(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2020 年 8 月 20 日 (20.08.2020)

WIPO | PCT

(10) 国际公布号
WO 2020/164267 A1

(51) 国际专利分类号:
G06F 16/35 (2019.01)

(21) 国际申请号: PCT/CN2019/117225

(22) 国际申请日: 2019 年 11 月 11 日 (11.11.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910113183.0 2019年2月13日 (13.02.2019) CN

(71) 申请人: 平安科技（深圳）有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN]; 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(72) 发明人: 徐亮 (XU, Liang); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。 金戈 (JIN, Ge); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。 肖京 (XIAO, Jing); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市世纪恒程知识产权代理事务所 (CENFO INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市南山区粤海街道高新技术产业园北区松坪山路3号奥特讯电力大厦201, Guangdong 518057 (CN)。

(81) 指定国 (除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(54) **Title:** TEXT CLASSIFICATION MODEL CONSTRUCTION METHOD AND APPARATUS, AND TERMINAL AND STORAGE MEDIUM

(54) **发明名称**: 文本分类模型构建方法、装置、终端及存储介质

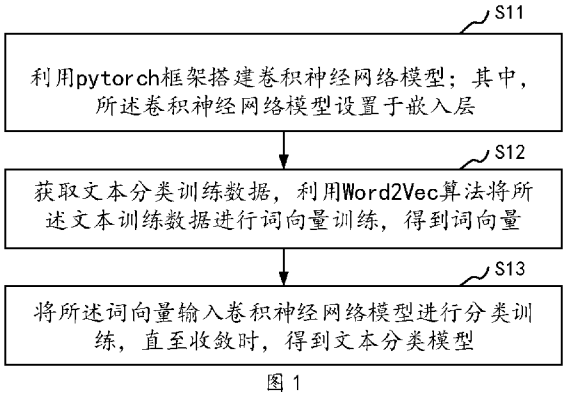

图 1

S11 Construct a convolutional neural network model by using a PyTorch framework, wherein the convolutional neural network model is arranged on an embedding layer

S12 Obtain text classification training data, and perform word vector training on the text training data by using a Word2Vec algorithm to obtain word vectors

S13 Input the word vectors into the convolutional neural network model for classification training, and obtain a text classification model when convergence occurs

(57) **Abstract:** A text classification model construction method and apparatus, and a terminal and a storage medium. The text classification model construction method comprises: constructing a convolutional neural network model by using a PyTorch framework, wherein the convolutional neural network model is arranged on an embedding layer (S11); obtaining text classification training data, and performing word vector training on the text training data by using a Word2Vec algorithm to obtain word vectors (S12); and inputting the word vectors into the convolutional neural network model for classification training, and obtaining a text classification model when convergence occurs (S13). A PyTorch framework is used for the text classification model. Due to the fact that the object-oriented interface design of the PyTorch framework comes from the Torch, the interface design of the Torch has the advantages of being flexible and easy to use, the PyTorch framework can print calculation results layer by layer to facilitate debugging, and therefore, the constructed text classification model is easier to maintain and debug.

WO 2020/164267 A1

[见续页]

(84) 指定国(除另有指明，要求每一种可提供的地区保护)：ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布：

— 包括国际检索报告(条约第21条(3))。

(57) 摘要：一种文本分类模型构建方法、装置、终端及存储介质。所述文本分类模型构建方法包括：利用pytorch框架搭建卷积神经网络模型；其中，所述卷积神经网络模型设置于嵌入层（S11）；获取文本分类训练数据，利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量（S12）；将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到文本分类模型（S13）。所述文本分类模型采用Pytorch框架，由于Pytorch框架面向对象的接口设计来源于torch，而torch的接口设计具有灵活易用的特点，并且PyTorch框架能够逐层打印出计算结果以便于调试，因此构建的文本分类模型更易于维护和调试。

文本分类模型构建方法、装置、终端及存储介质

[1] 本申请要求于2019年02月13日提交中国专利局、申请号为

[2] CN201910113183.0、申请名称为“文本分类模型构建方法、装置、终端及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在申请中。

[3] 技术领域

[4] 本申请涉及神经网络技术领域，尤其涉及一种文本分类模型构建方法、装置、终端及存储介质。

[5] 背景技术

[6] 随着移动互联网时代的到来，内容的生产和传播都发生了深刻的变化，为了满足信息爆炸背景下用户的多样化需求，迫切需要对文本信息进行有效的组织，文本分类是数据挖掘和信息检索领域研究的热点和核心技术。

[7] 现有的文本分类模型，如果要在文本分类问题上采用神经网络算法一般会采用tensorflow框架。但tensorflow框架代码冗长、接口设计过于晦涩难懂，导致构建的文本分类模型维护困难、调试不便、不易于操作。

[8] 发明内容

[9] 本申请提供一种文本分类模型构建方法、装置、终端及存储介质，以解决当前采用tensorflow框架构建文本分类模型时，因tensorflow框架代码冗长、接口设计过于晦涩难懂而导致构建的文本分类模型维护困难、调试不便、不易于操作的问题。

[10] 为解决上述问题，本申请采用如下技术方案：

[11] 本申请提供一种文本分类模型构建方法，包括如下步骤：

[12] 利用pytorch框架搭建卷积神经网络模型；其中，所述卷积神经网络模型设置于嵌入层；

[13] 获取文本分类训练数据，利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量；

[14] 将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到文本分

类模型。

[15] 在一实施例中，所述利用Word2Vec算法将所述文本训练数据进行词向量训练之前，还包括：

[16] 根据正则表达式匹配规则去除文本训练数据的停用词和符号；

[17] 利用结巴分词库将去除停用词和符号的文本训练数据进行中文分词。

[18] 在一实施例中，所述利用结巴分词库将去除停用词和符号的文本训练数据进行中文分词的步骤，包括：

[19] 利用结巴分词库确定文本训练数据中汉字之间的相关度；

[20] 将相关度大于预设值的汉字组成分词，得到分词结果。

[21] 在一实施例中，所述将所述词向量输入卷积神经网络模型进行分类训练的步骤，包括：

[22] 根据所述词向量并通过交叉熵损失函数和ADAM优化算法对卷积神经网络模型进行分类训练。

[23] 在一实施例中，所述利用Word2Vec算法将所述文本训练数据进行词向量训练的步骤，包括：

[24] 利用Word2Vec算法对大型语料数据进行词向量训练，得到词向量字典；

[25] 根据所述词向量字典将文本训练数据转化为词向量。

[26] 在一实施例中，所述利用pytorch框架搭建卷积神经网络模型之后，还包括：

[27] 在所述卷积神经网络模型上建立位置注意力机制和通道注意力机制；其中，所述位置注意力机制和通道注意力机制的输入与卷积神经网络模型的激活层的输出连接，所述位置注意力机制和通道注意力机制的输出与卷积神经网络模型的全连接层的输入连接。

[28] 在一实施例中，所述将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到文本分类模型的步骤，包括：

[29] 根据分类训练结果计算卷积神经网络模型的分类准确率；

[30] 当分类准确率低于预设值时，调整卷积神经网络模型的参数，利用词向量对所述卷积神经网络模型重新训练，直至收敛时，得到文本分类模型。

[31] 本申请提供的一种文本分类模型构建装置，包括：

[32] 搭建模块，用于利用pytorch框架搭建卷积神经网络模型；其中，所述卷积神经网络模型设置于嵌入层；

[33] 获取模块，用于获取文本分类训练数据，利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量；

[34] 训练模块，用于将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到文本分类模型。

[35] 本申请提供一种终端，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行如上任一项所述的文本分类模型构建方法的步骤。

[36] 本申请提供一种存储介质，其上存储有计算机可读指令，所述计算机可读指令被处理器执行时，实现如上任一项所述的文本分类模型构建方法。

[37] 相对于现有技术，本申请的技术方案至少具备如下优点：

[38] 本申请提供的文本分类模型构建方法，通过利用pytorch框架搭建卷积神经网络模型，然后获取文本分类训练数据，利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量；将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到文本分类模型。本申请的文本分类模型采用Pytorch框架，由于Pytorch框架面向对象的接口设计来源于torch，而torch的接口设计具有灵活易用的特点,并且PyTorch框架能够逐层打印出计算结果以便于调试，因此构建的文本分类模型更易于维护和调试。

[39] 本申请附加的方面和优点将在下面的描述中部分给出，这些将从下面的描述中变得明显，或通过本申请的实践了解到。

[40] 附图说明

[41] 图1本申请文本分类模型构建方法一种实施例流程框图；

[42] 图2为本申请文本分类模型构建装置一种实施例模块框图；

[43] 图3为本申请一个实施例中终端的内部结构框图。

[44] 本申请目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

[45] 具体实施方式

[46] 为了使本技术领域的人员更好地理解本申请方案，下面将结合本申请实施例中

的附图，对本申请实施例中的技术方案进行清楚、完整地描述。

[47] 在本申请的说明书和权利要求书及上述附图中的描述的一些流程中，包含了按照特定顺序出现的多个操作，但是应该清楚了解，这些操作可以不按照其在本文中出现的顺序来执行或并行执行，操作的序号如S11、S12等，仅仅是用于区分开各个不同的操作，序号本身不代表任何的执行顺序。另外，这些流程可以包括更多或更少的操作，并且这些操作可以按顺序执行或并行执行。需要说明的是，本文中的“第一”、“第二”等描述，是用于区分不同的消息、设备、模块等，不代表先后顺序，也不限定“第一”和“第二”是不同的类型。

[48] 本领域普通技术人员可以理解，除非特意声明，这里使用的单数形式“一”、“一个”、“所述”和“该”也可包括复数形式。应该进一步理解的是，本申请的说明书中使用的措辞“包括”是指存在所述特征、整数、步骤、操作、元件和/或组件，但是并不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、组件和/或它们的组。应该理解，当我们称元件被“连接”或“耦接”到另一元件时，它可以直接连接或耦接到其他元件，或者也可以存在中间元件。此外，这里使用的“连接”或“耦接”可以包括无线连接或无线耦接。这里使用的措辞“和/或”包括一个或更多个相关联的列出项的全部或任一单元和全部组合。

[49] 本领域普通技术人员可以理解，除非另外定义，这里使用的所有术语（包括技术术语和科学术语），具有与本申请所属领域中的普通技术人员的一般理解相同的意义。还应该理解的是，诸如通用字典中定义的那些术语，应该被理解为具有与现有技术的上下文中的意义一致的意义，并且除非像这里一样被特定定义，否则不会用理想化或过于正式的含义来解释。

[50] 下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。显然，所描述的实施例仅仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

[51] 请参阅图1，本申请提供了一种文本分类模型构建方法，以解决当前采用tensorflow框架构建文本分类模型时，因tensorflow框架代码冗长、接口设计过于晦涩

难懂而导致构建的文本分类模型维护困难、调试不便、不易于操作的问题。其中一种实施方式中，所述文本分类模型构建方法包括如下步骤:

[52] S11、利用pytorch框架搭建卷积神经网络模型；其中，所述卷积神经网络模型设置于嵌入层；

[53] 本实施例的卷积神经网络模型基于pytorch框架搭建，所述pytorch框架是一个python优先的深度学习框架，相对于Tensorflow深度学习框架，其具有更好的灵活性，可以在执行时动态构建或调整计算图，以便于训练过程直接打印隐含变量数值，用于进行调试。而Tensorflow深度学习框架运行时必须提前建好静态计算图，然后通过feed和run重复执行建好的图，因图是静态的，网络结构需要预先建立编译，然后进行训练，训练过程中，每一隐含变量无法直接打印，而是需要重新载入数据进行相关输出，操作不便。所述嵌入层具有降维的作用，输入到卷积神经网络模型的向量往往是高维度数据，比如8000维，嵌入层可以将其降到比如100维度的空间下进行运算，可以在压缩数据的同时让信息损失最小化，从而提高运算效率。

[54] S12、获取文本分类训练数据，利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量；

[55] 在本实施例中，可通过爬虫技术自动抓取互联网信息，从互联网信息中提取出文本分类训练数据，然后利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量。其中，所述word2vec算法基于一个浅层神经网络，其可以在百万数量级的词典和上亿的数据集上进行高效地训练，得到训练结果——词向量，并可以良好地度量词与词之间的相似性。

[56] S13、将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到文本分类模型。

[57] 本实施例将训练得到的词向量输入预先搭建好的卷积神经网络模型中，以对卷积神经网络模型进行分类训练，直至其收敛时，也即训练结果满足要求时，得到训练合格的文本分类模型，以后续用于对文本数据进行分类，如新闻标题分类、评论情感分类等。需要说明的是，输入卷积神经网络模型的词向量越多，则训练得到的文本分类模型的分类准确性越高。

[58] 本申请提供的文本分类模型构建方法，通过利用pytorch框架搭建卷积神经网络模型，然后获取文本分类训练数据，利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量；最后将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到训练合格的文本分类模型，以提高文本分类模型的分类准确率；同时，由于本申请的文本分类模型采用Pytorch框架，Pytorch框架面向对象的接口设计来源于torch，而torch的接口设计具有灵活易用的特点,并且PyTorch框架能够逐层打印出计算结果以便于调试，因此构建的文本分类模型更易于维护和调试。

[59] 在一实施例中，步骤S12的利用Word2Vec算法将所述文本训练数据进行词向量训练之前，还可包括：

[60] 根据正则表达式匹配规则去除文本训练数据的停用词和符号；

[61] 在本实施例中，正则表达式是由一些具有特殊含义的字符组成的字符串，多用于查找、替换符合规则的字符串。所述正则表达式匹配规则可以对字符串进行操作，以简化对字符串的复杂操作，其主要功能有匹配、切割、替换、获取。本实施例可以利用正则表达式匹配规则去除文本训练数据的停用词和符号，如文本中标点符号的删除，以得到有效文本训练数据。

[62] 利用结巴分词库将去除停用词和符号的文本训练数据进行中文分词。

[63] 在本实施例中，可根据结巴分词库中字词的搭配频率对文本训练数据进行中文分词，经过中文分词后的文本训练数据，在利用Word2Vec算法对其进行词向量训练时，训练效率更高，训练结果更佳。例如，当对“燕子去了，有再来的时候；杨柳枯了，有再青的时候；桃花谢了，有再开的时候”这段文本进行分词时，可根据结巴分词库中字词的搭配频率分为“燕子/去了，有/再来/的/时候；杨柳/枯了，有/再/青/的/时候；桃花/谢了，有/再开/的/时候”。当然，所述文本训练数据还可通过其他方式进行分词，在此不做具体限定。

[64] 在一实施例中，所述利用结巴分词库将去除停用词和符号的文本训练数据进行中文分词的步骤，包括：

[65] 利用结巴分词库确定文本训练数据中汉字之间的相关度；

[66] 将相关度大于预设值的汉字组成分词，得到分词结果。

[67] 在本实施例中，可通过计算文本训练数据中相邻汉字之间的相关度，并将相关度高的汉字组成分词，从而得到分词结果，以提高分词的准确性。例如，在对“枯藤老树昏鸦，小桥流水人家”这段文本进行分词时，根据结巴分词库确定“枯”和“藤”的相关度比“藤”和“老”的相关度要高，因此可以组成分词“枯藤”，依此类推，得到这段文本的分词结果为“枯藤”、“老树”、“昏鸦”、“小桥”、“流水”、“人家”。其中，所述相关度的预设值可根据需要灵活调整。

[68] 在一实施例中，步骤S13的将所述词向量输入卷积神经网络模型进行分类训练的步骤，包括：

[69] 根据所述词向量并通过交叉熵损失函数和ADAM优化算法对卷积神经网络模型进行分类训练。

[70] 在本实施例中，所述交叉熵损失函数可用来评估当前训练得到的概率分布与真实分布的差异情况，以了解文本分类模型的分类准确率，以对文本分类模型的相关参数实时调整，直至训练合格。其中，ADAM优化算法是在带动量的梯度下降法的基础上融合了一种加速梯度下降算法而形成的。相较于带动量的梯度下降法，所述ADAM优化算法可对分类训练结果进行偏差纠正，以提高分类准确性。

[71] 在一实施例中，步骤S12的利用Word2Vec算法将所述文本训练数据进行词向量训练的步骤，可具体包括：

[72] 利用Word2Vec算法对大型语料数据进行词向量训练，得到词向量字典；

[73] 本实施例可通过Word2Vec算法将大型语料数据进行词向量训练，得到词向字典量。这一步骤可通过Python中的gensim库实现，Gensim库是一个基于python的自然语言处理库，能够利用TF-IDF、 LDA,、LSI 等模型将文本转化成向量模式，以便进行进一步的处理。

[74] 根据所述词向量字典将文本训练数据转化为词向量。

[75] 本实施例可通过训练得到的词向量字典将文本训练数据转化为词向量，其中，文本训练数据中每一个字词在词向量字典中都有相应的词向量，从而得到文本训练数据中所有字词的词向量。

[76] 在一实施例中，步骤S11的利用pytorch框架搭建卷积神经网络模型之后，还可

包括：

[77] 在所述卷积神经网络模型上建立位置注意力机制和通道注意力机制；其中，所述位置注意力机制和通道注意力机制的输入与卷积神经网络模型的激活层的输出连接，所述位置注意力机制和通道注意力机制的输出与卷积神经网络模型的全连接层的输入连接。

[78] 在本实施例中，位置注意力机制与通道注意力机制的输入来源于卷积神经网络的激活层输出。卷积神经网络模型的输出可以为384*100*1的三维矩阵，对于位置注意力机制，可先将卷积神经网络模型的输出三维矩阵转化为384*100的矩阵，并通过两个并行全连接层输出100*384与384*100的矩阵，然后进行矩阵乘法及softmax映射，得到100*100的位置注意力矩阵。在此基础上，通过另一并行全连接层输出384*100的矩阵与位置注意力矩阵进行矩阵乘法，得到384*100的矩阵并将其转化为384*100*1的三维矩阵，并与卷积神经网络模型的输出加和，得到位置注意力机制的输出结果。

[79] 对于通道注意力，可首先将卷积神经网络模型的输出三维矩阵转化为384*100的矩阵，并通过两个并行全连接层输出384*100与100*384的矩阵，然后进行矩阵乘法及softmax映射，得到384*384的通道注意力矩阵。在此基础上通过另一并行全连接层输出100*384的矩阵与通道注意力矩阵进行矩阵乘法，得到100*384的矩阵并将其转化为384*100*1的三维矩阵，并与卷积神经网络模型输出加和，得到通道注意力机制输出结果。最后将位置注意力机制和通道注意力机制输出，并输入全连接层，从而完成整个卷积神经网络模型的输出。

[80] 其中，所述全连接层在整个卷积神经网络模型中起到“分类器”的作用。如果说卷积神经网络模型的卷积层、池化层和激活层等操作是将原始数据映射到隐层特征空间的话，全连接层则起到将学到的“分布式特征表示”映射到样本标记空间的作用。

[81] 在一实施例中，卷积神经网络模型的卷积层含高度为1、3、5的一维卷积和128条通道（通过padding实现卷积层输入输出维度一致），激活层函数可以为ReLU，其收敛更快，并且能保持同样效果。

[82] 在一实施例中，步骤S13的将所述词向量输入卷积神经网络模型进行分类训练

，直至收敛时，得到文本分类模型的步骤，可具体包括：

[83] 根据分类训练结果计算卷积神经网络模型的分类准确率；

[84] 当分类准确率低于预设值时，调整卷积神经网络模型的参数，利用词向量对所述卷积神经网络模型重新训练，直至收敛时，得到文本分类模型。

[85] 本实施例对卷积神经网络模型的分类准确率进行计算，并判断卷积神经网络模型的分类准确率是否低于预设值，若是，则调整卷积神经网络模型的参数，利用词向量对所述卷积神经网络模型重新训练，直至卷积神经网络模型的分类准确率高于预设值时，得到训练合格的文本分类模型，从而保证训练得到的文本分类模型有较好的分类效果。

[86] 请参考图2，本申请的实施例还提供了一种文本分类模型构建装置，一种本实施例中，包括搭建模块21、获取模块22和训练模块23。其中，

[87] 搭建模块21，用于利用pytorch框架搭建卷积神经网络模型；其中，所述卷积神经网络模型设置于嵌入层；

[88] 本实施例的卷积神经网络模型基于pytorch框架搭建，所述pytorch框架是一个python优先的深度学习框架，相对于Tensorflow深度学习框架，其具有更好的灵活性，可以在执行时动态构建或调整计算图，以便于训练过程直接打印隐含变量数值，用于进行调试。而Tensorflow深度学习框架运行时必须提前建好静态计算图，然后通过feed和run重复执行建好的图，因图是静态的，网络结构需要预先建立编译，然后进行训练，训练过程中，每一隐含变量无法直接打印，而是需要重新载入数据进行相关输出，操作不便。所述嵌入层具有降维的作用，输入到卷积神经网络模型的向量往往是高维度数据，比如8000维，嵌入层可以将其降到比如100维度的空间下进行运算，可以在压缩数据的同时让信息损失最小化，从而提高运算效率。

[89] 获取模块22，用于获取文本分类训练数据，利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量；

[90] 在本实施例中，可通过爬虫技术自动抓取互联网信息，从互联网信息中提取出文本分类训练数据，然后利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量。其中，所述word2vec算法基于一个浅层神经网络，其可以在百

万数量级的词典和上亿的数据集上进行高效地训练，得到训练结果——词向量，并可以良好地度量词与词之间的相似性。

[91] 训练模块23，用于将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到文本分类模型。

[92] 本实施例将训练得到的词向量输入预先搭建好的卷积神经网络模型中，以对卷积神经网络模型进行分类训练，直至其收敛时，也即训练结果满足要求时，得到训练合格的文本分类模型，以后续用于对文本数据进行分类，如新闻标题分类、评论情感分类等。需要说明的是，输入卷积神经网络模型的词向量越多，则训练得到的文本分类模型的分类准确性越高。

[93] 本申请提供的文本分类模型构建装置，通过利用pytorch框架搭建卷积神经网络模型，然后获取文本分类训练数据，利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量；最后将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到训练合格的文本分类模型，以提高文本分类模型的分类准确率；同时，由于本申请的文本分类模型采用Pytorch框架，Pytorch框架面向对象的接口设计来源于torch，而torch的接口设计具有灵活易用的特点,并且PyTorch框架能够逐层打印出计算结果以便于调试，因此构建的文本分类模型更易于维护和调试。

[94] 关于上述实施例中的装置，其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述，此处将不做详细阐述说明。

[95] 本申请提供的一种终端，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行如上任一项所述的文本分类模型构建方法的步骤。

[96] 在一实施例中，所述终端为一种计算机设备，如图3所示。本实施例所述的计算机设备可以是服务器、个人计算机以及网络设备等设备。所述计算机设备包括处理器302、存储器303、输入单元304以及显示单元305等器件。本领域技术人员可以理解，图3示出的设备结构器件并不构成对所有设备的限定，可以包括比图示更多或更少的部件，或者组合某些部件。存储器303可用于存储计算机可读指令301以及各功能模块，处理器302运行存储在存储器303的计算机可读指令

301，从而执行设备的各种功能应用以及数据处理。存储器可以是内存储器或外存储器，或者包括内存储器和外存储器两者。内存储器可以包括只读存储器(ROM)、可编程ROM(PROM)、电可编程ROM(EPROM)、电可擦写可编程ROM(EEPROM)、快闪存储器、或者随机存储器。外存储器可以包括硬盘、软盘、ZIP盘、U盘、磁带等。本申请所公开的存储器包括但不限于这些类型的存储器。本申请所公开的存储器只作为例子而非作为限定。

[97] 输入单元304用于接收信号的输入，以及接收用户输入的关键字。输入单元304可包括触控面板以及其它输入设备。触控面板可收集用户在其上或附近的触摸操作（比如用户使用手指、触笔等任何适合的物体或附件在触控面板上或在触控面板附近的操作），并根据预先设定的程序驱动相应的连接装置；其它输入设备可以包括但不限于物理键盘、功能键（比如播放控制按键、开关按键等）、轨迹球、鼠标、操作杆等中的一种或多种。显示单元305可用于显示用户输入的信息或提供给用户的信息以及计算机设备的各种菜单。显示单元305可采用液晶显示器、有机发光二极管等形式。处理器302是计算机设备的控制中心，利用各种接口和线路连接整个电脑的各个部分，通过运行或执行存储在存储器302内的软件程序和/或模块，以及调用存储在存储器内的数据，执行各种功能和处理数据。

[98] 作为一个实施例，所述计算机设备包括：一个或多个处理器302，存储器303，一个或多个计算机可读指令301，其中所述一个或多个计算机可读指令301被存储在存储器303中并被配置为由所述一个或多个处理器302执行，所述一个或多个计算机可读指令301配置用于执行以上实施例所述的文本分类模型构建方法。

[99] 在一个实施例中，本申请还提出了一种存储有计算机可读指令的存储介质，该计算机可读指令的存储介质可以为非易失性可读存储介质，该计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器执行上述文本分类模型构建方法。例如，所述存储介质可以是ROM、随机存取存储器（RAM）、CD-ROM、磁带、软盘和光数据存储设备等。

[100] 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，是可以通过计算机可读指令来指令相关的硬件来完成，该计算机可读指令可存储于

一存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，前述的存储介质可为磁碟、光盘、只读存储记忆体（Read-Only Memory，ROM）等非易失性存储介质，或随机存储记忆体（Random Access Memory，RAM）等。

[101] 综合上述实施例可知，本申请在于：

[102] 本申请提供的文本分类模型构建方法、装置、终端及存储介质，通过利用pytorch框架搭建卷积神经网络模型，然后获取文本分类训练数据，利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量；最后将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到训练合格的文本分类模型，以提高文本分类模型的分类准确率；同时，由于本申请的文本分类模型采用Pytorch框架，Pytorch框架面向对象的接口设计来源于torch，而torch的接口设计具有灵活易用的特点,并且PyTorch框架能够逐层打印出计算结果以便于调试，因此构建的文本分类模型更易于维护和调试。

[103] 以上所述实施例的各技术特征可以进行任意的组合，为使描述简洁，未对上述实施例中的各个技术特征所有可能的组合都进行描述，然而，只要这些技术特征的组合不存在矛盾，都应当认为是本说明书记载的范围。

[104] 以上所述实施例仅表达了本申请的几种实施方式，其描述较为具体和详细，但并不能因此而理解为对本申请专利范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本申请构思的前提下，还可以做出若干变形和改进，这些都属于本申请的保护范围。因此，本申请专利的保护范围应以所附权利要求为准。

权利要求书

[权利要求 1] 一种文本分类模型构建方法，其中，包括：

利用pytorch框架搭建卷积神经网络模型；其中，所述卷积神经网络模型设置于嵌入层，在所述卷积神经网络模型上建立位置注意力机制和通道注意力机制；其中，所述位置注意力机制和通道注意力机制的输入与卷积神经网络模型的激活层的输出连接，所述位置注意力机制和通道注意力机制的输出与卷积神经网络模型的全连接层的输入连接；

获取文本分类训练数据，利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量；

将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到文本分类模型。

[权利要求 2] 根据权利要求1所述的文本分类模型构建方法，其中，所述利用Word2Vec算法将所述文本训练数据进行词向量训练之前，还包括：

根据正则表达式匹配规则去除文本训练数据的停用词和符号；

利用结巴分词库将去除停用词和符号的文本训练数据进行中文分词。

[权利要求 3] 根据权利要求2所述的文本分类模型构建方法，其中，所述利用结巴分词库将去除停用词和符号的文本训练数据进行中文分词的步骤，包括:

利用结巴分词库确定文本训练数据中汉字之间的相关度；

将相关度大于预设值的汉字组成分词，得到分词结果。

[权利要求 4] 根据权利要求1所述的文本分类模型构建方法，其中，所述将所述词向量输入卷积神经网络模型进行分类训练的步骤，包括：

根据所述词向量并通过交叉熵损失函数和ADAM优化算法对卷积神经网络模型进行分类训练。

[权利要求 5] 根据权利要求1所述的文本分类模型构建方法，其中，所述利用Word2Vec算法将所述文本训练数据进行词向量训练的步骤，包括：

利用Word2Vec算法对大型语料数据进行词向量训练，得到词向量字典；

根据所述词向量字典将文本训练数据转化为词向量。

[权利要求 6] 根据权利要求1所述的文本分类模型构建方法，其中，所述将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到文本分类模型的步骤，包括：

根据分类训练结果计算卷积神经网络模型的分类准确率；

当分类准确率低于预设值时，调整卷积神经网络模型的参数，利用词向量对所述卷积神经网络模型重新训练，直至收敛时，得到文本分类模型。

[权利要求 7] 一种文本分类模型构建装置，其中，包括：

搭建模块，用于利用pytorch框架搭建卷积神经网络模型；其中，所述卷积神经网络模型设置于嵌入层，在所述卷积神经网络模型上建立位置注意力机制和通道注意力机制；其中，所述位置注意力机制和通道注意力机制的输入与卷积神经网络模型的激活层的输出连接，所述位置注意力机制和通道注意力机制的输出与卷积神经网络模型的全连接层的输入连接；

获取模块，用于获取文本分类训练数据，利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量；

训练模块，用于将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到文本分类模型。

[权利要求 8] 如权利要求7所述的文本分类模型构建装置，其中，所述的文本分类模型构建装置还包括：

去除模块，用于根据正则表达式匹配规则去除文本训练数据的停用词和符号；

分词模块，用于利用结巴分词库将去除停用词和符号的文本训练数据进行中文分词。

[权利要求 9] 如权利要求8所述的文本分类模型构建装置，其中，所述的分词模块还用于：

利用结巴分词库确定文本训练数据中汉字之间的相关度；

将相关度大于预设值的汉字组成分词，得到分词结果。

[权利要求 10] 如权利要求7所述的文本分类模型构建装置，其中，所述的训练模块还用于：

根据所述词向量并通过交叉熵损失函数和ADAM优化算法对卷积神经网络模型进行分类训练。

[权利要求 11] 如权利要求7所述的文本分类模型构建装置，其中，所述的获取模块还用于：

利用Word2Vec算法对大型语料数据进行词向量训练，得到词向量字典；

根据所述词向量字典将文本训练数据转化为词向量。

[权利要求 12] 如权利要求7所述的文本分类模型构建装置，其中，所述的训练模块还用于：

根据分类训练结果计算卷积神经网络模型的分类准确率；

当分类准确率低于预设值时，调整卷积神经网络模型的参数，利用词向量对所述卷积神经网络模型重新训练，直至收敛时，得到文本分类模型。

[权利要求 13] 一种终端，其中，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，使得所述处理器执行以下步骤：

利用pytorch框架搭建卷积神经网络模型；其中，所述卷积神经网络模型设置于嵌入层，在所述卷积神经网络模型上建立位置注意力机制和通道注意力机制；其中，所述位置注意力机制和通道注意力机制的输入与卷积神经网络模型的激活层的输出连接，所述位置注意力机制和通道注意力机制的输出与卷积神经网络模型的全连接层的输入连接；

获取文本分类训练数据，利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量；

将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到文本分类模型。

[权利要求 14] 如权利要求13所述的终端，其中，所述利用Word2Vec算法将所述文本训练数据进行词向量训练之前，还包括：

根据正则表达式匹配规则去除文本训练数据的停用词和符号；

利用结巴分词库将去除停用词和符号的文本训练数据进行中文分词。

[权利要求 15] 如权利要求14所述的终端，其中，所述利用结巴分词库将去除停用词和符号的文本训练数据进行中文分词的步骤，包括:

利用结巴分词库确定文本训练数据中汉字之间的相关度；

将相关度大于预设值的汉字组成分词，得到分词结果。

[权利要求 16] 如权利要求13所述的终端，其中，所述将所述词向量输入卷积神经网络模型进行分类训练的步骤，包括：

根据所述词向量并通过交叉熵损失函数和ADAM优化算法对卷积神经网络模型进行分类训练。

[权利要求 17] 如权利要求13所述的终端，其中，所述利用Word2Vec算法将所述文本训练数据进行词向量训练的步骤，包括：

利用Word2Vec算法对大型语料数据进行词向量训练，得到词向量字典；

根据所述词向量字典将文本训练数据转化为词向量。

[权利要求 18] 如权利要求13所述的终端，其中，所述将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到文本分类模型的步骤，包括：

根据分类训练结果计算卷积神经网络模型的分类准确率；

当分类准确率低于预设值时，调整卷积神经网络模型的参数，利用词向量对所述卷积神经网络模型重新训练，直至收敛时，得到文本分类模型。

[权利要求 19] 一种存储介质，其上存储有计算机可读指令，其中，所述计算机可读指令被处理器执行时，实现以下步骤：

利用pytorch框架搭建卷积神经网络模型；其中，所述卷积神经网络模型设置于嵌入层，在所述卷积神经网络模型上建立位置注意力机制和

通道注意力机制；其中，所述位置注意力机制和通道注意力机制的输入与卷积神经网络模型的激活层的输出连接，所述位置注意力机制和通道注意力机制的输出与卷积神经网络模型的全连接层的输入连接；

获取文本分类训练数据，利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量；

将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到文本分类模型。

[权利要求 20] 如权利要求19所述的存储介质，其中，所述利用Word2Vec算法将所述文本训练数据进行词向量训练之前，还包括：

根据正则表达式匹配规则去除文本训练数据的停用词和符号；

利用结巴分词库将去除停用词和符号的文本训练数据进行中文分词。

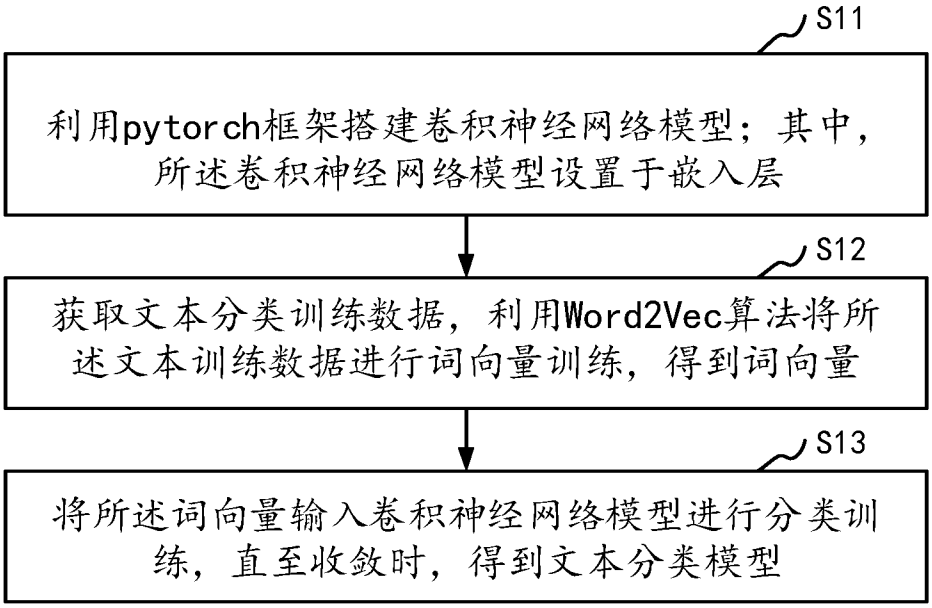
S11
利用pytorch框架搭建卷积神经网络模型；其中，所述卷积神经网络模型设置于嵌入层
S12
获取文本分类训练数据，利用Word2Vec算法将所述文本训练数据进行词向量训练，得到词向量
S13
将所述词向量输入卷积神经网络模型进行分类训练，直至收敛时，得到文本分类模型

图 1

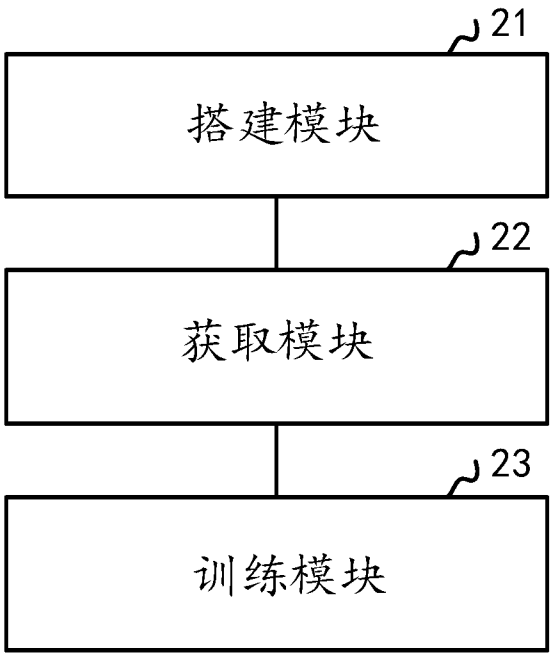
21
搭建模块
22
获取模块
23
训练模块

图 2

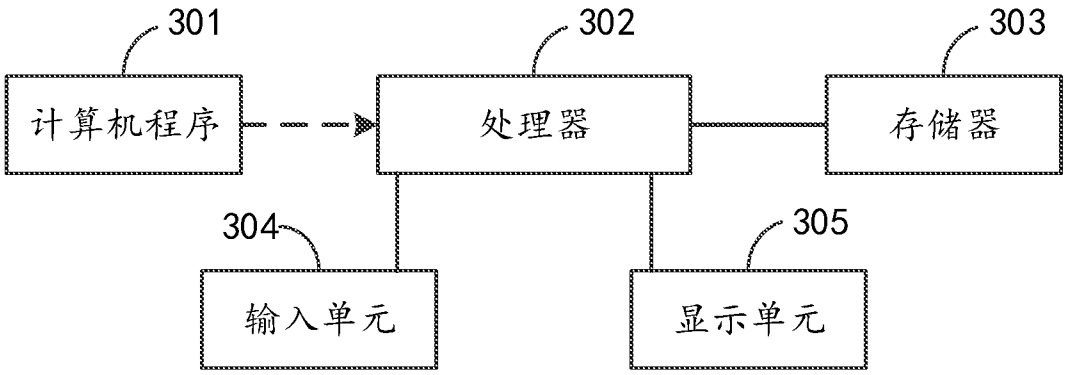

图 3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2019/117225

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/35(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, WPI, EPODOC, CNKI, IEEE: 文本, 分类, 神经网络, 嵌入层, 激活层, 全连接层, 向量, 训练, 模型, text, classification, model, CNN, pytorch, layer, training, vector

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109960726 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 02 July 2019 (2019-07-02) claims 1-20, description, paragraphs [0044]-[0095], and figures 1-3	1-20
Y	CN 108717439 A (HARBIN UNIVERSITY OF SCIENCE AND TECHNOLOGY) 30 October 2018 (2018-10-30) description, paragraphs [0031]-[0096], and figures 1 and 2	1-20
Y	CN 108520535 A (TIANJIN UNIVERSITY) 11 September 2018 (2018-09-11) description, paragraphs [0055] and [0056]	1-20
A	US 2018181864 A1 (TEXAS INSTRUMENTS INCORPORATED) 28 June 2018 (2018-06-28) entire document	1-20

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 December 2019	**27 February 2020**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088** **China**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2019/117225

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 109960726 A	02 July 2019	None	
CN 108717439 A	30 October 2018	None	
CN 108520535 A	11 September 2018	None	
US 2018181864 A1	28 June 2018	US 2018181857 A1	28 June 2018

国际检索报告

国际申请号
PCT/CN2019/117225

A. 主题的分类

G06F 16/35(2019.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

CNPAT, WPI, EPODOC, CNKI, IEEE: 文本，分类，神经网络，嵌入层，激活层，全连接层，向量，训练，模型，text, classification, model, CNN, pytorch, layer, training, vector

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
PX	CN 109960726 A (平安科技深圳有限公司) 2019年 7月 2日 (2019-07-02) 权利要求1-20，说明书第[0044]-[0095]段，附图1-3	1-20
Y	CN 108717439 A (哈尔滨理工大学) 2018年 10月 30日 (2018-10-30) 说明书第[0031]-[0096]段，附图1-2	1-20
Y	CN 108520535 A (天津大学) 2018年 9月 11日 (2018-09-11) 说明书第[0055]-[0056]段	1-20
A	US 2018181864 A1 (TEXAS INSTRUMENTS INCORPORATED) 2018年 6月 28日 (2018-06-28) 全文	1-20

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2019年 12月 19日	2020年 2月 27日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 周亚楠 电话号码 86-(10)-53961530

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2019/117225

检索报告引用的专利文件	公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN 109960726 A	2019年 7月 2日	无	
CN 108717439 A	2018年 10月 30日	无	
CN 108520535 A	2018年 9月 11日	无	
US 2018181864 A1	2018年 6月 28日	US 2018181857 A1	2018年 6月 28日