



## (12) 发明专利申请

(10) 申请公布号 CN 117688428 A

(43) 申请公布日 2024. 03. 12

(21) 申请号 202311834345.2

(22) 申请日 2023.12.28

(71) 申请人 杭州富算科技有限公司

地址 310051 浙江省杭州市滨江区西兴街  
道缤纷街615号4楼401室(72) 发明人 李腾飞 陈立峰 孙小超 赵华宇  
卫骞 杜浩 尤志强 卞阳  
张伟奇(74) 专利代理机构 北京慧加伦知识产权代理有  
限公司 16035

专利代理师 白玉娜

(51) Int. Cl.

G06F 18/24 (2023.01)

G06F 18/213 (2023.01)

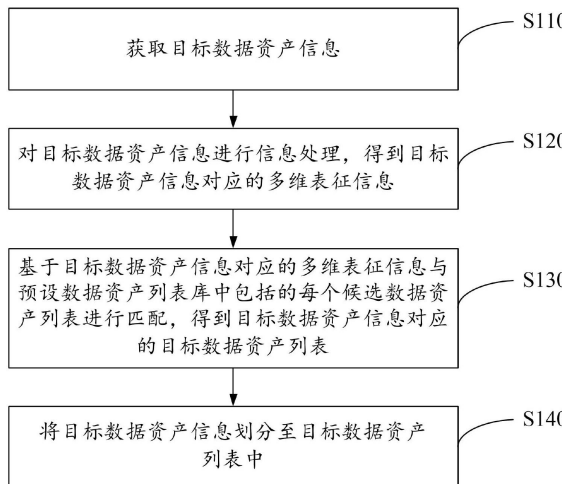
权利要求书3页 说明书9页 附图2页

## (54) 发明名称

多维度数据资产分类方法、装置、计算机设  
备和介质

## (57) 摘要

本公开提供一种多维度数据资产分类方法、装置、计算机设备和介质,包括:获取目标数据资产信息,目标数据资产信息包括第一维度数据资产、第二维度数据资产和第三维度数据资产;对目标数据资产信息进行信息处理,得到目标数据资产信息对应的多维表征信息;基于目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每个候选数据资产列表进行匹配,得到目标数据资产信息对应的目标数据资产列表;将目标数据资产信息划分至目标数据资产列表中。从而,通过数据资产的多维度信息进行信息处理,便于采用处理得到的多维度表征信息与多个数据资产进行匹配,有效识别出目标数据资产信息所属的多维分类,实现数据资产的精细化分类。



1. 一种多维度数据资产分类方法,其特征在于,包括:

获取目标数据资产信息,所述目标数据资产信息包括第一维度数据资产、第二维度数据资产和第三维度数据资产,所述第一维度数据资产用于描述数据特征信息,所述第二维度数据资产用于描述数据参与方提供的数据属性信息,所述第三维度数据资产用于描述数据分级分类信息;

对所述目标数据资产信息进行信息处理,得到所述目标数据资产信息对应的多维表征信息,所述多维表征信息用于表征所述第一维度数据资产、所述第二维度数据资产和所述第三维度数据资产;

基于所述目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每个候选数据资产列表进行匹配,得到所述目标数据资产信息对应的目标数据资产列表,所述预设数据资产列表中存储着多个多维度的候选数据资产列表,每个所述候选数据资产列表对应一种多维分类的数据资产;

将所述目标数据资产信息划分至所述目标数据资产列表中。

2. 根据权利要求1所述的方法,其特征在于,所述对所述目标数据资产信息进行信息处理,得到所述目标数据资产信息对应的多维表征信息,包括:

获取信息预处理模版;

基于所述信息预处理模版对所述目标数据资产信息进行处理,得到标签提示词,所述标签提示词用于描述所述第一维度数据资产、所述第二维度数据资产和所述第三维度数据资产中的部分数据资产或全部数据资产;

基于目标大语言模型对所述标签提示词进行信息处理,得到所述目标数据资产信息对应的多维表征信息。

3. 根据权利要求1所述的方法,其特征在于,所述预设数据资产列表库中包括的每个候选数据资产列表对应有标识表征信息;

所述基于所述目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每个候选数据资产列表进行匹配,得到所述目标数据资产信息对应的目标数据资产列表,包括:

匹配所述目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每个候选数据资产列表对应的标识表征信息,得到所述目标数据资产信息对应的多维表征信息与每个所述候选数据资产列表对应的标识表征信息之间的信息匹配度;

基于所述目标数据资产信息对应的多维表征信息与每个所述候选数据资产列表对应的标识表征信息之间的信息匹配度,从多个所述候选数据资产列表中确定所述目标数据资产信息对应的所述目标数据资产列表。

4. 根据权利要求3所述的方法,其特征在于,所述基于所述目标数据资产信息对应的多维表征信息与每个所述候选数据资产列表对应的标识表征信息之间的信息匹配度,从多个所述候选数据资产列表中确定所述目标数据资产信息对应的所述目标数据资产列表,包括:

获取所述目标数据资产信息对应的多维表征信息与每个所述候选数据资产列表对应的标识表征信息之间的信息匹配度中,最大信息匹配度对应的候选数据资产列表的数量;

若所述最大信息匹配度对应的候选数据资产列表的数量为至少两个,则从至少两个候

选数据资产列表中随机选取一个候选数据资产列表确定为所述目标数据资产信息对应的所述目标数据资产列表；

若所述最大信息匹配度对应的候选数据资产列表的数量为一个，则将所述最大信息匹配度对应的候选数据资产列表确定为所述目标数据资产信息对应的所述目标数据资产列表。

5. 根据权利要求4所述的方法，其特征在于，所述目标数据资产信息对应的多维表征信息与每个所述候选数据资产列表对应的标识表征信息之间的信息匹配度中，最大信息匹配度对应的候选数据资产列表的数量之后，还包括：

若所述最大信息匹配度对应的候选数据资产列表的数量为至少两个，则获取至少两个候选数据资产列表中包括的每个候选数据资产列表对应的所属领域；

确定与所述目标数据资产信息匹配度最大的所属领域对应的候选数据资产列表，为所述目标数据资产信息对应的所述目标数据资产列表。

6. 根据权利要求1所述的方法，其特征在于，所述将所述目标数据资产信息划分至所述目标数据资产列表中之后，还包括：

向所述目标数据资产信息对应的用户设备发送所述目标数据资产信息的多维分类结果；

响应于所述用户设备对所述目标数据资产列表的数据资产访问请求，基于所述目标数据资产信息的数据资产等级，确定所述用户设备可访问的目标数据资产；

向所述用户设备发送所述目标数据资产。

7. 根据权利要求1所述的方法，其特征在于，所述数据特征信息包括如下至少一项：数据样本量、数据特征数、数据特征名称、特征数值类型和特征统计性数据；

所述数据属性信息包括如下至少一项：数据领域、数据内容和数据交易场景；

所述数据分级分类信息如下至少一项：数据类型和数据敏感程度。

8. 一种多维度数据资产分类装置，其特征在于，包括：

获取模块，用于获取目标数据资产信息，所述目标数据资产信息包括第一维度数据资产、第二维度数据资产和第三维度数据资产，所述第一维度数据资产用于描述数据特征信息，所述第二维度数据资产用于描述数据参与方提供的数据属性信息，所述第三维度数据资产用于描述数据分级分类信息；

处理模块，用于对所述目标数据资产信息进行信息处理，得到所述目标数据资产信息对应的多维表征信息，所述多维表征信息用于表征所述第一维度数据资产、所述第二维度数据资产和所述第三维度数据资产；

匹配模块，用于基于所述目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每个候选数据资产列表进行匹配，得到所述目标数据资产信息对应的目标数据资产列表，所述预设数据资产列表中存储着多个多维度的候选数据资产列表，每个所述候选数据资产列表对应一种多维分类的数据资产；

发送模块，用于将所述目标数据资产信息划分至所述目标数据资产列表中。

9. 一种计算机设备，其特征在于，包括存储器和处理器，存储器中存储有计算机程序，处理器执行计算机程序时实现如权利要求1~7中任一所述的多维度数据资产分类方法。

10. 一种计算机可读存储介质，其上存储有计算机程序，其特征在于，计算机程序被处

理器执行时实现如权利要求1~7中任一所述的多维度数据资产分类方法。

## 多维度数据资产分类方法、装置、计算机设备和介质

### 技术领域

[0001] 本公开的实施例涉及数据资产管理技术领域,具体地,涉及适用于一种多维度数据资产分类方法、装置、计算机设备和介质。

### 背景技术

[0002] 数据资产是指由企业拥有或者控制的、且能够为企业未来带来经济利益的、并以物理方式或电子方式记录的数据资源,如文件资料、电子数据等。

[0003] 在企业中,并非所有的数据都构成数据资产,数据资产是能够为企业产生价值的数据资源,对数据资产进行有效分类能够便于企业统一管理众多数据资产。相关技术中,数据资产分类主要是采用数据资产本身单一维度下的信息对多个数据资产进行分类的。

[0004] 然而,采用上述方式,难以实现对数据资产的精细化分类。

### 发明内容

[0005] 本文中描述的实施例提供了一种多维度数据资产分类方法、装置、计算机设备和介质,克服了上述问题。

[0006] 第一方面,根据本公开的内容,提供了一种多维度数据资产分类方法,包括:

[0007] 获取目标数据资产信息,所述目标数据资产信息包括第一维度数据资产、第二维度数据资产和第三维度数据资产,所述第一维度数据资产用于描述数据特征信息,所述第二维度数据资产用于描述数据参与方提供的数据属性信息,所述第三维度数据资产用于描述数据分级分类信息;

[0008] 对所述目标数据资产信息进行信息处理,得到所述目标数据资产信息对应的多维表征信息,所述多维表征信息用于表征所述第一维度数据资产、所述第二维度数据资产和所述第三维度数据资产;

[0009] 基于所述目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每个候选数据资产列表进行匹配,得到所述目标数据资产信息对应的目标数据资产列表,所述预设数据资产列表中存储着多个多维度的候选数据资产列表,每个所述候选数据资产列表对应一种多维分类的数据资产;

[0010] 将所述目标数据资产信息划分至所述目标数据资产列表中。

[0011] 第二方面,根据本公开的内容,提供了一种多维度数据资产分类装置,包括:

[0012] 获取模块,用于获取目标数据资产信息,所述目标数据资产信息包括第一维度数据资产、第二维度数据资产和第三维度数据资产,所述第一维度数据资产用于描述数据特征信息,所述第二维度数据资产用于描述数据参与方提供的数据属性信息,所述第三维度数据资产用于描述数据分级分类信息;

[0013] 处理模块,用于对所述目标数据资产信息进行信息处理,得到所述目标数据资产信息对应的多维表征信息,所述多维表征信息用于表征所述第一维度数据资产、所述第二维度数据资产和所述第三维度数据资产;

[0014] 匹配模块,用于基于所述目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每个候选数据资产列表进行匹配,得到所述目标数据资产信息对应的目标数据资产列表,所述预设数据资产列表中存储着多个多维度的候选数据资产列表,每个所述候选数据资产列表对应一种多维分类的数据资产;

[0015] 发送模块,用于将所述目标数据资产信息划分至所述目标数据资产列表中。

[0016] 第三方面,提供了一种计算机设备,包括存储器和处理器,存储器中存储有计算机程序,处理器执行计算机程序时实现如以上任意一个实施例中多维度数据资产分类方法的步骤。

[0017] 第四方面,提供了一种计算机可读存储介质,计算机可读存储介质上存储有计算机程序,计算机程序被处理器执行时实现如以上任意一个实施例中多维度数据资产分类方法的步骤。

[0018] 本申请实施例提供的多维度数据资产分类方法,获取目标数据资产信息,目标数据资产信息包括第一维度数据资产、第二维度数据资产和第三维度数据资产,第一维度数据资产用于描述数据特征信息,第二维度数据资产用于描述数据参与方提供的数据属性信息,第三维度数据资产用于描述数据分级分类信息;对目标数据资产信息进行信息处理,得到目标数据资产信息对应的多维表征信息,多维表征信息用于表征第一维度数据资产、第二维度数据资产和第三维度数据资产;基于目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每个候选数据资产列表进行匹配,得到目标数据资产信息对应的目标数据资产列表,预设数据资产列表中存储着多个多维度的候选数据资产列表,每个候选数据资产列表对应一种多维分类的数据资产;将目标数据资产信息划分至目标数据资产列表中。如此,通过数据资产的多维度信息进行信息处理,便于采用处理得到的多维度表征信息与多个数据资产进行匹配,有效识别出目标数据资产信息所属的多维分类,实现数据资产的精细化分类。

[0019] 上述说明仅是本申请实施例技术方案的概述,为了能够更清楚了解本申请实施例的技术手段,而可依照说明书的内容予以实施,并且为了让本申请实施例的上述和其它目的、特征和优点能够更明显易懂,以下特举本申请的具体实施方式。

## 附图说明

[0020] 为了更清楚地说明本公开的实施例的技术方案,下面将对实施例的附图进行简要说明,应当知道,以下描述的附图仅仅涉及本公开的一些实施例,而非对本公开的限制,其中:

[0021] 图1是本公开提供的一种多维度数据资产分类方法的流程示意图。

[0022] 图2是本公开提供的一种多维度数据资产分类装置的结构示意图。

[0023] 图3是本公开提供的一种计算机设备的结构示意图。

[0024] 需要注意的是,附图中的元素是示意性的,没有按比例绘制。

## 具体实施方式

[0025] 为了使本公开的实施例的目的、技术方案和优点更加清楚,下面将结合附图,对本公开的实施例的技术方案进行清楚、完整的描述。显然,所描述的实施例是本公开的一部分

实施例,而不是全部的实施例。基于所描述的本公开的实施例,本领域技术人员在无需创造性劳动的前提下所获得的所有其它实施例,也都属于本公开保护的范围。

[0026] 除非另外定义,否则在此使用的所有术语(包括技术和科学术语)具有与本公开主题所属领域的技术人员所通常理解的相同含义。进一步将理解的是,诸如在通常使用的词典中定义的那些的术语应解释为具有与说明书上下文和相关技术中它们的含义一致的含义,并且将不以理想化或过于正式的形式来解释,除非在此另外明确定义。如在此所使用的,将两个或更多部分“连接”或“耦接”到一起的陈述应指这些部分直接结合到一起或通过一个或多个中间部件结合。

[0027] 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“实施例”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。

[0028] 本文中术语“和/或”,仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,例如,A和/或B,可以表示:存在A,同时存在A和B,存在B这三种情况。另外,本文中字符“/”,一般表示前后关联对象是一种“或”的关系。诸如“第一”和“第二”的术语仅用于将一个部件(或部件的一部分)与另一个部件(或部件的另一部分)区分开。

[0029] 在本申请的描述中,除非另有说明,“多个”的含义是指两个以上(包括两个),同理,“多组”指的是两组以上(包括两组)。

[0030] 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。

[0031] 图1是本公开实施例提供的一种多维度数据资产分类方法的流程示意图。

[0032] 如图1所示,多维度数据资产分类方法的具体过程包括:

[0033] S110、获取目标数据资产信息。

[0034] 其中,在获取数据交易参与方同意的前提下,数据要素流通平台会获取数据资产的相关信息,即目标数据资产信息,并将面部数据资产信息存储至相应数据库中。获取目标数据资产信息,即表示从数据要素流通平台的数据库中获取目标数据资产信息。

[0035] 目标数据资产信息包括第一维度数据资产、第二维度数据资产和第三维度数据资产,第一维度数据资产用于描述数据特征信息,第二维度数据资产用于描述数据参与方提供的属性信息,第三维度数据资产用于描述数据分级分类信息。

[0036] 数据特征信息包括如下至少一项:数据样本量、数据特征数、数据特征名称、特征数值类型和特征统计性数据。数据分级分类信息如下至少一项:数据类型和数据敏感程度。数据属性信息包括如下至少一项:数据领域、数据内容和数据交易场景,数据领域用于描述数据的行业背景,数据内容用于描述数据内容的概括特征,数据属性信息还可以包括数据来源,如提供数据的某个数据交易方提供方。

[0037] S120、对目标数据资产信息进行信息处理,得到目标数据资产信息对应的多维表征信息。

[0038] 其中,多维表征信息用于表征第一维度数据资产、第二维度数据资产和第三维度数据资产。多维表征信息可为目标数据资产的特点所对应的高维数值表征。

[0039] 一些实施例中,对目标数据资产信息进行信息处理,得到目标数据资产信息对应

的多维表征信息,包括:

[0040] 获取信息预处理模版;基于信息预处理模版对目标数据资产信息进行处理,得到标签提示词;基于目标大语言模型对标签提示词进行信息处理,得到目标数据资产信息对应的多维表征信息。

[0041] 其中,标签提示词用于描述第一维度数据资产、第二维度数据资产和第三维度数据资产中的部分数据资产或全部数据资产。即,标签提示词可用于描述第一维度数据资产,或第二维度数据资产,或第三维度数据资产,或者第一维度数据资产、第二维度数据资产以及第三维度数据资产中任意两个或任意三个数据资产。

[0042] 从而,能够基于信息预处理模版得到目标数据资产信息的标签提示词,便于对大语言模型进行精准输入,提升大语言模型的处理效率。

[0043] 另外,对目标数据资产信息进行信息处理,得到目标数据资产信息对应的多维表征信息,还可以包括:

[0044] 获取信息预处理模版;基于信息预处理模版优化初始大语言模型,得到目标大语言模型;基于目标大语言模型对目标数据资产信息进行信息处理,得到目标数据资产信息对应的多维表征信息。

[0045] 其中,目标大语言模型用于进行资产信息的多维表征处理。

[0046] 基于信息预处理模版优化初始大语言模型,得到目标大语言模型,可包括:将信息预处理模版形成信息处理指令用于初始大语言模型学习,得到目标大语言模型,或者,将信息预处理模版形成信息处理指令嵌入初始大语言模型中,得到目标大语言模型。

[0047] 可针对不同数据领域,提供不同的信息预处理模版。数据交易参与方可自定义信息预处理模版。

[0048] 从而,通过目标大语言模型对目标数据资产信息进行信息处理,得到目标数据资产信息对应的多维表征信息,便于快速实现目标数据资产信息的有效处理,提升信息处理效率。

[0049] S130、基于目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每个候选数据资产列表进行匹配,得到目标数据资产信息对应的目标数据资产列表。

[0050] 其中,预设数据资产列表中存储着多个多维度的候选数据资产列表,每个候选数据资产列表对应一种多维分类的数据资产。

[0051] 多维分类的数据资产可包括但不限于是:第一维度数据资产、第二维度数据资产和第三维度数据资产,第一维度数据资产用于描述数据特征信息,第二维度数据资产用于描述数据参与方提供的数据属性信息,第三维度数据资产用于描述数据分级分类信息。

[0052] 数据特征信息包括如下至少一项:数据样本量、数据特征数、数据特征名称、特征数值类型和特征统计性数据。数据分级分类信息如下至少一项:数据类型和数据敏感程度。数据属性信息包括如下至少一项:数据领域、数据内容和数据交易场景,数据领域用于描述数据的行业背景,数据内容用于描述数据内容的概括特征,数据属性信息还可以包括数据来源,如提供数据的某个数据交易方提供方。

[0053] 一些实施例,预设数据资产列表库中包括的每个候选数据资产列表对应有标识表征信息;

[0054] 基于目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每



个候选数据资产列表进行匹配,得到目标数据资产信息对应的目标数据资产列表,包括:

[0055] 匹配目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每个候选数据资产列表对应的标识表征信息,得到目标数据资产信息对应的多维表征信息与每个候选数据资产列表对应的标识表征信息之间的信息匹配度;基于目标数据资产信息对应的多维表征信息与每个候选数据资产列表对应的标识表征信息之间的信息匹配度,从多个候选数据资产列表中确定目标数据资产信息对应的目标数据资产列表。

[0056] 其中,标识表征信息用于描述候选数据资产列表对应的分类/领域标识特征。

[0057] 匹配目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每个候选数据资产列表对应的标识表征信息,得到目标数据资产信息对应的多维表征信息与每个候选数据资产列表对应的标识表征信息之间的信息匹配度,即为,将目标数据资产信息对应的多维表征信息分别与预设数据资产列表库中包括的每个候选数据资产列表对应的标识表征信息采用相似性计算(如余弦相似度算法)的方式进行相似性匹配,得到目标数据资产信息对应的多维表征信息与每个候选数据资产列表对应的标识表征信息之间的信息匹配度。

[0058] 从而,基于目标数据资产信息对应的多维表征信息与每个候选数据资产列表对应的标识表征信息之间的信息匹配度,从多个候选数据资产列表中有效确定出目标数据资产信息对应的目标数据资产列表。

[0059] 一些实施例中,基于目标数据资产信息对应的多维表征信息与每个候选数据资产列表对应的标识表征信息之间的信息匹配度,从多个候选数据资产列表中确定目标数据资产信息对应的目标数据资产列表,包括:

[0060] 获取目标数据资产信息对应的多维表征信息与每个候选数据资产列表对应的标识表征信息之间的信息匹配度中,最大信息匹配度对应的候选数据资产列表的数量;若最大信息匹配度对应的候选数据资产列表的数量为至少两个,则从至少两个候选数据资产列表中随机选取一个候选数据资产列表确定为目标数据资产信息对应的目标数据资产列表;若最大信息匹配度对应的候选数据资产列表的数量为一个,则将最大信息匹配度对应的候选数据资产列表确定为目标数据资产信息对应的目标数据资产列表。从而,有效确定出与目标数据资产信息相匹配的目标数据资产列表。

[0061] 一些实施例中,目标数据资产信息对应的多维表征信息与每个候选数据资产列表对应的标识表征信息之间的信息匹配度中,最大信息匹配度对应的候选数据资产列表的数量之后,还包括:

[0062] 若最大信息匹配度对应的候选数据资产列表的数量为至少两个,则获取至少两个候选数据资产列表中包括的每个候选数据资产列表对应的所属领域;确定与目标数据资产信息匹配度最大的所属领域对应的候选数据资产列表,为目标数据资产信息对应的目标数据资产列表。

[0063] 其中,在确定与目标数据资产信息匹配度最大的所属领域对应的候选数据资产列表时,若与目标数据资产信息匹配度最大的所属领域对应的候选数据资产列表的数量为至少一个,则将匹配度最大的所属领域对应的候选数据资产列表确定为目标数据资产信息对应的目标数据资产列表。

[0064] 若与目标数据资产信息匹配度最大的所属领域对应的候选数据资产列表的数量

为至少两个,则选择候选数据资产列表中存储数据资产的数量最小的一个候选数据资产列表为目标数据资产信息对应的目标数据资产列表,便于均衡每个候选数据资产列表中的数据存储量。

[0065] S140、将目标数据资产信息划分至目标数据资产列表中。

[0066] 其中,目标数据资产列表中可按照数据资产信息的存入时间从近至远的顺序依次排列多个数据资产,或者,按照优先等级从高到低的顺序的依次排列多个数据资产。

[0067] 将目标数据资产信息划分至目标数据资产列表中,即包括:将目标数据资产信息存储至目标数据资产列表的列表前部位置中,便于访问交易方能够直观查看到最新数据资产信息。或者,按照目标数据资产信息的优先等级将其存储至目标数据资产列表的相应存储位置处,便于访问交易方能够查看到优先等级较高的数据资产信息。

[0068] 本实施例中,通过获取目标数据资产信息,目标数据资产信息包括第一维度数据资产、第二维度数据资产和第三维度数据资产,第一维度数据资产用于描述数据特征信息,第二维度数据资产用于描述数据参与方提供的数据属性信息,第三维度数据资产用于描述数据分级分类信息;对目标数据资产信息进行信息处理,得到目标数据资产信息对应的多维表征信息,多维表征信息用于表征第一维度数据资产、第二维度数据资产和第三维度数据资产;基于目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每个候选数据资产列表进行匹配,得到目标数据资产信息对应的目标数据资产列表,预设数据资产列表中存储着多个多维度的候选数据资产列表,每个候选数据资产列表对应一种多维分类的数据资产;将目标数据资产信息划分至目标数据资产列表中。如此,通过数据资产的多维度信息进行信息处理,便于采用处理得到的多维度表征信息与多个数据资产进行匹配,有效识别出目标数据资产信息所属的多维分类,实现数据资产的精细化分类。

[0069] 一些实施例中,将目标数据资产信息划分至目标数据资产列表中之后,本实施例方法还可以包括:

[0070] 向目标数据资产信息对应的用户设备发送目标数据资产信息的多维分类结果;响应于用户设备对目标数据资产列表的数据资产访问请求,基于目标数据资产信息的数据资产等级,确定用户设备可访问的目标数据资产;向用户设备发送目标数据资产。

[0071] 其中,目标数据资产信息的多维分类结果用于描述目标数据资产信息划分至目标数据资产列表中。

[0072] 从而,通过向目标数据资产信息对应的用户设备发送目标数据资产信息的多维分类结果,便于及时告知用户目标数据资产信息的多维分类状态,同时,在用户访问目标数据资产列表时,根据目标数据资产信息的数据资产等级确定用户设备可访问的目标数据资产,有效保证目标数据资产列表中不可访问的数据资产的存储安全。

[0073] 图2为本实施例提供的一种多维度数据资产分类装置的结构示意图。多维度数据资产分类装置可以包括:获取模块210、处理模块220、匹配模块230和发送模块240。

[0074] 获取模块210,用于获取目标数据资产信息,目标数据资产信息包括第一维度数据资产、第二维度数据资产和第三维度数据资产,第一维度数据资产用于描述数据特征信息,第二维度数据资产用于描述数据参与方提供的数据属性信息,第三维度数据资产用于描述数据分级分类信息。

[0075] 处理模块220,用于对目标数据资产信息进行信息处理,得到目标数据资产信息对

应的多维表征信息,多维表征信息用于表征第一维度数据资产、第二维度数据资产和第三维度数据资产。

[0076] 匹配模块230,用于基于目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每个候选数据资产列表进行匹配,得到目标数据资产信息对应的目标数据资产列表,预设数据资产列表中存储着多个多维度的候选数据资产列表,每个候选数据资产列表对应一种多维分类的数据资产。

[0077] 发送模块240,用于将目标数据资产信息划分至目标数据资产列表中。

[0078] 在本实施例中,可选的,处理模块220,具体用于:

[0079] 获取信息预处理模版;基于信息预处理模版对目标数据资产信息进行处理,得到标签提示词,标签提示词用于描述第一维度数据资产、第二维度数据资产和第三维度数据资产中的部分数据资产或全部数据资产;基于目标大语言模型对标签提示词进行信息处理,得到目标数据资产信息对应的多维表征信息。

[0080] 在本实施例中,可选的,预设数据资产列表库中包括的每个候选数据资产列表对应标识表征信息。

[0081] 匹配模块230,包括:匹配单元和确定单元。

[0082] 匹配单元,用于匹配目标数据资产信息对应的多维表征信息与预设数据资产列表库中包括的每个候选数据资产列表对应的标识表征信息,得到目标数据资产信息对应的多维表征信息与每个候选数据资产列表对应的标识表征信息之间的信息匹配度。

[0083] 确定单元,用于基于目标数据资产信息对应的多维表征信息与每个候选数据资产列表对应的标识表征信息之间的信息匹配度,从多个候选数据资产列表中确定目标数据资产信息对应的目标数据资产列表。

[0084] 在本实施例中,可选的,确定单元,具体用于:

[0085] 获取目标数据资产信息对应的多维表征信息与每个候选数据资产列表对应的标识表征信息之间的信息匹配度中,最大信息匹配度对应的候选数据资产列表的数量;若最大信息匹配度对应的候选数据资产列表的数量为至少两个,则从至少两个候选数据资产列表中随机选取一个候选数据资产列表确定为目标数据资产信息对应的目标数据资产列表;若最大信息匹配度对应的候选数据资产列表的数量为一个,则将最大信息匹配度对应的候选数据资产列表确定为目标数据资产信息对应的目标数据资产列表。

[0086] 在本实施例中,可选的,确定单元,具体用于:

[0087] 若最大信息匹配度对应的候选数据资产列表的数量为至少两个,则获取至少两个候选数据资产列表中包括的每个候选数据资产列表对应的所属领域;确定与目标数据资产信息匹配度最大的所属领域对应的候选数据资产列表,为目标数据资产信息对应的目标数据资产列表。

[0088] 在本实施例中,可选的,还包括:确定模块。

[0089] 发送模块240,还用于向目标数据资产信息对应的用户设备发送目标数据资产信息的多维分类结果。

[0090] 确定模块,用于响应于用户设备对目标数据资产列表的数据资产访问请求,基于目标数据资产信息的数据资产等级,确定用户设备可访问的目标数据资产。

[0091] 发送模块240,还用于向用户设备发送目标数据资产。

[0092] 在本实施例中,可选的,数据特征信息包括如下至少一项:数据样本量、数据特征数、数据特征名称、特征数值类型和特征统计性数据;数据属性信息包括如下至少一项:数据领域、数据内容和数据交易场景;数据分级分类信息如下至少一项:数据类型和数据敏感程度。

[0093] 本公开提供的多维度数据资产分类装置,可执行上述方法实施例,其具体实现原理和技术效果,可参见上述方法实施例,本公开此处不再赘述。

[0094] 本申请实施例还提供了一种计算机设备。具体请参阅图3,图3为本实施例计算机设备基本结构框图。

[0095] 计算机设备包括通过系统总线相互通信连接存储器310和处理器320。需要指出的是,图中仅示出了具有存储器310和处理器320的计算机设备,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 GateArray,FPGA)、数字处理器(Digital Signal Processor,DSP)、嵌入式设备等。

[0096] 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。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。

[0097] 存储器310至少包括一种类型的可读存储介质,可读存储介质包括非易失性存储器(non-volatile memory)或易失性存储器,例如,闪存(flashmemory)、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(random access memory,RAM)、只读存储器(read-only memory,ROM)、可擦写可编程只读存储器(erasable programmable read-onlymemory,EPR0M)、电可擦写可编程只读存储器(electrically erasable programmable read-onlymemory,EEPROM)、可编程只读存储器(programmable read-only memory,PROM)、磁性存储器、磁盘、光盘等,RAM可以包括静态RAM或动态RAM。在一些实施例中,存储器310可以是计算机设备的内部存储单元,例如,该计算机设备的硬盘或内存。在另一些实施例中,存储器310也可以是计算机设备的外部存储设备,例如该计算机设备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)卡或闪存卡(Flash Card)等。当然,存储器310还可以既包括计算机设备的内部存储单元也包括其外部存储设备。本实施例中,存储器310通常用于存储安装于计算机设备的操作系统和各类应用软件,例如上述方法的程序代码等。此外,存储器310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。

[0098] 处理器320通常用于执行计算机设备的总体操作。本实施例中,存储器310用于存储程序代码或指令,程序代码包括计算机操作指令,处理器320用于执行存储器310存储的程序代码或指令或者处理数据,例如运行上述方法的程序代码。

[0099] 本文中,总线可以是工业标准体系结构(Industry StandardArchitecture,ISA)总线、外设部件互连标准(Peripheral Component Interconnect,PCI)总线或扩展工业标准结构(Extended Industry StandardArchitecture,EISA)总线等。该总线系统可以分为地址总线、数据总线、控制总线等。为便于表示,图中仅用一条粗线表示,但并不表示仅有一根总线或一种类型的总线。

[0100] 本申请的另一实施例还提供一种计算机可读介质,计算机可读介质可以是计算机可读信号介质或者计算机可读介质。计算机中的处理器读取存储在计算机可读介质中的计算机可读程序代码,使得处理器能够执行在上述方法中每个步骤、或各步骤的组合中规定的功能动作;生成实施在框图的每一块、或各块的组合中规定的功能动作的装置。

[0101] 计算机可读介质包含但不限于电子、磁性、光学、电磁、红外的存储器或半导体系统、设备或者装置,或者前述的任意适当组合,存储器用于存储程序代码或指令,程序代码包括计算机操作指令,处理器用于执行存储器存储的上述方法的程序代码或指令。

[0102] 存储器和处理器的定义,可以参考前述计算机设备实施例的描述,在此不再赘述。

[0103] 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、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,模块或单元的划分,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,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0104] 在本申请各个实施例中的各功能单元或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,也可以采用软件功能单元的形式实现。

[0105] 集成的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。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或者该技术方案的全部或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,或者网络设备等)或处理器(processor)执行本申请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(Read-Only Memory, ROM)、随机存取存储器(Random Access Memory, RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0106] 在权利要求中,不应将位于括号之间的任何参考符号构造成对权利要求的限制。本申请描述的“包含”不排除存在未列在权利要求中的元件或步骤。位于元件之前的单词“一”或“一个”不排除存在多个这样的元件。本申请可以借助于包括有若干不同元件的硬件以及借助于适当编程的计算机来实现。在列举了装置若干的单元权利要求中,这些装置中的若干个单元可以通过同一个硬件项来具体体现。第一、第二、以及第三等的使用不表示任何顺序,可将这些单词解释为名称。上述实施例中的步骤,除有特殊说明外,不应理解为对执行顺序的限定。

[0107] 以上实施例仅用以说明本申请的技术方案,而非对其限制;尽管参照前述实施例对本申请进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例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同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围。

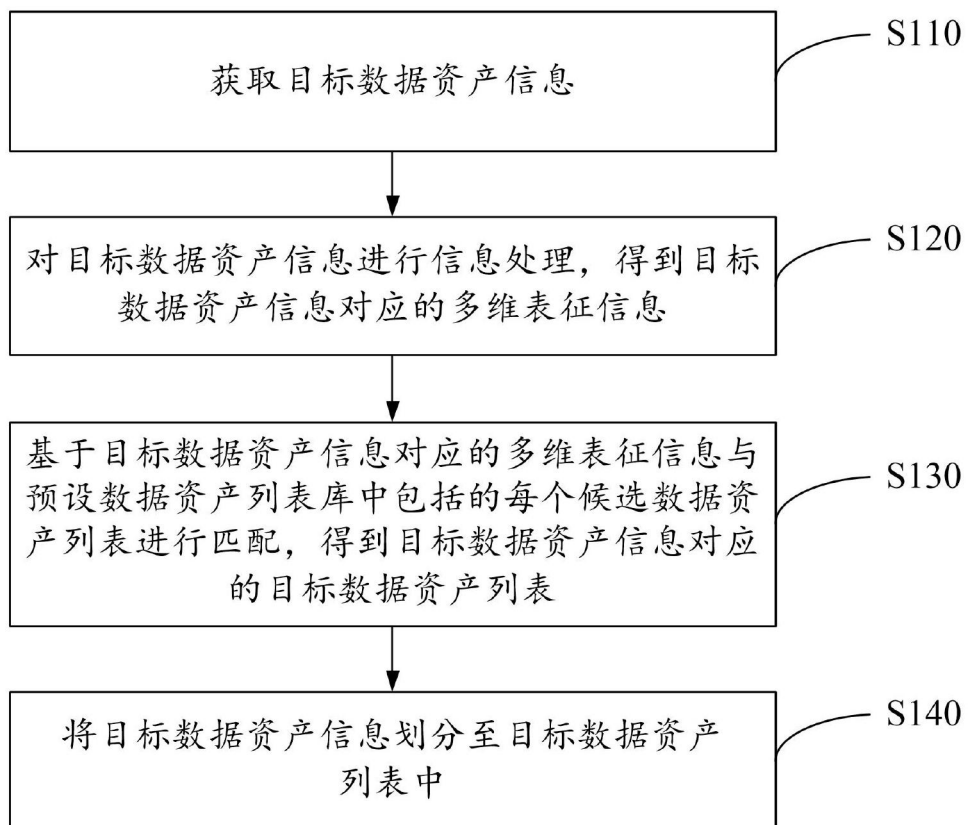


图1

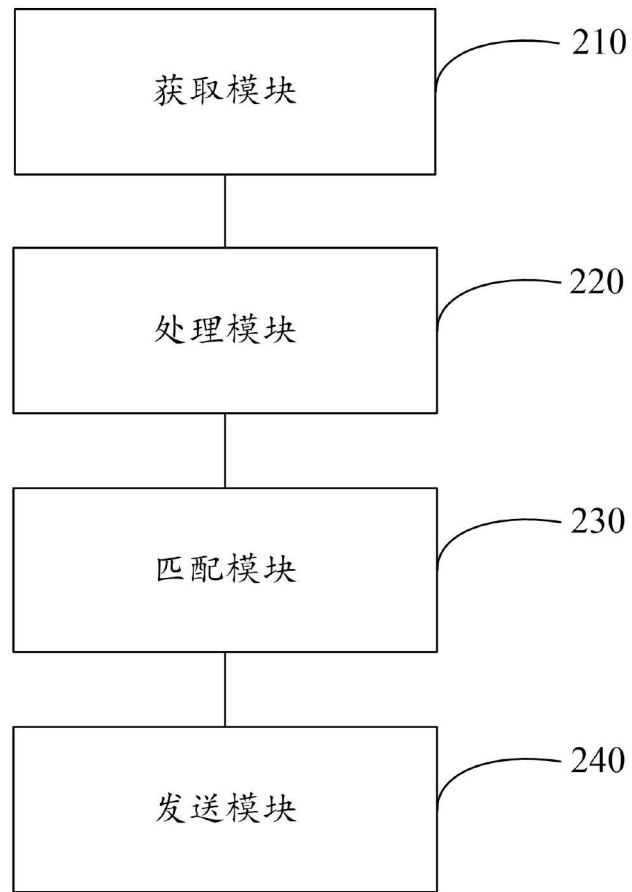


图2

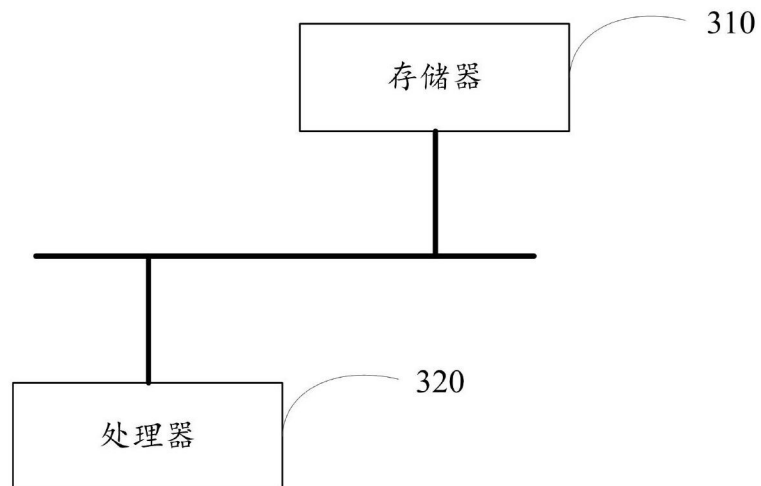


图3