-IDF 模型，处理被检索文本和检索词；

6，计算文本相似度。

经过计算相似度计算，其相似度值在 0 到 1 之间，若相似度值越高，则说明，答复和留言越相似，越有相关性。观察结果发现，有的问题答复与问题之间的相

似度值为 0，有的可以达到 0.4424。提取出问题和答案相似度为 0 的回答与问题（以第 21 条数据为例)。其问题为：

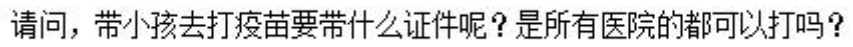


图 6 相似度值为 0 的留言

其答复为：

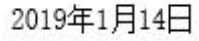


图 7 相似度值为 0 的留言答复

显然这样的回答是十分不合理，我们可以认为这样的回答是无效的。提取出问题和答案相似度为 0.442 的回答与问题（以第 27 条数据为例)。

其问题为：

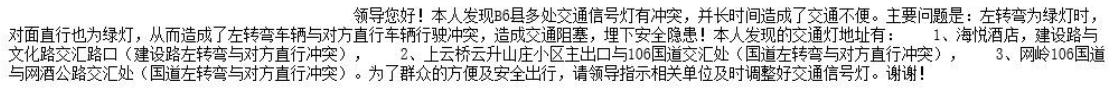


图 8 相似度值为 0.422 的留言

其回复为：

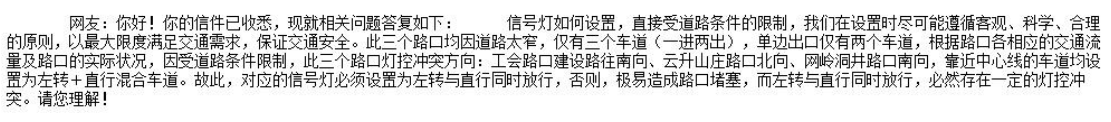


图 9 相似度为 0.422 的留言回复

这样的问题是合情合理的，可以认为这样的回复是能满足留言者的。

3.计算语句通顺性

在中文语法中，对词语有动词名词等词性分类。一个通顺的句子，必定是符合语法的，按照合理的词性词语搭配而组成的。例如“我才饭吃了”和“我才吃

了饭”相比较，后者是符合语法的。我们引入深度神经网络（DNN），判断问题的回答是否符合自然语言表达的习惯。我们将“我才饭吃了”和“我才吃了饭”两个句子作 DNN 打分：

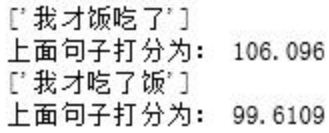


图 10 DNN 示例打分

从结果来看，打分越低，则句子的通顺性越好，越容易理解。留言回复的示例打分结果如下：

网友“A00077538”：您好！针对您反映A3区含浦镇马路卫生很差的问题，A3区学士街道、含浦街道高度重视，现回复如下：您留言中反映的含浦镇在2013年已经析出两个街道，分别是学士街道和含浦街道，鉴于您问题中没有说明卫生较差的具体路段，也没有相应的参照物，同时您也未留下联系方式，请您看到回复后，致电学士街道0731-0000-00000000或者含浦街道0731-0000-00000000反映相关问题。感谢您对我们工作的关心、监督与支持。2019年4月24日
该回复的DNN模型打分为： 1382.58

网友“A000100804”：您好！针对您反映A3区教师村小区盼望早日安装电梯的问题，A3区住建局高度重视，立即组织精干力量调查处理，现回复如下：为了完善住宅使用功能，提高我区既有多层住宅居民的宜居水平，2018年6月7日，A市A3区人民政府办公室下发了《关于A市A3区既有多层住宅增设电梯实施方案》的通知。该方案明确了增设电梯条件及流程。详情可咨询政务服务中心住建局受理申请老旧小区加装电梯窗口，咨询电话：0000-00000000。感谢您对我们工作的关心、监督与支持。2019年4月2日
该回复的DNN模型打分为： 1183.66

网友：您好！现将您反映的情况回复如下：学校高三年级学习任务重、距离高考时间短。根据《西地省教育厅印发〈关于进一步规范普通中小学办学行为的规定〉的通知》（楚教发[2009]36号）第十五条规定“严格规范补课、办班行为。高三年级为适应高考时间提前要求，可以在寒暑假共增加不超过一个月的时间用于完成教学任务，但须征得学生及其家长的同意；但不得收取任何费用”。在上期期末召开了相关学生家长会，充分征求了家长与学生意见后，在自愿的基础上利用暑假弥补教学时间，学校未收取费用。拟定于8月16日—25日组织学生来校，利用10天时间赶教学进度，便于后面有一定的时间用于复习。不存在威胁学生来B7县二中。2018年8月14日
该回复的DNN模型打分为： 1121.39

网民您好：您反映的问题，经过处理，现予以回复：1、B市B1区快乐休闲网吧于1998年开始营业，统一社会信用代码为9430202ma4leqa393，位于B市B1区601厂单身公寓前。该网吧有57台机子，东面有3个换气扇，2台3匹空调。2、请当事人双方以及茨菇塘办事处工作人员到现场开展噪声监测。现场监测昼间噪声为55.3db（a），背景值为54.1db（a），投诉人住在钻石新村20栋201，室内客厅监测值为49.3db（a），符合二类商住混合区排放标准。投诉人拒绝了卧室监测。2018年8月16日
该回复的DNN模型打分为： 3962.48

网民您好：您反映的问题建议您拨打12345市长热线，可以就有关问题进咨询，谢谢！
该回复的DNN模型打分为： 960.5

图 11 DNN 留言详情打分结果示例

在示例数据中，第四条回复的打分最高，说明回复语句逻辑以及通顺性在这五条回复中最低。例如:回复中的“背景值为 54.1db（a)”难以理解；“请当事人双方以及茨菇塘办事处工作人员到现场开展噪声监测”，缺少主语，有语病错误；“三个换气扇”应该是“三台换气扇”...而示例中的第五条回复打分最低，最便于理解，回复语句逻辑以及通顺性在这五条回复中最高。

4.答复意见和留言详情分类

借助第一题“群众留言分类”的思想，我们想到，将所给出的“附件 4”中的答复意见和问题详情进行分类。测试结果如下（数据过多，仅展示部分数据）：

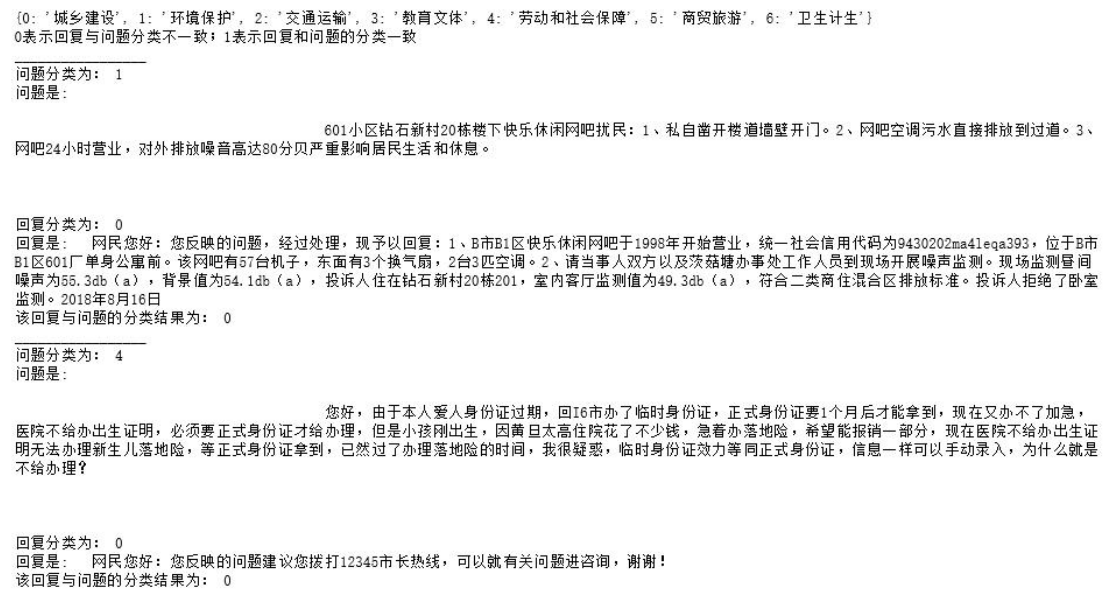


图 12 答复意见和留言分类结果

可以看到，一些留言和其答复是不属于同一类的。上述的第二条回复十分笼统，在回复上未直接了当地做出回答。

（二）评价结果

通过以上四个指标，我们建立起了一套评价体系。将时间及时性划分为四个等级：7 天内回复，定义为回复效率高；8 到 30 天内回复，定义为回复效率一般；31 到 60 天内回复，为回复效率低；61 天以上回复，定义为回复效率极低。从语句通顺性我们划分了三个层面：打分在 1200 以下 ，定义为“回复十分便于理解”；打分在 1200 -4000 ，定义为“回复能理解”；打分在 4000 以上，

定义为“回复基本能理解”。结合答复意见及问题详情的分类和回答与问题的相似度：我们将答复的相关性划分为“回复质量合格”，“回复质量良好”，“回复质量一般”，“回复质量优秀”四个类别。

下面是我们建立起的评价模型的部分评价：

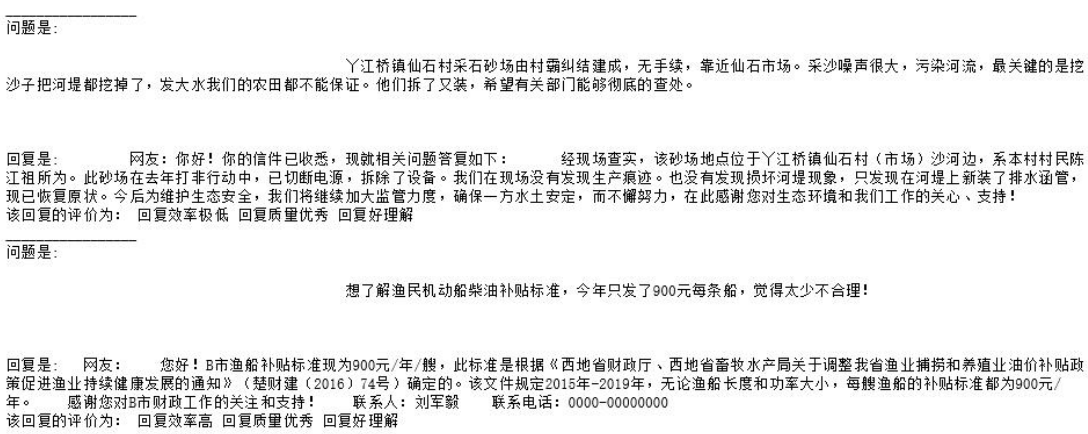


图 13 模型评价结果

（三）小结

我们通过以上四个方法，建立了三个指标来评测留言答复的完整性。从答复的相关性，效率，可解释性（语义通顺性）等角度，对答复做出了评价。通过人为观察答复和留言，我们的评测结果是合理的。