“智慧政务”——基于 BERT 的群众留言分类与基于增量聚类的热点发现

摘要

智慧城市的建设是我国重要的国家战略之一，也是解决城市问题的有效方法。随着智慧城市的逐渐落地、智慧应用的日趋多元化，加强对于智慧城市发展建设情况的监测与理解、提供基于数据的精准化规划指导变得更加重要。在此背景下，本文围绕我国智慧城市建设的热点开展实证研究，帮助政府部门更好的服务社会，服务人民。

针对问题一：首先对原始文本进行数据筛选和文本预处理操作将文本映射到向量空间中，然后构建基于 Bert 的深度学习模型进行训练，并进行参数调整，最终得到一个群众留言分类模型。

针对问题二：通过对数据中问题描述的基本清洗，进一步将问题描述处理成向量形式，并以此为基础进行文本聚类操作，生成若干个聚类中心即是热点问题，与此同时以这些聚类中心为基础，在一个热点问题的一定距离半径内的问题描述可以归类为同一热点问题，并以问题描述的数量作为该热点问题的热度指数。

针对问题三：通过给定的相关部门对留言的答复意见，我们需要对答复意见进行量化分析，通过推理规则为每一条答复意见给定一个分值，该分值通过题干中给出的相关性、完整性和可解释三种评价指标进行评定。首先对答复意见进行大致分类；然后细分评价规则；最后对每一类答复意见选定合适的评分规则进行评分。

另外，为了完成上述实验，本文所有方法均在 Python3.7 软件环境下编程实现。

关键词： 自注意力机制，深度学习， 文本聚类，词向量，量化分析

Abstract

The construction of smart cities is one of China's important national strategies and an effective way to solve urban problems. With the gradual implementation of smart cities and the increasing diversification of smart applications, it has become more important to strengthen the monitoring and understanding of the development and construction of smart cities and provide accurate data-based planning guidance. In this context, this article conducts empirical research around the hotspots of the construction of smart cities in China to help government departments better serve society and serve the people.

For problem one: First, data filtering and text preprocessing operations are performed on the original text to map the text into a vector space, and then a deep learning model based on Bert is constructed for training and parameter adjustment, and finally a mass message classification model is obtained.

For problem two: through basic cleaning of the problem description in the data, the problem description is further processed into a vector form, and text clustering operations are performed on this basis, and generating several clustering centers is a hot issue. At the same time, these Based on the clustering center, problem descriptions within a certain distance radius of a hotspot problem can be classified as the same hotspot problem, and the number of problem descriptions is used as the hot index of the hotspot problem.

For problem three: through the response opinions of the given relevant departments to the message, we need to quantify the response opinions, and give each reply opinion a score through inference rules, and the score is passed through the relevant Three evaluation indicators of sex, completeness and interpretability are used for evaluation. First classify the reply opinions roughly; then subdivide the evaluation rules; and finally select the appropriate scoring rules for each type of reply opinions to score.

In addition, in order to complete the above experiments, all the methods in this article are programmed in the Python3.7 software environment.

Keywords： Self-attention mechanism, deep learning, text clustering, word vectors, quantitative analysis

目录

1 挖掘背景与目标

1.1 挖掘背景

随着科技的发展和社会的进步，智慧政务理念逐渐开始进入人们的视线。利用先进的人工智能技术提高政府部门在办公、监管、服务和决策等多方面的智能化水平，就是政务智能化服务的典型应用，其目标之一就是希望通过一个统一的对外服务窗口，面向群众和企业提供各类政务服务事项的网上办理和预约服务，提高政府服务的办事效率和用户体验，使人民群众少跑腿，逐步做到一网受理、只跑一次、一次办成。

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。

人工智能在智慧政务中将得到充分应用。当前，“让信息多跑路，让群众少跑腿”，已经成为全国的一个口号。在 5G 到来后，未来政务信息化将会出现新的目标，即“智能化管理，智慧化服务”，这意味着人工智能等新技术会在智慧政务中得到充分的应用。从智能化管理来看，人工智能将在交通、环保、市场监管、公共安全等领域获得新的突破。另外，市场监管精准化的要求，有可能促进人工智能在市场监管乃至社会管理的某些特定领域发挥更大的作用。当然，在推广人工智能应用时，不仅要考虑成本，还要考虑个人隐私保护、公民个人权利尊重等，因此，要防止不惜一切成本的新技术的滥用所产生的不良社会后果。

因此，如何从海量的、复杂的数据中利用自然语言处理技术，解决群众留言分类、热点问题挖掘等问题，是提高政府部门工作效率的一种关键技术。为此，利用数据挖掘技术研究“智慧政务”中的留言答复问题具有重要的意义。

1.2 挖掘目标

(1) 为处理网络问政平台群众留言过多的问题，首先要按照一定的划分体系为留言内容进行分类，目的是便于将群众留言分派至相关部门进行处理。然而，目前为止大部分电子政务系统留言区仍然依靠人工经验进行分类处理，存在工作量大、效率底下的问题，我们需要根据留言内容建立标

签分类模型。

(2) 热点问题可被解释为某一时段内群众集中反映的问题。及时发现热点问题有助于相关部门及时处理，提高服务效率。为达到上述目的，我们根据某一时段内反映在特定地点或特定人群之间的留言进行分类处理，并给出合理的热度评价指标及评价结果。

(3) 政府部门对留言的答复意见体现了以人民为中心的发展理念，相关部门的留言回复时刻体现着政府和国家队人民群众的关心情况，但何种答复意见才是有意义、有价值的呢？我们将从答复的相关性、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。

2 问题分析

2.1 问题一的分析

群众留言分类属于自然语言处理中的一个重要领域，即文本分类。通过文本内容和人工标注的标签来训练分类模型。我们首先对原始内容进行数据筛选和文本预处理操作并通过词嵌入技术将文本映射到向量空间，然后使用深度学习网络进行训练，最终得到一个文本分类模型。

本题的主要难点在于深度学习模型的构建和超参数的调整。因此需要设计实验来评估模型的效果并对比不同参数对分类精度的影响。

2.2 问题二的分析

通过对问题的理解，问题的目的在于找出热点的问题，这可以看作是一个文本数据的聚类的问题，所有数据没有给出固定的标签，而是需要程序自动的将某段时间内群众反映的问题汇总到一起，不仅如此，程序并不能只停留在将问题汇总到一起，更重要的一点是要区分出哪些是热点的问题，是群众更关注的问题。根据附件 3 给出的数据的格式，我们可以将“问题描述”部分单独抽取出来，采用递增聚类的策略，首先读取第一个问题描述，并生成第一个热点问题，然后读取第二个问题描述，并与热点问题计算相关性，我们可以找出与问题描述最相关的热点问题且这个相关性大于一定的阈值，我们就可以认为这个问题描述属于该热点问题，如果一个问题描述不满足上述的要求，则以这个问题描述为种子，再次生成一个新的热点问题。

整个处理问题的流程会涉及到问题描述向量化的问题、问题描述与热点问题

相关性的计算以及相关性阈值的选取和热点问题数目的限定等等。不仅上文提到的这些，为了抓住热点问题，我们可以用聚类结果中距离聚类中心点一定范围内相关的问题数量作为评价当前热点问题中心的重要标准，以此来对多个热点问题来进行排序。

2.3 问题三的分析

通过根据附件 4 给定的相关部门对留言的答复意见，我们需要对答复意见进行量化分析，为每一条答复意见给定一个分值，该分值通过题干中给出的相关性、完整性和可解释三种评价指标进行评定。首先对答复意见进行大致分类；然后细分评价规则；最后对每一类答复意见选定合适的评分规则进行评分。我们采用“满足则加分，不满足不加不扣分”的方式量化评估，分数越高评定质量越好。

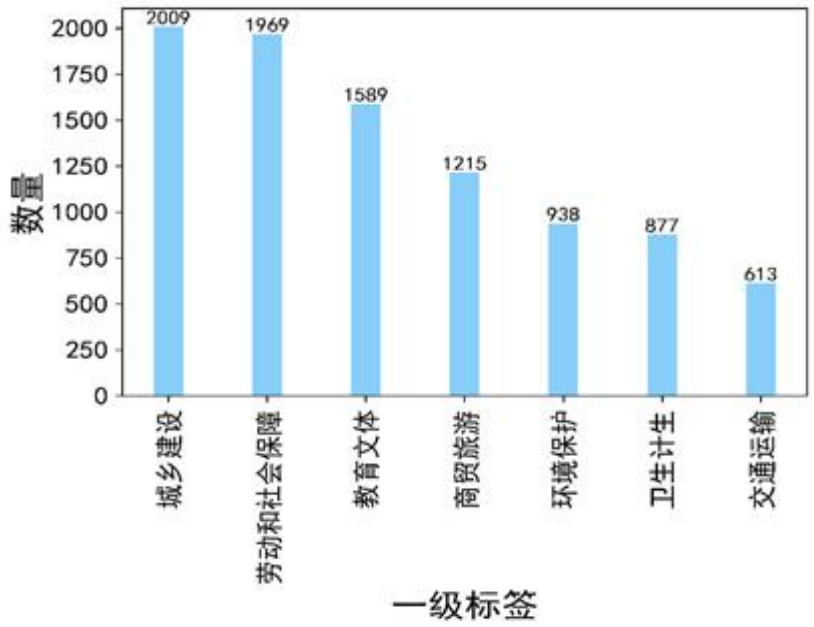
3 分析方法与过程

3.1 问题一分析方法与过程

本章针对群众留言分类的问题，首先在原始数据中选择相关文本，然后对文本进行预处理，最后使用基于深度学习的转移双向编码表示模型（BERT）训练文本分类器。

3.1.1 数据的选取

针对群众留言分类问题，本文在原始数据中选择了留言主题和留言详情作为我们的输入数据，这主要考虑到部分群众没有在留言主题中清楚的表述出自己想要解决的问题，且留言主题的文本长度太短，无法蕴含足够的语义信息。因此，我们将留言主题和留言详情合并形成新的文本作为我们模型的输入数据。



如图所示，原始数据中共包含 7 种一级标签分类，它们的占比不均衡，其中城乡建设样本共有 2009 条，而交通运输只有 613 条。这种类别分布不均衡会对后续的模型训练造成一定的影响。因此，我们对数量较少的类别实行过采样，从而使得不同类别的占比均衡，提高分类准确度。

3.1.2 数据预处理

本节详尽介绍了我们对文本内容的预处理操作和对标签的预处理。

3.1.2.1 文本内容预处理

(1) 分词。不同于英文句子，中文句子需要进行分词操作，即将一句话分离成单独的词语。分词的效果会直接影响模型的性能。因此，我们使用了业界公认效果较好的 jieba 分词工具。

(2) 去停用词。在信息检索中，为节省存储空间和提高搜索效率，在处理自然语言数据（或文本）之前或之后会自动过滤掉某些字或词，比如“的”、“是”、“而且”、“但是”、”非常“等。这些字或词即被称为 Stop Words（停用词）。这些词语并没有携带明显的语义特征，因此，我们在分词的同时，将包含在停用词表中的单词去掉。

(3)词嵌入。由于单词并不能直接作为模型的输入，所以我们需要将单词转换成具有语义信息的向量空间中。由于我们采用的是 BERT 模型，BERT 模型已经在预训练的过程中学习到了良好的单词表征，因此我们并没有自行训练词向量，

而是在 BERT 训练的词向量的基础上进行微调，这样不仅能够节约大量的训练时间，还能大大提升模型的分类性能。

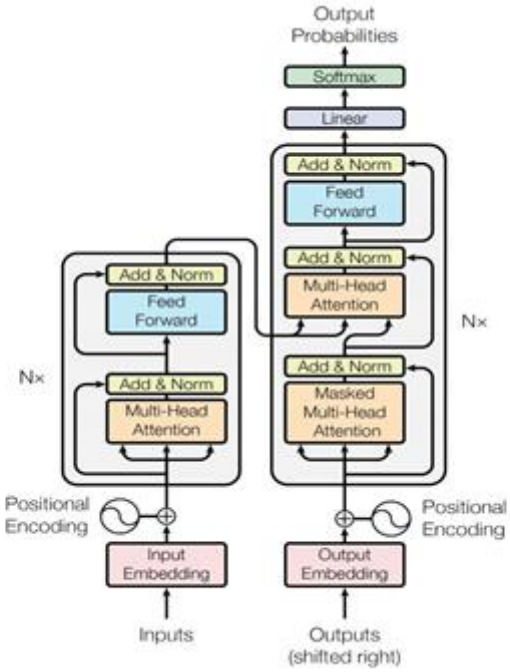
3.1.2.2 标签的预处理

本任务中共有 7 个标签，因此我们将 7 种标签映射到数字空间中，简单的使用数字 0-6 来作为标签的表示。

3.1.3 模型的构建

文本分类是自然语言处理中应用最广泛的任务之一。随着深度学习的不断发展，各种优秀的神经网络模型被提出来进行自然语言处理任务，循环神经网络（RNN）和卷积神经网络（CNN）是其中的典型代表，但由于循环神经网络无法获得长距离的语义信息,其升级版长短期记忆力模型(LSTM）和 门控循环单元(GRU）又被提出并迅速成为自然语言处理任务的首选。

近几年谷歌提出的 Transformer 又将自然语言处理的效果提上了一个新的台阶。Transformer 模型没有使用任何 RNN 结构，而是直接采用了自注意力机制来学习单词的语义信息并创新性的提出多头注意力机制来学习不同空间中单词的语义信息。其在翻译任务中达到了业界最高水平。



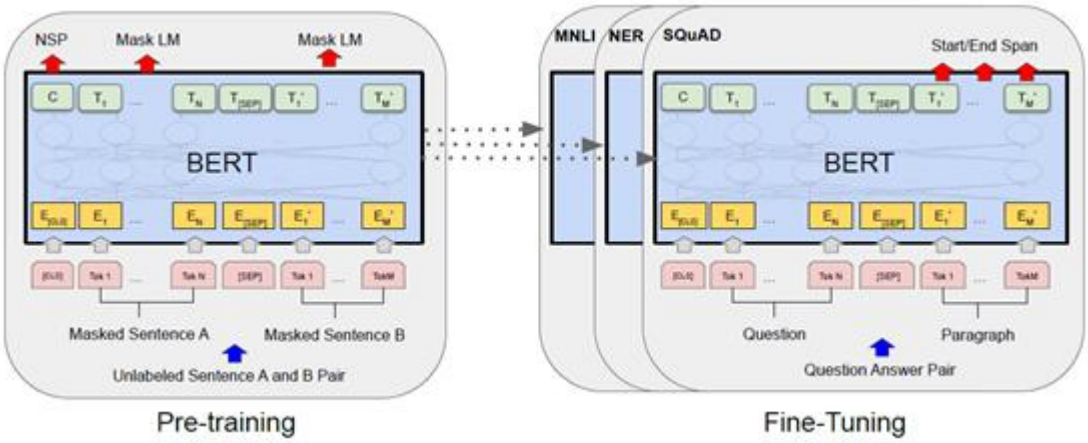
如上图所示，Transformer 共有两个部分，第一部分是编码器，主要负责将学习输入数据的语义信息，第二部分是解码器，负责学习生成句子和输入数据之

间的关系。多头注意力是通过 h 个不同的线性变换对 Q，K，V 进行投影，最后将不同的 attention 结果拼接起来，多头注意力和自注意力机制计算公式为：

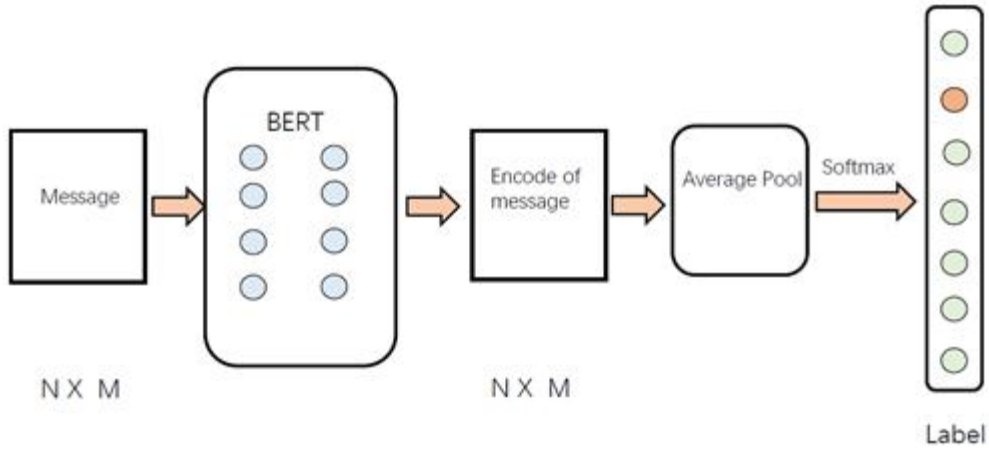
$$\mathrm{MultiHead}(Q,K,V)=\mathrm{Concat}(\mathrm{head}_1,\ldots,\mathrm{head}_h)W^O$$

$$\mathrm{head}_i=\mathrm{Attention}\left(QW_i^Q,KW_i^K,VW_i^V\right)$$

而 BERT (Bidirectional Encoder Representations from Transformers) 模型的基本结构就是多层的 Transformer 编码部分。BERT 是一种强大的语言表示模型，它利用了双向 Transformer 来对语言进行预训练，并能够在下游任务中进行微调从而实现良好的效果。



如图所示，BERT 共有两部，第一步是预训练部分，其使用无标签数据在不同的训练任务中进行训练。第二步是微调，其在预训练的基础上针对不同的下游任务进行参数微调，从而达到良好的效果。



3.1.4 实验评估

本节通过实验对所提出的模型进行验证。通过与其他传统的循环神经网络和

卷积神经网络模型进行对比，证明我们模型的有效性。同时，我们还进行了模型的调参优化并对部分重要参数进行了实验分析。

3.1.4.1 实验环境

硬件环境：

CPU	Intel(R) Core(TM) i7-8750H
GPU	GTX1080Ti
内存	16G

软件环境：

操作系统	Windows10 专业版
Python	3.6.1
CUDA	9.1
Pytorch	1.4

3.1.4.2 评估指标

本题采用 Macro-f1 来作为分类精度指标。计算公式如下：

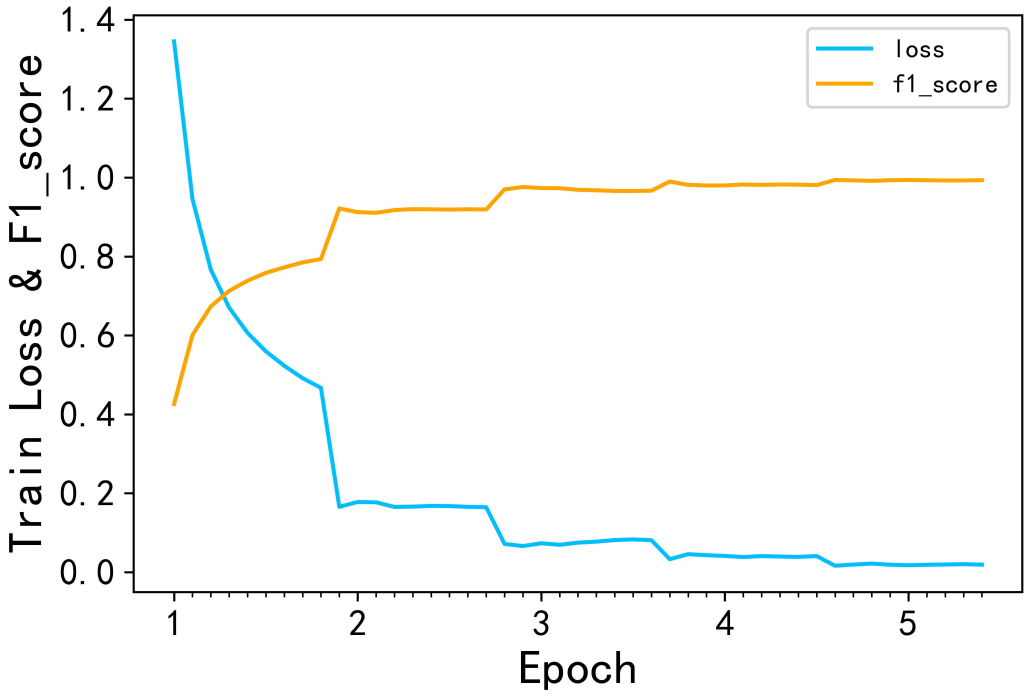
$$F_1 = \frac{1}{n}\sum_{i=1}^{n}\frac{2P_iR_i}{P_i + R_i}$$

3.1.3.1 实验参数设置及结果

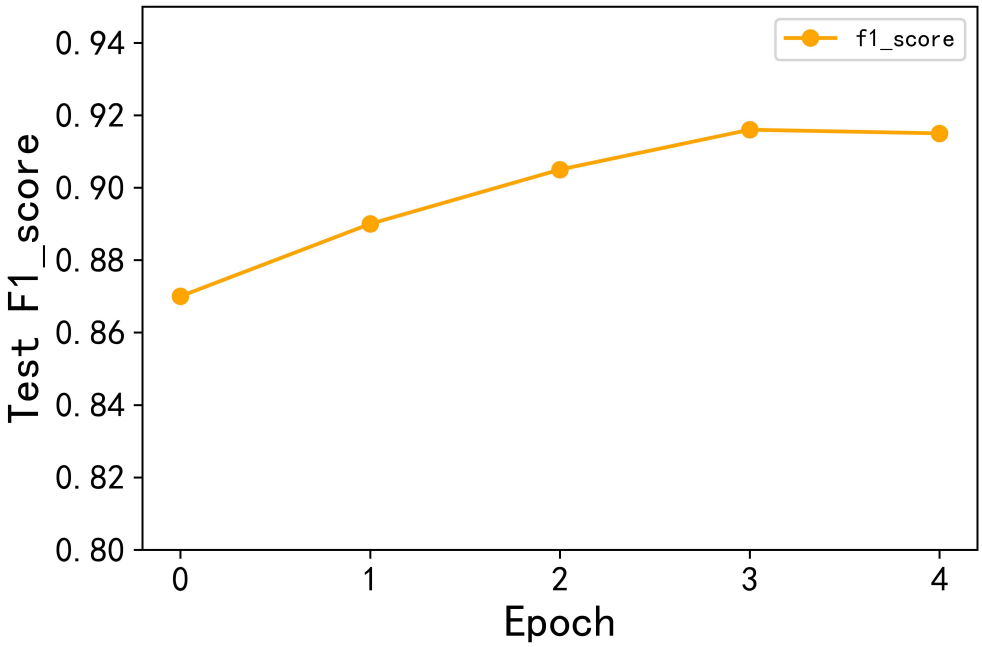
模型的相关参数如下表所示：

句子最大长度	512
单词嵌入维度	768
学习率	2e-5
Batch_size	16
迭代次数	5

在训练过程中，我们将原数据的 80%作为训练集，20%作为测试集。实验结果如下图所示。



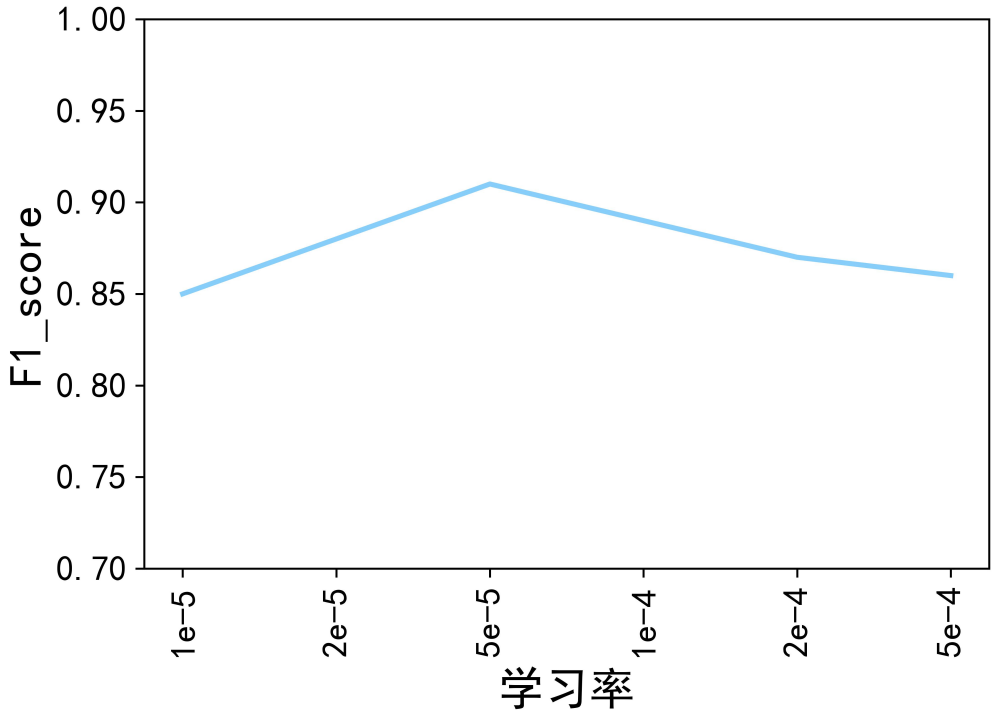
训练集上模型的损失值和 F1_score



测试集上模型的 F1_score

如图所示，在训练集上模型的 F1_score 可以达到 99.3%。测试集上模型的 F1_score 在第四个 epoch 上达到最大值，最大为 91.5%。

3.1.3.2　不同实验参数的效果对比分析



为了选择最优的学习率，我们设置了一组学习率，并在测试集上进行了对比验证。如图所示，当学习率为 5e-5 时，模型的效果最好，可以达到 91%的 F1_score。因此我们最终确定将学习率设置为 5e-5。

3.2 问题二分析方法与过程

3.2.1 分词与去除停用词

中文分词是中文自然语言处理的第一步，一个优秀的分词系统取决于足够的语料和完善的模型，很多机构和公司也都会开发和维护自己的分词系统。中文分词模型主要是从两个方面来做工作的，一方面是基于规则的分词方法，基于规则是指根据一个已有的词典，采用前向最大匹配、后向最大匹配、双向最大匹配等人工设定的规则来进行分词。当然，分词所使用的规则可以设计得更复杂，从而使分词效果更理想。但是由于中文博大精深、语法千变万化，很难设计足够全面而通用的规则，并且具体的上下文语境、词语之间的搭配组合也都会影响到最终的分词结果，这些挑战都使得基于规则的分词模型愈发力不从心。另一方面则是基于统计的分词方法。基于统计是从大量人工标注语料中总结词的概率分布以及

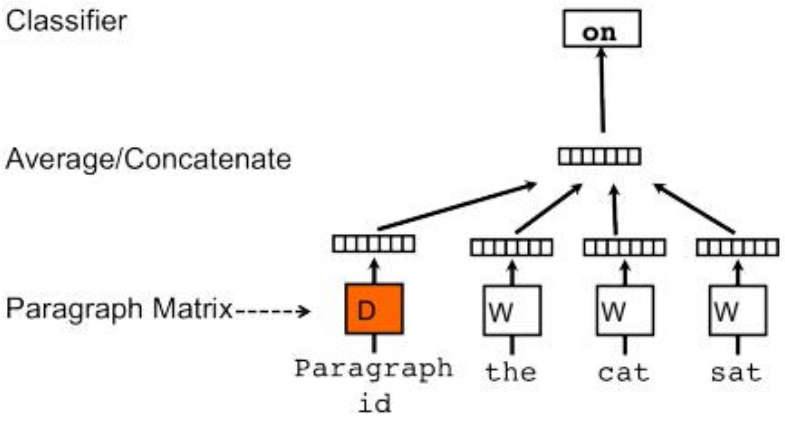
词之间的常用搭配，使用有监督学习训练分词模型。这里，我们使用的是 Python 的库文件——jieba 分词来对问题描述进行处理。

此外，为了减少一些常用的、无意义的词对最终结果的影响，所以我们添加了停用词表，来去除这些停用词，最终每个问题描述可以通过数个词语来作为整个问题描述的概括表示。

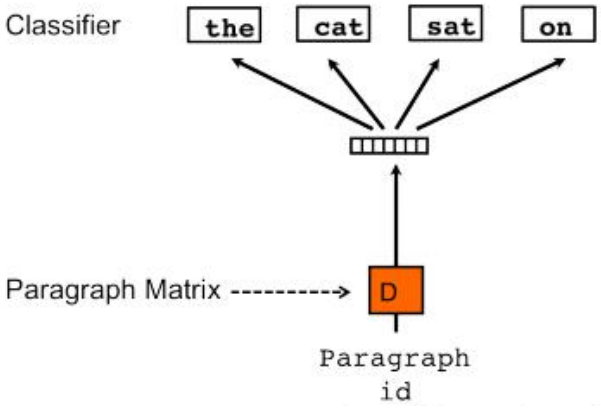
3.2.2 生成词向量和文档向量

Word2Vec[5]是 Google 在 2013 年开源的一款词向量计算工具，它的特点是将所有的词向量化，这样词与词之间就可以定量的去度量他们之间的关系，挖掘词之间的联系。Word2Vec 表示的词向量不仅考虑了词之间的语义信息，还压缩了维度。但是，有时候当我们需要得到 Sentence/Document 的向量表示，虽然可以直接将 Sentence/Document 中所有词的向量取均值作为 Sentence/Document 的向量表示，但是这样会忽略单词之间的排列顺序对句子或文本信息的影响。Doc2vec 是在 Word2vec 的基础上做出的改进，它不仅考虑了词和词之间的语义，也考虑了词序。

Doc2Vec[6]有两种模型，分别为：句向量的分布记忆模型（PV-DM: Distributed Memory Model of Paragraph Vectors）和句向量的分布词袋（PV-DBOW: Distributed Bag of Words version of Paragraph Vector）。DM 模型在给定上下文和文档向量的情况下预测单词的概率。即在训练时，首先将每个文档 ID 和语料库中的所有词初始化一个 K 维的向量，然后将文档向量和上下文词的向量输入模型，隐层将这些向量累加（或取均值、或直接拼接起来）得到中间向量，作为输出层 softmax 的输入。在一个文档的训练过程中，文档 ID 保持不变，共享着同一个文档向量，相当于在预测单词的概率时，都利用了这个句子的语义。



DBOW 模型在给定文档向量的情况下预测文档中一组随机抽样的单词的概率。



事实上，Doc2vec 的 DM 模型跟 Word2vec 的 CBOW 很像，DBOW 模型跟 Word2vec 的 Skip-gram 很像。Doc2Vec 为不同长度的段落训练出同一长度的向量；不同段落的词向量不共享；训练集训练出来的词向量意思一致，可以共享。

3.2.3 计算问题描述与热点问题相关性

首先，我们可以抽取出问题描述中相对重要的关键词语，我们可以使用前面生成的词语的词向量作为基础，来合并为热点问题的嵌入表示，实际的语言表达过程中，并不是所有词语在表达中占有同样的重要程度，所以，通过使用自适应的权重来协调多个词语，进而最终生成热点问题的向量形式的表示：

$$sen\_vec = \frac{\sum_{i=1}^{m} vec_i * e^{weight(i)}}{m}$$

其中，sen_vec 表示热点问题的向量表示形式，m 表示每个样本中的词的个数，ver_i 表示每个词的词向量，weight(i)则表示每个词的权重，权重可以根据 TF-IDF 或者是信息增益的方法求得。这里我们采用统计的整个数据的 TF-IDF 作为单个词语的权重。

3.2.4 计算热点问题热度

关于热点问题的热度指标，我们使用聚类完成后得到的各个热点问题的热度中心[7]，去计算热点问题的向量表示形式和问题描述的向量表示形式的余弦相似度，余弦相似度的定义公式表达如下：

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i-1}^{n} (A_i)^2} \times \sqrt{\sum_{i=1}^{n} (B_i)^2}}.$$

通过余弦相似度的计算，可以计算出各个问题到热点问题中心的距离，因此可以判断一个问题描述属于哪个热点问题，从而对属于热点问题的问题描述进行数量统计，进一步可以以统计结果作为热点问题热度的标识。

3.3 问题三分析方法与过程

3.3.1 数据分析与属性选取

根据附件 4 中的信息，我们首先从中选取合适的属性构建评价方案，在“留言主题”、“留言详情”、“答复意见”和“留言/答复时间”四个属性中进行筛选，并选取“留言详情”和“答复意见”作为评价答复意见的重要因素，筛选原因如下：

- **不将“留言主题”作为评价值质量标准的原因：**留言主题为留言者问题的高度概括，但不够具体。比如留言者有多条问题一次提出，答复意见需逐条予以回复，此时的“留言主题”将无法起到评价答复意见质量的作用。
- **不将“留言/答复时间”作为评价质量标准的原因：**工作人员需时间去调查了解留言者所提出的问题，不同问题难易程度、时间复杂度不同，所以时间不能作为评价标准。

然而，只确定好需要评定的属性不足以给出一套对答复意见的评价方案。为了明确 “答复意见”中内容的优劣情况，通过查看“答复意见”的内容，我们将其分为如下四类，回复质量从高到低为：

1. **直接回复：**相关部分经调研后对留言者进行相关内容回复，直接对留言者的问题进行解决结果回复，或经过转移至相关部门后给予了回复信息。
2. **继续处理回复：**由于留言者本身描述不够清晰，导致相关部分无法给出回复，故表示需要与留言者进一步进行沟通。
3. **间接回复：**给出国家方案、大体流程等信息后，并给予留言问题更具相关性部分的联系方式，以便让留言者进一步咨询。
4. **转移回复：**根据留言相关性，将问题转移至相关部门进一步处理，回复中无后续处理内容。此类回复并没有为留言者带来实际意义。

但是，只从上述 4 类回复质量进行质量判断仍然太过粗糙，我们还需要对每一个答复内容进行量化分析，即给出具体的评分准则。在本次实验过程中用到的评分指标如下：

- **相关性：**答复意见是否与留言内容密切相关，是否会出现答非所问的情况。
- **完整性：**答复意见是否解决留言者所提出的全部问题，是否存在漏答的情况。
- **可解释性：**答复意见是否解决了留言者提出的问题，答复意见是否具有合理性。

通过以上分析，我们主要完成了三部分内容：(1) 决定采用的评价属性：留言详情”和“答复意见”；(2) 对答复意见内容进行分类：直接回复、继续处理回复、间接回复和转移回复；(3) 确定所采用的评分指标：相关性、完整性和可解释性。

3.3.2 回复质量量化分析

由于不同类别的“答复意见”质量不同，所以它们本身就是一种评估方式，这种情况下不能所有类别的“答复意见”都共享同一套推理规则。我们为此建立了权重体系，使得不同质量下的“答复意见”拥有不同的权重。在本文中我们使得的权重系数如下表所示：

直接回复	继续处理回复	间接回复	转移回复
0.8	0.6	0.5	0

注意：转移回复后，得到了转移后的答复意见，则该“答复意见”自动转为直接回复类；而转移后并没有得到任何相关部分的解答，此时该回复内容是没有任何意义的(如附件四第 2813 行)，因此该类“答复意见”权重为 0.

基于上面的这些提示信息，我们给出一些推理规则用以判断该回复内容是否应该得分。以 5 分满分制为基础，如果该回复内容在某规则下得分，则将该内容的全部得分线性相加，并乘以该回复内容所属类别对应的权重，最终该“答复意见”评定得分公式为：

$$score = 2\sum_{i=1}^{n} w_t \, I(T \in R_i)$$

其中 n 为质量评价规则的数量，wt 为答复意见所代表的归类类别权重，T 代表该答复意见的文本内容，Ri 为当前所对比的得分判断规则，I(.)为指示函数，若

括号内内容为真，则返回值为根据判断规则拟定的值，否则，返回值为0。

推理规则如下所示：

- R1：解决了留言中提出的全部问题。
- R2：解决了留言中提出的一半以上的问题。
- R3：给出了进一步解决问题的咨询电话。
- R4：没有答非所问、跑题的情况。
- R5：问题转移到其他相关部门后得到了回复。
- R6：对于描述不清楚的留言有给出进一步沟通的联系方式。
- R7：给出相关政策及法律法规。

对于不同类型的答复内容，我们使用不同的推理规则：

	R1	R2	R3	R4	R5	R6	R7
直接回复(*0.8)	√	√	√	√	√		√
继续处理回复 (*0.6)			√			√ √	√
间接回复(*0.5)			√ √	√			√
转移回复(*0)							

3.3.3 例证分析

——直接回复：

“留言详情”:“梅溪湖至今没有一个图书馆，这与梅溪湖品位极不相称。建议在艺术中心先期借一个小馆开办读书馆。方便住在梅溪湖的市民借阅。”

“答复意见”:“网友“UU008706”您好！您的留言已收悉。现将有关情况回复如下：梅溪湖一期引进A市图书馆分馆，位于梅溪湖创新中心，已开馆营业。梅溪湖二期金菊路与雪松路东南角规划有西地省图书馆新馆，目前正在进行前期筹备工作，具体开馆时间待定。感谢您对我们工作的支持、理解与监督！2019年1月9日”

*得分计算：score=2*0.8*(R1+R2+R4) = 2*0.8*3=4.8*

——继续处理回复：

“留言详情”:“12月16日上午，我来到A3区桐梓坡路益丰大药房，参加五折最高减30元的活动，当时也问了他们还有名额，所以用两个帐户迅速付款一共购买六盒210克的云南白药牙膏，付完款后再去查帐户，才发现两个帐户都没享受活动价格，便马上提出退款，他们便以各种理由拒绝退货，并将牙膏都已装入塑料袋中，回家后再清点牙膏时，这才发现他们将一盒180克的牙膏做210克牙膏的价格卖给了我，于是马上又赶到他们店，并对他们这种做法提出质疑，他

们却没有半点愧疚，只是说没库存了。我又再一次提出退货要求，他们还是拒绝。”

“答复意见”：“网友“UU0081194”您好！您的留言已收悉。现将有关情况回复如下：因您未留下联系方式及投诉的相关证据材料，市工商局无法根据您提供的信息进行投诉信息的登记分送和处理。您可直接拨打我局消费者投诉举报热线 0731-12315 进行反映。感谢您对我们工作的支持、理解与监督！2018 年 12 月 28 日”

*得分计算：score = 2*0.6(R3+R6+R7) = 2*0.6(1+2+1) = 4.8*

——间接回复：

“留言详情”：“尊敬的胡书记：您好！过去在小区买房是为了自己，买的便宜的 7 楼，现在接了 80 多岁的老母亲来住，上楼下楼十分不方便，作为工薪阶层，重新买房不现实。期待党的好政策快快落到实处。我们知道 A 市过去也曾作了很大努力，已经在不少小区先试先行安装了电梯，方便了群众，经验很成功，值得推广。期待 A 市住建、财政、国土、安全等部门尽早研究出台为老社区惠民装电梯的规范性文件，特别对于个别不肯安装电梯、影响到整栋楼装不了电梯的用户，应该采取哪些鼓励或鞭策措施等，使党的这一好政策，四面开花，结下硕果。我说的小区在 A 市 A3 区教师村。前两年搭帮政府关心爱护，小区长年积水问题已经得到解决。现在又要恳求党和政府领导，想群众之所想，急群众之所急，解决新的装电梯问题。期待回复。谢谢尊敬的胡书记!”

“答复意见”：“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”

*得分计算：score=2*0.5*(R3+R4+R7) = 2*0.5*(2+1+1) = 4*

——转移回复：

“留言详情”：“强烈反对 I 市 9 路公交车改线路获悉从 7 月 1 日起 9 路公交车要更改线路，滨江路将被弃线。滨江路乘车群众表示强烈反对.....”

“答复意见”：“UU008194”您的留言已收悉。关于您反映的问题，已转市交通运输局调查处理。”

得分计算：0

4 总结

本文的主要目的利用数据挖掘与数学建模技术实现“智慧政务”中群众留言的处理问题，并提高相关部门的办公效率，减少工作人员负担和人工经验引起的误差。

我们对第一问的分类算法采用构建基于 Bert 的深度学习模型，将原始文本和预处理后的数据映射到向量空间中并进行训练，最终得到一个群众留言分类模型；对第二问，我们将问题描述处理成向量形式，并以此作为基础进行文本聚类操作，在一个热点问的一定距离半径内的问题描述可以归类为同一热点问题，并以问题描述的数量作为该热点问题的热度指数；对于第三问我们对答复意见大致分为 4 类，由类别质量为每一个类别赋予权重，然后细分评判规则，对每一类答复意见选定合适的评分规则进行评分。

参考文献

[1] 温有奎，温浩，乔晓东．让知识产生智慧——基于人工智能的文本挖掘与问答技术研究[J]．情报学报, 2019(7).

[2] 袁野，于敏敏，陶于祥, 等．基于文本挖掘的我国人工智能产业政策量化研究[J]．中国电子科学研究院学报，2018, 13(06):43-48.

[3] 李彬．基于文本挖掘的政府工作报告研究综述[J]．福建质量管理，2018(13).

[4] 张惠 王冰．基于文本挖掘的政府公共价值测度与比较[J]．安徽理工大学学报:社会科学版, 2015(17):39.

[5] Tomas Mikolov, Kai Chen, Greg Corrado and Jeffrey Dean. Efficient Estimation of Word Representations in Vector Space[C]. 1st International Conference on Learning Representations, {ICLR} 2013.

[6] Quoc V. Le, Tomas Mikolov. Distributed Representations of Sentences and Documents. Proceedings of the 31th International Conference on Machine Learning, {ICML} 2014, Beijing, China.

[7] Wiwatcharakoses C , Hasegawa O . Two-Pass Clustering Technique for Orientation-Invariant and Language-Independent Text Localization[C]// 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA). IEEE, 2017.

[8] Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." arXiv preprint arXiv:1810.04805 (2018).

[9] Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones,

Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. "Attention is all you need." In Advances in neural information processing systems, pp. 5998-6008. 2017.

[10] 黄培松，黄沛杰，丁健德，艾文程，and 章锦川．″基于隐含主题协同注意力网络的领域分类方法．″ 中文信息学报 34, no. 2 (2020): 73-79.

[11] 王丽亚，刘昌辉，蔡敦波，赵彤洲，and 王梦．″基于字符级联合网络特征融合的中文文本情感分析．″ 微电子学与计算机 1 (2020): 13.