



(12) 发明专利申请

(10) 申请公布号 CN 116629254 A

(43) 申请公布日 2023. 08. 22

(21) 申请号 202310493599.6

(22) 申请日 2023.05.05

(71) 申请人 杭州正策信息科技有限公司  
地址 310000 浙江省杭州市西湖区文三路  
199号1幢5楼501室

(72) 发明人 叶雄 叶琼 丁瑞 宣琪杰  
徐宏淼

(74) 专利代理机构 杭州新维鹰知识产权代理有  
限公司 33474  
专利代理师 张玉平

(51) Int. Cl.  
G06F 40/289 (2020.01)  
G06F 40/216 (2020.01)  
G06N 20/00 (2019.01)

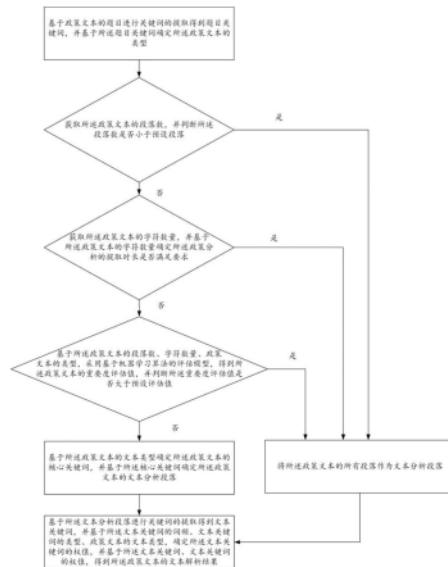
权利要求书2页 说明书7页 附图4页

(54) 发明名称

一种基于文本分析和识别的政策文本分析方法

(57) 摘要

本发明提供一种基于文本分析和识别的政策文本分析方法,属于数据处理技术领域,具体包括:基于政策文本的题目确定政策文本的类型;获取政策文本的段落数,基于政策文本的段落数、字符数量、类型得到重要度评估值,并判断其是否大于预设评估值,若是,则将政策文本的所有段落作为文本分析段落,若否,则基于文本类型确定政策文本的核心关键词,并基于核心关键词确定文本分析段落;基于文本分析段落进行关键词的提取得到文本关键词,并基于文本关键词的词频、文本关键词的类型、政策文本的文本类型确定其权值,并基于文本关键词、文本关键词的权值得到政策文本的文本解析结果,从而进一步提升了文本解析的准确性和可靠性。



CN 116629254 A

1. 一种基于文本分析和识别的政策文本分析方法,其特征在于,具体包括:

S11基于政策文本的题目进行关键词的提取得到题目关键词,并基于所述题目关键词确定所述政策文本的类型;

S12获取所述政策文本的段落数,并判断所述段落数是否小于预设段落,若是,则将所述政策文本的所有段落作为文本分析段落,并进入步骤S15;若否,则进入步骤S13;

S13获取所述政策文本的字符数量,并基于所述政策文本的字符数量确定所述政策分析的提取时长是否满足要求,若是,则将所述政策文本的所有段落作为文本分析段落,并进入步骤S15;若否,则进入步骤S14;

S14基于所述政策文本的段落数、字符数量、政策文本的类型,采用基于机器学习算法的评估模型,得到所述政策文本的重要度评估值,并判断所述重要度评估值是否大于预设评估值,若是,则将所述政策文本的所有段落作为文本分析段落,并进入步骤S15,若否,则基于所述政策文本的文本类型确定所述政策文本的核心关键词,并基于所述核心关键词确定所述政策文本的文本分析段落;

S15基于所述文本分析段落进行关键词的提取得到文本关键词,并基于所述文本关键词的词频、文本关键词的类型、政策文本的文本类型,确定所述文本关键词的权值,并基于所述文本关键词、文本关键词的权值,得到所述政策文本的文本解析结果。

2. 如权利要求1所述的政策文本分析方法,其特征在于,所述题目关键词根据所述政策文本的题目,采用基于TF-IDF算法的关键词提取模型进行提取,并基于所述题目关键词的匹配结果确定所述政策文本的类型。

3. 如权利要求1所述的政策文本分析方法,其特征在于,所述预设段落根据所述政策文本的类型、所述政策文本的颁布日期进行确定,其中所述政策文本的类型越重要,所述政策文本的下发日期越近,则所述预设段落越小。

4. 如权利要求1所述的政策文本分析方法,其特征在于,所述政策文本的重要度评估值构建的具体步骤为:

基于所述政策文本的类型,确定所述政策文本的类型是否属于指定文本类型,若是,则将所述政策文本的重要度评估值设置为1,若否,则进入下一步骤;

基于所述政策文本的段落数和所述字符数量,采用基于机器学习算法的文本提取难度评估模型,得到所述政策文本的提取难度评估值,并基于所述提取难度评估值确定所述政策文本的提取时长是否无法满足要求,若是,则进入下一步骤,若否,则将所述政策文本的重要度评估值设置为1;

基于所述政策文本的提取难度评估值、所述政策文本的类型,采用基于机器学习算法的评估模型,得到所述政策文本的基础重要度评估值;

基于所述政策文本的颁布时间对所述基础重要度评估值进行修正,得到所述政策文本的重要度评估值。

5. 如权利要求4所述的政策文本分析方法,其特征在于,所属指定文本类型包括通告、通知、通报、批复。

6. 如权利要求1所述的政策文本分析方法,其特征在于,基于所述核心关键词进行所述政策文本的文本分析段落的确定,具体包括:

基于所述核心关键词和所述政策文本,确定所述政策文本中存在所述核心关键词的段

落,并将所述政策文本中存在所述核心关键词的段落作为所述政策文本的文本分析段落。

7.如权利要求1所述的政策文本分析方法,其特征在于,文本关键词的权值构建的具体步骤为:

基于所述政策文本的文本类型对所述文本关键词进行筛选,得到所述文本关键词中的重要关键词,并将所述重要关键词的权值设置为1,并将除所述重要关键词外的文本关键词作为其它关键词;

基于所述其它关键词的类型对所述其它关键词进行筛选,得到所述其它关键词中的重要类型关键词,并将所述重要类型关键词的权值设置为1,并将除所述重要类型关键词外的其它关键词作为筛选关键词;

基于所述筛选关键词的词频、筛选关键词的类型、政策文本的文本类型,采用基于机器学习算法的权值评估模型,确定所述筛选关键词的权值。

8.如权利要求1所述的政策文本分析方法,其特征在于,基于所述文本关键词的权值进行文本解析关键词的确认,并基于所述文本解析关键词生成所述政策文本的文本解析结果。

9.一种计算机系统,包括:通信连接的存储器和处理器,以及存储在所述存储器上并能够在所述处理器上运行的计算机程序,其特征在于:所述处理器运行所述计算机程序时执行权利要求1-8任一项所述的一种基于文本分析和识别的政策文本分析方法。

10.一种计算机存储介质,其上存储有计算机程序,当所述计算机程序在计算机中执行时,令计算机执行权利要求1-8任一项所述的一种基于文本分析和识别的政策文本分析方法。

## 一种基于文本分析和识别的政策文本分析方法

### 技术领域

[0001] 本发明属于文本分析技术领域,尤其涉及一种基于文本分析和识别的政策文本分析方法。

### 背景技术

[0002] 为了实现对政策文本的分析,在授权发明专利授权公告号CN112906382B《基于图神经网络的政策文本多标签标注方法及系统》中通过获取待标注的政策文本;对待标注的政策文本进行预处理,对预处理后的政策文本进行分词;将分词得到的单词和预先得到的加权单词向量,输入到训练后的全连接神经网络中,输出待标注政策文本的多标签,但是却存在以下技术问题:

[0003] 1、未考虑结合政策文本的类型进行分词的权值的确定,对于不同类型的政策文本,例如通知、决定、意见等等,若采用相同的文本解析模型和权值确定方式,则会导致最终的文本分析结果都不够准确。

[0004] 2、未考虑结合关键词的识别结果进行文本分析的段落的确定,对于政策文本,特别是篇幅较大时,若不能结合关键词的识别结果进行文本分析的段落的定位,则会导致最终的文本分析的效率会有所影响。

[0005] 针对上述技术问题,本发明提供了一种基于文本分析和识别的政策文本分析方法。

### 发明内容

[0006] 为实现本发明目的,本发明采用如下技术方案:

[0007] 根据本发明的一个方面,提供了一种基于文本分析和识别的政策文本分析方法。

[0008] 一种基于文本分析和识别的政策文本分析方法,其特征在于,具体包括:

[0009] S11基于政策文本的题目进行关键词的提取得到题目关键词,并基于所述题目关键词确定所述政策文本的类型;

[0010] S12获取所述政策文本的段落数,并判断所述段落数是否小于预设段落,若是,则将所述政策文本的所有段落作为文本分析段落,并进入步骤S15;若否,则进入步骤S13;

[0011] S13获取所述政策文本的字符数量,并基于所述政策文本的字符数量确定所述政策分析的提取时长是否满足要求,若是,则将所述政策文本的所有段落作为文本分析段落,并进入步骤S15;若否,则进入步骤S14;

[0012] S14基于所述政策文本的段落数、字符数量、政策文本的类型,采用基于机器学习算法的评估模型,得到所述政策文本的重要度评估值,并判断所述重要度评估值是否大于预设评估值,若是,则将所述政策文本的所有段落作为文本分析段落,并进入步骤S15,若否,则基于所述政策文本的文本类型确定所述政策文本的核心关键词,并基于所述核心关键词确定所述政策文本的文本分析段落;

[0013] S15基于所述文本分析段落进行关键词的提取得到文本关键词,并基于所述文本

关键词的词频、文本关键词的类型、政策文本的文本类型,确定所述文本关键词的权值,并基于所述文本关键词、文本关键词的权值,得到所述政策文本的文本解析结果。

[0014] 通过基于题目关键词进行政策文本的类型的确定,从而使得对于政策文本的文本分析能够结合政策文本的类型,从而进一步提升了文本解析的准确性和可靠性。

[0015] 通过预设段落、设定字符数量的设置,从而使得能够结合政策文本的段落数以及字符数量实现了对提取难度较小的政策文本的筛选,在保证政策文本的文本解析的处理效率的基础上,保证了政策文本的文本解析的准确性,同时也降低了需要进行重要度评估值的评估的政策文本的数量,提升了系统的处理效率。

[0016] 通过基于政策文本的段落数、字符数量、政策文本的类型,采用基于机器学习算法的评估模型,得到所述政策文本的重要度评估值,从而实现了从多角度对政策文本的重要度的评估,保证了重要的政策文本的文本解析的重要性的评估的准确性,也保证了重要度较高的政策文本的文本解析的准确性。

[0017] 通过基于文本关键词的词频、文本关键词的类型、政策文本的文本类型,确定文本关键词的权值,从而实现了从多角度对文本关键词的权值的评估,避免了原来的仅仅依靠词频导致的权值的评估的准确性较差的技术问题,进一步提升了文本解析的准确性。

[0018] 另一方面,本申请实施例中提供一种计算机系统,包括:通信连接的存储器和处理器,以及存储在所述存储器上并能够在所述处理器上运行的计算机程序,其特征在于:所述处理器运行所述计算机程序时执行上述的一种基于文本分析和识别的政策文本分析方法。

[0019] 另一方面,本发明提供了一种计算机存储介质,其上存储有计算机程序,当所述计算机程序在计算机中执行时,令计算机执行上述的一种基于文本分析和识别的政策文本分析方法。

[0020] 其他特征和优点将在随后的说明书中阐述,并且,部分地从说明书中变得显而易见,或者通过实施本发明而了解。本发明的目的和其他优点在说明书以及附图中所特别指出的结构来实现和获得。

[0021] 为使本发明的上述目的、特征和优点能更明显易懂,下文特举较佳实施例,并配合所附附图,作详细说明如下。

### 附图说明

[0022] 通过参照附图详细描述其示例实施方式,本发明的上述和其它特征及优点将变得更加明显。

[0023] 图1是根据实施例1的一种基于文本分析和识别的政策文本分析方法的流程图;

[0024] 图2是根据实施例1的政策文本的重要度评估值构建的具体步骤的流程图;

[0025] 图3是根据实施例1的文本关键词的权值构建的具体步骤的流程图;

[0026] 图4是根据实施例3的一种计算机存储介质的结构图。

### 具体实施方式

[0027] 现在将参考附图更全面地描述示例实施方式。然而,示例实施方式能够以多种形式实施,且不应被理解为限于在此阐述的实施方式;相反,提供这些实施方式使得本发明将全面和完整,并将示例实施方式的构思全面地传达给本领域的技术人员。图中相同的附图

标记表示相同或类似的结构,因而将省略它们的详细描述。

[0028] 用语“一个”、“一”、“该”、“所述”用以表示存在一个或多个要素/组成部分/等;用语“包括”和“具有”用以表示开放式的包括在内的意思并且是指除了列出的要素/组成部分/等之外还可存在另外的要素/组成部分/等。

[0029] 申请人发现,在进行政府下发的与企业相关的政策文本的解析过程中,现有技术中未考虑结合政策文本的类型进行分词的权值的确定,从而导致最终的文本分析结果不够准确;未考虑结合关键词的识别结果进行文本分析的段落的确定,对于政策文本,特别是篇幅较大时,若不能结合关键词的识别结果进行文本分析的段落的定位,则会导致最终的文本分析的效率会有所影响。。

[0030] 实施例1

[0031] 为解决上述问题,根据本发明的一个方面,如图1所示,提供了一种基于文本分析和识别的政策文本分析方法,其特征在于,具体包括:

[0032] S11基于政策文本的题目进行关键词的提取得到题目关键词,并基于所述题目关键词确定所述政策文本的类型;

[0033] 具体的,所述题目关键词根据所述政策文本的题目,采用基于TF-IDF算法的关键词提取模型进行提取,并基于所述题目关键词的匹配结果确定所述政策文本的类型。

[0034] 在另外一种可能的实施例中,若提取得到的题目关键词为通知,则根据通知的匹配结果确定政策文本的类型为通知。

[0035] 通过基于题目关键词进行政策文本的类型的确定,从而使得对于政策文本的文本分析能够结合政策文本的类型,从而进一步提升了文本解析的准确性和可靠性。

[0036] S12获取所述政策文本的段落数,并判断所述段落数是否小于预设段落,若是,则将所述政策文本的所有段落作为文本分析段落,并进入步骤S15;若否,则进入步骤S13;

[0037] 具体的,所述预设段落根据所述政策文本的类型、所述政策文本的颁布日期进行确定,其中所述政策文本的类型越重要,所述政策文本的下发日期越近,则所述预设段落越小。

[0038] 具体的举个例子,若政策文本的段落数为30段,预设段落为29段,此时的段落数量较少,则将所述政策文本的所有段落作为文本分析段落。

[0039] S13获取所述政策文本的字符数量,并基于所述政策文本的字符数量确定所述政策分析的提取时长是否满足要求,若是,则将所述政策文本的所有段落作为文本分析段落,并进入步骤S15;若否,则进入步骤S14;

[0040] 具体的举个例子,若政策文本的字符数量为6457个字符,而设定字符数量为5000字,此时的提取时长无法满足要求,则需要进入步骤S14进行评估,若政策文本的字符数量为4785个字符,则将政策文本的所有段落作为文本分析段落。

[0041] 通过预设段落、设定字符数量的设置,从而使得能够结合政策文本的段落数以及字符数量实现了对提取难度较小的政策文本的筛选,在保证政策文本的文本解析的处理效率的基础上,保证了政策文本的文本解析的准确性,同时也降低了需要进行重要度评估值的评估的政策文本的数量,提升了系统的处理效率。

[0042] S14基于所述政策文本的段落数、字符数量、政策文本的类型,采用基于机器学习算法的评估模型,得到所述政策文本的重要度评估值,并判断所述重要度评估值是否大于

预设评估值,若是,则将所述政策文本的所有段落作为文本分析段落,并进入步骤S15,若否,则基于所述政策文本的文本类型确定所述政策文本的核心关键词,并基于所述核心关键词确定所述政策文本的文本分析段落;

[0043] 具体的,如图2所示,所述政策文本的重要度评估值构建的具体步骤为:

[0044] S21基于所述政策文本的类型,确定所述政策文本的类型是否属于指定文本类型,若是,则将所述政策文本的重要度评估值设置为1,若否,则进入步骤S22;

[0045] 具体的举个例子,政策文本包括政策文件名称,政策类型、发文文号、颁布部门、效力级别、时效性、颁布日期、实施日期、内容主题等多个变量;政策类型包括决议、决定、命令、公报、公告、通告、意见、通知、通报、报告、请示、批复、议案、函、纪要等。

[0046] 具体的,所属指定文本类型包括通告、通知、通报、批复。

[0047] 具体的举个例子,政策文本的重要度评估值的取值范围在0到1之间,其中所述政策文本的重要度评估值越高,则说明该政策文本的重要程度越高。

[0048] S22基于所述政策文本的段落数和所述字符数量,采用基于机器学习算法的文本提取难度评估模型,得到所述政策文本的提取难度评估值,并基于所述提取难度评估值确定所述政策文本的提取时长是否无法满足要求,若是,则进入步骤S23,若否,则将所述政策文本的重要度评估值设置为1;

[0049] 具体的举个例子,当政策文本的提取难度评估值为0.6,当其小于初始设定的值0.7时,则说明该政策文本的提取难度评估值并不高,因此将政策文本的重要度评估值设置为1,从而实现对提取难度较低的政策文本的提取。

[0050] 具体的举个例子,所述基于机器学习算法的文本提取难度评估模型采用基于GS0-BPNN算法的文本提取难度评估模型,其中所述文本提取难度评估模型构建的具体步骤为:

[0051] 步骤1:根据需要优化的参数,确定BP神经网络的结构,运用改进萤火虫算法进行实数编码,将BP神经网络中的连接权值 $\omega_{ih}$ , $\omega_{hq}$ 和阈值 $\theta_\varepsilon$ 作为一个整体参数进行编码,其中每一个个体都表示BPNN的全部权值和阈值,并且表示一个BP神经网络的结构。

[0052] 步骤2:随机生成N个萤火虫作为算法中的初始萤火虫种群,设置每只萤火虫的初始荧光素 $l_0$ ,感知半径 $R_s$ ,初始步长 $s$ ,最大、最小步长 $s_{\max}$ 和 $s_{\min}$ ,荧光素的更新率 $\gamma$ 和挥发系数 $\rho$ ,同时令初始迭代 $t=0$ ,设定最大迭代次数 $t_{\max}$ 。

[0053] 步骤3:选择适当的适应度函数,本发明选取的适应度函数为正则化均方根误差函数,根据公式

[0054]  $l_i(t+1) = (1-\rho) l_i(t) + \gamma f(x_i(t+1))$

[0055] 进行萤火虫的荧光素更新,计算萤火虫在t代的荧光素值。

[0056] 步骤4:利用轮盘对赌法选择个体j,当萤火虫i和j两者之间的距离小于决策半径时,根据公式

[0057] 
$$P_{ij}(t) = \frac{l_j(t) - l_i(t)}{\sum_{k \in N_i(t)} l_k(t) - l_i(t)}$$

[0058] 计算出萤火虫i向萤火虫j移动的概率,并且用公式

[0059] 
$$x_i(t+1) = x_i(t) + L \frac{x_j(t) - x_i(t)}{\|x_j(t) - x_i(t)\|}$$

[0060] 进行位置更新,通过适应度函数

$$[0061] \quad f = \left\{ \frac{1}{(S-1)\sigma^2} \sum_{\epsilon=1}^S [t_p - y_p]^2 \right\}^{\frac{1}{2}}$$

[0062] 来计算更新位置后的目标函数值,进而更新全局最优值。

[0063] 步骤5:进行决策半径的更新。

[0064] 步骤6:若迭代达到最大设定数 $t_{\max}$ 或者精度满足用户要求,则程序完成;否则返回步骤3。

[0065] 步骤7:通过改进GSO算法优化得到的结果确定BPNN的权值和阈值,作为BP神经网络预测模型的初始连接权值和阈值。

[0066] 步骤8:根据确定的BP神经网络的结构,即确定的输入层节点数、隐含层节点数和输出层节点数,选定合适的学习速率 $\eta$ 和惯性系数 $\alpha$ ,再根据改进萤火虫算法得到的BPNN的连接权值和阈值对BPNN预测网络进行训练,并基于训练完成的BPNN预测网络进行文本提取难度的评估。

[0067] 具体的举个例子,基本萤火虫算法与其它新兴群算法相同,具有一系列问题,其中最容易出现的问题是在迭代后期存在局部或者全局极值附近反复震荡,在本文中可以借鉴粒子群算法中引入惯性权重的方法来解决此问题。

[0068] 由于迭代过后萤火虫之间彼此吸引而造成距离减小,个体间相对吸引力增大,从而使移动的距离过大而无法到达稳定的最终极值位置;惯性权重因子可以通过调节萤火虫个体函数值加强搜索能力,减小算法误差,因此引入惯性权重因子

$$[0069] \quad \omega = \omega_{\max} - (\omega_{\max} - \omega_{\min}) \cdot \frac{1}{2} \left( \frac{1}{t+1} \left( \frac{t}{t_{\max}} \right)^2 + \left( 1 - \frac{1}{t+1} \right) \sin \left( \frac{\pi}{2} \left( 1 - \frac{t}{t_{\max}} \right) \right) \right)$$

[0070] 式中, $\omega_{\max}$ 和 $\omega_{\min}$ 分别代表最大权重和最小权重; $t$ 和 $t_{\max}$ 代表当前迭代次数和最大迭代次数。

[0071] 因此改进后的位置更新公式如下:

$$[0072] \quad x_i(t+1) = \omega x_i(t) + L \left( 1 - \frac{\omega}{\omega^2 + 1} \right) \frac{x_j(t) - x_i(t)}{\|x_j(t) - x_i(t)\|}$$

[0073] 其中 $x_i(t)$ 为萤火虫 $i$ 在第 $t$ 次迭代时的位置, $L$ 为萤火虫 $i$ 的移动步长, $x_j(t)$ 为萤火虫 $j$ 在第 $t$ 次迭代时的位置。

[0074] S23基于所述政策文本的提取难度评估值、所述政策文本的类型,采用基于机器学习算法的评估模型,得到所述政策文本的基础重要度评估值;

[0075] S24基于所述政策文本的颁布时间对所述基础重要度评估值进行修正,得到所述政策文本的重要度评估值。

[0076] 具体的,基于所述核心关键词进行所述政策文本的文本分析段落的确定,具体包括:

[0077] 基于所述核心关键词和所述政策文本,确定所述政策文本中存在所述核心关键词的段落,并将所述政策文本中存在所述核心关键词的段落作为所述政策文本的文本分析段落。

[0078] 通过基于政策文本的段落数、字符数量、政策文本的类型,采用基于机器学习算法



的评估模型,得到所述政策文本的重要度评估值,从而实现了从多角度对政策文本的重要度的评估,保证了重要的政策文本的文本解析的重要性的评估的准确性,也保证了重要度较高的政策文本的文本解析的准确性。

[0079] S15基于所述文本分析段落进行关键词的提取得到文本关键词,并基于所述文本关键词的词频、文本关键词的类型、政策文本的文本类型,确定所述文本关键词的权值,并基于所述文本关键词、文本关键词的权值,得到所述政策文本的文本解析结果。

[0080] 具体的,如图3所示,文本关键词的权值构建的具体步骤为:

[0081] S31基于所述政策文本的文本类型对所述文本关键词进行筛选,得到所述文本关键词中的重要关键词,并将所述重要关键词的权值设置为1,并将除所述重要关键词外的文本关键词作为其它关键词;

[0082] S32基于所述其它关键词的类型对所述其它关键词进行筛选,得到所述其它关键词中的重要类型关键词,并将所述重要类型关键词的权值设置为1,并将除所述重要类型关键词外的其它关键词作为筛选关键词;

[0083] S33基于所述筛选关键词的词频、筛选关键词的类型、政策文本的文本类型,采用基于机器学习算法的权值评估模型,确定所述筛选关键词的权值。

[0084] 具体的,所述文本关键词的权值的取值范围在0到1之间,其中所述文本关键词的权值越大,则所述文本关键词的重要程度越高。

[0085] 具体的,基于所述文本关键词的权值进行文本解析关键词的确认,并基于所述文本解析关键词生成所述政策文本的文本解析结果。

[0086] 具体的举个例子,一般选取权值较大的文本关键词作为文本解析关键词。

[0087] 具体的举个例子,当等到文本解析关键词后,将所述文本解析关键词进行输出作为所述政策文本的文本解析结果,或者将所述文本解析关键词标注出来,从而实现对政策文本的文本解析。

[0088] 通过基于文本关键词的词频、文本关键词的类型、政策文本的文本类型,确定文本关键词的权值,从而实现了从多角度对文本关键词的权值的评估,避免了原来的仅仅依靠词频导致的权值的评估的准确性较差的技术问题,进一步提升了文本解析的准确性。

[0089] 实施例2

[0090] 本申请实施例中提供一种计算机系统,包括:通信连接的存储器和处理器,以及存储在所述存储器上并能够在所述处理器上运行的计算机程序,其特征在于:所述处理器运行所述计算机程序时执行上述的一种基于文本分析和识别的政策文本分析方法。

[0091] 具体的,本实施例还提供了一种计算机系统,该计算机系统包括通过系统总线连接的处理器、存储器、网络接口和数据库;其中,该计算机系统的处理器用于提供计算和控制能力;该计算机系统的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机程序和数据库。该内存储器为非易失性存储介质中的操作系统和计算机程序的运行提供环境。该计算机设备网络接口用于与外部的终端通过网络连接通信。该计算机程序被处理器执行时以实现上述的一种基于文本分析和识别的政策文本分析方法。

[0092] 实施例3

[0093] 如图4所示,本发明提供了一种计算机存储介质,其上存储有计算机程序,当所述

计算机程序在计算机中执行时,令计算机执行上述的一种基于文本分析和识别的政策文本分析方法。

[0094] 具体的,本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,是可以通过计算机程序来指令相关的硬件来完成,所述的计算机程序可存储于一非易失性计算机可读取存储介质中,该计算机程序在执行时,可包括如上述各方法的实施例的流程。其中,本申请所提供的各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用,均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器(ROM)、可编程ROM(PROM)、电可编程ROM(EPROM)、电可擦除可编程ROM(EEPROM)或闪存。易失性存储器可包括随机存取存储器(RAM)或者外部高速缓冲存储器。作为说明而非局限,RAM以多种形式可得,诸如静态RAM(SRAM)、动态RAM(DRAM)、同步DRAM(SDRAM)、双数据率SDRAM(DDRSDRAM)、增强型SDRAM(ESDRAM)、同步链路(Synchlink)DRAM(SLDRAM)、存储器总线(Rambus)直接RAM(RDRAM)、直接存储器总线动态RAM(DRDRAM)、以及存储器总线动态RAM(RDRAM)等。

[0095] 在本申请所提供的几个实施例中,应该理解到,所揭露的系统和方法,也可以通过其它的方式实现。以上所描述的系统实施例仅仅是示意性的,例如,附图中的流程图和框图显示了根据本发明的多个实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段或代码的一部分,模块、程序段或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现方式中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个连续的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或动作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0096] 另外,在本发明各个实施例中的各功能模块可以集成在一起形成一个独立的部分,也可以是各个模块单独存在,也可以两个或两个以上模块集成形成一个独立的部分。

[0097] 功能如果以软件功能模块的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备)执行本发明各个实施例方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0098] 以上所述依据本发明的理想实施例为启示,通过上述的说明内容,相关工作人员完全可以在不偏离本项发明技术思想的范围内,进行多样的变更以及修改。本项发明的技术性范围并不局限于说明书上的内容,必须要根据权利要求范围来确定其技术性范围。

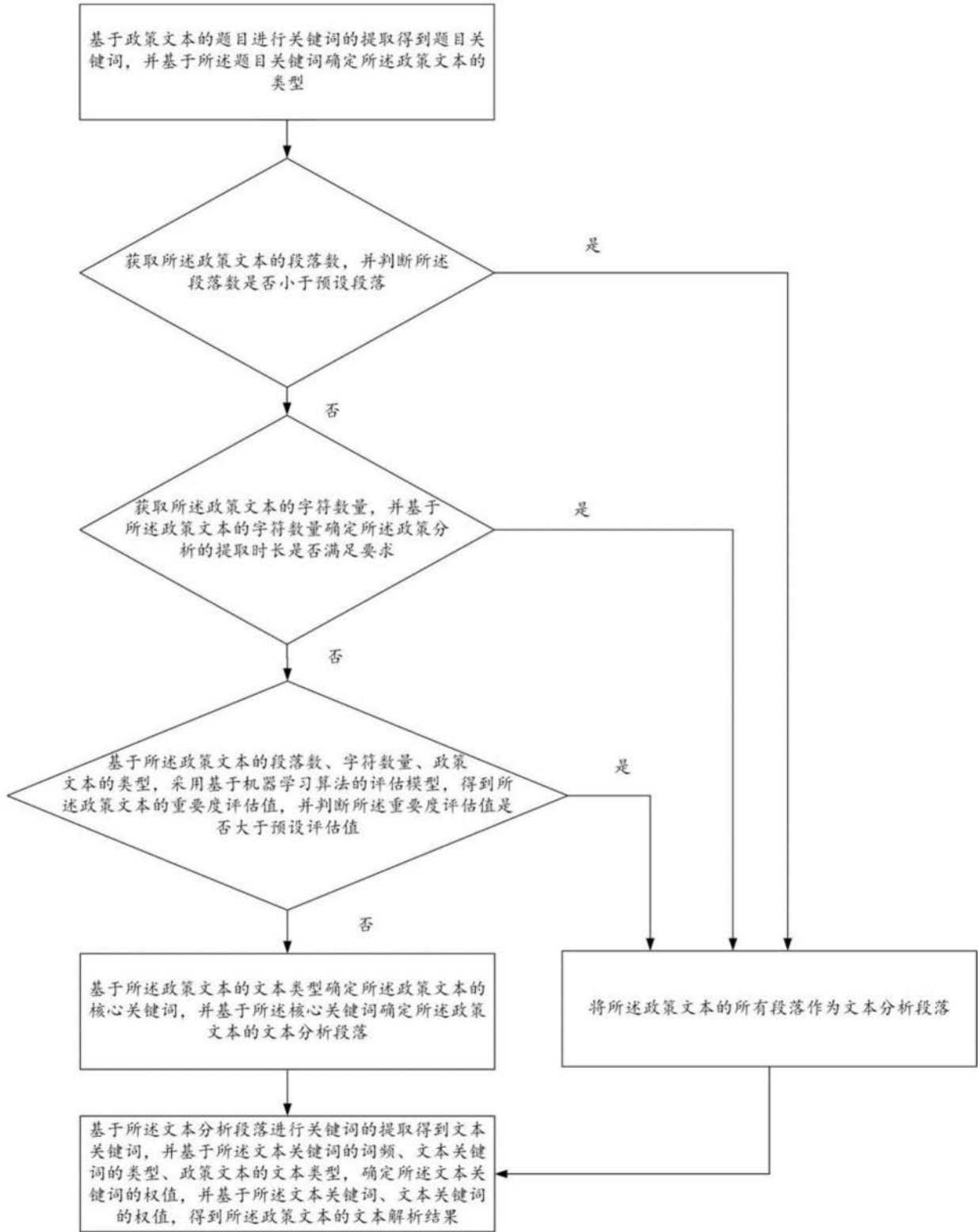


图1

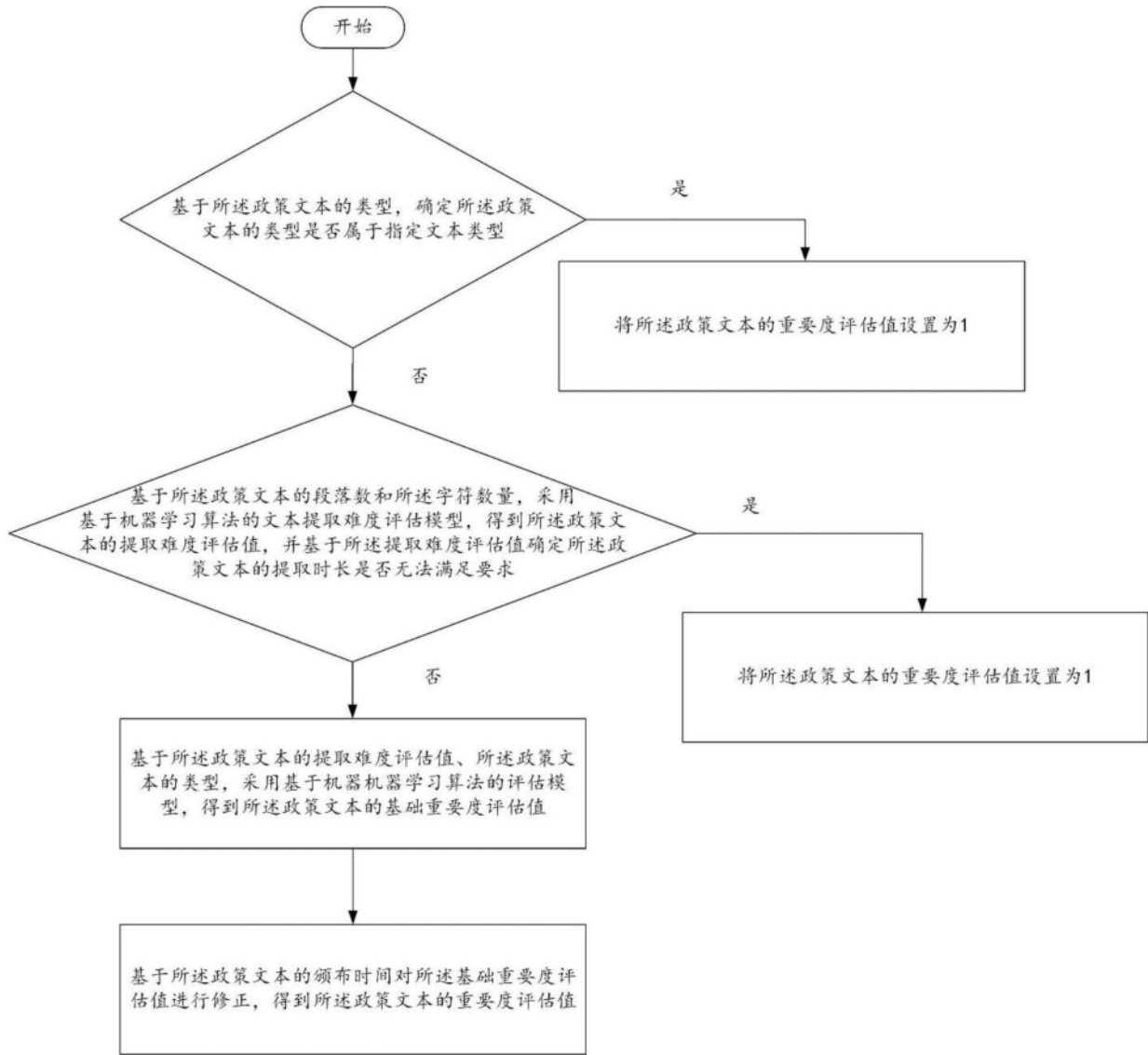


图2

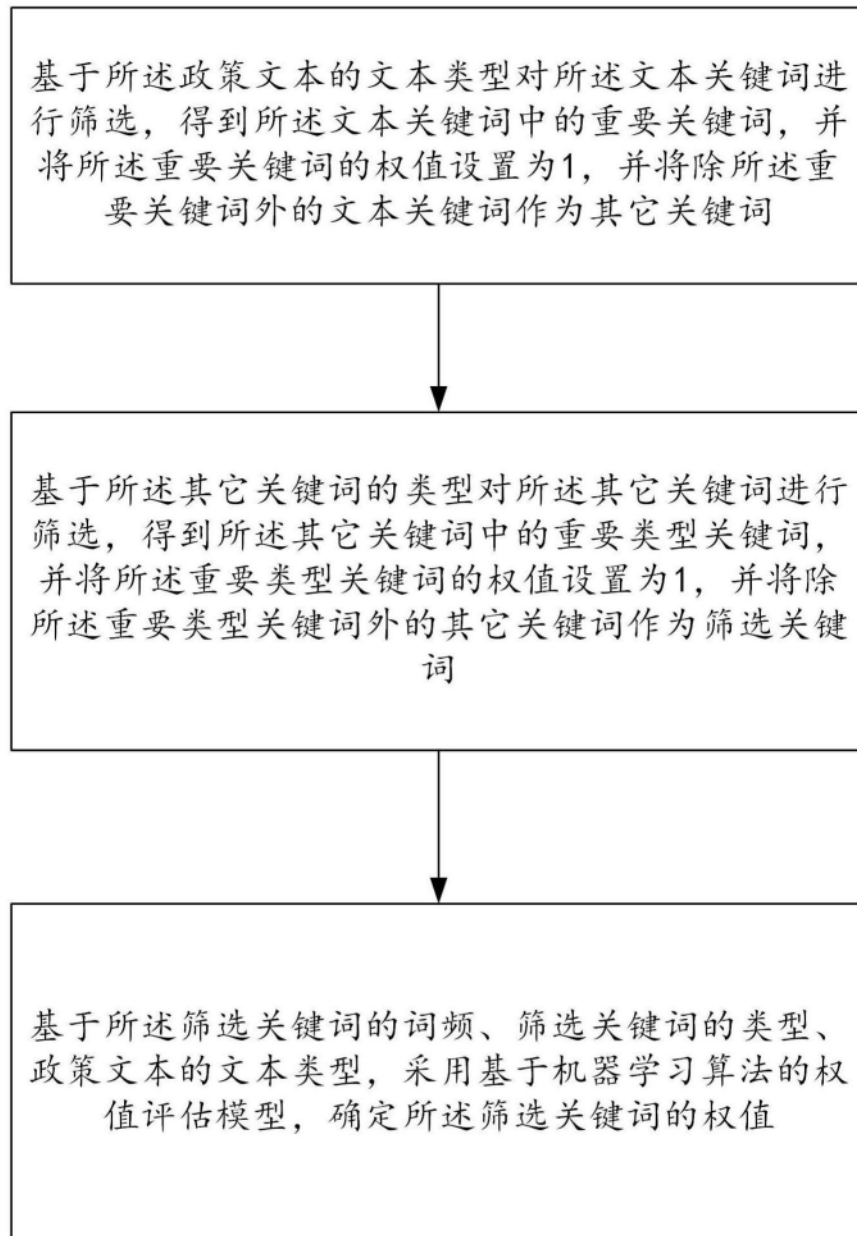


图3

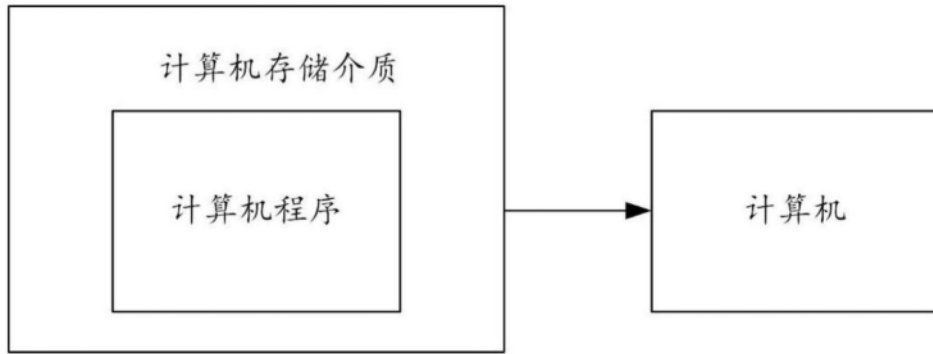


图4