

(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
2022 年 7 月 7 日 (07.07.2022)



(10) 国际公布号
WO 2022/142010 A1

(51) 国际专利分类号:
G06F 16/35 (2019.01)

(21) 国际申请号: PCT/CN2021/090432

(22) 国际申请日: 2021 年 4 月 28 日 (28.04.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202011617671.4 2020 年 12 月 30 日 (30.12.2020) 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: 陈昊(CHEN, Hao); 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: 深圳市世联合知识产权代理有限公司(SL INTELLECTUAL PROPERTY CO., LTD.);
中国广东省深圳市福田区沙头街道天安社区泰然四路 29 号天安创新科技广场一期 A 座 805E, Guangdong 518000 (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, IT, 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,

(54) Title: METHOD AND APPARATUS FOR CLASSIFYING UNBALANCED SAMPLES, DEVICE AND STORAGE MEDIUM

(54) 发明名称: 一种非均衡样本的分类方法、装置、设备及存储介质

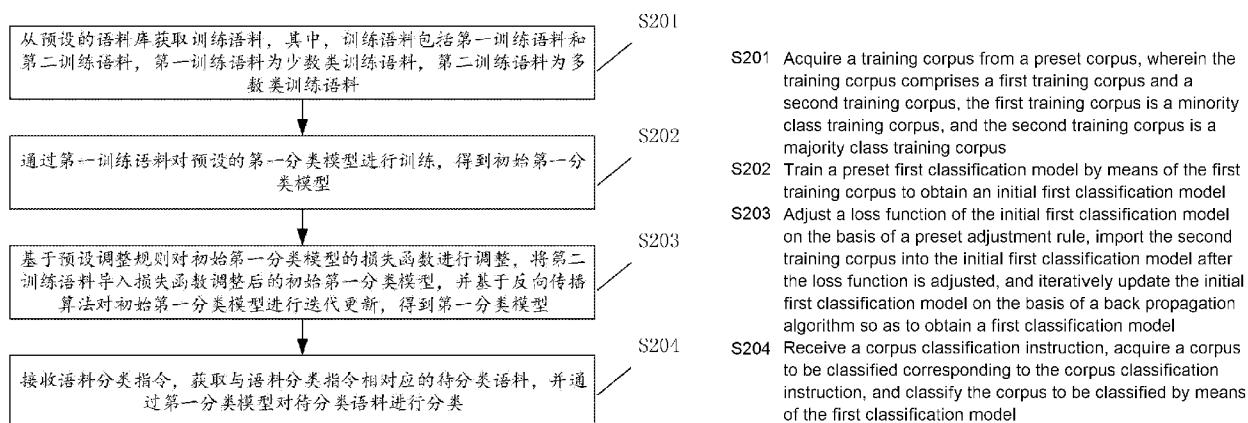


图 2

(57) Abstract: A method and an apparatus for classifying unbalanced samples, a device and a storage medium, which belong to the machine learning technology in the technical field of artificial intelligence. Said method comprises: acquiring a training corpus, wherein the training corpus comprises a first training corpus and a second training corpus, the first training corpus is a minority class training corpus, and the second training corpus is a majority class training corpus; performing training by means of the first training corpus to obtain an initial first classification model; performing inversion by means of a loss function, then importing the second training corpus into the initial first classification model, and performing iteration processing on the initial first classification model to obtain a first classification model; and acquiring a corpus to be classified, and classifying the corpus to be classified by means of the first classification model. In addition, the method also relates to the blockchain technology, and the corpus to be classified can be stored in a blockchain. According to the method, the loss function of a classification model is inverted by means of a preset adjustment rule, and then an exclusive classification model is obtained by means of training, increasing the precision of classification models of unbalanced samples.

ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, ZA, ZM, ZW。

- (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: 一种非均衡样本的分类方法、装置、设备和存储介质, 属于人工智能技术领域中的机器学习技术, 该方法通过获取训练语料, 其中, 训练语料包括第一训练语料和第二训练语料, 第一训练语料为少数类训练语料, 第二训练语料为多数类训练语料, 通过第一训练语料训练得到初始第一分类模型, 通过损失函数进行反转, 然后将第二训练语料导入初始第一分类模型, 并对初始第一分类模型进行迭代, 得到第一分类模型, 获取待分类语料, 并通过第一分类模型对待分类语料进行分类。此外, 该方法还涉及区块链技术, 待分类语料可存储于区块链中。该方法通过预设调整规则对分类模型的损失函数进行反转, 然后训练得到具有排他性的分类模型, 提高了非均衡样本分类模型的精度。

一种非均衡样本的分类方法、装置、设备及存储介质

本申请要求于 2020 年 12 月 30 日提交中国专利局、申请号为 202011617671.4，发明名称为“一种非均衡样本的分类方法、装置、设备及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请属于人工智能技术领域，具体涉及一种非均衡样本的分类方法、装置、设备和存储介质。

背景技术

在机器学习领域中，绝大多数模型是基于正常分布的数据集训练得到的，适用于有偏数据集的模型设计以及获取方法非常少。但是在某些特定场景下更多的是有偏数据集的情况，比如在自然语言分类的领域中，有侮辱攻击性的语言占整体数据的比例非常少，但是在实际业务场景中，又必须对侮辱攻击性的语言进行准确的分类。

针对有偏数据集的分类模型设计以及训练一直是学界与业界的难点，当前较为常用的方法是，在模型训练的损失函数上做一定的改进，即对于某些数量很少的训练样本类别加入额外的权重，相当于强制模型更加关注这些类别的训练样本。由于这些类别的训练样本数量较少，通过设置相应权重，总体说来，可以获得一个较为稳健的分类模型，使得该分类模型在分类任务获得一个较好的表现。但是，在非均衡样本的分类过程中，发明人意识到样本分类是多种多样的，使得样本权重值计算极其复杂，我们很难去平衡各个种类样本之间的权重。而且针对特征特别相近的训练样本，上述通过设置权重的方式训练的得到分类模型难以对其进行分类，即通过设置权重的方式来实现有偏数据集的分类在很多场景下分类精度达不到要求。

发明内容

本申请实施例的目的在于提出一种非均衡样本的分类方法、装置、计算机设备及存储介质，以解决现有的有偏数据集的分类方式中分类过程复杂，以及分类精度达不到要求要求的技术问题。

为了解决上述技术问题，本申请实施例提供一种非均衡样本的分类方法，采用了如下所述的技术方案：

一种非均衡样本的分类方法，包括：

从预设的语料库获取训练语料，其中，训练语料包括第一训练语料和第二训练语料，第一训练语料为少数类训练语料，第二训练语料为多数类训练语料；

通过第一训练语料对预设的第一分类模型进行训练，得到初始第一分类模型；

基于预设调整规则对初始第一分类模型的损失函数进行调整，将第二训练语料导入损失函数调整后的初始第一分类模型，并基于反向传播算法对初始第一分类模型进行迭代更新，得到第一分类模型；

接收语料分类指令，获取与语料分类指令相对应的待分类语料，并通过第一分类模型对待分类语料进行分类。

进一步地，训练语料的语种类型为第一语种，在从预设的语料库获取训练语料的步骤之后，还包括：

计算第一训练语料和第二训练语料的语料相似度，并比对语料相似度与预设相似度阈值的大小；

若语料相似度大于或等于预设相似度阈值，则将训练语料翻译为第二语种，其中，在第二语种的语种环境下，第一训练语料和第二训练语料的语料相似度小于预设相似度阈值。

进一步地，预设的第一分类模型为包括编码层和解码层，通过第一训练语料对预设的第一分类模型进行训练，得到初始第一分类模型的步骤，具体包括：

通过编码层提取第一训练语料的语料特征，并对语料特征进行矢量编码，得到特征矢量；

将特征矢量与预先存储在解码层中的特征标签进行特征映射，获取特征映射结果；

基于特征映射结果对预设的第一分类模型进行迭代，得到初始第一分类模型。

进一步地，基于特征映射结果对预设的第一分类模型进行迭代，得到初始第一分类模型的步骤，具体包括：

构建初始第一分类模型的损失函数，得到第一损失函数，其中，第一损失函数包含对抗因子；

基于第一损失函数计算特征映射结果与预设映射结果的误差，得到映射误差；

基于映射误差和反向传播算法对预设的第一分类模型进行迭代，得到初始第一分类模型。

进一步地，基于预设调整规则对初始第一分类模型的损失函数进行调整，将第二训练语料导入损失函数调整后的初始第一分类模型，并基于反向传播算法对初始第一分类模型进行迭代更新，得到第一分类模型的步骤，具体包括：

基于预设调整规则对初始第一分类模型的损失函数进行反转，得到第二损失函数；

将第二训练语料导入损失函数反转后的初始第一分类模型，获取分类结果；

基于第二损失函数计算分类结果与预设分类结果的误差，得到分类误差；

将分类误差与预设分类误差阈值进行比较，若分类误差小于或等于预设分类误差阈值，则通过反向传播算法对损失函数反转后的初始第一分类模型进行迭代更新，直到分类误差大于预设分类误差阈值为止；

输出分类误差大于预设分类误差阈值的第一分类模型。

进一步地，在基于预设调整规则对初始第一分类模型的损失函数进行调整，将第二训练语料导入损失函数调整后的初始第一分类模型，并基于反向传播算法对初始第一分类模型进行迭代更新，得到第一分类模型的步骤之后，还包括：

对第一训练语料进行矢量化处理，得到第一训练语料的语料矢量；

对语料矢量和特征矢量进行矢量拼接，得到语料特征矩阵；

将语料特征矩阵导入预设的第二分类模型，通过第二分类模型的卷积核对语料特征矩阵进行卷积运算，获取卷积运算结果；

基于卷积运算结果对第二分类模型进行迭代更新，输出训练完成的第二分类模型。

进一步地，基于卷积运算结果对第二分类模型进行迭代更新，输出训练完成的第二分类模型的步骤，具体包括：

基于交叉熵损失函数和莱温斯坦距离函数构建第二分类模型的损失函数，得到初始第三损失函数，其中，初始第三损失函数包括交叉熵因子和莱温斯坦因子；

为初始第三损失函数的交叉熵因子和莱温斯坦因子分别赋予相同的初始权重值；

基于卷积运算结果对交叉熵因子和莱温斯坦因子的初始权重值进行调整，直至初始第三损失函数的输出达到最小值，得到第三损失函数；

通过第三损失函数计算卷积运算结果的运算误差，基于运算误差采用反向传播算法对第二分类模型进行迭代更新，得到训练完成的第二分类模型。

为了解决上述技术问题，本申请实施例还提供一种非均衡样本的分类装置，采用了如下所述的技术方案：

一种非均衡样本的分类装置，包括：

语料获取模块，用于从预设的语料库获取训练语料，其中，训练语料包括第一训练语料和第二训练语料，第一训练语料为少数类训练语料，第二训练语料为多数类训练语料；

第一模型训练模块，用于通过第一训练语料对预设的第一分类模型进行训练，得到初始第一分类模型；

反转训练模块，用于基于预设调整规则对初始第一分类模型的损失函数进行调整，将第二训练语料导入损失函数调整后的初始第一分类模型，并基于反向传播算法对初始第一分类模型进行迭代更新，得到第一分类模型；

语料分类模块，用于接收语料分类指令，获取与语料分类指令相对应的待分类语料，并通过第一分类模型对待分类语料进行分类。

为了解决上述技术问题，本申请实施例还提供一种计算机设备，采用了如下所述的技术方案：

一种设备，其特征在于，包括存储器和处理器，存储器中存储有计算机可读指令，处理器执行所述计算机可读指令时实现如上述的非均衡样本的分类方法的步骤。

为了解决上述技术问题，本申请实施例还提供一种计算机可读存储介质，采用了如下所述的技术方案：

一种计算机可读存储介质，其特征在于，计算机可读存储介质上存储有计算机可读指令，计算机可读指令被处理器执行时实现如上述所述的非均衡样本的分类方法的步骤。

与现有技术相比，本申请实施例主要有以下有益效果：

本申请公开了一种非均衡样本的分类方法、装置、设备和存储介质，属于人工智能技术领域，本申请通过获取训练语料，其中，训练语料包括第一训练语料和第二训练语料，第一训练语料为少数类训练语料，第二训练语料为多数类训练语料，通过少数类训练语料训练得到初始第一分类模型，通过预设调整规则对初始第一分类模型的损失函数进行反转，然后将多数类训练语料导入损失函数反转后的初始第一分类模型，并对初始第一分类模型进行迭代，得到第一分类模型，获取待分类语料，并通过第一分类模型对待分类语料进行分类。本申请通过预设调整规则对分类模型的损失函数进行反转，然后训练得到一个具有排他性的分类模型，在进行分类任务时，该分类模型仅能对少数类训练语料对应数据类型做出相应响应和输出，而对于多数类训练语料对应数据类型则无法响应，提高了非均衡样本分类模型的精度。

附图说明

为了更清楚地说明本申请中的方案，下面将对本申请实施例描述中所需要使用的附图作一个简单介绍，显而易见地，下面描述中的附图是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图1示出了本申请可以应用于其中的示例性系统架构图；

图2示出了根据本申请的非均衡样本的分类方法的一个实施例的流程图；

图3示出了根据本申请的非均衡样本的分类装置的一个实施例的结构示意图；

图4示出了根据本申请的计算机设备的一个实施例的结构示意图。

具体实施方式

除非另有定义，本文所使用的所有的技术和科学术语与属于本申请的技术领域的技术人员通常理解的含义相同；本文中在申请的说明书中所使用的术语只是为了描述具体的实施例的目的，不是旨在于限制本申请；本申请的说明书和权利要求书及上述附图说明中的术语“包括”和“具有”以及它们的任何变形，意图在于覆盖不排他的包含。本申请的说明书和权利要求书或上述附图中的术语“第一”、“第二”等是用于区别不同对象，而不是用于描述特定顺序。

在本文中提及“实施例”意味着，结合实施例描述的特定特征、结构或特性可以包含在本申请的至少一个实施例中。在说明书中的各个位置出现该短语并不一定均是指相同的实施例，也不是与其它实施例互斥的独立的或备选的实施例。本领域技术人员显式地和隐式地理解的是，本文所描述的实施例可以与其它实施例相结合。

为了使本技术领域的人员更好地理解本申请方案，下面将结合附图，对本申请实施例中的技术方案进行清楚、完整地描述。

如图1所示，系统架构100可以包括终端设备101、102、103，网络104和服务器105。网络104用以在终端设备101、102、103和服务器105之间提供通信链路的介质。网络104可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

用户可以使用终端设备101、102、103通过网络104与服务器105交互，以接收或发送消息等。终端设备101、102、103上可以安装有各种通讯客户端应用，例如网页浏览器应用、购物类应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等。

终端设备101、102、103可以是具有显示屏并且支持网页浏览的各种电子设备，包括但不限于智能手机、平板电脑、电子书阅读器、MP3播放器(Moving Picture Experts Group Audio Layer III, 动态影像专家压缩标准音频层面3)、MP4(Moving Picture Experts Group Audio Layer IV, 动态影像专家压缩标准音频层面4)播放器、膝上型便携计算机和台式计算机等等。

服务器105可以是提供各种服务的服务器，例如对终端设备101、102、103上显示的页面提供支持的后台服务器。

需要说明的是，本申请实施例所提供的非均衡样本的分类方法一般由服务器执行，相应地，非均衡样本的分类装置一般设置于服务器中。

应该理解，图1中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要，可以具有任意数目的终端设备、网络和服务器的。

继续参考图2，示出了根据本申请的非均衡样本的分类的方法的一个实施例的流程图。所述的非均衡样本的分类方法，包括以下步骤：

S201，从预设的语料库获取训练语料，其中，训练语料包括第一训练语料和第二训练语料，第一训练语料为少数类训练语料，第二训练语料为多数类训练语料。

具体的，从预设的语料库获取训练语料，其中，训练语料包括第一训练语料和第二训练语料，第一训练语料为少数类训练语料，第二训练语料为多数类训练语料。在本申请具体的实施例中，多数类训练语料可以达到10万条以上，少数类训练语料可能只有10条。训练语料可以是任意一种语种的语料，在本申请具体的实施例中，训练语料为中文语料。

S202，通过第一训练语料对预设的第一分类模型进行训练，得到初始第一分类模型。

其中，针对上述样本严重失衡的训练样本的分类场景，传统通过设置分类权重，然后将多数类训练语料和少数类训练语料随机组合的方式训练出来的分类模型的分类精度可能达不到要求，无法实现精准分类。本申请通过训练得到一个具有排他性的分类模型，在进行分类任务时，该分类模型仅能对少数类训练语料对应数据类型做出相应响应和输出，而对于多数类训练语料对应数据类型则无法响应，提高了非均衡样本分类模型的精度。

具体的，预设的第一分类模型可以是经典的transformer模型，先通过少数类训练语料对transformer模型进行训练，得到初始第一分类模型。其中，transformer模型结构包括encoder(编码层)和decoder(解码层)。通过编码层提取少数类训练语料的语料特

征,并对语料特征进行矢量编码,得到特征矢量,将特征矢量与预先存储在解码层中的特征标签进行特征映射,获取特征映射结果,基于特征映射结果对预设的第一分类模型进行迭代,得到初始第一分类模型,可以针对少数类训练语料对应数据类型做出相应响应和输出。

S203,基于预设调整规则对初始第一分类模型的损失函数进行调整,将第二训练语料导入损失函数调整后的初始第一分类模型,并基于反向传播算法对初始第一分类模型进行迭代更新,得到第一分类模型。

其中,基于预设调整规则对初始第一分类模型的损失函数进行调整是指对初始第一分类模型的损失函数反转,例如,初始第一分类模型的损失函数即第一损失函数为 L_1 ,在本申请具体的实施例中,反转后初始第一分类模型的损失函数 L_2 为 $1-L_1$ 。

具体的,当初始第一分类模型训练完毕后,基于预设调整规则对初始第一分类模型的损失函数进行反转,将多数类训练语料导入损失函数反转后的初始第一分类模型,并基于反向传播算法对损失函数反转后的初始第一分类模型反向迭代,得到第一分类模型,即利用反转后的损失函数鼓励初始第一分类模型朝着排除多数类训练语料的方向学习,得到一个具有排他性的第一分类模型,第一分类模型对于多数类训练语料对应数据类型无法响应,通过第一分类模型少数类训练语料对应数据类型和多数类训练语料对应数据类型,提高了非均衡样本分类模型的精度。

S204,接收语料分类指令,获取与语料分类指令相对应的待分类语料,并通过第一分类模型对待分类语料进行分类。

具体的,当接收到用户的语料分类指令时,获取与语料分类指令相对应的待分类语料,并通过训练好的第一分类模型对待分类语料进行分类。如果第一分类模型没有输出,即认为输入语料属于多数类数据对应数据类型,如果第一分类模型存在输出,即认为输入语料属于少数类数据对应数据类型,通过第一分类模型少数类训练语料对应数据类型和多数类训练语料对应数据类型。

在本实施例中,非均衡样本的分类方法运行于其上的电子设备(例如图1所示的服务器/终端设备)可以通过有线连接方式或者无线连接方式接收语料分类指令。需要指出的是,上述无线连接方式可以包括但不限于3G/4G连接、WiFi连接、蓝牙连接、WiMAX连接、Zigbee连接、UWB(ultra wideband)连接、以及其他现在已知或将来开发的无线连接方式。

本申请公开了一种非均衡样本的分类方法,属于人工智能技术领域,本申请通过获取训练语料,其中,训练语料包括第一训练语料和第二训练语料,第一训练语料为少数类训练语料,第二训练语料为多数类训练语料,通过少数类训练语料训练得到初始第一分类模型,通过预设调整规则对初始第一分类模型的损失函数进行反转,然后将多数类训练语料导入损失函数反转后的初始第一分类模型,并对初始第一分类模型进行迭代,得到第一分类模型,获取待分类语料,并通过第一分类模型对待分类语料进行分类。本申请通过预设调整规则对分类模型的损失函数进行反转,然后训练得到一个具有排他性的分类模型,在进行分类任务时,该分类模型仅能对少数类训练语料对应数据类型做出相应响应和输出,而对于多数类训练语料对应数据类型则无法响应,提高了非均衡样本分类模型的精度。

进一步地,训练语料的语种类型为第一语种,在从预设的语料库获取训练语料的步骤之后,还包括:

计算第一训练语料和第二训练语料的语料相似度,并比对语料相似度与预设相似度阈值的大小;

若语料相似度大于或等于预设相似度阈值,则将训练语料翻译为第二语种,其中,在第二语种的语种环境下,第一训练语料和第二训练语料的语料相似度小于预设相似度阈值。

其中,在本申请具体的实施例中,训练数据集可以是样本严重失衡的数据集,比如说有10万条语料是A类别,而与此同时,只有几条活着几十条语料是B类别,并且A类别中可能可能存在某些语料与B类别中的语料在语法结构上非常接近。例如,

语料一：“今天天气很热”，是指陈述语料类型。

语料二：“今天天气很热吗”，这是疑问语料类型。

针对上述情况，在传统分类模型进行训练时，由于两个语料相似度较高的原因，分类模型可能难以获得分辨上述两个语料的能力，导致模型精度不够，而在本申请中，可以通过语料翻译来解决上述问题。

首先，通过先将汉语语料翻译为成能够拉大语料特征区别的语言，在本申请具体的实施例中，将汉语语料翻译为德语语料。

举例说明，上述的两个语料翻译成德文分别是“Es ist heute sehr heiß”、“Wie heiß es heute ist”，从这里可以看出，将语料翻译成德文，将在确保原有信息的同时，有效拉大两者的区别。

然后，将分类问题转变为相似度衡量的问题，即将原本传统的分类问题，变成一个衡量输入语料经过处理后，同某一类型语料是否相似的问题，举例说明：

语料二：“今天天气很热吗？”

语料三：“今天你过得快乐吗？”

语料二和语料三为未分类的语料，但是已分类语料中有语料四：“今天天气很冷吗？”，很显然语料二相对于语料三更加接近语料四，因而语料二应当属于语料四所在的数据集合，而语料三不属于语料四所在的数据集合。

在本申请具体的实施例中，服务器内预设有语种翻译映射表，语种翻译映射表中记录了第一语种和第二语种的翻译映射关系，例如“中文-德文”等等，第一语种和第二语种的翻译映射关系可以根据各个语种之间的语种结构差异确定，且具有翻译映射关系的第一语种和第二语种在语种结构上差异较大。在本申请一种具体的实施例中，服务器可以在识别训练语料的第一语种后，通过在语种翻译映射表中查询第一语种所对应的第二语种，然后将获取到的第一语种翻译为查询到的第二语种。举例说明，服务器识别训练语料的第一语种为中文，通过在语种翻译映射表中查询到“中文”语种对应的第二语种为“德文”，然后服务器将中文的训练语料翻译为德文。

在上述实施例中，通过对训练语料进行翻译处理，以拉大训练语料中多数类训练语料和少数类训练语料之间的特征差异，使得训练完成的第一分类模型仅得到对少数类训练语料对应类型的数据响应。在本申请具体的实施例中，由于这里输入的语料是汉语，因此需要挑选相对于汉语在语法结构等方面差异性比较大的语言进行翻译，本申请采用的是将汉语翻译成德语，将汉语翻译成德语以拉大训练语料中多数类训练语料和少数类训练语料之间的特征差异。

进一步地，预设的第一分类模型为包括编码层和解码层，通过第一训练语料对预设的第一分类模型进行训练，得到初始第一分类模型的步骤，具体包括：

通过编码层提取第一训练语料的语料特征，并对语料特征进行矢量编码，得到特征矢量；

将特征矢量与预先存储在解码层中的特征标签进行特征映射，获取特征映射结果；

基于特征映射结果对预设的第一分类模型进行迭代，得到初始第一分类模型。

具体的，预设的第一分类模型为包括编码层和解码层，第一分类模型设计的核心思路是使用编码层作对输入语料的矢量进行编码，将其变成特征矢量，而解码层用来构建起这些特征矢量同解码层中的特征标签进行特征映射。

在上述实施例中，如果仅仅使用少数类别语料进行训练，那么 transformer 模型的编码层和解码层仅会学习适用于这些少数类别的特征矢量提取方法和映射，当输入的语料不属于少数类别时，那么编码层获取不了特征矢量，也无法构建起合理映射的条件。

进一步地，基于特征映射结果对预设的第一分类模型进行迭代，得到初始第一分类模型的步骤，具体包括：

构建初始第一分类模型的损失函数，得到第一损失函数，其中，第一损失函数包含对抗因子；

基于第一损失函数计算特征映射结果与预设映射结果的误差，得到映射误差；

基于映射误差和反向传播算法对预设的第一分类模型进行迭代，得到初始第一分类模型。

其中，构建初始第一分类模型的损失函数，得到第一损失函数，其中，第一损失函数包含对抗因子，这里加入对抗因子的目的是方便后续对初始第一分类模型进行反向迭代更新时，拉大映射误差大于预设误差阈，使得初始第一分类模型朝着排除多数类训练语料的方向学习，得到一个具有排他性的第一分类模型，第一分类模型对于多数类训练语料对应数据类型无法响应。

第一损失函数表达式如下：

$$L_1 = L_{ori} + \lambda D$$

其中，这里的 L_{ori} 表示的是标准的 transformer 损失函数， λ 是常数， D 是经典的对抗损失函数，在本申请中对抗损失函数 D 根据下述公式进行构建：

$$\min \max V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_{noise}(z)} [\log(1 - D(G(z)))]$$

其中， D 是对抗损失网络中的决定器， G 指的是对抗损失网络中的生成器，在本申请中， G 可以是上述的 transformer 模型， x 和 z 是输入的训练数据， $x \sim p_{data}(x)$ 、 $z \sim p_{noise}(z)$ 指的是训练数据 x 和 z 服从某种分布， E 表示取抗损失网络输出的均值。

具体的，基于构建的第一损失函数计算特征映射结果与预设映射结果的误差，得到映射误差，比对映射误差与预设的预设映射误差阈值，将映射误差与预设映射误差阈值进行比较，若映射误差大于预设映射误差阈值，则通过反向传播算法对初始第一映射误差模型进行迭代更新，直到映射误差小于或等于预设映射误差阈值为止，输出映射误差大于预设映射误差阈值的初始第一分类模型。

进一步地，基于预设调整规则对初始第一分类模型的损失函数进行调整，将第二训练语料导入损失函数调整后的初始第一分类模型，并基于反向传播算法对初始第一分类模型进行迭代更新，得到第一分类模型的步骤，具体包括：

基于预设调整规则对初始第一分类模型的损失函数进行反转，得到第二损失函数；

将第二训练语料导入损失函数反转后的初始第一分类模型，获取分类结果；

基于第二损失函数计算分类结果与预设分类结果的误差，得到分类误差；

将分类误差与预设分类误差阈值进行比较，若分类误差小于或等于预设分类误差阈值，则通过反向传播算法对损失函数反转后的初始第一分类模型进行迭代更新，直到分类误差大于预设分类误差阈值为止；

输出分类误差大于预设分类误差阈值的初始第一分类模型。

其中，基于预设调整规则对初始第一分类模型的损失函数进行调整是指对初始第一分类模型的损失函数反转，例如，第一损失函数为 L_1 ，在本申请具体的实施例，反转后初始第一分类模型的损失函数为 $L_2 = 1 - L_1$ 。

具体的，基于预设调整规则对初始第一分类模型的损失函数进行反转，得到第二损失函数；

将第二训练语料导入损失函数反转后的初始第一分类模型，获取分类结果，基于第二损失函数计算分类结果与预设分类结果的误差，得到分类误差，将分类误差与预设分类误差阈值进行比较，若分类误差小于或等于预设分类误差阈值，则通过反向传播算法对损失

函数反转后的初始第一分类模型进行迭代更新，直到分类误差大于预设分类误差阈值为止，输出分类误差大于预设分类误差阈值的的第一分类模型。

在上述实施例中，通过对初始第一分类模型的损失函数进行反转，使得初始第一分类模型朝着排除多数类训练语料的方向学习，得到一个具有排他性的第一分类模型，第一分类模型对于多数类训练语料对应数据类型无法响应，通过第一分类模型少数类训练语料对应数据类型和多数类训练语料对应数据类型，提高了非均衡样本分类模型的精度。

进一步地，在基于预设调整规则对初始第一分类模型的损失函数进行调整，将第二训练语料导入损失函数调整后的初始第一分类模型，并基于反向传播算法对初始第一分类模型进行迭代更新，得到第一分类模型的步骤之后，还包括：

对第一训练语料进行矢量化处理，得到第一训练语料的语料矢量；

对语料矢量和特征矢量进行矢量拼接，得到语料特征矩阵；

将语料特征矩阵导入预设的第二分类模型，通过第二分类模型的卷积核对语料特征矩阵进行卷积运算，获取卷积运算结果；

基于卷积运算结果对第二分类模型进行迭代更新，输出训练完成的第二分类模型。

其中，根据高维流形分类原理，如果在低维情况下无法线性可分的少数类训练语料，在高维情况下可以做到线性可分，这既是说明，高维情况下的样本更容易实现分类。由于少数类训练语料通常数量较少，难以构成训练一个完整分类器的充分数据集，但是在高维条件下，以这些样本作为聚类中心的初始点，应该较为容易找到少数类训练语料的聚类中心。在上述方案中，通过升维操作在一个不完备的数据集上训练出一个单纯的分类型模型，至于聚类中心点这是由于在模型的损失函数指导下的模型自适应行为。

具体的，在得到第一分类模型之后，还可以继续通过第二分类模型对少数类训练语料进行分类，第二分类模型可以是 CNN 深度神经网络模型。对第二分类模型进行训练时，输入的矢量包括少数类训练语料对应的语料矢量和上述 transformer 模型的编码层输出的特征矢量，这里使用特征矢量的原因就在于利用特征矢量获得更加高维的特征，而高维特征有利于少数类训练语料的分类。

在上述实施例中，通过对第一训练语料进行矢量化处理，得到第一训练语料的语料矢量，对语料矢量和特征矢量进行矢量拼接，得到语料特征矩阵，将语料特征矩阵导入预设的第二分类模型，通过第二分类模型的卷积核对语料特征矩阵进行卷积运算，获取卷积运算结果，基于卷积运算结果对第二分类模型进行迭代更新，输出训练完成的第二分类模型，第二分类模型可以对少数类训练语料进行分类。

进一步地，基于卷积运算结果对第二分类模型进行迭代更新，输出训练完成的第二分类模型的步骤，具体包括：

基于交叉熵损失函数和莱温斯坦距离函数构建第二分类模型的损失函数，得到初始第三损失函数，其中，初始第三损失函数包括交叉熵因子和莱温斯坦因子；

为初始第三损失函数的交叉熵因子和莱温斯坦因子分别赋予相同的初始权重值；

基于卷积运算结果对交叉熵因子和莱温斯坦因子的初始权重值进行调整，直至初始第三损失函数的输出达到最小值，得到第三损失函数；

通过第三损失函数计算卷积运算结果的运算误差，基于运算误差采用反向传播算法对第二分类模型进行迭代更新，得到训练完成的第二分类模型。

具体的，初始第三损失函数的形式如下：

$$L_3 = \alpha L_{cls} + \beta L_{lev}$$

这里的 α 与 β 是权重系数，在本申请具体的实施例中， α 与 β 的和为 1， α 与 β 的初始值可以设为 0.5 和 0.5。 L_{cls} 是标准的交叉熵损失函数，而 L_{lev} 是标准的莱温斯坦距离函数，其中，交叉熵损失函数用于指导分类任务，而莱温斯坦距离函数用于确保聚类

和形成聚类中心。在第二分类模型训练过程中，可以跟据卷积运算结果持续调整权重系数 α 与 β ，直至初始第三损失函数的输出达到最小值，将输出达到最小值时的初始第三损失函数确定为第二分类模型的损失函数，即第三损失函数，通过第三损失函数计算卷积运算结果的运算误差，基于运算误差采用反向传播算法对第二分类模型进行迭代更新，得到训练完成的第二分类模型。

在实际应用过程中，样本分类是多种多样的，使得样本权重值计算极其复杂，我们很难去平衡各个种类样本之间的权重。而且针对特征特别相近的训练样本，上述通过设置权重的方式训练的得到分类模型难以对其进行分类，即通过设置权重的方式来实现有偏数据集的分类在很多场景下分类精度达不到要求，因此，现有的通过为不同种类样本进行赋权的非均衡样本的分类存在一定的缺陷。

针对上述技术问题，本申请公开了一种非均衡样本的分类方法、装置、设备和存储介质，属于人工智能技术领域，本申请通过获取训练语料，其中，训练语料包括第一训练语料和第二训练语料，第一训练语料为少数类训练语料，第二训练语料为多数类训练语料，通过少数类训练语料训练得到初始第一分类模型，通过预设调整规则对初始第一分类模型的损失函数进行反转，然后将多数类训练语料导入损失函数反转后的初始第一分类模型，并对初始第一分类模型进行迭代，得到第一分类模型，获取待分类语料，并通过第一分类模型对待分类语料进行分类。本申请通过预设调整规则对分类模型的损失函数进行反转，然后训练得到一个具有排他性的分类模型，在进行分类任务时，该分类模型仅能对少数类训练语料对应数据类型做出相应响应和输出，而对于多数类训练语料对应数据类型则无法响应，提高了非均衡样本分类模型的精度。

需要强调的是，为进一步保证上述待分类语料的私密和安全性，上述待分类语料还可以存储于一区块链的节点中。

本申请所指区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链(Blockchain)，本质上是一个去中心化的数据库，是一串使用密码学方法相关联产生的数据块，每一个数据块中包含了一批次网络交易的信息，用于验证其信息的有效性(防伪)和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层等。

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计算机可读存储介质中，该计算机可读指令在执行时，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。

应该理解的是，虽然附图的流程图中的各个步骤按照箭头的指示依次显示，但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明，这些步骤的执行并没有严格的顺序限制，其可以以其他的顺序执行。而且，附图的流程图中的至少一部分步骤可以包括多个子步骤或者多个阶段，这些子步骤或者阶段并不必然是在同一时刻执行完成，而是可以在不同的时刻执行，其执行顺序也不必然是依次进行，而是可以与其他步骤或者其他步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

进一步参考图3，作为对上述图2所示方法的实现，本申请提供了一种非均衡样本的分类装置的一个实施例，该装置实施例与图2所示的方法实施例相对应，该装置具体可以应用于各种电子设备中。

如图3所示，本实施例所述的非均衡样本的分类装置包括：

语料获取模块301，用于从预设的语料库获取训练语料，其中，训练语料包括第一训练语料和第二训练语料，第一训练语料为少数类训练语料，第二训练语料为多数类训练语料；

第一模型训练模块 302, 用于通过第一训练语料对预设的第一分类模型进行训练, 得到初始第一分类模型;

反转训练模块 303, 用于基于预设调整规则对初始第一分类模型的损失函数进行调整, 将第二训练语料导入损失函数调整后的初始第一分类模型, 并基于反向传播算法对初始第一分类模型进行迭代更新, 得到第一分类模型;

语料分类模块 304, 用于接收语料分类指令, 获取与语料分类指令相对应的待分类语料, 并通过第一分类模型对待分类语料进行分类。

进一步地, 训练语料的语种类型为第一语种, 该非均衡样本的分类装置还包括:

语料相似度计算模块, 用于计算第一训练语料和第二训练语料的语料相似度, 并比对语料相似度与预设相似度阈值的大小;

语料翻译模块, 用于当语料相似度大于或等于预设相似度阈值时, 将训练语料翻译为第二语种, 其中, 在第二语种的语种环境下, 第一训练语料和第二训练语料的语料相似度小于预设相似度阈值。

进一步地, 预设的第一分类模型为包括编码层和解码层, 第一模型训练模块 302 具体包括:

特征提取单元, 用于通过编码层提取第一训练语料的语料特征, 并对语料特征进行矢量编码, 得到特征矢量;

特征映射单元, 用于将特征矢量与预先存储在解码层中的特征标签进行特征映射, 获取特征映射结果;

第一迭代单元, 用于基于特征映射结果对预设的第一分类模型进行迭代, 得到初始第一分类模型。

进一步地, 第一迭代单元具体包括:

第一损失函数构建子单元, 用于构建初始第一分类模型的损失函数, 得到第一损失函数, 其中, 第一损失函数包含对抗因子;

映射误差计算子单元, 用于基于第一损失函数计算特征映射结果与预设映射结果的误差, 得到映射误差;

第一迭代子单元, 用于基于映射误差和反向传播算法对预设的第一分类模型进行迭代, 得到初始第一分类模型。

进一步地, 反转训练模块 303 具体包括:

函数反转单元, 用于基于预设调整规则对初始第一分类模型的损失函数进行反转, 得到第二损失函数;

语料分类单元, 用于将第二训练语料导入损失函数反转后的初始第一分类模型, 获取分类结果;

分类误差计算单元, 用于基于第二损失函数计算分类结果与预设分类结果的误差, 得到分类误差;

第二迭代单元, 用于将分类误差与预设分类误差阈值进行比较, 若分类误差小于或等于预设分类误差阈值, 则通过反向传播算法对损失函数反转后的初始第一分类模型进行迭代更新, 直到分类误差大于预设分类误差阈值为止;

模型输出单元, 用于输出分类误差大于预设分类误差阈值的第一分类模型。

进一步地, 该非均衡样本的分类装置还包括:

矢量化模块, 用于对第一训练语料进行矢量化处理, 得到第一训练语料的语料矢量;

矢量拼接模块, 用于对语料矢量和特征矢量进行矢量拼接, 得到语料特征矩阵;

卷积运算模块, 用于将语料特征矩阵导入预设的第二分类模型, 通过第二分类模型的卷积核对语料特征矩阵进行卷积运算, 获取卷积运算结果;

模型迭代模块, 用于基于卷积运算结果对第二分类模型进行迭代更新, 输出训练完成的第二分类模型。

进一步地，模型迭代模块具体包括：

第三损失函数构建单元，用于基于交叉熵损失函数和莱温斯坦距离函数构建第二分类模型的损失函数，得到初始第三损失函数，其中，初始第三损失函数包括交叉熵因子和莱温斯坦因子；

权重赋值单元，用于为初始第三损失函数的交叉熵因子和莱温斯坦因子分别赋予相同的初始权重值；

第三损失函数优化单元，用于基于卷积运算结果对交叉熵因子和莱温斯坦因子的初始权重值进行调整，直至初始第三损失函数的输出达到最小值，得到第三损失函数；

第三迭代单元，用于通过第三损失函数计算卷积运算结果的运算误差，基于运算误差采用反向传播算法对第二分类模型进行迭代更新，得到训练完成的第二分类模型。

本申请公开了一种非均衡样本的分类装置，属于人工智能技术领域，本申请通过获取训练语料，其中，训练语料包括第一训练语料和第二训练语料，第一训练语料为少数类训练语料，第二训练语料为多数类训练语料，通过少数类训练语料训练得到初始第一分类模型，通过预设调整规则对初始第一分类模型的损失函数进行反转，然后将多数类训练语料导入损失函数反转后的初始第一分类模型，并对初始第一分类模型进行迭代，得到第一分类模型，获取待分类语料，并通过第一分类模型对待分类语料进行分类。本申请通过预设调整规则对分类模型的损失函数进行反转，然后训练得到一个具有排他性的分类模型，在进行分类任务时，该分类模型仅能对少数类训练语料对应数据类型做出相应响应和输出，而对于多数类训练语料对应数据类型则无法响应，提高了非均衡样本分类模型的精度。

为解决上述技术问题，本申请实施例还提供计算机设备。具体请参阅图4，图4为本实施例计算机设备基本结构框图。

所述计算机设备4包括通过系统总线相互通信连接存储器41、处理器42、网络接口43。需要指出的是，图中仅示出了具有组件41-43的计算机设备4，但是应理解的是，并不要求实施所有示出的组件，可以替代的实施更多或者更少的组件。其中，本技术领域技术人员可以理解，这里的计算机设备是一种能够按照事先设定或存储的指令，自动进行数值计算和/或信息处理的设备，其硬件包括但不限于微处理器、专用集成电路(Application Specific Integrated Circuit, ASIC)、可编程门阵列(Field-Programmable Gate Array, FPGA)、数字处理器(Digital Signal Processor, DSP)、嵌入式设备等。

所述计算机设备可以是桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备。所述计算机设备可以与用户通过键盘、鼠标、遥控器、触摸板或声控设备等方式进行人机交互。

所述存储器41至少包括一种类型的可读存储介质，所述可读存储介质包括闪存、硬盘、多媒体卡、卡型存储器（例如，SD或DX存储器等）、随机访问存储器(RAM)、静态随机访问存储器(SRAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、可编程只读存储器(PROM)、磁性存储器、磁盘、光盘等。在一些实施例中，所述存储器41可以是所述计算机设备4的内部存储单元，例如该计算机设备4的硬盘或内存。在另一些实施例中，所述存储器41也可以是所述计算机设备4的外部存储设备，例如该计算机设备4上配备的插接式硬盘，智能存储卡(Smart Media Card, SMC)，安全数字(Secure Digital, SD)卡，闪存卡(Flash Card)等。当然，所述存储器41还可以既包括所述计算机设备4的内部存储单元也包括其外部存储设备。本实施例中，所述存储器41通常用于存储安装于所述计算机设备4的操作系统和各类应用软件，例如非均衡样本的分类方法的计算机可读指令等。此外，所述存储器41还可以用于暂时地存储已经输出或者将要输出的各类数据。

所述处理器42在一些实施例中可以是中央处理器(Central Processing Unit, CPU)、控制器、微控制器、微处理器、或其他数据处理芯片。该处理器42通常用于控制所述计

计算机设备 4 的总体操作。本实施例中，所述处理器 42 用于运行所述存储器 41 中存储的计算机可读指令或者处理数据，例如运行所述非均衡样本的分类方法的计算机可读指令。

所述网络接口 43 可包括无线网络接口或有线网络接口，该网络接口 43 通常用于在所述计算机设备 4 与其他电子设备之间建立通信连接。

本申请公开了一种非均衡样本的分类设备，属于人工智能技术领域，本申请通过获取训练语料，其中，训练语料包括第一训练语料和第二训练语料，第一训练语料为少数类训练语料，第二训练语料为多数类训练语料，通过少数类训练语料训练得到初始第一分类模型，通过预设调整规则对初始第一分类模型的损失函数进行反转，然后将多数类训练语料导入损失函数反转后的初始第一分类模型，并对初始第一分类模型进行迭代，得到第一分类模型，获取待分类语料，并通过第一分类模型对待分类语料进行分类。本申请通过预设调整规则对分类模型的损失函数进行反转，然后训练得到一个具有排他性的分类模型，在进行分类任务时，该分类模型仅能对少数类训练语料对应数据类型做出相应响应和输出，而对于多数类训练语料对应数据类型则无法响应，提高了非均衡样本分类模型的精度。

本申请还提供了另一种实施方式，即提供一种计算机可读存储介质，所述计算机可读存储介质可以是非易失性，也可以是易失性，所述计算机可读存储介质存储有计算机可读指令，所述计算机可读指令可被至少一个处理器执行，以使所述至少一个处理器执行如上述的非均衡样本的分类方法的步骤。

本申请公开了一种非均衡样本的分类存储介质，属于人工智能技术领域，本申请通过获取训练语料，其中，训练语料包括第一训练语料和第二训练语料，第一训练语料为少数类训练语料，第二训练语料为多数类训练语料，通过少数类训练语料训练得到初始第一分类模型，通过预设调整规则对初始第一分类模型的损失函数进行反转，然后将多数类训练语料导入损失函数反转后的初始第一分类模型，并对初始第一分类模型进行迭代，得到第一分类模型，获取待分类语料，并通过第一分类模型对待分类语料进行分类。本申请通过预设调整规则对分类模型的损失函数进行反转，然后训练得到一个具有排他性的分类模型，在进行分类任务时，该分类模型仅能对少数类训练语料对应数据类型做出相应响应和输出，而对于多数类训练语料对应数据类型则无法响应，提高了非均衡样本分类模型的精度。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献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（如 ROM/RAM、磁碟、光盘）中，包括若干指令用以使得一台终端设备（可以是手机，计算机，服务器，空调器，或者网络设备等）执行本申请各个实施例所述的方法。

显然，以上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，附图中给出了本申请的较佳实施例，但并不限制本申请的专利范围。本申请可以以许多不同的形式来实现，相反地，提供这些实施例的目的是使对本申请的公开内容的理解更加透彻全面。尽管参照前述实施例对本申请进行了详细的说明，对于本领域的技术人员而言，其依然可以对前述各具体实施方式所记载的技术方案进行修改，或者对其中部分技术特征进行等效替换。凡是利用本申请说明书及附图内容所做的等效结构，直接或间接运用在其他相关的技术领域，均同理在本申请专利保护范围之内。

权 利 要 求 书

1. 一种非均衡样本的分类方法，包括：

从预设的语料库获取训练语料，其中，所述训练语料包括第一训练语料和第二训练语料，所述第一训练语料为少数类训练语料，所述第二训练语料为多数类训练语料；

通过所述第一训练语料对预设的第一分类模型进行训练，得到初始第一分类模型；

5 基于预设调整规则对所述初始第一分类模型的损失函数进行调整，将所述第二训练语料导入损失函数调整后的所述初始第一分类模型，并基于反向传播算法对初始第一分类模型进行迭代更新，得到第一分类模型；

接收语料分类指令，获取与所述语料分类指令相对应的待分类语料，并通过所述第一分类模型对所述待分类语料进行分类。

10 2. 如权利要求 1 所述的非均衡样本的分类方法，其中，所述训练语料的语种类型为第一语种，在所述从预设的语料库获取训练语料的步骤之后，还包括：

计算所述第一训练语料和所述第二训练语料的语料相似度，并比对所述语料相似度与预设相似度阈值的大小；

15 若所述语料相似度大于或等于预设相似度阈值，则将所述训练语料翻译为第二语种，其中，在所述第二语种的语种环境下，所述第一训练语料和所述第二训练语料的语料相似度小于预设相似度阈值。

3. 如权利要求 1 所述的非均衡样本的分类方法，其中，所述预设的第一分类模型为包括编码层和解码层，所述通过所述第一训练语料对预设的第一分类模型进行训练，得到初始第一分类模型的步骤，具体包括：

20 通过所述编码层提取所述第一训练语料的语料特征，并对所述语料特征进行矢量编码，得到特征矢量；

将所述特征矢量与预先存储在所述解码层中的特征标签进行特征映射，获取特征映射结果；

25 基于所述特征映射结果对所述预设的第一分类模型进行迭代，得到初始第一分类模型。

4. 如权利要求 3 所述的非均衡样本的分类方法，其中，所述基于所述特征映射结果对所述预设的第一分类模型进行迭代，得到初始第一分类模型的步骤，具体包括：

构建所述初始第一分类模型的损失函数，得到第一损失函数，其中，所述第一损失函数包含对抗因子；

基于第一损失函数计算所述特征映射结果与预设映射结果的误差，得到映射误差；

30 基于所述映射误差和反向传播算法对所述预设的第一分类模型进行迭代，得到初始第一分类模型。

5. 如权利要求 4 所述的非均衡样本的分类方法，其中，基于预设调整规则对所述初始第一分类模型的损失函数进行调整，将所述第二训练语料导入损失函数调整后的所述初始第一分类模型，并基于反向传播算法对初始第一分类模型进行迭代更新，得到第一分类模型的步骤，具体包括：

基于预设调整规则对所述初始第一分类模型的损失函数进行反转，得到第二损失函数；

将所述第二训练语料导入损失函数反转后的所述初始第一分类模型，获取分类结果；

基于第二损失函数计算所述分类结果与预设分类结果的误差，得到分类误差；

40 将分类误差与预设分类误差阈值进行比较，若分类误差小于或等于预设分类误差阈值，则通过反向传播算法对损失函数反转后的初始第一分类模型进行迭代更新，直到分类误差大于预设分类误差阈值为止；

输出分类误差大于预设分类误差阈值的第一分类模型。

6. 如权利要求 3 所述的非均衡样本的分类方法，其中，在所述基于预设调整规则对所述初始第一分类模型的损失函数进行调整，将所述第二训练语料导入损失函数调整后的所述初始第一分类模型，并基于反向传播算法对初始第一分类模型进行迭代更新，得到第一分类模型的步骤之后，还包括：

对所述第一训练语料进行矢量化处理,得到所述第一训练语料的语料矢量;

对所述语料矢量和所述特征矢量进行矢量拼接,得到语料特征矩阵;

将所述语料特征矩阵导入预设的第二分类模型,通过所述第二分类模型的卷积核对所
述语料特征矩阵进行卷积运算,获取卷积运算结果;

5 基于所述卷积运算结果对所述第二分类模型进行迭代更新,输出训练完成的所述第二
分类模型。

7.如权利要求6所述的非均衡样本的分类方法,其中,所述基于所述卷积运算结果对
所述第二分类模型进行迭代更新,输出训练完成的所述第二分类模型的步骤,具体包括:

10 基于交叉熵损失函数和莱温斯坦距离函数构建所述第二分类模型的损失函数,得到初
始第三损失函数,其中,所述初始第三损失函数包括交叉熵因子和莱温斯坦因子;

为所述初始第三损失函数的交叉熵因子和莱温斯坦因子分别赋予相同的初始权重值;

基于所述卷积运算结果对所述交叉熵因子和所述莱温斯坦因子的初始权重值进行调
整,直至所述初始第三损失函数的输出达到最小值,得到第三损失函数;

15 通过所述第三损失函数计算所述卷积运算结果的运算误差,基于所述运算误差采用反
向传播算法对所述第二分类模型进行迭代更新,得到训练完成的所述第二分类模型。

8.一种非均衡样本的分类装置,包括:

语料获取模块,用于从预设的语料库获取训练语料,其中,所述训练语料包括第一训
练语料和第二训练语料,所述第一训练语料为少数类训练语料,所述第二训练语料为多数
类训练语料;

20 第一模型训练模块,用于通过所述第一训练语料对预设的第一分类模型进行训练,得
到初始第一分类模型;

反转训练模块,用于基于预设调整规则对所述初始第一分类模型的损失函数进行调整,
将所述第二训练语料导入损失函数调整后的所述初始第一分类模型,并基于反向传播算法
对初始第一分类模型进行迭代更新,得到第一分类模型;

25 语料分类模块,用于接收语料分类指令,获取与所述语料分类指令相对应的待分类语
料,并通过所述第一分类模型对所述待分类语料进行分类。

9.一种设备,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,所述处理
器执行所述计算机可读指令时实现如下所述的非均衡样本的分类方法:

30 从预设的语料库获取训练语料,其中,所述训练语料包括第一训练语料和第二训练语
料,所述第一训练语料为少数类训练语料,所述第二训练语料为多数类训练语料;

通过所述第一训练语料对预设的第一分类模型进行训练,得到初始第一分类模型;

基于预设调整规则对所述初始第一分类模型的损失函数进行调整,将所述第二训练语
料导入损失函数调整后的所述初始第一分类模型,并基于反向传播算法对初始第一分类模
型进行迭代更新,得到第一分类模型;

35 接收语料分类指令,获取与所述语料分类指令相对应的待分类语料,并通过所述第
一分类模型对所述待分类语料进行分类。

10.如权利要求9所述的设备,其中,所述训练语料的语种类型为第一语种,在所述
从预设的语料库获取训练语料的步骤之后,还包括:

40 计算所述第一训练语料和所述第二训练语料的语料相似度,并比对所述语料相似度与
预设相似度阈值的大小;

若所述语料相似度大于或等于预设相似度阈值,则将所述训练语料翻译为第二语种,
其中,在所述第二语种的语种环境下,所述第一训练语料和所述第二训练语料的语料相似
度小于预设相似度阈值。

45 11.如权利要求9所述的设备,其中,所述预设的第一分类模型为包括编码层和解码
层,所述通过所述第一训练语料对预设的第一分类模型进行训练,得到初始第一分类模型
的步骤,具体包括:

通过所述编码层提取所述第一训练语料的语料特征，并对所述语料特征进行矢量编码，得到特征矢量；

将所述特征矢量与预先存储在所述解码层中的特征标签进行特征映射，获取特征映射结果；

5 基于所述特征映射结果对所述预设的第一分类模型进行迭代，得到初始第一分类模型。

12. 如权利要求 11 所述的设备，其中，所述基于所述特征映射结果对所述预设的第一分类模型进行迭代，得到初始第一分类模型的步骤，具体包括：

构建所述初始第一分类模型的损失函数，得到第一损失函数，其中，所述第一损失函数包含对抗因子；

10 基于第一损失函数计算所述特征映射结果与预设映射结果的误差，得到映射误差；

基于所述映射误差和反向传播算法对所述预设的第一分类模型进行迭代，得到初始第一分类模型。

13. 如权利要求 12 所述的设备，其中，基于预设调整规则对所述初始第一分类模型的损失函数进行调整，将所述第二训练语料导入损失函数调整后的所述初始第一分类模型，并基于反向传播算法对初始第一分类模型进行迭代更新，得到第一分类模型的步骤，具体包括：

基于预设调整规则对所述初始第一分类模型的损失函数进行反转，得到第二损失函数；

将所述第二训练语料导入损失函数反转后的所述初始第一分类模型，获取分类结果；

基于第二损失函数计算所述分类结果与预设分类结果的误差，得到分类误差；

20 将分类误差与预设分类误差阈值进行比较，若分类误差小于或等于预设分类误差阈值，则通过反向传播算法对损失函数反转后的初始第一分类模型进行迭代更新，直到分类误差大于预设分类误差阈值为止；

输出分类误差大于预设分类误差阈值的所述第一分类模型。

14. 如权利要求 11 所述的设备，其中，在所述基于预设调整规则对所述初始第一分类模型的损失函数进行调整，将所述第二训练语料导入损失函数调整后的所述初始第一分类模型，并基于反向传播算法对初始第一分类模型进行迭代更新，得到第一分类模型的步骤之后，还包括：

对所述第一训练语料进行矢量化处理，得到所述第一训练语料的语料矢量；

对所述语料矢量和所述特征矢量进行矢量拼接，得到语料特征矩阵；

30 将所述语料特征矩阵导入预设的第二分类模型，通过所述第二分类模型的卷积核对所述语料特征矩阵进行卷积运算，获取卷积运算结果；

基于所述卷积运算结果对所述第二分类模型进行迭代更新，输出训练完成的所述第二分类模型。

15. 一种计算机可读存储介质，所述计算机可读存储介质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下所述的非均衡样本的分类方法：

从预设的语料库获取训练语料，其中，所述训练语料包括第一训练语料和第二训练语料，所述第一训练语料为少数类训练语料，所述第二训练语料为多数类训练语料；

通过所述第一训练语料对预设的第一分类模型进行训练，得到初始第一分类模型；

40 基于预设调整规则对所述初始第一分类模型的损失函数进行调整，将所述第二训练语料导入损失函数调整后的所述初始第一分类模型，并基于反向传播算法对初始第一分类模型进行迭代更新，得到第一分类模型；

接收语料分类指令，获取与所述语料分类指令相对应的待分类语料，并通过所述第一分类模型对所述待分类语料进行分类。

16. 如权利要求 15 所述的计算机可读存储介质，其中，所述训练语料的语种类型为第一语种，在所述从预设的语料库获取训练语料的步骤之后，还包括：

计算所述第一训练语料和所述第二训练语料的语料相似度，并比对所述语料相似度与预设相似度阈值的大小；

若所述语料相似度大于或等于预设相似度阈值，则将所述训练语料翻译为第二语种，其中，在所述第二语种的语种环境下，所述第一训练语料和所述第二训练语料的语料相似度小于预设相似度阈值。

17. 如权利要求 15 所述的计算机可读存储介质，其中，所述预设的第一分类模型为包括编码层和解码层，所述通过所述第一训练语料对预设的第一分类模型进行训练，得到初始第一分类模型的步骤，具体包括：

通过所述编码层提取所述第一训练语料的语料特征，并对所述语料特征进行矢量编码，得到特征矢量；

将所述特征矢量与预先存储在所述解码层中的特征标签进行特征映射，获取特征映射结果；

基于所述特征映射结果对所述预设的第一分类模型进行迭代，得到初始第一分类模型。

18. 如权利要求 17 所述的计算机可读存储介质，其中，所述基于所述特征映射结果对所述预设的第一分类模型进行迭代，得到初始第一分类模型的步骤，具体包括：

构建所述初始第一分类模型的损失函数，得到第一损失函数，其中，所述第一损失函数包含对抗因子；

基于第一损失函数计算所述特征映射结果与预设映射结果的误差，得到映射误差；

基于所述映射误差和反向传播算法对所述预设的第一分类模型进行迭代，得到初始第一分类模型。

19. 如权利要求 18 所述的计算机可读存储介质，其中，基于预设调整规则对所述初始第一分类模型的损失函数进行调整，将所述第二训练语料导入损失函数调整后的所述初始第一分类模型，并基于反向传播算法对初始第一分类模型进行迭代更新，得到第一分类模型的步骤，具体包括：

基于预设调整规则对所述初始第一分类模型的损失函数进行反转，得到第二损失函数；

将所述第二训练语料导入损失函数反转后的所述初始第一分类模型，获取分类结果；

基于第二损失函数计算所述分类结果与预设分类结果的误差，得到分类误差；

将分类误差与预设分类误差阈值进行比较，若分类误差小于或等于预设分类误差阈值，则通过反向传播算法对损失函数反转后的初始第一分类模型进行迭代更新，直到分类误差大于预设分类误差阈值为止；

输出分类误差大于预设分类误差阈值的第一分类模型。

20. 如权利要求 17 所述的计算机可读存储介质，其中，在所述基于预设调整规则对所述初始第一分类模型的损失函数进行调整，将所述第二训练语料导入损失函数调整后的所述初始第一分类模型，并基于反向传播算法对初始第一分类模型进行迭代更新，得到第一分类模型的步骤之后，还包括：

对所述第一训练语料进行矢量化处理，得到所述第一训练语料的语料矢量；

对所述语料矢量和所述特征矢量进行矢量拼接，得到语料特征矩阵；

将所述语料特征矩阵导入预设的第二分类模型，通过所述第二分类模型的卷积核对所述语料特征矩阵进行卷积运算，获取卷积运算结果；

基于所述卷积运算结果对所述第二分类模型进行迭代更新，输出训练完成的所述第二分类模型。

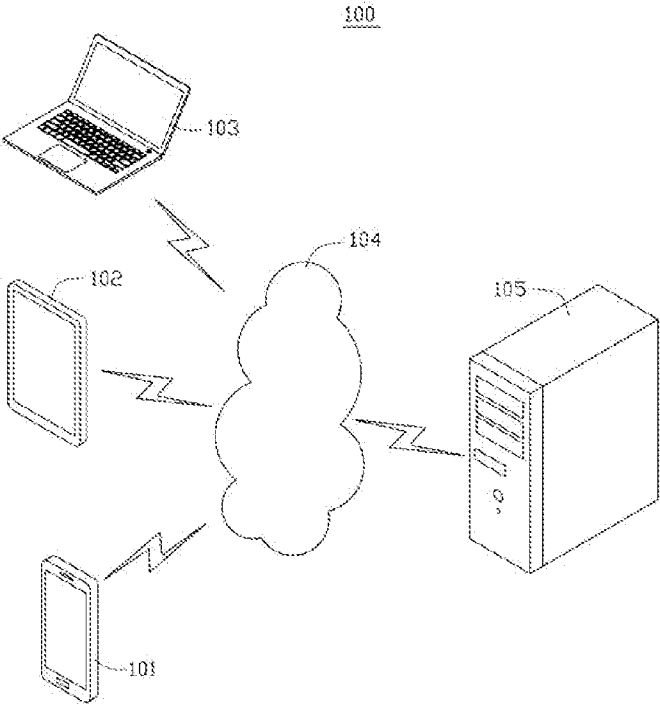


图 1

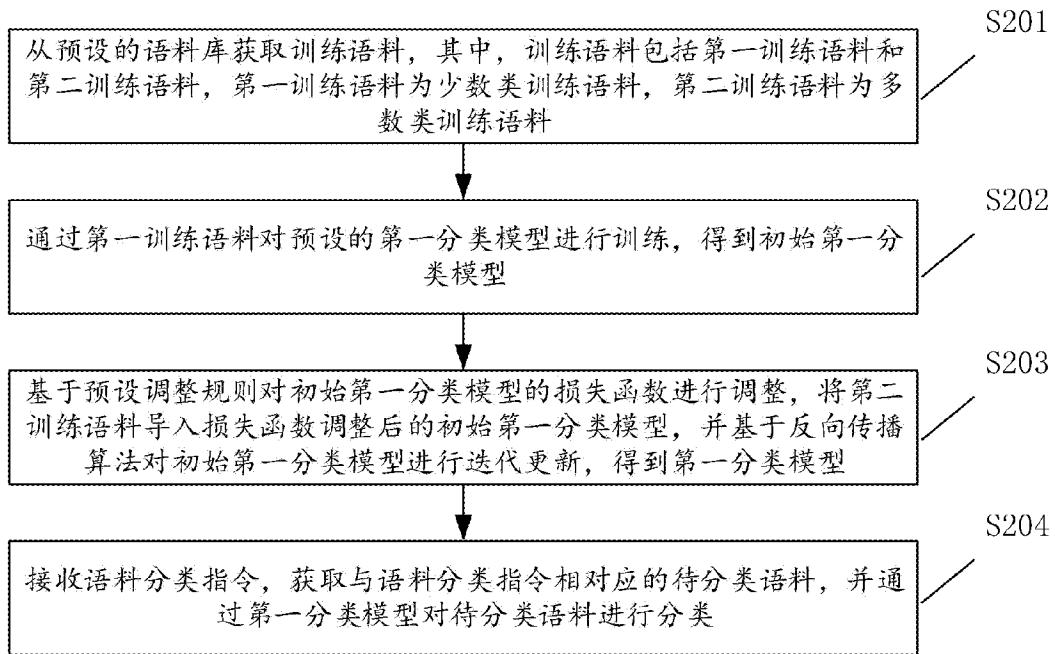


图 2

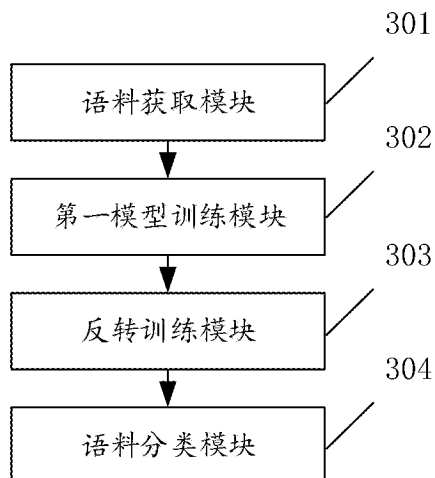


图 3

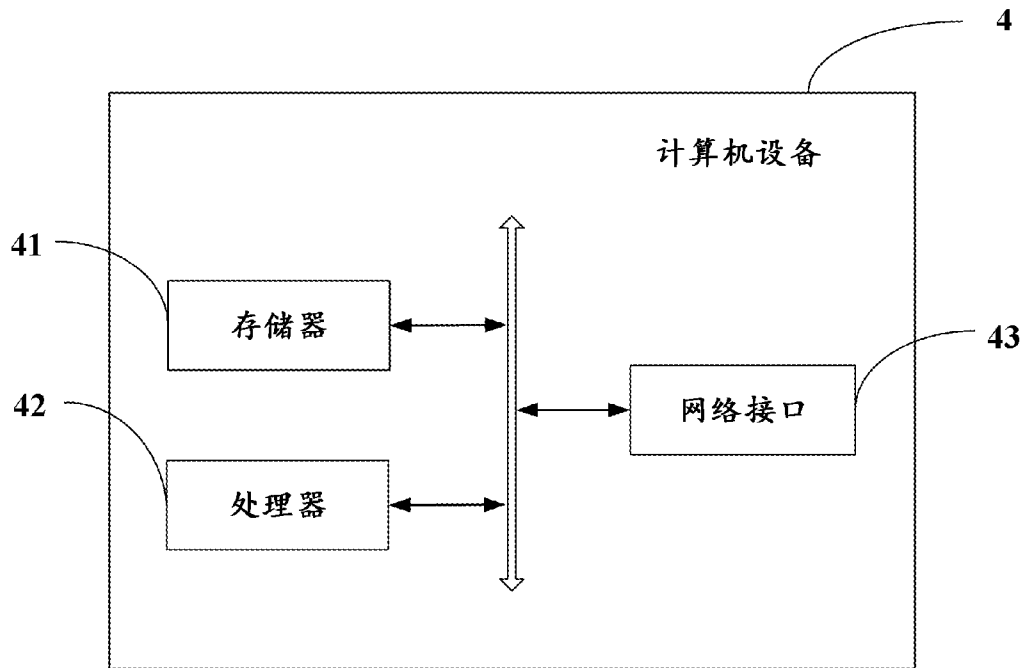


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/090432

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: 样本, 非均衡, 不平衡, 训练, 模型, 分类, 语料, 少数, 多数, 损失函数, 调整, 反向传播算法, sample, inequality, training, model, classify, corpus, minority, majority, loss function, adjust, BP

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 110276075 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 24 September 2019 (2019-09-24) description paragraphs [0052]-[0076], [0160]	1-20
A	CN 106095928 A (NATIONAL COMPUTER NETWORK AND INFORMATION SECURITY MANAGEMENT CENTER) 09 November 2016 (2016-11-09) entire document	1-20
A	CN 110569508 A (CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 13 December 2019 (2019-12-13) entire document	1-20
A	CN 111310814 A (ZHIFUBAO (HANGZHOU) INFORMATION TECHNOLOGY CO., LTD.) 19 June 2020 (2020-06-19) entire document	1-20
A	CN 111400443 A (BEIJING PINECONE ELECTRONICS CO., LTD.) 10 July 2020 (2020-07-10) 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

08 September 2021

Date of mailing of the international search report

28 September 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088, China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/090432

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2020034482 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 30 January 2020 (2020-01-30) entire document	1-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/090432

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110276075	A	24 September 2019	None			
CN	106095928	A	09 November 2016	None			
CN	110569508	A	13 December 2019	None			
CN	111310814	A	19 June 2020	None			
CN	111400443	A	10 July 2020	None			
US	2020034482	A1	30 January 2020	WO	2020219369	A1	29 October 2020

国际检索报告

国际申请号

PCT/CN2021/090432

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, EPDOC, CNKI, IEEE: 样本, 非均衡, 不平均, 训练, 模型, 分类, 语料, 少数, 多数, 损失函数, 调整, 反向传播算法, sample, inequality, training, model, classify, corpus, minority, majority, loss function, adjust, BP		
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
A	CN 110276075 A (腾讯科技深圳有限公司) 2019年 9月 24日 (2019 - 09 - 24) 说明书第[0052]-[0076]、[0160]段	1-20
A	CN 106095928 A (国家计算机网络与信息安全管理中心) 2016年 11月 9日 (2016 - 11 - 09) 全文	1-20
A	CN 110569508 A (重庆邮电大学) 2019年 12月 13日 (2019 - 12 - 13) 全文	1-20
A	CN 111310814 A (支付宝杭州信息技术有限公司) 2020年 6月 19日 (2020 - 06 - 19) 全文	1-20
A	CN 111400443 A (北京松果电子有限公司) 2020年 7月 10日 (2020 - 07 - 10) 全文	1-20
A	US 2020034482 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 2020年 1月 30日 (2020 - 01 - 30) 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 2021年 9月 8日		国际检索报告邮寄日期 2021年 9月 28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 董显彬 电话号码 86-(10)-53961407

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/090432

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	110276075	A	2019年 9月 24日	无			
CN	106095928	A	2016年 11月 9日	无			
CN	110569508	A	2019年 12月 13日	无			
CN	111310814	A	2020年 6月 19日	无			
CN	111400443	A	2020年 7月 10日	无			
US	2020034482	A1	2020年 1月 30日	WO	2020219369	A1	2020年 10月 29日