



(12) 发明专利申请

(10) 申请公布号 CN 114996386 A

(43) 申请公布日 2022. 09. 02

(21) 申请号 202210583187.7

(22) 申请日 2022.05.25

(71) 申请人 中国平安财产保险股份有限公司
地址 518000 广东省深圳市福田区益田
路5033号平安金融中心12、13、38、39、
40、62层

(72) 发明人 李玉青

(74) 专利代理机构 深圳市沃德知识产权代理事
务所(普通合伙) 44347
专利代理师 高杰 于志光

(51) Int. Cl.

G06F 16/31 (2019.01)

G06F 16/35 (2019.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

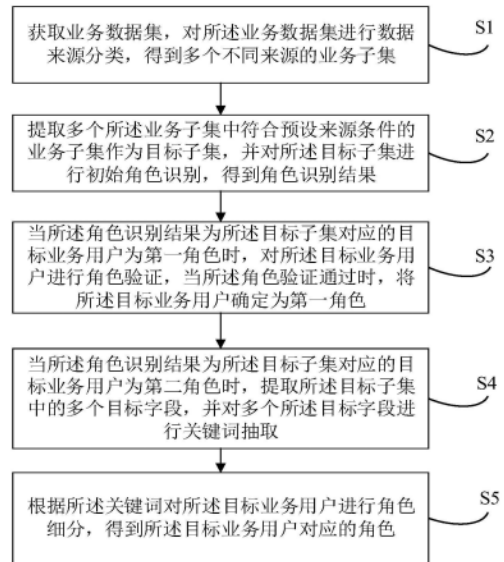
权利要求书2页 说明书12页 附图2页

(54) 发明名称

业务角色识别方法、装置、设备及存储介质

(57) 摘要

本发明涉及人工智能技术,揭露一种业务角色识别方法,包括:对业务数据集进行来源分类,得到多个不同来源的业务子集;提取多个业务子集中符合预设来源条件的业务子集作为目标子集,对目标子集进行初始角色识别的角色识别结果为目标业务用户为第一角色,对目标业务用户进行角色验证,当角色验证通过时,将目标业务用户确定为第一角色;角色识别结果为目标业务用户为第二角色,抽取多个目标子集中的目标字段的关键词;根据关键词对目标业务用户进行角色细分,得到目标业务用户的角色。此外,本发明还涉及区块链技术,目标子集可存储于区块链的节点。本发明还提出一种业务角色识别装置、电子设备以及存储介质。本发明可以提高业务角色识别的效率。



1. 一种业务角色识别方法,其特征在于,所述方法包括:

获取业务数据集,对所述业务数据集进行数据来源分类,得到多个不同来源的业务子集;

提取多个所述业务子集中符合预设来源条件的业务子集作为目标子集,并对所述目标子集进行初始角色识别,得到角色识别结果;

当所述角色识别结果为所述目标子集对应的目标业务用户为第一角色时,对所述目标业务用户进行角色验证,当所述角色验证通过时,将所述目标业务用户确定为第一角色;

当所述角色识别结果为所述目标子集对应的目标业务用户为第二角色时,提取所述目标子集中的多个目标字段,并对多个所述目标字段进行关键词抽取;

根据所述关键词对所述目标业务用户进行角色细分,得到所述目标业务用户对应的角色。

2. 如权利要求1所述的业务角色识别方法,其特征在于,所述对所述业务数据集进行数据来源分类,得到多个不同来源的业务子集,包括:

对所述业务数据集进行标准化处理,得到标准数据集;

利用预训练好的来源识别模型对所述标准数据集进行来源识别,得到所述标准数据集中多个标准数据对应的来源;

将所述来源相同的数据进行聚类,得到多个不同来源的业务子集。

3. 如权利要求2所述的业务角色识别方法,其特征在于,所述对所述业务数据集进行标准化处理,得到标准数据集,包括:

对所述业务数据集中的业务数据进行数据清洗,得到清洗数据集;

保留所述清洗数据集中符合预设的配置规则的数据作为标准数据集。

4. 如权利要求1所述的业务角色识别方法,其特征在于,所述根据所述关键词对所述目标业务用户进行角色细分,得到所述目标业务用户对应的角色,包括:

基于所述关键词对所述目标业务用户进行关键词标签标记;

在预设的关键词角色库中索引出所述关键词标签对应的关键角色,并将所述关键角色作为所述目标业务用户对应的角色。

5. 如权利要求1所述的业务角色识别方法,其特征在于,所述对多个所述目标字段进行关键词抽取,包括:

采集训练数据集,并构建深度神经网络;

利用所述训练数据集对所述深度神经网络进行训练,得到训练好的关键词抽取模型;

将所述目标字段输出至所述关键词抽取模型中,得到多个目标字段对应的关键词。

6. 如权利要求1所述的业务角色识别方法,其特征在于,所述对所述目标业务用户进行角色验证,包括:

基于所述第一角色在预设的角色信息参考库中获取对应的角色信息;

根据所述角色信息对所述目标业务用户进行信息验证,并在所述信息验证通过时将所述目标业务用户确定为第一角色。

7. 如权利要求1至6中任一项所述的业务角色识别方法,其特征在于,所述对所述目标子集进行初始角色识别,得到角色识别结果,包括:

将所述目标子集与预设的个人信息库中的数据进行比对,将与所述目标子集一致的数

据对应的个人角色作为所述目标子集对应的目标业务用户的初始角色；

根据所述初始角色的相关信息对所述目标子集进行初始验证，当所述初始验证通过时，确定所述初始角色为所述目标子集对应的目标业务用户对应的角色。

8. 一种业务角色识别装置，其特征在于，所述装置包括：

来源分类模块，用于获取业务数据集，对所述业务数据集进行数据来源分类，得到多个不同来源的业务子集；

初始识别模块，用于提取多个所述业务子集中符合预设来源条件的业务子集作为目标子集，并对所述目标子集进行初始角色识别，得到角色识别结果；

角色验证模块，用于当所述角色识别结果为所述目标子集对应的目标业务用户为第一角色时，对所述目标业务用户进行角色验证，当所述角色验证通过时，将所述目标业务用户确定为第一角色；

角色细分模块，用于当所述角色识别结果为所述目标子集对应的目标业务用户为第二角色时，提取所述目标子集中的多个目标字段，并对多个所述目标字段进行关键词抽取，根据所述关键词对所述目标业务用户进行角色细分，得到所述目标业务用户对应的角色。

9. 一种电子设备，其特征在于，所述电子设备包括：

至少一个处理器；以及，

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的计算机程序，所述计算机程序被所述至少一个处理器执行，以使所述至少一个处理器能够执行如权利要求1至7中任意一项所述的业务角色识别方法。

10. 一种计算机可读存储介质，存储有计算机程序，其特征在于，所述计算机程序被处理器执行时实现如权利要求1至7中任意一项所述的业务角色识别方法。

业务角色识别方法、装置、设备及存储介质

技术领域

[0001] 本发明涉及人工智能技术领域,尤其涉及一种业务角色识别方法、装置、电子设备及计算机可读存储介质。

背景技术

[0002] 在不同的业务领域中,出单都是业务中的一个关键部分,通过简化出单流程,减少出单操作人的工作量,提高出单效率,可以达到提高出单人员产能,降低公司经营成本的目的,而准确地识别出单人员的角色是简化出单流程的首要条件。

[0003] 传统识别业务人员角色的方法是通过业务角色数据库进行查询和比对,得到当前业务人员所属的角色,这种方法需要将当前业务人员与所述业务角色数据库中的数据进行逐个比对,进而确认业务角色,进而导致业务角色识别的效率较低。

发明内容

[0004] 本发明提供一种业务角色识别方法、装置及计算机可读存储介质,其主要目的在于提高业务角色识别的效率。

[0005] 为实现上述目的,本发明提供的一种业务角色识别方法,包括:

[0006] 获取业务数据集,对所述业务数据集进行数据来源分类,得到多个不同来源的业务子集;

[0007] 提取多个所述业务子集中符合预设来源条件的业务子集作为目标子集,并对所述目标子集进行初始角色识别,得到角色识别结果;

[0008] 当所述角色识别结果为所述目标子集对应的目标业务用户为第一角色时,对所述目标业务用户进行角色验证,当所述角色验证通过时,将所述目标业务用户确定为第一角色;

[0009] 当所述角色识别结果为所述目标子集对应的目标业务用户为第二角色时,提取所述目标子集中的多个目标字段,并对多个所述目标字段进行关键词抽取;

[0010] 根据所述关键词对所述目标业务用户进行角色细分,得到所述目标业务用户对应的角色。

[0011] 可选地,所述对所述业务数据集进行数据来源分类,得到多个不同来源的业务子集,包括:

[0012] 对所述业务数据集进行标准化处理,得到标准数据集;

[0013] 利用预训练好的来源识别模型对所述标准数据集进行来源识别,得到所述标准数据集中多个标准数据对应的来源;

[0014] 将所述来源相同的数据进行聚类,得到多个不同来源的业务子集。

[0015] 可选地,所述对所述业务数据集进行标准化处理,得到标准数据集,包括:

[0016] 对所述业务数据集中的业务数据进行数据清洗,得到清洗数据集;

[0017] 保留所述清洗数据集中符合预设的配置规则的数据作为标准数据集。

- [0018] 可选地,所述根据所述关键词对所述目标业务用户进行角色细分,得到所述目标业务用户对应的角色,包括:
- [0019] 基于所述关键词对所述目标业务用户进行关键词标签标记;
- [0020] 在预设的关键词角色库中索引出所述关键词标签对应的关键角色,并将所述关键角色作为所述目标业务用户对应的角色。
- [0021] 可选地,所述对多个所述目标字段进行关键词抽取,包括:
- [0022] 采集训练数据集,并构建深度神经网络;
- [0023] 利用所述训练数据集对所述深度神经网络进行训练,得到训练好的关键词抽取模型;
- [0024] 将所述目标字段输出至所述关键词抽取模型中,得到多个目标字段对应的关键词。
- [0025] 可选地,所述对所述目标业务用户进行角色验证,包括:
- [0026] 基于所述第一角色在预设的角色信息参考库中获取对应的角色信息;
- [0027] 根据所述角色信息对所述目标业务用户进行信息验证,并在所述信息验证通过时将所述目标业务用户确定为第一角色。
- [0028] 可选地,所述对所述目标子集进行初始角色识别,得到角色识别结果,包括:
- [0029] 将所述目标子集与预设的个人信息库中的数据进行比对,将与所述目标子集一致的数据对应的个人角色作为所述目标子集对应的目标业务用户的初始角色;
- [0030] 根据所述初始角色的相关信息对所述目标子集进行初始验证,当所述初始验证通过时,确定所述初始角色为所述目标子集对应的目标业务用户对应的角色。
- [0031] 为了解决上述问题,本发明还提供一种业务角色识别装置,所述装置包括:
- [0032] 来源分类模块,用于获取业务数据集,对所述业务数据集进行数据来源分类,得到多个不同来源的业务子集;
- [0033] 初始识别模块,用于提取多个所述业务子集中符合预设来源条件的业务子集作为目标子集,并对所述目标子集进行初始角色识别,得到角色识别结果;
- [0034] 角色验证模块,用于当所述角色识别结果为所述目标子集对应的目标业务用户为第一角色时,对所述目标业务用户进行角色验证,当所述角色验证通过时,将所述目标业务用户确定为第一角色;
- [0035] 角色细分模块,用于当所述角色识别结果为所述目标子集对应的目标业务用户为第二角色时,提取所述目标子集中的多个目标字段,并对多个所述目标字段进行关键词抽取,根据所述关键词对所述目标业务用户进行角色细分,得到所述目标业务用户对应的角色。
- [0036] 为了解决上述问题,本发明还提供一种电子设备,所述电子设备包括:
- [0037] 至少一个处理器;以及,
- [0038] 与所述至少一个处理器通信连接的存储器;其中,
- [0039] 所述存储器存储有可被所述至少一个处理器执行的计算机程序,所述计算机程序被所述至少一个处理器执行,以使所述至少一个处理器能够执行上述所述的业务角色识别方法。
- [0040] 为了解决上述问题,本发明还提供一种计算机可读存储介质,所述计算机可读存

储介质中存储有至少一个计算机程序,所述至少一个计算机程序被电子设备中的处理器执行以实现上述所述的业务角色识别方法。

[0041] 本发明实施例通过对业务数据集进行数据来源分类,得到多个不同来源的业务子集,所述数据来源分类可以作为业务数据集进行初步划分的依据,并提取多个所述业务子集中符合预设来源条件的业务子集作为目标子集,保证了目标子集的来源的准确性。对目标子集进行初始角色识别,得到角色识别结果,并根据角色识别结果分别进行角色验证或者角色细分,可以在保证角色识别的准确度的情况下提高角色识别的效率。因此本发明提出的业务角色识别方法、装置、电子设备及计算机可读存储介质,可以实现解决业务角色识别的效率不够高的问题。

附图说明

[0042] 图1为本发明一实施例提供的业务角色识别方法的流程示意图;

[0043] 图2为本发明一实施例提供的业务角色识别装置的功能模块图;

[0044] 图3为本发明一实施例提供的实现所述业务角色识别方法的电子设备的结构示意图。

[0045] 本发明目的的实现、功能特点及优点将结合实施例,参照附图做进一步说明。

具体实施方式

[0046] 应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0047] 本申请实施例提供一种业务角色识别方法。所述业务角色识别方法的执行主体包括但不限于服务端、终端等能够被配置为执行本申请实施例提供的该方法的电子设备中的至少一种。换言之,所述业务角色识别方法可以由安装在终端设备或服务端设备的软件或硬件来执行,所述软件可以是区块链平台。所述服务端包括但不限于:单台服务器、服务器集群、云端服务器或云端服务器集群等。所述服务器可以是独立的服务器,也可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、内容分发网络(Content Delivery Network, CDN)、以及大数据和人工智能平台等基础云计算服务的云服务器。

[0048] 参照图1所示,为本发明一实施例提供的业务角色识别方法的流程示意图。在本实施例中,所述业务角色识别方法包括:

[0049] S1、获取业务数据集,对所述业务数据集进行数据来源分类,得到多个不同来源的业务子集。

[0050] 本发明实例中,所述业务数据集中包含来自公司内部、公司外部对接平台以及集体内对接平台等不同系统来源的出单数据。由于不同系统来源的数据由不同的业务人员操作,因此可以将系统来源作为角色识别的初始参考。

[0051] 具体地,所述对所述业务数据集进行数据来源分类,得到多个不同来源的业务子集,包括:

[0052] 对所述业务数据集进行标准化处理,得到标准数据集;

[0053] 利用预训练好的来源识别模型对所述标准数据集进行来源识别,得到所述标准数据集中多个标准数据对应的来源;

- [0054] 将所述来源相同的数据进行聚类,得到多个不同来源的业务子集。
- [0055] 进一步地,所述对所述业务数据集进行标准化处理,得到标准数据集,包括:
- [0056] 对所述业务数据集中的业务数据进行数据清洗,得到清洗数据集;
- [0057] 保留所述清洗数据集中符合预设的配置规则的数据作为标准数据集。
- [0058] 详细地,所述数据清洗包括对所述业务数据集中的异常值执行删除操作,对所述业务数据集中的缺值执行补全操作。其中,可以利用具有异常检测的java语句对所述业务数据进行异常检测。当所述业务数据集中存在缺失值时,本发明实施例可以采用现有的缺失值填充方法对所述业务数据集进行补全操作,现有的缺失值填充方法包括但不限于填充默认值、均值、众数、KNN填充。
- [0059] 其中,所述预设的配置规则是根据具体情况指定的,在本方案中,所述配置规则是将业务数据中的空白字段执行删除。
- [0060] 具体地,所述利用预训练好的来源识别模型对所述标准数据集进行来源识别之前,所述方法还包括:
- [0061] 将所述标准数据集输入至预设的来源识别模型中进行来源预测,得到预测来源标签;
- [0062] 将所述预测来源标签和预设的真实来源标签进行比对;
- [0063] 当所述预测来源标签和所述真实来源标签一致时,将所述来源识别模型输出为预训练好的来源识别模型;
- [0064] 当所述预测来源标签和所述真实来源标签不一致时,对所述来源识别模型进行模型参数调整,得到调整后的来源识别模型;
- [0065] 将所述标准数据集输入至调整后的来源识别模型中进行来源预测,得到新的预测来源标签,当所述新的预测来源标签与所述真实来源标签一致时,将调整后的来源识别模型为训练好的来源识别模型。
- [0066] 详细地,所述预设的来源识别模型可以为卷积神经网络模型。
- [0067] 具体地,所述将所述标准数据集输入至预设的来源识别模型中进行来源预测,得到预测来源标签,包括:
- [0068] 利用所述来源识别模型中的卷积层和池化层对所述标准数据集进行卷积处理及池化处理,得到特征数据集;
- [0069] 将所述特征数据集输入至预设的激活函数中,得到所述特征数据集对应的激活概率;
- [0070] 根据所述激活概率和预设的标签参考表得到对应的预测来源标签。
- [0071] 进一步地,将所述来源相同的数据进行聚类,得到多个不同来源的业务子集。
- [0072] S2、提取多个所述业务子集中符合预设来源条件的业务子集作为目标子集,并对所述目标子集进行初始角色识别,得到角色识别结果。
- [0073] 本发明实施例中,所述预设的来源条件可以为公司内部,提取多个所述业务子集中符合预设来源条件的业务子集作为目标子集即提取多个所述业务子集中数据来源为公司内部的业务子集作为目标子集。
- [0074] 具体地,所述对所述目标子集进行初始角色识别,得到角色识别结果,包括:
- [0075] 将所述目标子集与预设的个人信息库中的数据进行比对,将与所述目标子集一致

的数据对应的个人角色作为所述目标子集对应的目标业务用户的初始角色；

[0076] 根据所述初始角色的相关信息对所述目标子集进行初始验证,当所述初始验证通过时,确定所述初始角色为所述目标子集对应的目标业务用户对应的角色。

[0077] 详细地,所述个人信息库中包含多个员工信息,所述初始角色的相关信息是有关于所述初始角色的基本身份信息等。

[0078] 例如,将所述目标子集与预设的个人信息库中的数据进行比对,员工信息中的员工所属业务条线判断员工是业务员还是内勤;分别用员工信息表中的员工UM账号、员工中文名称、关联录单操作人等相关信息优先判断录单人员的角色,比如是否为出单内勤,或者业务员。

[0079] S3、当所述角色识别结果为所述目标子集对应的目标业务用户为第一角色时,对所述目标业务用户进行角色验证,当所述角色验证通过时,将所述目标业务用户确定为第一角色。

[0080] 本发明实施例中,所述第一角色为业务员。即当所述角色识别结果为所述目标子集对应的目标业务用户为第一角色时,对所述目标业务用户进行角色验证。

[0081] 具体地,所述对所述目标业务用户进行角色验证,包括:

[0082] 基于所述第一角色在预设的角色信息参考库中获取对应的角色信息;

[0083] 根据所述角色信息对所述目标业务用户进行信息验证,并在所述信息验证通过时将所述目标业务用户确定为第一角色。

[0084] 详细地,所述角色信息参考库中包含不同角色及不同角色对应的角色信息。

[0085] S4、当所述角色识别结果为所述目标子集对应的目标业务用户为第二角色时,提取所述目标子集中的多个目标字段,并对多个所述目标字段进行关键词抽取。

[0086] 本发明实施例中,所述第二角色可以为客户。而由于客户又可以划分为代理人或者经纪人等不同角色,因此可以进行进一步地细分。

[0087] 具体地,所述提取所述目标子集中的多个目标字段,包括:

[0088] 查找到所述目标子集中的名称索引;

[0089] 将所述名称索引下的字段作为所述目标子集中的多个目标字段。

[0090] 详细地,将所述目标子集中的名称索引下面包含的数据作为所述目标子集中的多个目标字段,例如,所述名称索引为“客户名称”,所述目标字段为代理人张三。

[0091] 进一步地,所述对多个所述目标字段进行关键词抽取,包括:

[0092] 采集训练数据集,并构建深度神经网络;

[0093] 利用所述训练数据集对所述深度神经网络进行训练,得到训练好的关键词抽取模型;

[0094] 将所述目标字段输出至所述关键词抽取模型中,得到多个目标字段对应的关键词。

[0095] 具体地,可以利用爬虫技术采集得到训练数据集。将所述目标字段输出至所述关键词抽取模型中,得到多个目标字段对应的关键词,可以为关键词“代理”,关键词“经济”。

[0096] S5、根据所述关键词对所述目标业务用户进行角色细分,得到所述目标业务用户对应的角色。

[0097] 本发明实施例中,所述根据所述关键词对所述目标业务用户进行角色细分,得到

所述目标业务用户对应的角色,包括:

[0098] 基于所述关键词对所述目标业务用户进行关键词标签标记;

[0099] 在预设的关键词角色库中索引出所述关键词标签对应的关键角色,并将所述关键角色作为所述目标业务用户对应的角色。

[0100] 详细地,所述关键词可以为“代理”或者“经济”,则将关键词“代理”,关键词“经济”作为所述目标业务用户的关键词标签,所述关键词角色库中包含了多个涉及关键词标签的角色名,类似于“代理人”或者“经纪人”。在预设的关键词角色库中索引出所述关键词标签“代理”对应的关键角色“代理人”,并将所述关键角色“代理人”作为所述目标业务用户对应的角色。

[0101] 本发明实施例通过对业务数据集进行数据来源分类,得到多个不同来源的业务子集,所述数据来源分类可以作为业务数据集进行初步划分的依据,并提取多个所述业务子集中符合预设来源条件的业务子集作为目标子集,保证了目标子集的来源的准确性。对目标子集进行初始角色识别,得到角色识别结果,并根据角色识别结果分别进行角色验证或者角色细分,可以在保证角色识别的准确度的情况下提高角色识别的效率。因此本发明提出的业务角色识别方法可以实现解决业务角色识别的效率不够高的问题。

[0102] 如图2所示,是本发明一实施例提供的业务角色识别装置的功能模块图。

[0103] 本发明所述业务角色识别装置100可以安装于电子设备中。根据实现的功能,所述业务角色识别装置100可以包括来源分类模块101、初始识别模块102、角色验证模块103及角色细分模块104。本发明所述模块也可以称之为单元,是指一种能够被电子设备处理器所执行,并且能够完成固定功能的一系列计算机程序段,其存储在电子设备的存储器中。

[0104] 在本实施例中,关于各模块/单元的功能如下:

[0105] 所述来源分类模块101,用于获取业务数据集,对所述业务数据集进行数据来源分类,得到多个不同来源的业务子集;

[0106] 所述初始识别模块102,用于提取多个所述业务子集中符合预设来源条件的业务子集作为目标子集,并对所述目标子集进行初始角色识别,得到角色识别结果;

[0107] 所述角色验证模块103,用于当所述角色识别结果为所述目标子集对应的目标业务用户为第一角色时,对所述目标业务用户进行角色验证,当所述角色验证通过时,将所述目标业务用户确定为第一角色;

[0108] 所述角色细分模块104,用于当所述角色识别结果为所述目标子集对应的目标业务用户为第二角色时,提取所述目标子集中的多个目标字段,并对多个所述目标字段进行关键词抽取,根据所述关键词对所述目标业务用户进行角色细分,得到所述目标业务用户对应的角色。

[0109] 详细地,所述业务角色识别装置100各模块的具体实施方式如下:

[0110] 步骤一、获取业务数据集,对所述业务数据集进行数据来源分类,得到多个不同来源的业务子集。

[0111] 本发明实例中,所述业务数据集中包含来自公司内部、公司外部对接平台以及集体内对接平台等不同系统来源的出单数据。由于不同系统来源的数据由不同的业务人员操作,因此可以将系统来源作为角色识别的初始参考。

[0112] 具体地,所述对所述业务数据集进行数据来源分类,得到多个不同来源的业务子

集,包括:

[0113] 对所述业务数据集进行标准化处理,得到标准数据集;

[0114] 利用预训练好的来源识别模型对所述标准数据集进行来源识别,得到所述标准数据集中多个标准数据对应的来源;

[0115] 将所述来源相同的数据进行聚类,得到多个不同来源的业务子集。

[0116] 进一步地,所述对所述业务数据集进行标准化处理,得到标准数据集,包括:

[0117] 对所述业务数据集中的业务数据进行数据清洗,得到清洗数据集;

[0118] 保留所述清洗数据集中符合预设的配置规则的数据作为标准数据集。

[0119] 详细地,所述数据清洗包括对所述业务数据集中的异常值执行删除操作,对所述业务数据集中的缺值执行补全操作。其中,可以利用具有异常检测的java语句对所述业务数据进行异常检测。当所述业务数据集中存在缺失值时,本发明实施例可以采用现有的缺失值填充方法对所述业务数据集进行补全操作,现有的缺失值填充方法包括但不限于填充默认值、均值、众数、KNN填充。

[0120] 其中,所述预设的配置规则是根据具体情况指定的,在本方案中,所述配置规则是将业务数据中的空白字段执行删除。

[0121] 具体地,所述利用预训练好的来源识别模型对所述标准数据集进行来源识别之前,还执行:

[0122] 将所述标准数据集输入至预设的来源识别模型中进行来源预测,得到预测来源标签;

[0123] 将所述预测来源标签和预设的真实来源标签进行比对;

[0124] 当所述预测来源标签和所述真实来源标签一致时,将所述来源识别模型输出为预训练好的来源识别模型;

[0125] 当所述预测来源标签和所述真实来源标签不一致时,对所述来源识别模型进行模型参数调整,得到调整后的来源识别模型;

[0126] 将所述标准数据集输入至调整后的来源识别模型中进行来源预测,得到新的预测来源标签,当所述新的预测来源标签与所述真实来源标签一致时,将调整后的来源识别模型为训练好的来源识别模型。

[0127] 详细地,所述预设的来源识别模型可以为卷积神经网络模型。

[0128] 具体地,所述将所述标准数据集输入至预设的来源识别模型中进行来源预测,得到预测来源标签,包括:

[0129] 利用所述来源识别模型中的卷积层和池化层对所述标准数据集进行卷积处理及池化处理,得到特征数据集;

[0130] 将所述特征数据集输入至预设的激活函数中,得到所述特征数据集对应的激活概率;

[0131] 根据所述激活概率和预设的标签参考表得到对应的预测来源标签。

[0132] 进一步地,将所述来源相同的数据进行聚类,得到多个不同来源的业务子集。

[0133] 步骤二、提取多个所述业务子集中符合预设来源条件的业务子集作为目标子集,并对所述目标子集进行初始角色识别,得到角色识别结果。

[0134] 本发明实施例中,所述预设的来源条件可以为公司内部,提取多个所述业务子集

中符合预设来源条件的业务子集作为目标子集即提取多个所述业务子集中数据来源为公司内部的业务子集作为目标子集。

[0135] 具体地,所述对所述目标子集进行初始角色识别,得到角色识别结果,包括:

[0136] 将所述目标子集与预设的个人信息库中的数据进行比对,将与所述目标子集一致的数据对应的个人角色作为所述目标子集对应的目标业务用户的初始角色;

[0137] 根据所述初始角色的相关信息对所述目标子集进行初始验证,当所述初始验证通过时,确定所述初始角色为所述目标子集对应的目标业务用户对应的角色。

[0138] 详细地,所述个人信息库中包含多个员工信息,所述初始角色的相关信息是有关于所述初始角色的基本身份信息等。

[0139] 例如,将所述目标子集与预设的个人信息库中的数据进行比对,员工信息中的员工所属业务条线判断员工是业务员还是内勤;分别用员工信息表中的员工UM账号、员工中文名称、关联录单操作人等相关信息优先判断录单人员的角色,比如是否为出单内勤,或者业务员。

[0140] 步骤三、当所述角色识别结果为所述目标子集对应的目标业务用户为第一角色时,对所述目标业务用户进行角色验证,当所述角色验证通过时,将所述目标业务用户确定为第一角色。

[0141] 本发明实施例中,所述第一角色为业务员。即当所述角色识别结果为所述目标子集对应的目标业务用户为第一角色时,对所述目标业务用户进行角色验证。

[0142] 具体地,所述对所述目标业务用户进行角色验证,包括:

[0143] 基于所述第一角色在预设的角色信息参考库中获取对应的角色信息;

[0144] 根据所述角色信息对所述目标业务用户进行信息验证,并在所述信息验证通过时将所述目标业务用户确定为第一角色。

[0145] 详细地,所述角色信息参考库中包含不同角色及不同角色对应的角色信息。

[0146] 步骤四、当所述角色识别结果为所述目标子集对应的目标业务用户为第二角色时,提取所述目标子集中的多个目标字段,并对多个所述目标字段进行关键词抽取。

[0147] 本发明实施例中,所述第二角色可以为客户。而由于客户又可以划分为代理人或者经纪人等不同角色,因此可以进行进一步地细分。

[0148] 具体地,所述提取所述目标子集中的多个目标字段,包括:

[0149] 查找到所述目标子集中的名称索引;

[0150] 将所述名称索引下的字段作为所述目标子集中的多个目标字段。

[0151] 详细地,将所述目标子集中的名称索引下面包含的数据作为所述目标子集中的多个目标字段,例如,所述名称索引为“客户名称”,所述目标字段为代理人张三。

[0152] 进一步地,所述对多个所述目标字段进行关键词抽取,包括:

[0153] 采集训练数据集,并构建深度神经网络;

[0154] 利用所述训练数据集对所述深度神经网络进行训练,得到训练好的关键词抽取模型;

[0155] 将所述目标字段输出至所述关键词抽取模型中,得到多个目标字段对应的关键词。

[0156] 具体地,可以利用爬虫技术采集得到训练数据集。将所述目标字段输出至所述关

关键词抽取模型中,得到多个目标字段对应的关键词,可以为关键词“代理”,关键词“经济”。

[0157] 步骤五、根据所述关键词对所述目标业务用户进行角色细分,得到所述目标业务用户对应的角色。

[0158] 本发明实施例中,所述根据所述关键词对所述目标业务用户进行角色细分,得到所述目标业务用户对应的角色,包括:

[0159] 基于所述关键词对所述目标业务用户进行关键词标签标记;

[0160] 在预设的关键词角色库中索引出所述关键词标签对应的关键角色,并将所述关键角色作为所述目标业务用户对应的角色。

[0161] 详细地,所述关键词可以为“代理”或者“经济”,则将关键词“代理”,关键词“经济”作为所述目标业务用户的关键词标签,所述关键词角色库中包含了多个涉及关键词标签的角色名,类似于“代理人”或者“经纪人”。在预设的关键词角色库中索引出所述关键词标签“代理”对应的关键角色“代理人”,并将所述关键角色“代理人”作为所述目标业务用户对应的角色。

[0162] 本发明实施例通过对业务数据集进行数据来源分类,得到多个不同来源的业务子集,所述数据来源分类可以作为业务数据集进行初步划分的依据,并提取多个所述业务子集中符合预设来源条件的业务子集作为目标子集,保证了目标子集的来源的准确性。对目标子集进行初始角色识别,得到角色识别结果,并根据角色识别结果分别进行角色验证或者角色细分,可以在保证角色识别的准确度的情况下提高角色识别的效率。因此本发明提出的业务角色识别装置可以实现解决业务角色识别的效率不够高的问题。

[0163] 如图3所示,是本发明一实施例提供的实现业务角色识别方法的电子设备的结构示意图。

[0164] 所述电子设备1可以包括处理器10、存储器11、通信总线12以及通信接口13,还可以包括存储在所述存储器11中并可在所述处理器10上运行的计算机程序,如业务角色识别程序。

[0165] 其中,所述处理器10在一些实施例中可以由集成电路组成,例如可以由单个封装的集成电路所组成,也可以是由多个相同功能或不同功能封装的集成电路所组成,包括一个或者多个中央处理器(Central Processing unit,CPU)、微处理器、数字处理芯片、图形处理器及各种控制芯片的组合等。所述处理器10是所述电子设备的控制核心(Control Unit),利用各种接口和线路连接整个电子设备的各个部件,通过运行或执行存储在所述存储器11内的程序或者模块(例如执行业务角色识别程序等),以及调用存储在所述存储器11内的数据,以执行电子设备的各种功能和处理数据。

[0166] 所述存储器11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、移动硬盘、多媒体卡、卡型存储器(例如:SD或DX存储器等)、磁性存储器、磁盘、光盘等。所述存储器11在一些实施例中可以是电子设备的内部存储单元,例如该电子设备的移动硬盘。所述存储器11在另一些实施例中也可以是电子设备的外部存储设备,例如电子设备上配备的插接式移动硬盘、智能存储卡(Smart Media Card,SMC)、安全数字(Secure Digital,SD)卡、闪存卡(Flash Card)等。进一步地,所述存储器11还可以既包括电子设备的内部存储单元也包括外部存储设备。所述存储器11不仅可以用于存储安装于电子设备的应用软件及各类数据,例如业务角色识别程序的代码等,还可以用于暂时地存储已经输出或者将要输出

的数据。

[0167] 所述通信总线12可以是外设部件互连标准(peripheral component interconnect,简称PCI)总线或扩展工业标准结构(extended industry standard architecture,简称EISA)总线等。该总线可以分为地址总线、数据总线、控制总线等。所述总线被设置为实现所述存储器11以及至少一个处理器10等之间的连接通信。

[0168] 所述通信接口13用于上述电子设备与其他设备之间的通信,包括网络接口和用户接口。可选地,所述网络接口可以包括有线接口和/或无线接口(如WI-FI接口、蓝牙接口等),通常用于在该电子设备与其他电子设备之间建立通信连接。所述用户接口可以是显示器(Display)、输入单元(比如键盘(Keyboard)),可选地,用户接口还可以是标准的有线接口、无线接口。可选地,在一些实施例中,显示器可以是LED显示器、液晶显示器、触控式液晶显示器以及OLED(Organic Light-Emitting Diode,有机发光二极管)触摸器等。其中,显示器也可以适当的称为显示屏或显示单元,用于显示在电子设备中处理的信息以及用于显示可视化的用户界面。

[0169] 图3仅示出了具有部件的电子设备,本领域技术人员可以理解的是,图3示出的结构并不构成对所述电子设备1的限定,可以包括比图示更少或者更多的部件,或者组合某些部件,或者不同的部件布置。

[0170] 例如,尽管未示出,所述电子设备还可以包括给各个部件供电的电源(比如电池),优选地,电源可以通过电源管理装置与所述至少一个处理器10逻辑相连,从而通过电源管理装置实现充电管理、放电管理、以及功耗管理等功能。电源还可以包括一个或一个以上的直流或交流电源、再充电装置、电源故障检测电路、电源转换器或者逆变器、电源状态指示器等任意组件。所述电子设备还可以包括多种传感器、蓝牙模块、Wi-Fi模块等,在此不再赘述。

[0171] 应该了解,所述实施例仅为说明之用,在专利申请范围上并不受此结构的限制。

[0172] 所述电子设备1中的所述存储器11存储的业务角色识别程序是多个指令的组合,在所述处理器10中运行时,可以实现:

[0173] 获取业务数据集,对所述业务数据集进行数据来源分类,得到多个不同来源的业务子集;

[0174] 提取多个所述业务子集中符合预设来源条件的业务子集作为目标子集,并对所述目标子集进行初始角色识别,得到角色识别结果;

[0175] 当所述角色识别结果为所述目标子集对应的目标业务用户为第一角色时,对所述目标业务用户进行角色验证,当所述角色验证通过时,将所述目标业务用户确定为第一角色;

[0176] 当所述角色识别结果为所述目标子集对应的目标业务用户为第二角色时,提取所述目标子集中的多个目标字段,并对多个所述目标字段进行关键词抽取;

[0177] 根据所述关键词对所述目标业务用户进行角色细分,得到所述目标业务用户对应的角色。

[0178] 具体地,所述处理器10对上述指令的具体实现方法可参考附图对应实施例中相关步骤的描述,在此不赘述。

[0179] 进一步地,所述电子设备1集成的模块/单元如果以软件功能单元的形式实现并作

为独立的产品销售或使用，可以存储在一个计算机可读存储介质中。所述计算机可读存储介质可以是易失性的，也可以是非易失性的。例如，所述计算机可读介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器 (ROM, Read-Only Memory)。

[0180] 本发明还提供一种计算机可读存储介质，所述可读存储介质存储有计算机程序，所述计算机程序在被电子设备的处理器所执行时，可以实现：

[0181] 获取业务数据集，对所述业务数据集进行数据来源分类，得到多个不同来源的业务子集；

[0182] 提取多个所述业务子集中符合预设来源条件的业务子集作为目标子集，并对所述目标子集进行初始角色识别，得到角色识别结果；

[0183] 当所述角色识别结果为所述目标子集对应的目标业务用户为第一角色时，对所述目标业务用户进行角色验证，当所述角色验证通过时，将所述目标业务用户确定为第一角色；

[0184] 当所述角色识别结果为所述目标子集对应的目标业务用户为第二角色时，提取所述目标子集中的多个目标字段，并对多个所述目标字段进行关键词抽取；

[0185] 根据所述关键词对所述目标业务用户进行角色细分，得到所述目标业务用户对应的角色。

[0186] 在本发明所提供的几个实施例中，应该理解到，所揭露的设备、装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述模块的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式。

[0187] 所述作为分离部件说明的模块可以是或者也可以不是物理上分开的，作为模块显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。

[0188] 另外，在本发明各个实施例中的各功能模块可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用硬件加软件功能模块的形式实现。

[0189] 对于本领域技术人员而言，显然本发明不限于上述示范性实施例的细节，而且在不背离本发明的精神或基本特征的情况下，能够以其他的具体形式实现本发明。

[0190] 因此，无论从哪一点来看，均应将实施例看作是示范性的，而且是非限制性的，本发明的范围由所附权利要求而不是上述说明限定，因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本发明内。不应将权利要求中的任何附关联图标记视为限制所涉及的权利要求。

[0191] 本发明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。

[0192] 本申请实施例可以基于人工智能技术对相关的数据进行获取和处理。其中，人工

智能(Artificial Intelligence, AI) 是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能,感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。

[0193] 此外,显然“包括”一词不排除其他单元或步骤,单数不排除复数。系统权利要求中陈述的多个单元或装置也可以由一个单元或装置通过软件或者硬件来实现。第一、第二等词语用来表示名称,而并不表示任何特定的顺序。

[0194] 最后应说明的是,以上实施例仅用以说明本发明的技术方案而非限制,尽管参照较佳实施例对本发明进行了详细说明,本领域的普通技术人员应当理解,可以对本发明的技术方案进行修改或等同替换,而不脱离本发明技术方案的精神和范围。

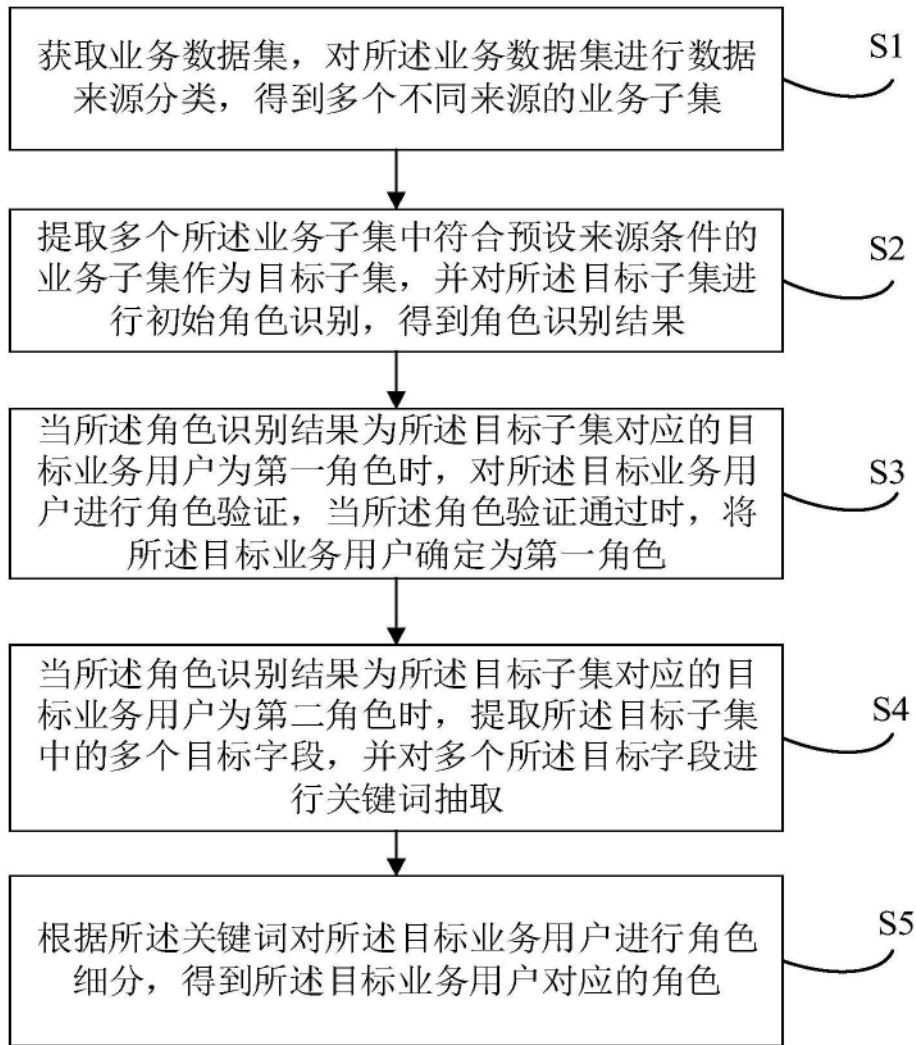


图1

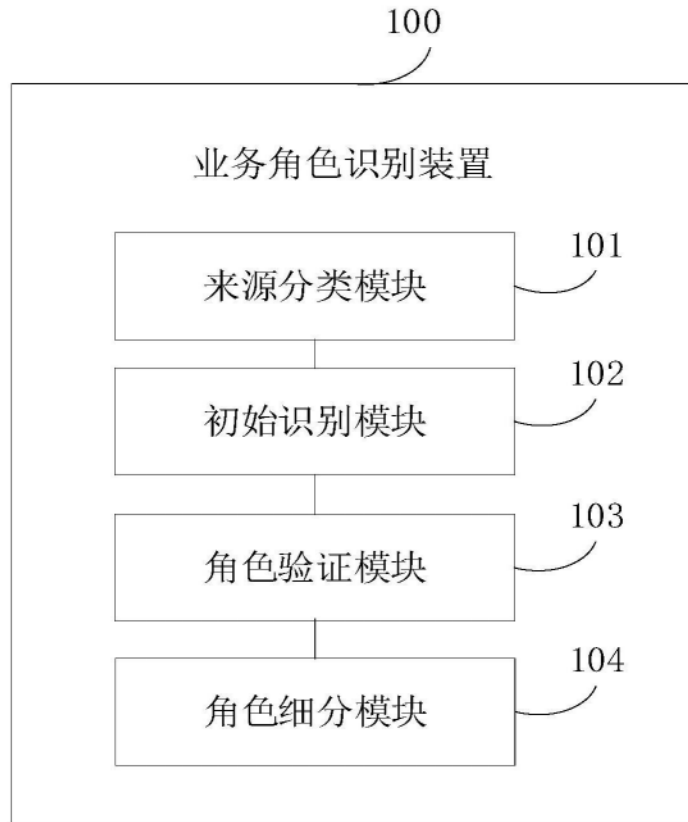


图2

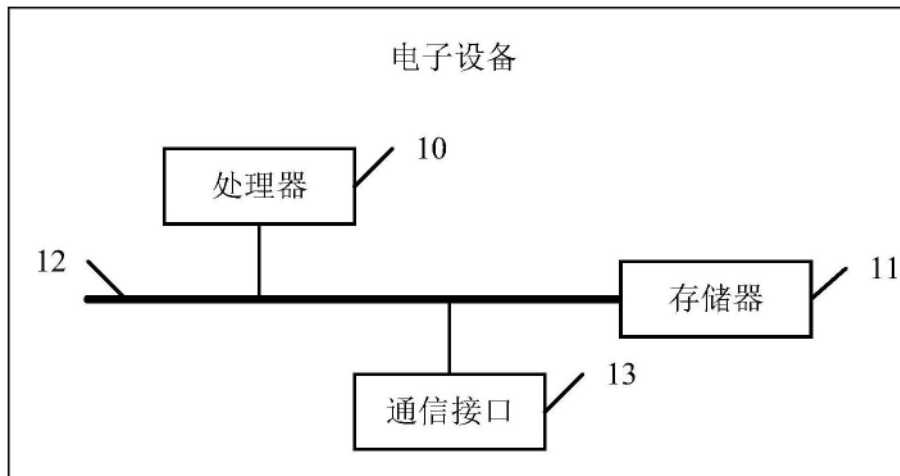


图3