

(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
2021 年 7 月 29 日 (29.07.2021)



(10) 国际公布号
WO 2021/147486 A1

(51) 国际专利分类号:
G06Q 30/02 (2012.01)

(21) 国际申请号: PCT/CN2020/129123

(22) 国际申请日: 2020 年 11 月 16 日 (16.11.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010071525.X 2020 年 1 月 21 日 (21.01.2020) CN

(71) 申请人: 深圳前海微众银行股份有限公司
(**WEBANK CO., LTD.**) [CN/CN]; 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518027 (CN)。

(72) 发明人: 衣志昊(**YI, Zhihao**); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong

518027 (CN)。程勇(**CHENG, Yong**); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518027 (CN)。刘洋(**LIU, Yang**); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518027 (CN)。陈天健(**CHEN, Tianjian**); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518027 (CN)。

(74) 代理人: 北京同达信恒知识产权代理有限公司
(**TDIP & PARTNERS**); 中国北京市西城区裕民路 18 号北环中心 A 座 2002, Beijing 100029 (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,

(54) **Title:** DATA PROCESSING METHOD AND APPARATUS

(54) 发明名称: 一种数据处理方法及装置

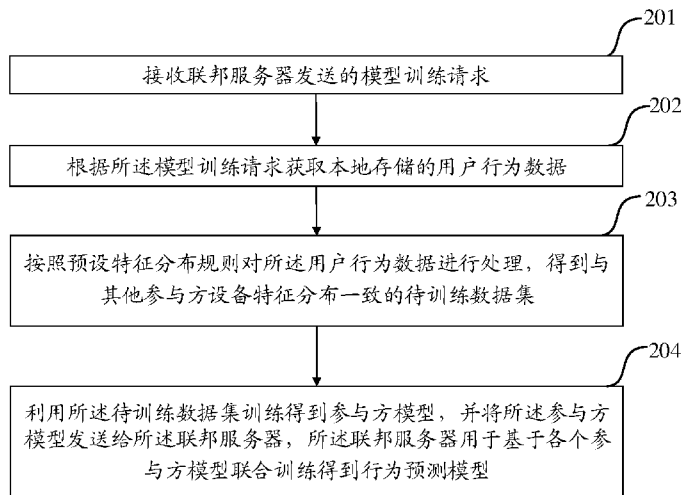


图 2

- 201 Receive a model training request sent by a federation server
202 Acquire locally stored user behavior data according to the model training request
203 Process the user behavior data according to a preset feature distribution rule to obtain a set of data to be trained which has a feature distribution consistent with that of other participant devices
204 Obtain, by means of training, a participant model by using the set of data to be trained, and send the participant model to the federation server, wherein the federation server is used for obtaining, by means of joint training, a behavior prediction model on the basis of each participant model

(57) **Abstract:** A data processing method and apparatus, which relate to the field of Fintech, and are used for solving the problem in the prior art of it not being possible to obtain, by means of training, a universal behavior prediction model. The method comprises: a participant device receiving a model training request sent by a federation server; acquiring locally stored user behavior data according to the model training request (202); processing the user behavior data according to a preset feature distribution rule to obtain a set of data to be trained which has a feature distribution consistent with that of other participant devices (203); and the participant devices using the set of data to be trained which has the consistent feature distribution to obtain, by means of training, participant models with a consistent model structure, so that the federation server can obtain, by means of training, a behavior prediction model on the basis of the participant models with the consistent model structure. Since a behavior prediction model combines behavior data features of participant devices, the model can be used for predicting the behavior of users in the participant devices, so that the universality of the model is relatively good, and the accuracy thereof is relatively high.

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,
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: 一种数据处理方法及装置, 涉及金融科技Fintech领域, 用以解决现有技术无法训练得到通用的行为预测模型的问题。其中方法包括: 参与方设备接收联邦服务器发送的模型训练请求, 根据模型训练请求获取本地存储的用户行为数据(202), 并按照预设特征分布规则处理用户行为数据得到与其他参与方设备特征分布一致的待训练数据集(203), 通过各个参与方设备使用特征分布一致的待训练数据集训练得到模型结构一致的各个参与方模型, 使得联邦服务器能够基于模型结构一致的各个参与方模型训练得到行为预测模型。由于行为预测模型结合了各个参与方设备的行为数据特征, 因此该模型能够用于预测各个参与方设备中的用户的行为, 模型的通用性较好, 且准确率较高。

一种数据处理方法及装置

相关申请的交叉引用

本申请要求在2020年01月21日提交中国专利局、申请号为202010071525.X、申请名称为“一种数据处理方法及装置”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

本发明涉及金融科技（Fintech）技术领域，尤其涉及一种数据处理方法及装置。

背景技术

随着计算机技术的发展，越来越多的技术应用在金融领域，传统金融业正在逐步向金融科技（Fintech）转变。然而，由于金融行业的安全性和实时性要求较高，因此金融科技领域也对技术提出了更高的要求。在金融领域中，推广用户之前通常需要对用户的行为进行预测，比如通过预测用户对某商品感兴趣的概率，能够预先将该商品不感兴趣的用戶排除，从而降低无用的推广操作，提高推广效率。

现阶段，各个参与方仅能使用本地存储的用户行为数据训练得到行为预测模型。然而，由于不同参与方中的用户不同，对象也不同，因此某一参与方训练的行为预测模型仅能适用于对本参与方中的用户的行为进行预测，而无法对其它参与方中的用户的行为进行预测，从而导致行为预测模型的通用性较差，预测的准确率较低。

发明内容

本发明提供一种数据处理方法及装置，用以训练得到通用的行为预测模型，以实现对各个参与方中的用户的行为进行预测，进而能够提升用户行为预测的准确率。

第一方面，本发明提供一种数据处理方法，该方法应用于参与方设备，该方法包括：参与方设备接收联邦服务器发送的模型训练请求，根据该模型训练请求获取本地存储的用户行为数据，并按照预设特征分布规则对所述用户行为数据进行处理，得到与其他参与方设备特征分布一致的待训练数据集，之后利用该待训练数据集训练得到参与方模型，并将该参与方模型发送给所述联邦服务器，以使联邦服务器基于各个参与方模型联合训练得到行为预测模型。

在一种可能的实现方式中，预设特征分布规则为按照共用的用户行为特征维度和对象行为特征维度生成特征分布一致的待训练数据的规则。这种情况下，参与方设备按照预设特征分布规则对用户行为数据进行处理，得到与其他参与方设备特征分布一致的训练数据集，包括：参与方设备按照共用的用户行为特征维度，从用户行为数据中提取出任一用户的用户行为特征，再按照共用的对象行为特征维度，从用户行为数据中提取出任一对象的对象行为特征，之后构建任一用户的用户行为特征和任一对象的对象行为特征之间的特征对，基于该特征对，生成与其他参与方设备特征分布一致的待训练数据，并基于各个用户与各个对象构成的各个特征对对应的待训练数据，构建待训练数据集。

在一种可能的实现方式中，参与方设备按照共用的用户行为特征维度，从用户行为数据中提取出任一用户的用户行为特征，包括：参与方设备从用户行为数据中提取出用户在

任一共用的用户行为特征维度下的特征值，根据该用户在各个共用的用户行为特征维度下的特征值，构建得到该用户的用户行为特征。相应地，参与方设备按照共用的对象行为特征维度，从用户行为数据中提取出任一对象的对象行为特征，包括：参与方设备从用户行为数据中提取出对象在任一共用的对象行为特征维度下的特征值，根据对象在各个共用的对象行为特征维度下的特征值，构建得到对象的对象行为特征。

在一种可能的实现方式中，参与方设备基于征对，生成与其他参与方设备特征分布一致的待训练数据，包括：参与方设备按照共用的交互行为特征维度，从用户行为数据中提取出用户与对象的交互特征，根据用户与对象的交互特征，确定特征对对应的标签，之后将拼接用户的用户行为特征、对象的对象行为特征、交互特征及标签得到的特征向量作为特征对对应的待训练数据。

在一种可能的实现方式中，参与方设备根据各个特征对对应的待训练数据，构建待训练数据集，包括：参与方设备根据各个特征对对应的标签，确定属于正负样本的待训练数据的比例是否符合预设范围，若不符合预设范围，则对标签为负样本的待训练数据进行下采样，或者对标签为正样本的待训练数据进行上采样处理。若符合预设范围，则基于各个待训练数据构建待训练数据集。

在一种可能的实现方式中，参与方设备将参与方模型发送给联邦服务器之后，还可以接收联邦服务器发送的行为预测模型，并将待测特征对对应的特征信息输入行为预测模型中进行预测，得到待测特征对对应的预测标签。其中，待测特征对对应的特征信息包括待测用户的用户行为特征、待测对象的对象行为特征和待测用户与待测对象的交互特征中的任意一项或任意多项。预测标签用于确定待测用户是否会对待测对象执行预设行为。

在一种可能的实现方式中，共用的用户行为特征维度可以包括用户执行预设行为的时间和/或用户在各时段内执行预设行为的次数。相应地，共用的对象行为特征维度可以包括对象被执行预设行为的时间和/或对象在各时段内被执行预设行为的次数。

在一种可能的实现方式中，交互行为特征维度可以包括用户对对象执行预设行为的时间和/或用户在各时段内对对象执行预设行为的次数。

第二方面，本发明提供一种数据处理装置，该装置包括：收发模块，用于接收联邦服务器发送的模型训练请求；获取模块，用于根据模型训练请求获取本地存储的用户行为数据；处理模块，用于按照预设特征分布规则对用户行为数据进行处理，得到与其他参与方设备特征分布一致的待训练数据集；训练模块，用于利用待训练数据集训练得到参与方模型，并将参与方模型发送给联邦服务器，联邦服务器用于基于各个参与方模型联合训练得到行为预测模型。

在一种可能的实现方式中，预设特征分布规则为按照共用的用户行为特征维度和对象行为特征维度生成特征分布一致的待训练数据的规则。在这种情况下，处理模块具体用于：按照共用的用户行为特征维度，从用户行为数据中提取出任一用户的用户行为特征，按照共用的对象行为特征维度，从用户行为数据中提取出任一对象的对象行为特征，构建任一用户的用户行为特征和任一对象的对象行为特征之间的特征对，基于特征对，生成与其他参与方设备特征分布一致的待训练数据，基于各个用户与各个对象构成的各个特征对对应的待训练数据，构建待训练数据集。

在一种可能的实现方式中，处理模块具体用于：从用户行为数据中提取出用户在任一共用的用户行为特征维度下的特征值，根据用户在各个共用的用户行为特征维度下的特征

值，构建得到用户的用户行为特征，从用户行为数据中提取出对象在任一共用的对象行为特征维度下的特征值，根据对象在各个共用的对象行为特征维度下的特征值，构建得到对象的对象行为特征。

在一种可能的实现方式中，处理模块具体用于：按照共用的交互行为特征维度，从用户行为数据中提取出用户与对象的交互特征，根据用户与对象的交互特征，确定特征对对应的标签，将拼接用户的用户行为特征、对象的对象行为特征、交互特征及标签得到的特征向量作为特征对对应的待训练数据。

在一种可能的实现方式中，处理模块具体用于：根据各个特征对对应的标签，确定属于正负样本的待训练数据的比例是否符合预设范围，若不符合预设范围，则对标签为负样本的待训练数据进行下采样处理，或者对标签为正样本的待训练数据进行上采样处理。若符合预设范围，则基于各个待训练数据构建待训练数据集。

在一种可能的实现方式中，该装置还包括预测模块。在收发模块将参与方模型发送给联邦服务器之后，收发模块还可以接收联邦服务器发送的行为预测模型。预测模块用于：将待测特征对对应的特征信息输入行为预测模型中进行预测，得到待测特征对对应的预测标签。其中，待测特征对对应的特征信息包括待测用户的用户行为特征、待测对象的对象行为特征和待测用户与待测对象的交互特征中的任意一项或任意多项。预测标签用于确定待测用户是否会对待测对象执行预设行为。

在一种可能的实现方式中，共用的用户行为特征维度可以包括用户执行预设行为的时间和/或用户在各时段内执行预设行为的次数。相应地，共用的对象行为特征维度可以包括对象被执行预设行为的时间和/或对象在各时段内被执行预设行为的次数。

在一种可能的实现方式中，交互行为特征维度可以包括用户对对象执行预设行为的时间和/或用户在各时段内对对象执行预设行为的次数。

第三方面，本发明提供一种计算设备，包括至少一个处理单元以及至少一个存储单元，其中，所述存储单元存储有计算机程序，当所述程序被所述处理单元执行时，使得所述处理单元执行上述第一方面任意所述的数据处理方法。

第四方面，本发明提供一种计算机可读存储介质，其存储有可由计算设备执行的计算机程序，当所述程序在所述计算设备上运行时，使得所述计算设备执行上述第一方面任意所述的数据处理方法。

本发明提供的数据处理方法及装置，通过各个参与方设备按照预设特征分布规则对用户行为数据进行处理，得到与其他参与方设备特征分布一致的待训练数据集，使得各个参与方设备能够使用特征分布一致的待训练数据集训练得到模型结构一致的参与方模型，如此，联邦服务器能够基于各个参与方设备训练得到的模型结构一致的参与方模型训练得到行为预测模型，由于该行为预测模型结合了各个参与方设备的行为数据特征，因此能够用于对各个参与方设备中的用户的行为进行预测，行为预测模型的通用性较好，且还能够提升用户行为预测的准确率。

本发明的这些方面或其他方面在以下实施例的描述中会更加简明易懂。

附图说明

为了更清楚地说明本发明实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简要介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域

域的普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 为本发明实施例提供的一种适用的系统架构示意图；

图 2 为本发明实施例中提供的一种数据处理方法的流程示意图；

5 图 3 为本发明实施例中提供的一种数据处理方式的执行流程示意图；

图 4 为本发明实施例提供的一种模型训练方法的交互流程示意图；

图 5 为本发明实施例提供的一种数据处理装置的结构示意图；

图 6 为本发明实施例提供的一种前端设备的结构示意图；

图 7 为本发明实施例提供的一种后端设备的结构示意图。

10

具体实施方式

为了使本发明的目的、技术方案和优点更加清楚，下面将结合附图对本发明作进一步地详细描述，显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其它实施例，都属于本发明保护的范围。

15

图 1 为本发明实施例提供的一种适用的系统架构示意图，如图 1 所示，该系统架构中可以包括联邦服务器 110 和至少两个参与方设备，比如参与方设备 121、参与方设备 122 和参与方设备 123。其中，联邦服务器 110 可以与每个参与方设备连接，比如可以通过有线方式连接，也可以通过无线方式连接，具体不作限定。

20

基于图 1 所示意的系统架构，图 2 为本发明实施例提供的一种数据处理方法对应的流程示意图，该方法可以应用于任一参与方设备，比如图 1 所示意的参与方设备 121、参与方设备 122 或参与方设备 123，具体不作限定。如图 2 所示，该方法包括：

步骤 201，参与方设备接收联邦服务器发送的模型训练请求。

步骤 202，参与方设备根据模型训练请求获取本地存储的用户行为数据。

25

步骤 203，参与方设备按照预设特征分布规则对用户行为数据进行处理，得到与其他参与方设备特征分布一致的待训练数据集。

步骤 204，参与方设备利用待训练数据集训练得到参与方模型，并将参与方模型发送给所述联邦服务器，以使联邦服务器基于各个参与方模型联合训练得到行为预测模型。

30

本发明的上述实施例中，通过各个参与方设备按照预设特征分布规则对所述用户行为数据进行处理，得到与其他参与方设备特征分布一致的待训练数据集，使得各个参与方设备能够使用特征分布一致的待训练数据集训练得到模型结构一致的参与方模型，如此，联邦服务器能够基于各个参与方设备训练得到的模型结构一致的参与方模型训练得到行为预测模型，由于该行为预测模型结合了各个参与方设备的行为数据特征，因此能够用于对各个参与方设备中的用户的行为进行预测，行为预测模型的通用性较好。

35

在上述步骤 203 中，预设特征分布规则可以存在多种可能，比如：

可能一

在可能一中，预设特征分布规则可以设置为：联邦服务器 110 先统计各个参与方设备中的全部用户和全部对象，再使用 one-hot 算法对全部用户进行编码，将全部用户映射到第一向量空间，使用 one-hot 算法对全部对象进行编码，将全部对象映射到第二向量空间，进一步地，联邦服务器 110 将第一向量空间和第二向量空间同步给各个参与方设备，以使

40

各个参与方设备利用本地用户在第一向量空间中的映射值和本地对象在第二向量空间中的映射值构建参与方设备的待训练数据集。

举例来说，当用户行为数据为门禁数据时，若参与方设备 A 的用户包括用户 1 和用户 2，参与方设备 A 的对象包括门 1 和门 2，参与方设备 B 的用户包括用户 3 和用户 4，参与方设备 B 的对象包括门 3 和门 4，联邦服务器 110 通过 one-hot 算法将全部用户映射到第一向量空间[0 0 0 0]（第一向量空间中的各个向量位依次对应用户 1、用户 2、用户 3 和用户 4）中，将全部对象映射到第二向量空间[0 0 0 0]（第二向量空间中的各个向量位依次对应门 1、门 2、门 3 和门 4）中，则：当用户 1 在门 1 处刷卡、用户 2 未刷卡、用户 3 在门 4 处刷卡、用户 4 未刷卡时，参与方设备 A 构建得到的待训练数据集包括待训练数据[1 0 0 0 1 0 0 0]和待训练数据[0 1 0 0 0 0 0 0]，参与方设备 B 构建得到的待训练数据集包括待训练数据[0 0 1 0 0 0 0 1]和待训练数据[0 0 0 1 0 0 1 0]。

由此可知，通过设置可能一中的预设特征分布规则，使得各个参与方设备中的待训练数据集能够具有相同的特征分布。然而，可能一将各个参与方设备的全部用户和全部对象映射到对应的向量空间，而不同参与方设备的待训练数据通常集中分布在各自对应的映射值范围内，比如参与方设备 A 的待训练数据集中在第 1 位、第 2 位和第 5 位，而第 3 位、第 4 位、第 6 位和第 7 位均为零，参与方设备 B 的待训练数据集中在第 3 位、第 4 位、第 6 位和第 7 位，而第 1 位、第 2 位和第 5 位均为零。因此，采用可能一会使得参与方设备的待训练数据集中包含较多的无意义数据，当使用该种形式的待训练数据集进行模型训练时，这些无意义的数可能会对模型训练产生噪声影响，导致模型训练的效果较差。且，这些无意义的数还会增加模型训练时间，导致系统的性能损耗较大，模型训练的效率较差。

可能二

在可能二中，为了解决可能一存在的上述问题，预设特征分布规则可以设置为：联邦服务器 110 使用 embedding 算法对每个参与方设备中的用户进行编码，以将每个参与方设备中的用户映射到各个参与方设备对应的第一公共向量空间，并使用 embedding 算法对每个参与方设备中的对象进行编码，以将每个参与方设备中的对象映射到各个参与方设备对应的第二公共向量空间。进一步地，联邦服务器 110 将第一向量空间和第二向量空间同步给各个参与方设备，以使各个参与方设备基于相同的第一公共向量空间和第二公共向量空间构建待训练数据集。

举例来说，当用户行为特征为门禁数据时，若参与方设备 A 的用户包括用户 1 和用户 2，参与方设备 A 的对象包括门 1 和门 2，参与方设备 B 的用户包括用户 3 和用户 4，参与方设备 B 的对象包括门 3 和门 4，联邦服务器 110 通过 one-hot 算法将参与方设备 A 的用户 1 和用户 2（或者参与方设备 B 的用户 3 和用户 4）映射到第一公共向量空间[0 0]中，将参与方设备 A 的对象门 1 和门 2（或者参与方设备 B 的门 3 和门 4）映射到第二公共向量空间[0 0]中。其中，第一公共向量空间的第 1 位在参与方设备 A 中用于标识用户 1，在参与方设备 B 中用于标识用户 3，第一公共向量空间的第 2 位在参与方设备 A 中用于标识用户 2，在参与方设备 B 中用于标识用户 4；相应地，第二公共向量空间的第 1 位在参与方设备 A 中用于标识门 1，在参与方设备 B 中用于标识门 3，第二公共向量空间的第 2 位在参与方设备 A 中用于标识门 2，在参与方设备 B 中用于标识门 4。如此，若用户 1 在门 1 处刷卡，用户 2 未刷卡，用户 3 在门 4 处刷卡，用户 4 未刷卡，则参与方设备 A 构建得

到的待训练数据集中包括待训练数据[1 0 1 0]和待训练数据[0 1 0 0]，参与方设备 B 构建得到的待训练数据集中包括待训练数据[1 0 0 1]和待训练数据[0 1 1 0]。

由此可知，通过设置可能二中的预设特征分布规则，使得各个参与方设备中的待训练数据集能够具有相同的特征分布，且各个参与方设备的待训练数据中基本不包含无意义的数据，因此可以降低样本数据的数据量，提高模型训练的效率。然而，由于可能二将各个参与方设备的用户和对象分别映射到同一公共向量空间中，因此不同参与方设备的待训练数据中的相同位在不同参与方设备中表示不同的含义，如此，若使用具有相同形式但含义不同的待训练数据进行模型训练，会使得联邦服务器 110 无法提取出不同参与方设备的用户特征以及对象特征，导致模型的效果较差，可能会增大预测的误判概率。

基于此，本发明实施例提供了一种可能的数据处理方式，在该数据处理方式中，预设特征分布规则设置为各个参与方设备按照共用的用户行为特征维度和对象行为特征维度生成特征分布一致的待训练数据的规则。

图 3 为利用该预设特征分布规则处理得到待训练数据集的执行流程示意图，该方法适用于任一参与方设备，例如图 1 所示意出的参与方设备 121、参与方设备 122 或参与方设备 123。如图 3 所示，该方法包括：

步骤 301，参与方设备按照共用的用户行为特征维度，从用户行为数据中提取出任一用户的用户行为特征。

本发明实施例中，用户行为数据可以包括设定时段内参与方设备中的每个用户对一个或多个对象执行预设行为的行为数据。举例来说，若用户行为数据为社区门禁数据，设定时段为 10 月上旬，参与方设备为社区 A，则参与方设备的用户行为数据中可以包括社区 A 中的每个用户在 10 月上旬时在小区 A 的每个门处的刷卡数据。比如，若小区 A 中设置有门 1 和门 2，则参与方设备的用户行为数据可以包括用户在门 1 处刷卡的日期和时间、用户在门 2 处刷卡的日期和时间。在该示例中，预设行为是指刷卡行为，一个或多个对象是指小区 A 中的门 1 和门 2。再举一个例子，若用户行为数据为电商平台数据，参与方设备为书商 C，设定时段为 10 月上旬，则参与方设备的用户行为数据中可以包括书商 C 的每个用户 10 月上旬时在书商 C 的平台上购买书的消费数据。比如，若书商 C 的平台上发布有书 1 和书 2，则参与方设备的用户行为数据可以包括用户购买书 1 的日期和时间、用户购买书 2 的日期和时间。在该示例中，预设行为是指购买行为，一个或多个对象是指书商 C 的平台上发布的书 1 和书 2。

具体实施中，参与方设备可以按照共用的用户行为特征维度，从用户行为数据中提取出任一用户在任一共用的用户行为特征维度下的特征值，然后再根据该用户在各个共用的用户行为特征维度下的特征值，构建得到该用户的用户行为特征。其中，共用的用户行为特征维度用于指示用户对各个对象的执行情况，共用的用户行为特征维度可以包括用户执行预设行为的时间和/或用户在各时段内执行所述预设行为的次数，比如可以包括设定时段内用户每次执行预设行为的时间、设定时段内用户执行预设行为的总次数、设定时段的每个时间间隔内用户执行预设行为的次数，等等。举例来说，若用户行为数据包括社区 A 中的每个用户在 10 月上旬时在小区 A 的门 1、门 2 处的刷卡数据，则任一用户的用户行为特征可以包括 10 月上旬时该用户每次在门 1 或门 2 处刷卡的时间、10 月上旬时该用户在门 1 和门 2 处刷卡的总次数、10 月上旬的每周内该用户在门 1 和门 2 处刷卡的次数、10 月上旬的每天内该用户在门 1 和门 2 处刷卡的次数等。

在一个示例中，为了简化待训练数据集的数据量，并提高用户行为特征的时效性，共用的用户行为特征维度可以设置为包括用户最近一次执行预设行为的时间、用户最近一天执行预设行为的次数、用户最近一周执行预设行为的次数中的任意一项或任意多项。可以理解的，该示例中的时间信息（即最近一次、最近一天和最近一周）可以由本领域技术人员根据经验进行设置，比如若时间信息为最近五次、最近两天、最近两周，则共用的用户行为特征维度为用户最近五次执行预设行为的时间、用户最近两天执行预设行为的次数、用户最近两周执行预设行为的次数。

步骤 302，参与方设备按照共用的对象行为特征维度，从用户行为数据中提取出任一对象的对象行为特征。

具体实施中，参与方设备可以按照共用的对象行为特征维度，从用户行为数据中提取出任一对象在任一共用的对象行为特征维度下的特征值，然后再根据该对象在各个共用的对象行为特征维度下的特征值，构建得到该对象的对象行为特征。其中，共用的对象行为特征维度用于指示对象的被执行情况，共用的对象行为特征维度可以包括对象被执行预设行为的时间和/或对象在各时段内被执行预设行为的次数，比如可以包括设定时段内对象被执行预设行为的时间、设定时段内对象被执行预设行为的总次数、设定时段的每个时间间隔内对象被执行预设行为的次数，等等。举例来说，若用户行为数据包括社区 A 中的每个用户在 10 月上旬时在小区 A 的门 1、门 2 处的刷卡数据，对象为门 1，则该对象的对象行为特征可以包括 10 月上旬时门 1 被刷卡的时间、10 月上旬时门 1 被刷卡的总次数、10 月上旬的每周内门 1 被刷卡的次数、10 月上旬的每天内门 1 被刷卡的次数等。

在一个示例中，为了简化待训练数据集的数据量，提高对象行为特征的时效性，共用的对象行为特征特为可以设置为包括对象最近一次被执行预设行为的时间、对象最近一天被执行预设行为的次数、对象最近一周被执行预设行为的次数中的任意一项或任意多项。可以理解的，该示例中的时间信息（即最近一次、最近一天和最近一周）可以由本领域技术人员根据经验进行设置，比如若时间信息为最近三次、最近三天、最近两周，则共用的对象行为特征可以为对象最近三次被执行预设行为的时间、对象最近三天被执行预设行为的次数、对象最近两周被执行预设行为的次数。

步骤 303，参与方设备构建任一用户的用户行为特征和任一对象的对象行为特征之间的特征对，基于所述特征对，生成与其他参与方设备特征分布一致的待训练数据。

具体实施中，参与方设备可以先统计得到参与方设备中的全部用户和全部对象，然后针对任一用户和任一对象建立一个特征对。如此，参与方设备中的每个用户可以对应多个特征对，且参与方设备中的每个对象也可以对应多个特征对。举例来说，当小区 A 中的用户包括用户 1 和用户 2，对象包括门 1 和门 2 时，参与方设备可以根据小区 A 中的全部用户和全部对象构建四个特征对，分别为：user1-item1 特征对、user1-item2 特征对、user2-item1 特征对、user2-item2 特征对。其中，user1 是指用户 1，user2 是指用户 2，item1 是指对象 1，item2 是指对象 2。如此，通过为每个用户和每个对象设置一个特征对，能够使得每个用户和每个对象对应一个待训练数据，相比于现有技术直接使用每个用户的行为数据和每个对象的行为数据训练模型来说，可以降低待训练数据集的数据量，提高模型训练的效率。

在一个示例中，针对于任一特征对，参与方设备还可以按照共用的交互行为特征维度，从用户行为数据中提取出该特征对中的用户与该特征对中的对象的交互特征。其中，共用的交互行为特征维度用于指示用户对对象的执行情况，交互行为特征维度可以包括用户对

对象执行预设行为的时间和/或用户在各时段内对对象执行所述预设行为的次数，比如可以包括设定时段内用户对对象执行预设行为的时间、设定时段内用户对对象执行预设行为的总次数、设定时段的每个时间间隔内用户对对象执行预设行为的次数，等等。举例来说，若用户行为数据包括社区 A 中的每个用户在 10 月上旬时在小区 A 的门 1、门 2 处的刷卡数据，特征对中的用户和对象分别为用户 1 和门 1，则交互特征可以包括 10 月上旬时用户在门 1 处刷卡的时间、10 月上旬时用户在门 1 处刷卡的总次数、10 月上旬的每周内用户在门 1 处刷卡的次数、10 月上旬的每天内用户在门 1 处刷卡的次数等。

在一个示例中，为了简化待训练数据集的数据量，提高交互特征的时效性，共用的交互特征维度可以设置为如下任意一项或任意多项：用户最近一次对对象执行预设行为的时间、用户最近一周对对象执行预设行为的次数、用户对对象执行预设行为的次数在所有用户中的排名。可以理解的，该示例中的时间信息（即最近一次、最近一周）可以由本领域技术人员根据经验进行设置，比如若时间信息为最近三次、最近两周，则共用的交互特征维度也可以为用户最近三次对对象执行预设行为的时间、用户最近两周对对象执行预设行为的次数、用户对对象执行预设行为的次数在所有用户中的排名。

本发明实施例中，参与方设备可以采用监督式的机器学习算法执行模型训练。在监督式的机器学习算法中，每条待训练数据还需要设置对应的标签，标签可以根据行为预测模型的预测功能来设置，待训练数据对应的标签用于指示特征对中的用户和对象是否在设定时段内实现了行为预测模型的功能。

在一个示例中，参与方设备还可以根据交互特征确定特征对对应的标签，比如：若行为预测模型的功能为预测用户对对象刷卡的次数，则特征对对应的标签可以设置为设定时段内特征对中的用户对特征对中的对象刷卡的次数；若行为预测模型的功能为预测用户在某一时刻是否会在对象处刷卡，则特征对对应的标签可以基于设定时段的该时刻时特征对中的用户是否在特征对中的对象处刷卡来设置。比如若该时刻时特征对中的用户在特征对中的对象处刷卡，则可以设置特征对对应的标签为第一标签，若该时刻时特征对中的用户未在特征对中的对象处刷卡，则可以设置特征对对应的标签为第二标签。其中，第一标签和第二标签可以由本领域技术人员根据经验进行设置，比如第一标签可以设置为 1，第二标签可以设置为 0，或者第一标签可以设置为 0，第二标签可以设置为 1，具体不作限定。举例来说，若行为预测模型用于预测用户在未来时段的某一天是否会对某一对象执行预设行为，则当该特征对中的用户在设定时段（与未来时段时长相同）的某一天（与未来时段的某一天对应）对对象执行了预设行为时，该特征对对应的标签可以为第一标签，当该特征对中的用户在设定时段的某一天没有对对象执行预设行为时，该特征对对应的标签可以为第二标签。或者，若行为预测模型用于预测用户在某一时刻对对象执行预设行为，则当该特征对中的用户在设定时段的该时刻对对象执行了预设行为时，特征对对应的标签可以为第一标签，当该特征对中的用户在设定时段的该时刻没有对对象执行预设行为时，该特征对对应的标签可以为第二标签。

或者，特征对对应的标签也可以由特征向量来表示。比如，特征对对应的标签可以为基于设定时段的某一子时段内的用户行为数据提取得到的用户行为特征、对象行为特征和交互特征拼接得到的特征向量。举例来说，设定时段为第 1 天至第 7 天，若根据第 1 天至第 7 天的用户行为数据提取得到用户 1 刷卡的用户行为特征、门 1 被刷卡的对象行为特征和用户 1 在门 1 处刷卡的交互特征，则还可以根据第 8 天的用户行为数据提取得到第 8 天

中用户 1 在门 1 处刷卡的交互特征,然后将拼接第 8 天的交互特征作为特征对对应的标签。

本发明实施例中,通过提取特征对对应的标签,使得参与方设备能够基于监督式的机器学习算法实现模型训练,且行为预测模型更有针对性,预测效果较好。

进一步地,针对于任一特征对,参与方设备还可以对该特征对中的用户的用户行为特征、对象的对象行为特征、交互特征以及标签进行拼接,并将拼接得到的特征向量作为特征对对应的待训练数据。其中,拼接的顺序不作限定,比如可以按照用户的用户行为特征、对象的对象行为特征、交互特征、标签的顺序依次拼接,也可以按照对象的对象行为特征、用户的用户行为特征、交互特征、标签的顺序依次拼接,还可以按照交互特征、对象的对象行为特征、用户的用户行为特征、标签的顺序依次拼接,等等。

本发明实施例中,通过将用户执行预设行为的执行情况作为用户行为特征,将对象被执行预设行为的执行情况作为对象行为特征,将用户对对象执行行为的执行情况作为交互特征,可以使得不同参与方设备中用户的用户行为特征具有同一表达形式,不同参与方设备中对象的对象行为特征具有同一表达形式,不同参与方设备中用户与对象的交互特征也具有同一表达形式。如此,不同参与方设备可以根据与其他参与方设备具有同一表达形式的用户行为特征、与其他参与方设备具有同一表达形式的对象行为特征、与其他参与方设备具有同一表达形式的交互和标签构建特征分布一致的待训练数据。由于各个参与方设备的待训练数据具有相同的数据分布,且特征维度相同,因此不同参与方设备能够基于特征分布一致的待训练数据训练得到模型结构相同的参与方模型。

步骤 304,参与方设备基于各个用户与各个对象构成的各个特征对对应的待训练数据,构建所述待训练数据集。

本发明实施例中,每个待训练数据对应的标签可以指示该待训练数据为正样本数据还是负样本数据。比如当待训练数据对应的标签为第一标签时,待训练数据为正样本数据,当待训练数据对应的标签为第二标签时,待训练数据为负样本数据。或者,当待训练数据的标签指示在设定时段内实现了行为预测模型的功能时,待训练数据为正样本数据。当待训练数据中的标签指示在设定时段内未能实现行为预测模型的功能时,待训练数据为负样本数据。

在一种可能的场景中,由于用户具有惯性思维,因此用户可能会经常对某一对象执行预设行为,而很少对其它的对象执行预设行为。如此,参与方设备构建得到的各个待训练数据中会包含较少的正样本数据和较多的负样本数据。比如若小区中的门较多,而用户习惯性地小区的某一个门处刷卡,而很少在其它门处刷卡,因此在基于用户与小区中的各个门构建得到各个待训练数据后,各个待训练数据中的正样本数据的数量很少,而负样本数据的数量很多。如此,若参与方设备直接使用各个待训练数据进行模型训练,则数量较多的负样本数据可能会带偏模型参数,导致训练出的行为预测模型无法准确预测用户对对象执行预设行为的可能性,行为预测模型的效果较差。

为了解决上述问题,在一种可能的实现方式中,参与方设备在基于参与方设备中的全部用户与全部对象构建得到各个待训练数据后,可以先根据各个特征对对应的标签,确定出待训练数据集中的正样本数据和负样本数据,然后判断属于正负样本的待训练数据的比例是否符合预设范围,若不符合预设范围,则可以对负样本数据进行下采样处理,或者可以对正样本数据进行上采样处理,比如采用下采样方法来减少负样本数据的数量和权重,或者使用上采样方法来增多待训练数据中正样本数据的数量或权重。相应地,若符合预设

范围，则可以基于各个待训练数据构建待训练数据集。

下面分别介绍几种可能的上采样方法和下采样方法。

上采样方法一：参与方设备可以先从全部用户与全部对象对应的待训练数据中获取全部或部分正样本数据，然后对这全部或部分正样本数据进行复制，以通过复制的方式增多待训练数据中正样本数据的数量。

上采样方法二：参与方设备可以先从全部用户与全部对象对应的待训练数据中获取全部的正样本数据，然后每次从全部的正样本数据（或上次选取遗留的正样本数据）中选取至少两个正样本数据，拼接至少两个正样本数据中的不同部分得到一个新的正样本数据。比如，若选取了两个正样本数据，则可以将第一个正样本数据的前半部分和第二个正样本数据的后半部分拼接为一个新的正样本数据，或者可以将第一个正样本数据的后 1/3 部分和第二个正样本数据的前 2/3 部分拼接为一个新的正样本数据，具体不作限定。

上采样方法三：参与方设备可以先从全部用户与全部对象对应的待训练数据中获取全部或部分正样本数据，然后增大其中每个正样本数据在损失函数中的权重，不同正样本数据被增大的权重可以相同，也可以不同，不作限定。如此，参与方设备在使用正样本数据和负样本数据训练模型时，由于这些正样本数据的权重较大，因此可以降低负样本数据带偏模型参数的能力，降低模型训练的偏差，提高模型的效果。

下采样方法一：参与方设备可以先从全部用户与全部对象对应的待训练数据中获取全部的负样本数据，然后随机从全部的负样本数据中选取部分负样本数据进行删除，以通过随机删除的方式减少待训练数据中负样本数据的数量。

下采样方法二：参与方设备可以先从全部用户与全部对象对应的待训练数据中获取全部的负样本数据，然后计算得到每个负样本数据与其它负样本数据的相似度，若相似度小于预设相似度阈值，则删除该负样本数据，若相似度大于或等于预设相似度阈值，则保留该负样本数据。通过将与其它负样本数据的相似度较小的负样本删除，可以基于较为相似的负样本数据执行模型训练，从而提高行为预测模型识别不同待预测数据的能力，提高预测效果。

具体实施中，针对于任一负样本数据，参与方设备可以先计算该负样本数据与任一其它负样本数据的相似度，然后将该负样本数据与所有其它负待训练数据的相似度的平均值或加权平均值作为该负样本数据与其它负样本数据的相似度。

下采样方法三：参与方设备可以先从全部用户与全部对象对应的待训练数据中获取全部或部分负样本数据，然后减小其中每个负样本数据在损失函数中的权重，不同负样本数据被减小的权重可以相同，也可以不同，不作限定。如此，参与方设备在使用正样本数据和负样本数据训练模型时，由于这些负样本数据的权重较小，因此可以降低负样本数据带偏模型参数的能力，降低模型训练的偏差，提高模型的效果。

本发明实施例中，通过实验发现，若采用上采样方法一中的复制方式和/或上采样方法二中的拼接方式增加正样本数据的数量，或者，若采用下采样方法一中的删除方式和/或下采样方法二中的相似度删除方式降低负样本数据的数量，则在将待训练数据中正样本数据和负样本数据的数量比例调整至 1:3 时，模型训练的效果较好。相应地，若采用上采样方法三中的加权方式增加正样本数据的权重，或者，若采用下采样方法三中的降权方式降低负样本数据的权重，则在将待训练数据中正样本数据的权重和负样本数据的权重比例调整至 20:1 时，模型训练的效果较好。

在步骤 204 中, 参与方设备可以先将全部的待训练数据划分为训练数据、验证数据和测试数据, 训练数据用于参与方设备训练得到参与方模型, 验证数据用于参与方设备验证参与方模型的效果, 测试数据用于在模型训练结束后, 联邦服务器 110 验证行为预测模型的效果。

具体实施中, 参与方设备可以将训练得到的参与方模型的模型参数上报给联邦服务器 110, 参与方模型的模型参数用于联邦服务器 110 根据各个参与方模型的模型参数得到综合模型参数, 若确定满足模型训练的结束条件, 则根据综合模型参数得到行为预测模型, 若确定不满足模型训练的结束条件, 则将综合模型参数下发给各个参与方设备, 以联合各个参与方设备循环执行步骤 201~步骤 204。

作为一种示例, 参与方设备可以仅将参与方模型的模型参数上报给联邦服务器 110。如此, 在接收到各个参与方模型的模型参数后, 联邦服务器 110 可以基于各个模型参数计算得到平均模型参数, 并可以将平均模型参数确定为本次训练的综合模型参数。作为另一种示例, 参与方设备可以同时参与方模型的模型参数和损失函数发送给联邦服务器 110, 如此, 联邦服务器 110 在接收到各个模型参数和损失函数后, 可以先根据各个损失函数确定各个模型参数的权重, 再使用加权平均的方式对各个模型参数进行计算, 得到综合模型参数。其中, 若模型参数对应的损失函数越小, 说明参与方模型的效果越好, 因此可以为该模型参数分配较大的权重, 相应地, 若模型参数对应的损失函数越大, 说明参与方模型的效果越差, 因此可以为该模型参数分配较小的权重。具体实施中, 可以先按照由小到大的顺序对各个损失函数进行排序, 若损失函数的排序越靠后, 则可以设置该损失函数对应的模型参数的权重越小, 若损失函数的排序越靠前, 则可以设置该损失函数对应的模型参数的权重越大。举例来说, 若参与方设备 121~参与方设备 123 的损失函数分别为 0.05、0.30、0.15, 说明参与方设备 121~参与方设备 123 对应的参与方模型的效果排名为: 参与方设备 122 对应的参与方模型>参与方设备 123 对应的参与方模型>参与方设备 121 对应的参与方模型, 因此, 联邦服务器 110 可以设置参与方设备 121~参与方设备 123 对应的模型参数的权重分别为 10%、60%、30%。

本发明实施例中, 模型训练的结束条件可以包括以下任意一项或任意多项: 本次训练的综合模型参数收敛、已执行训练的次数大于或等于预设次数、已执行训练的时长大于或等于预设训练时长, 具体可以由本领域技术人员根据经验进行设置, 具体不作限定。具体实施中, 若模型训练的结束条件为已执行训练的次数大于或等于 5 次, 则当参与方设备依次训练了 5 次参与方模型 (即第 5 次训练结束) 后, 联邦服务器 110 可以确定第 5 次训练满足模型训练的结束条件。或者, 若模型训练的结束条件为已执行训练的时长大于或等于 5 分钟, 则从联邦服务器 110 下发模型训练请求开始至执行到第 5 分钟时 (若此时正在执行第 3 次训练), 联邦服务器 110 可以确定第 3 次训练满足模型训练的结束条件。相应的, 若模型训练的结束条件为本次训练的综合模型参数收敛, 则针对于本次训练, 联邦服务器 110 还可以根据本次训练中各参与方设备发送的损失函数计算得到本次训练的综合损失函数, 若确定本次训练的综合损失函数处于收敛状态 (比如本次训练的综合损失函数小于或等于某一阈值), 则可以确定本次训练满足模型训练的结束条件, 否则确定本次训练不满足模型训练的结束条件。

进一步地, 若本次训练满足模型训练的结束条件, 则联邦服务器 110 可以使用本次训练的综合模型参数构建得到行为预测模型, 若本次训练不满足模型训练的结束条件, 则联

邦服务器 110 可以将本次训练的综合模型参数下发给各个参与方设备，以使各个参与方设备基于本次训练的综合模型参数，使用各个参与方设备的待训练数据集重新执行下一次训练，直至满足模型训练的结束条件为止。

5 在一个示例中，联邦服务器 110 构建得到行为预测模型后，还可以将行为预测模型下发给各个参与方设备。如此，任一参与方设备接收到行为预测模型后，可以将待测特征对对应的特征信息输入行为预测模型中进行预测，得到待测特征对对应的预测标签。其中，待测特征对对应的特征信息可以包括待测用户的用户行为特征、待测对象的对象行为特征和待测用户与待测对象的交互特征中的任意一项或任意多项，预测标签用于确定待测用户是否会对待测对象执行预设行为。

10 本发明实施例中，通过联合各个参与方设备中的用户行为数据构建得到行为预测模型，使得行为预测模型能够用于对任一参与方设备中的用户的行为进行预测，行为预测模型具有较好的通用性。

下面从联邦服务器与各个参与方设备的交互角度描述本发明实施例中的数据处理方法。

15 图 4 为本发明实施例提供的一种模型训练的整体流程示意图，该方法包括：

步骤 401，联邦服务器 110 向各个参与方设备下发模型训练请求，模型训练请求中携带有初始模型参数。

20 步骤 402，任一参与方设备接收到联邦服务器下发的模型训练请求后，获取本地存储的用户行为数据，按照预设特征分布规则对用户行为数据进行处理，得到与其他参与方设备特征分布一致的待训练数据集。

步骤 403，任一参与方设备基于初始模型参数，利用待训练数据集训练得到与其他参与方设备模型结构一致的参与方模型。

步骤 404，任一参与方设备将参与方模型的模型参数上报给联邦服务器。

25 步骤 405，联邦服务器接收到各个参与方设备上报的参与方模型的模型参数后，根据各个参与方模型的模型参数得到综合模型参数。

步骤 406，联邦服务器确定是否满足模型训练的结束条件，若是，则执行步骤 407，若否，则执行步骤 408。

步骤 407，联邦服务器根据综合模型参数构建得到行为预测模型。

步骤 408，联邦服务器将综合模型参数下发给各个参与方设备。

30 步骤 409，任一参与方设备接收到联邦服务器下发的综合模型参数后，使用综合模型参数更新本地存储的初始模型参数，并执行步骤 403。

35 本发明的上述实施例中，参与方设备接收联邦服务器发送的模型训练请求，根据所述模型训练请求获取本地存储的用户行为数据，按照预设特征分布规则对所述用户行为数据进行处理，得到与其他参与方设备特征分布一致的待训练数据集，利用所述待训练数据集训练得到参与方设备模型，并将所述参与方设备模型发送给所述联邦服务器，以便于所述联邦服务器基于各个参与方设备模型联合训练得到行为预测模型。本发明实施例中，通过各个参与方设备按照预设特征分布规则对所述用户行为数据进行处理，得到与其他参与方设备特征分布一致的待训练数据集，使得各个参与方设备能够使用特征分布一致的待训练数据集训练得到模型结构一致的参与方设备模型，如此，联邦服务器能够基于各个参与方设备训练得到的模型结构一致的参与方设备模型训练得到行为预测模型，由于该行为预测

模型结合了各个参与方设备的行为数据特征，因此能够用于对各个参与方设备中的用户的行为进行预测，行为预测模型的通用性较好。

针对上述方法流程，本发明实施例还提供一种数据处理装置，该装置的具体内容可以参照上述方法实施。

图5为本发明实施例提供的一种数据处理装置的结构示意图，如图5所示，该装置包括：

收发模块 501，用于接收联邦服务器发送的模型训练请求；

获取模块 502，用于根据所述模型训练请求获取本地存储的用户行为数据；

处理模块 503，用于按照预设特征分布规则对所述用户行为数据进行处理，得到与其他参与方设备特征分布一致的待训练数据集；

训练模块 504，用于利用所述待训练数据集训练得到参与方模型，并将所述参与方模型发送给所述联邦服务器，所述联邦服务器用于基于各个参与方模型联合训练得到行为预测模型。

可选地，所述预设特征分布规则为按照共用的用户行为特征维度和对象行为特征维度生成特征分布一致的待训练数据的规则。在这种情况下，所述处理模块 503 具体用于：先按照所述共用的用户行为特征维度，从所述用户行为数据中提取出任一用户的用户行为特征，再按照所述共用的对象行为特征维度，从所述用户行为数据中提取出任一对象的对象行为特征，之后构建任一用户的用户行为特征和任一对象的对象行为特征之间的特征对，基于所述特征对，生成与其他参与方设备特征分布一致的待训练数据，最后基于各个用户与各个对象构成的各个特征对对应的待训练数据，构建所述待训练数据集。

可选地，所述处理模块 503 具体用于：先从所述用户行为数据中提取出所述用户在任一共用的用户行为特征维度下的特征值，再根据所述用户在各个共用的用户行为特征维度下的特征值，构建得到所述用户的用户行为特征。以及，从所述用户行为数据中提取出所述对象在任一共用的对象行为特征维度下的特征值，根据所述对象在各个共用的对象行为特征维度下的特征值，构建得到所述对象的对象行为特征。

可选地，所述处理模块 503 具体用于：先按照共用的交互行为特征维度，从所述用户行为数据中提取出所述用户与所述对象的交互特征，再根据所述用户与所述对象的交互特征，确定所述特征对对应的标签，之后将拼接所述用户的用户行为特征、所述对象的对象行为特征、所述交互特征及所述标签得到的特征向量作为所述特征对对应的待训练数据。

可选地，所述处理模块 503 具体用于：根据所述各个特征对对应的标签，确定属于正负样本的待训练数据的比例是否符合预设范围，若不符合预设范围，则对标签为负样本的待训练数据进行下采样处理，或者对标签为正样本的待训练数据进行上采样处理。若符合预设范围，则基于所述各个待训练数据构建所述待训练数据集。

可选地，所述装置还包括预测模块 505。在所述收发模块 501 将所述参与方模型发送给所述联邦服务器之后，所述收发模块 501 还接收所述联邦服务器发送的所述行为预测模型。所述预测模块 505 将待测特征对对应的特征信息输入所述行为预测模型中进行预测，得到所述待测特征对对应的预测标签。其中，所述待测特征对对应的特征信息包括待测用户的用户行为特征、待测对象的对象行为特征和所述待测用户与所述待测对象的交互特征中的任意一项或任意多项。所述预测标签用于确定所述待测用户是否会对所述待测对象执行预设行为。

可选地，所述共用的用户行为特征维度可以包括用户执行预设行为的时间和/或用户在

各时段内执行所述预设行为的次数。

可选地，所述共用的对象行为特征维度可以包括对象被执行所述预设行为的时间和/或对象在各时段内被执行所述预设行为的次数。

可选地，所述交互行为特征维度可以包括用户对对象执行预设行为的时间和/或用户在各时段内对对象执行所述预设行为的次数。

从上述内容可以看出：本发明的上述实施例中，参与方设备接收联邦服务器发送的模型训练请求，根据所述模型训练请求获取本地存储的用户行为数据，按照预设特征分布规则对所述用户行为数据进行处理，得到与其他参与方设备特征分布一致的待训练数据集，利用所述待训练数据集训练得到参与方模型，并将所述参与方模型发送给所述联邦服务器，以便于所述联邦服务器基于各个参与方模型联合训练得到行为预测模型。本发明实施例中，通过各个参与方设备按照预设特征分布规则对所述用户行为数据进行处理，得到与其他参与方设备特征分布一致的待训练数据集，使得各个参与方设备能够使用特征分布一致的待训练数据集训练得到模型结构一致的参与方模型，如此，联邦服务器能够基于各个参与方设备训练得到的模型结构一致的参与方模型训练得到行为预测模型，由于该行为预测模型结合了各个参与方设备的行为数据特征，因此能够用于对各个参与方设备中的用户的行为进行预测，行为预测模型的通用性较好。

基于同一发明构思，本发明实施例还提供了一种计算设备，包括至少一个处理单元以及至少一个存储单元，其中，所述存储单元存储有计算机程序，当所述程序被所述处理单元执行时，使得所述处理单元执行图 2 至图 4 任意所述的方法。

基于同一发明构思，本发明实施例还提供了一种计算机可读存储介质，其存储有可由计算设备执行的计算机程序，当所述程序在所述计算设备上运行时，使得所述计算设备执行图 2 至图 4 任意所述的方法。

基于同一发明构思，本发明实施例提供了一种终端设备，如图 6 所示，包括至少一个处理器 601，以及与至少一个处理器连接的存储器 602，本发明实施例中不限定处理器 601 与存储器 602 之间的具体连接介质，图 6 中处理器 601 和存储器 602 之间通过总线连接为例。总线可以分为地址总线、数据总线、控制总线等。

在本发明实施例中，存储器 602 存储有可被至少一个处理器 601 执行的指令，至少一个处理器 601 通过执行存储器 602 存储的指令，可以执行前述的数据处理方法中所包括的步骤。

其中，处理器 601 是终端设备的控制中心，可以利用各种接口和线路连接终端设备的各个部分，通过运行或执行存储在存储器 602 内的指令以及调用存储在存储器 602 内的数据，从而实现数据处理。可选的，处理器 601 可包括一个或多个处理单元，处理器 601 可集成应用处理器和调制解调处理器，其中，应用处理器主要处理操作系统、用户界面和应用程序等，调制解调处理器主要处理下发指令。可以理解的是，上述调制解调处理器也可以不集成到处理器 601 中。在一些实施例中，处理器 601 和存储器 602 可以在同一芯片上实现，在一些实施例中，它们也可以在独立的芯片上分别实现。

处理器 601 可以是通用处理器，例如中央处理器（CPU）、数字信号处理器、专用集成电路（Application Specific Integrated Circuit, ASIC）、现场可编程门阵列或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件，可以实现或者执行本发明实施例中公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者任何常规的处理器等。

结合数据处理实施例所公开的方法的步骤可以直接体现为硬件处理器执行完成，或者用处理器中的硬件及软件模块组合执行完成。

存储器 602 作为一种非易失性计算机可读存储介质，可用于存储非易失性软件程序、非易失性计算机可执行程序以及模块。存储器 602 可以包括至少一种类型的存储介质，例如可以包括闪存、硬盘、多媒体卡、卡型存储器、随机访问存储器(Random Access Memory, RAM)、静态随机访问存储器(Static Random Access Memory, SRAM)、可编程只读存储器(Programmable Read Only Memory, PROM)、只读存储器(Read Only Memory, ROM)、带电可擦除可编程只读存储器(Electrically Erasable Programmable Read-Only Memory, EEPROM)、磁性存储器、磁盘、光盘等等。存储器 602 是能够用于携带或存储具有指令或数据结构形式的期望的程序代码并能够由计算机存取的任何其他介质，但不限于此。本发明实施例中的存储器 602 还可以是电路或者其它任意能够实现存储功能的装置，用于存储程序指令和/或数据。

基于同一发明构思，本发明实施例提供了一种后端设备，如图 7 所示，包括至少一个处理器 701，以及与至少一个处理器连接的存储器 702，本发明实施例中不限定处理器 701 与存储器 702 之间的具体连接介质，图 7 中处理器 701 和存储器 702 之间通过总线连接为例。总线可以分为地址总线、数据总线、控制总线等。

在本发明实施例中，存储器 702 存储有可被至少一个处理器 701 执行的指令，至少一个处理器 701 通过执行存储器 702 存储的指令，可以执行前述的数据处理方法中所包括的步骤。

其中，处理器 701 是后端设备的控制中心，可以利用各种接口和线路连接后端设备的各个部分，通过运行或执行存储在存储器 702 内的指令以及调用存储在存储器 702 内的数据，从而实现数据处理。可选的，处理器 701 可包括一个或多个处理单元，处理器 701 可集成应用处理器和调制解调处理器，其中，应用处理器主要处理操作系统、应用程序等，调制解调处理器主要对接收到的指令进行解析以及对接收到的结果进行解析。可以理解的是，上述调制解调处理器也可以不集成到处理器 701 中。在一些实施例中，处理器 701 和存储器 702 可以在同一芯片上实现，在一些实施例中，它们也可以在独立的芯片上分别实现。

处理器 701 可以是通用处理器，例如中央处理器(CPU)、数字信号处理器、专用集成电路(Application Specific Integrated Circuit, ASIC)、现场可编程门阵列或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件，可以实现或者执行本发明实施例中公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者任何常规的处理器等。结合数据处理实施例所公开的方法的步骤可以直接体现为硬件处理器执行完成，或者用处理器中的硬件及软件模块组合执行完成。

存储器 702 作为一种非易失性计算机可读存储介质，可用于存储非易失性软件程序、非易失性计算机可执行程序以及模块。存储器 702 可以包括至少一种类型的存储介质，例如可以包括闪存、硬盘、多媒体卡、卡型存储器、随机访问存储器(Random Access Memory, RAM)、静态随机访问存储器(Static Random Access Memory, SRAM)、可编程只读存储器(Programmable Read Only Memory, PROM)、只读存储器(Read Only Memory, ROM)、带电可擦除可编程只读存储器(Electrically Erasable Programmable Read-Only Memory, EEPROM)、磁性存储器、磁盘、光盘等等。存储器 702 是能够用于携带或存储具有指令

或数据结构形式的期望的程序代码并能够由计算机存取的任何其他介质，但不限于此。本发明实施例中的存储器 702 还可以是电路或者其它任意能够实现存储功能的装置，用于存储程序指令和/或数据。

本领域内的技术人员应明白，本发明的实施例可提供为方法、或计算机程序产品。因此，本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

本发明是参照根据本发明实施例的方法、设备（系统）、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

尽管已描述了本发明的优选实施例，但本领域内的技术人员一旦得知了基本创造性概念，则可对这些实施例作出另外的变更和修改。所以，所附权利要求意欲解释为包括优选实施例以及落入本发明范围的所有变更和修改。

显然，本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围。这样，倘若本发明的这些修改和变型属于本发明权利要求及其等同技术的范围之内，则本发明也意图包含这些改动和变型在内。

权 利 要 求

1、一种数据处理方法，其特征在于，应用于参与方设备，所述方法包括：

接收联邦服务器发送的模型训练请求；

根据所述模型训练请求获取本地存储的用户行为数据；

5 按照预设特征分布规则对所述用户行为数据进行处理，得到与其他参与方设备特征分布一致的待训练数据集；

利用所述待训练数据集训练得到参与方模型，并将所述参与方模型发送给所述联邦服务器，所述联邦服务器用于基于各个参与方模型联合训练得到行为预测模型。

10 2、根据权利要求 1 所述的方法，其特征在于，所述预设特征分布规则为按照共用的用户行为特征维度和对象行为特征维度生成特征分布一致的待训练数据的规则；

所述按照预设特征分布规则对所述用户行为数据进行处理，得到与其他参与方设备特征分布一致的训练数据集，包括：

按照所述共用的用户行为特征维度，从所述用户行为数据中提取出任一用户的用户行为特征；

15 按照所述共用的对象行为特征维度，从所述用户行为数据中提取出任一对象的对象行为特征；

构建任一用户的用户行为特征和任一对象的对象行为特征之间的特征对，基于所述特征对，生成与其他参与方设备特征分布一致的待训练数据；

20 基于各个用户与各个对象构成的各个特征对对应的待训练数据，构建所述待训练数据集。

3、根据权利要求 2 所述的方法，其特征在于，所述按照所述共用的用户行为特征维度，从所述用户行为数据中提取出任一用户的用户行为特征，包括：

25 从所述用户行为数据中提取出所述用户在任一共用的用户行为特征维度下的特征值，根据所述用户在各个共用的用户行为特征维度下的特征值，构建得到所述用户的用户行为特征；

所述按照所述共用的对象行为特征维度，从所述用户行为数据中提取出任一对象的对象行为特征，包括：

30 从所述用户行为数据中提取出所述对象在任一共用的对象行为特征维度下的特征值，根据所述对象在各个共用的对象行为特征维度下的特征值，构建得到所述对象的对象行为特征。

4、根据权利要求 2 所述的方法，其特征在于，所述基于所述特征对，生成与其他参与方设备特征分布一致的待训练数据，包括：

按照共用的交互行为特征维度，从所述用户行为数据中提取出所述用户与所述对象的交互特征；

35 根据所述用户与所述对象的交互特征，确定所述特征对对应的标签；

将拼接所述用户的用户行为特征、所述对象的对象行为特征、所述交互特征及所述标签得到的特征向量作为所述特征对对应的待训练数据。

5、根据权利要求 4 所述的方法，其特征在于，所述根据所述各个特征对对应的待训练数据，构建所述待训练数据集，包括：

根据所述各个特征对对应的标签，确定属于正负样本的待训练数据的比例是否符合预设范围；

若不符合预设范围，则对标签为负样本的待训练数据进行下采样处理，或者对标签为正样本的待训练数据进行上采样处理；若符合预设范围，则基于所述各个待训练数据构建所述待训练数据集。

6、根据权利要求 5 所述的方法，其特征在于，

所述对标签为负样本的待训练数据进行下采样处理，包括：

减少标签为负样本的所述待训练数据的数量和权重；

所述对标签为正样本的待训练数据进行上采样处理，包括：

增加标签为正样本的所述待训练数据的数量或权重。

7、根据权利要求 1 至 6 中任一项所述的方法，其特征在于，所述将所述参与方模型发送给所述联邦服务器之后，还包括：

接收所述联邦服务器发送的所述行为预测模型；

将待测特征对对应的特征信息输入所述行为预测模型中进行预测，得到所述待测特征对对应的预测标签；所述待测特征对对应的特征信息包括待测用户的用户行为特征、待测对象的对象行为特征和所述待测用户与所述待测对象的交互特征中的任意一项或任意多项；所述预测标签用于确定所述待测用户是否会对所述待测对象执行预设行为。

8、根据权利要求 2 至 6 中任一项所述的方法，其特征在于，所述共用的用户行为特征维度包括用户执行预设行为的时间和/或用户在各时段内执行所述预设行为的次数；所述共用的对象行为特征维度包括对象被执行所述预设行为的时间和/或对象在各时段内被执行所述预设行为的次数。

9、根据权利要求 4 至 6 中任一项所述的方法，其特征在于，所述交互行为特征维度包括用户对对象执行预设行为的时间和/或用户在各时段内对对象执行所述预设行为的次数。

10、一种数据处理装置，其特征在于，所述装置包括：

收发模块，用于接收联邦服务器发送的模型训练请求；

获取模块，用于根据所述模型训练请求获取本地存储的用户行为数据；

处理模块，用于按照预设特征分布规则对所述用户行为数据进行处理，得到与其他参与方设备特征分布一致的待训练数据集；

训练模块，用于利用所述待训练数据集训练得到参与方模型，并将所述参与方模型发送给所述联邦服务器，所述联邦服务器用于基于各个参与方模型联合训练得到行为预测模型。

11、根据权利要求 10 所述的装置，其特征在于，所述预设特征分布规则为按照共用的用户行为特征维度和对象行为特征维度生成特征分布一致的待训练数据的规则；

所述处理模块具体用于：

按照所述共用的用户行为特征维度，从所述用户行为数据中提取出任一用户的用户行为特征；

按照所述共用的对象行为特征维度，从所述用户行为数据中提取出任一对象的对象行为特征；

构建任一用户的用户行为特征和任一对象的对象行为特征之间的特征对，基于所述特

征对,生成与其他参与方设备特征分布一致的待训练数据;

基于各个用户与各个对象构成的各个特征对对应的待训练数据,构建所述待训练数据集。

12、根据权利要求 11 所述的装置,其特征在于,所述处理模块具体用于:

5 从所述用户行为数据中提取出所述用户在任一共用的用户行为特征维度下的特征值,根据所述用户在各个共用的用户行为特征维度下的特征值,构建得到所述用户的用户行为特征;

所述按照所述共用的对象行为特征维度,从所述用户行为数据中提取出任一对象的对象行为特征,包括:

10 从所述用户行为数据中提取出所述对象在任一共用的对象行为特征维度下的特征值,根据所述对象在各个共用的对象行为特征维度下的特征值,构建得到所述对象的对象行为特征。

13、根据权利要求 11 所述的装置,其特征在于,所述处理模块具体用于:

15 按照共用的交互行为特征维度,从所述用户行为数据中提取出所述用户与所述对象的交互特征;

根据所述用户与所述对象的交互特征,确定所述特征对对应的标签;

将拼接所述用户的用户行为特征、所述对象的对象行为特征、所述交互特征及所述标签得到的特征向量作为所述特征对对应的待训练数据。

14、根据权利要求 13 所述的装置,其特征在于,所述处理模块具体用于:

20 根据所述各个特征对对应的标签,确定属于正负样本的待训练数据的比例是否符合预设范围;

若不符合预设范围,则对标签为负样本的待训练数据进行下采样处理,或者对标签为正样本的待训练数据进行上采样处理;若符合预设范围,则基于所述各个待训练数据构建所述待训练数据集。

25 15、根据权利要求 14 所述的装置,其特征在于,所述处理模块具体用于:

减少标签为负样本的所述待训练数据的数量和权重;或者,

增加标签为正样本的所述待训练数据的数量或权重。

16、根据权利要求 10 至 15 中任一项所述的装置,其特征在于,所述装置还包括预测模块;在所述收发模块将所述参与方模型发送给所述联邦服务器之后:

30 所述收发模块还用于:接收所述联邦服务器发送的所述行为预测模型;

所述预测模块用于:将待测特征对对应的特征信息输入所述行为预测模型中进行预测,得到所述待测特征对对应的预测标签;

35 其中,所述待测特征对对应的特征信息包括待测用户的用户行为特征、待测对象的对象行为特征和所述待测用户与所述待测对象的交互特征中的任意一项或任意多项;所述预测标签用于确定所述待测用户是否会对所述待测对象执行预设行为。

17、根据权利要求 11 至 15 中任一项所述的装置,其特征在于,所述共用的用户行为特征维度包括用户执行预设行为的时间和/或用户在各时段内执行所述预设行为的次数;所述共用的对象行为特征维度包括对象被执行所述预设行为的时间和/或对象在各时段内被执行所述预设行为的次数。

40 18、根据权利要求 13 至 15 中任一项所述的装置,其特征在于,所述交互行为特征维

度包括用户对对象执行预设行为的时间和/或用户在各时段内对对象执行所述预设行为的次数。

5 19、一种计算设备，其特征在于，包括至少一个处理单元以及至少一个存储单元，其中，所述存储单元存储有计算机程序，当所述程序被所述处理单元执行时，使得所述处理单元执行权利要求1~9任一权利要求所述的方法。

20、一种计算机可读存储介质，其特征在于，其存储有可由计算设备执行的计算机程序，当所述程序在所述计算设备上运行时，使得所述计算设备执行权利要求1~9任一权利要求所述的方法。

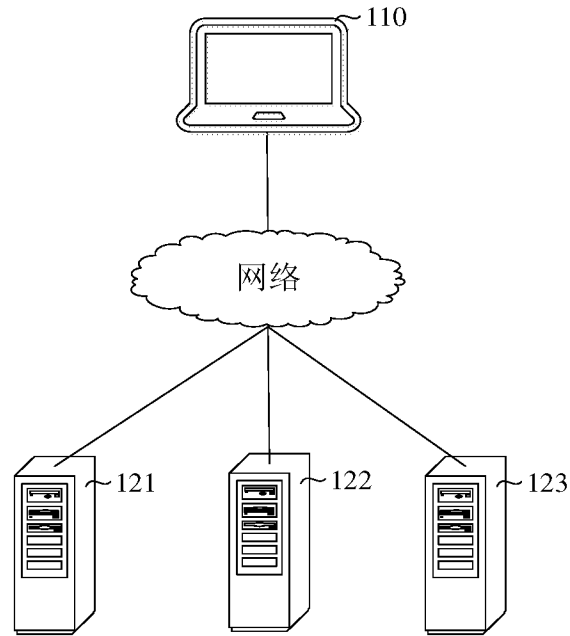


图 1

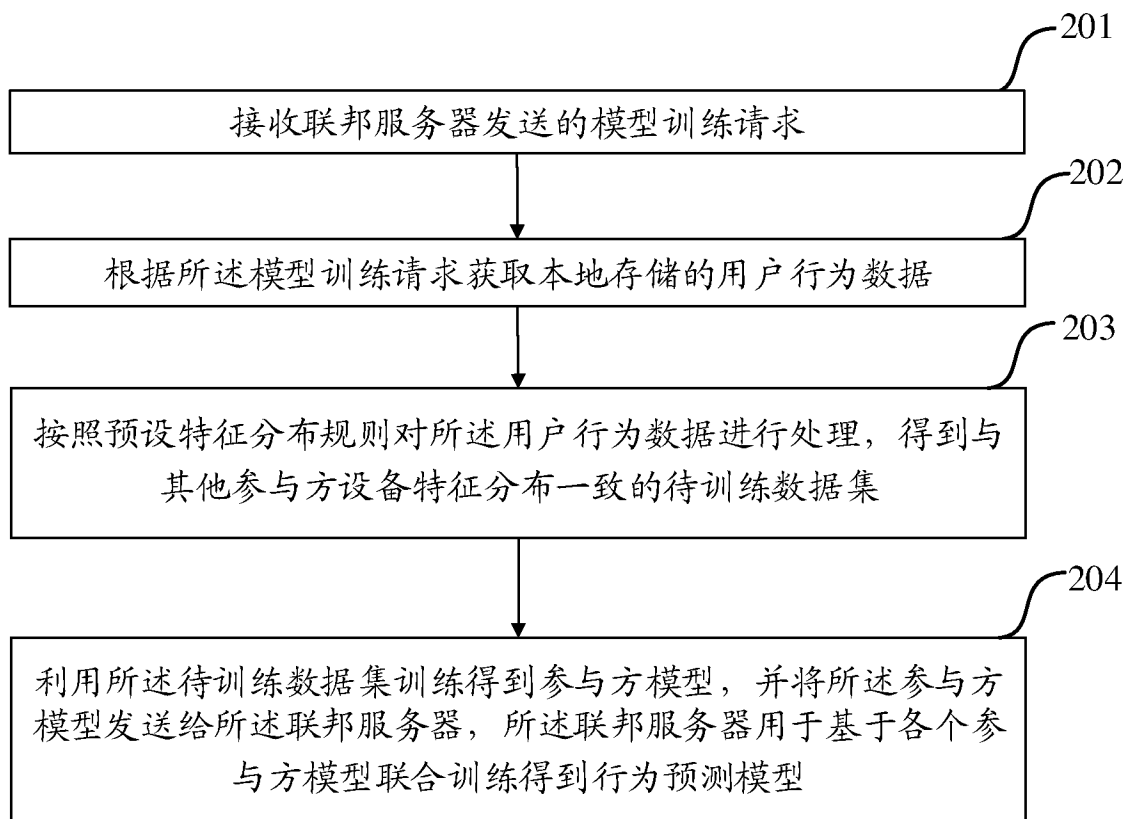


图 2

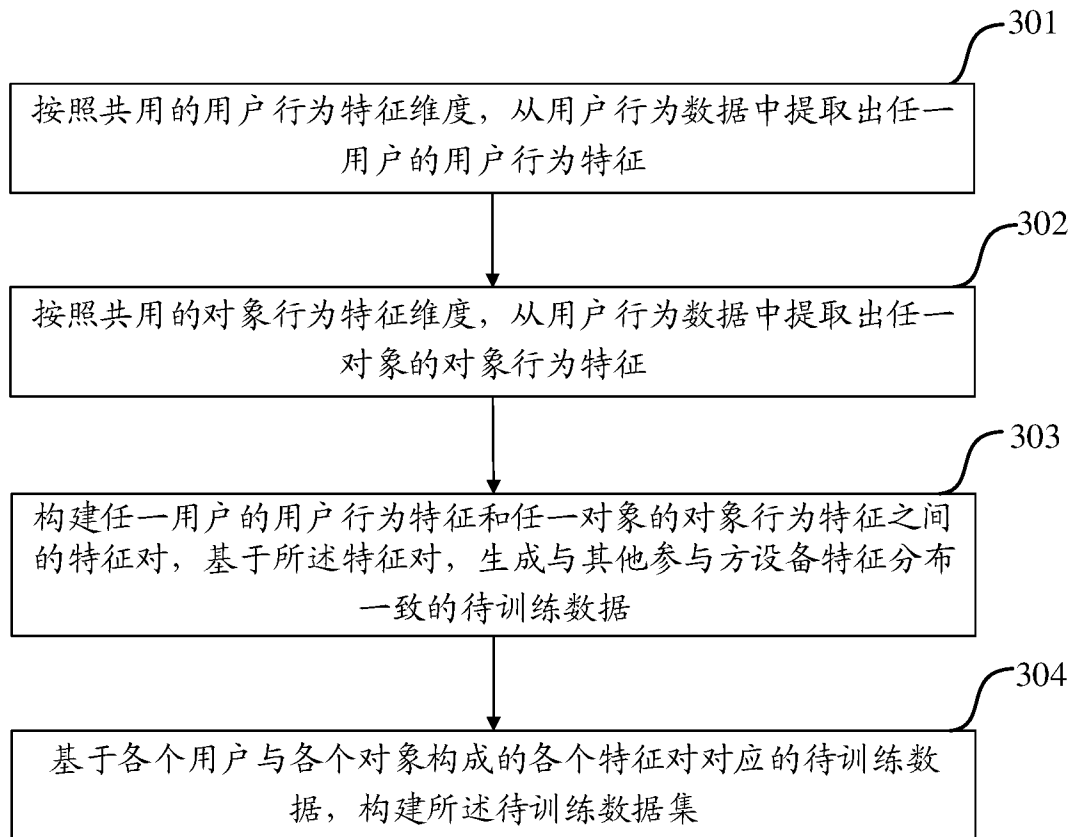


图 3

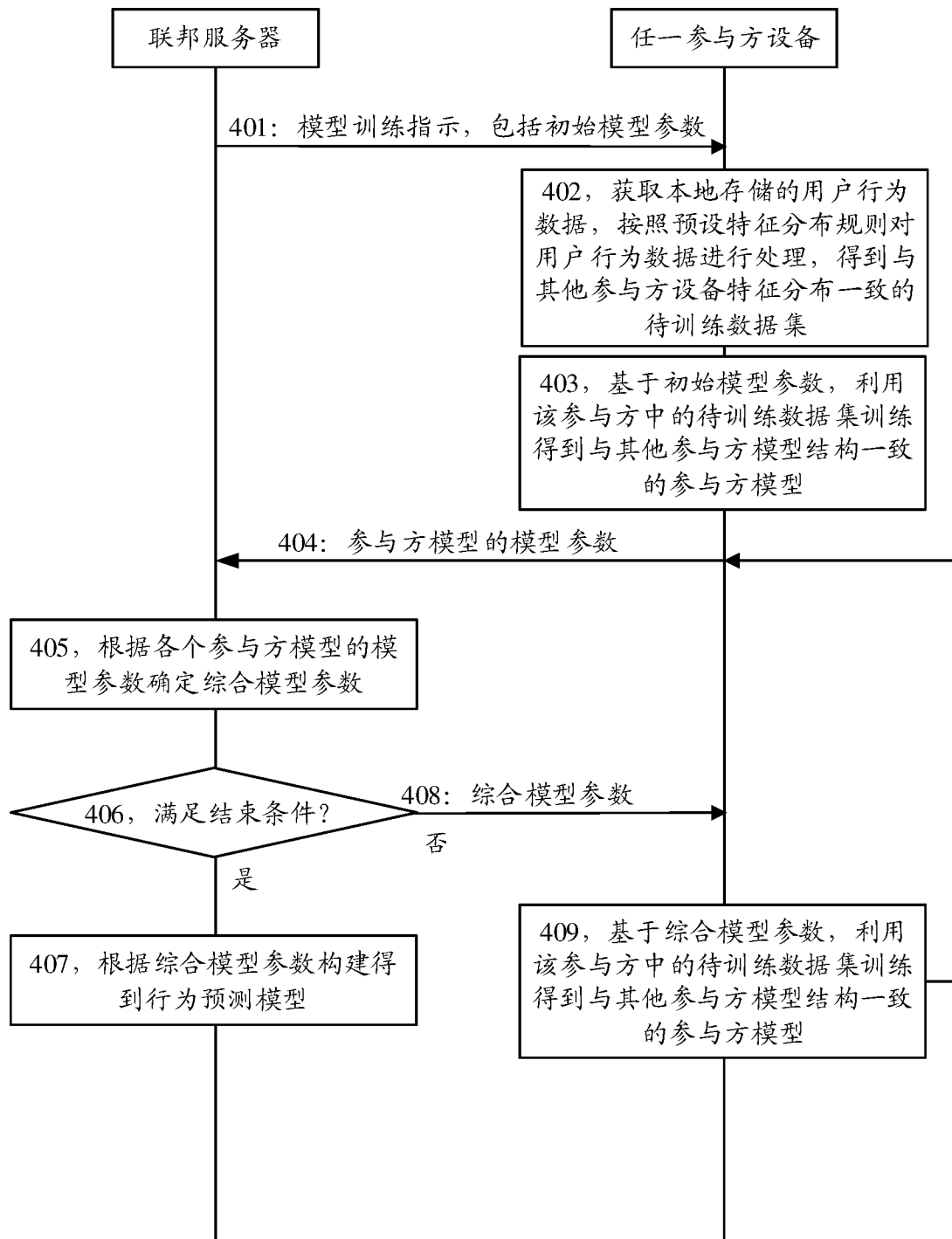


图 4

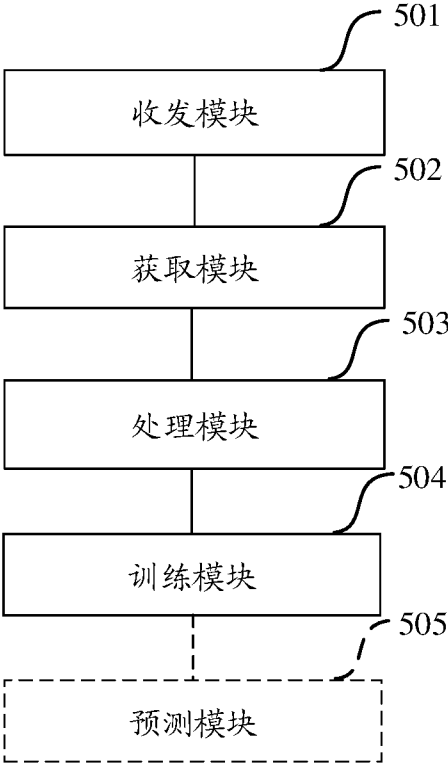


图 5

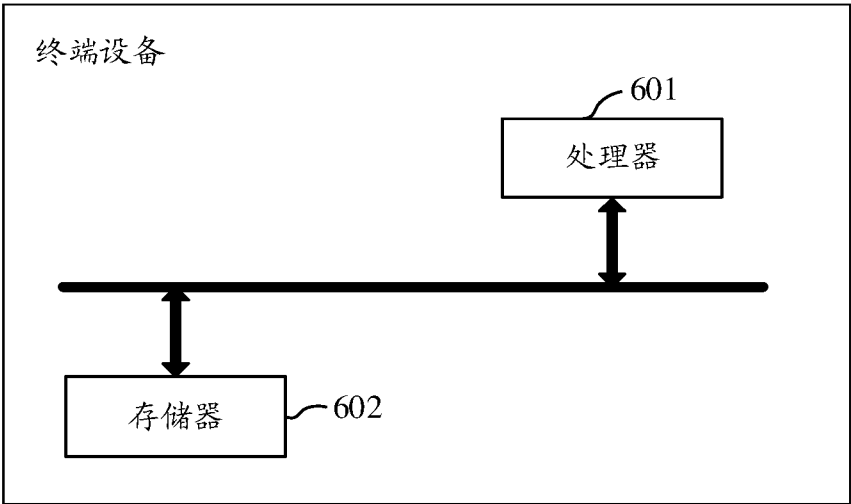


图 6

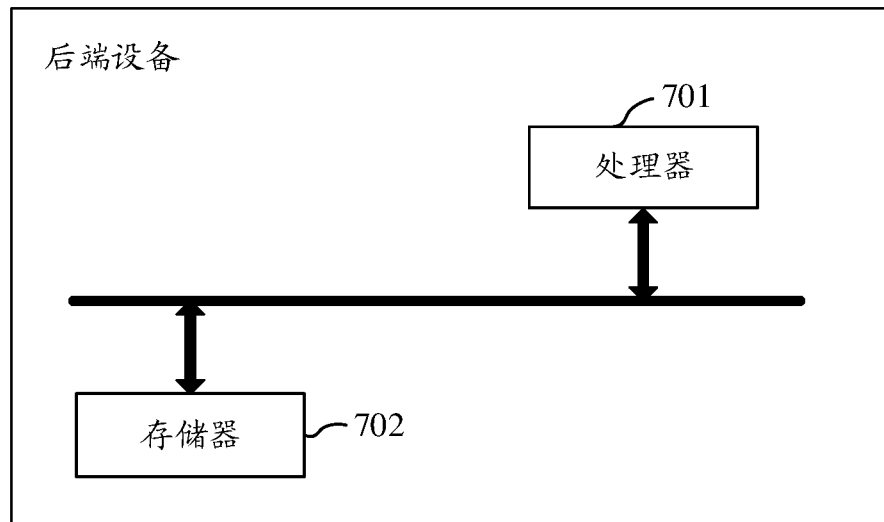


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/129123

A. CLASSIFICATION OF SUBJECT MATTER

G06Q 30/02(2012.01)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)

G06Q

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: 模型, 训练, 学习, 联邦, 机器, 参与, 参加, 节点, 用户, 特征, 预设, 预测, 预估, 维度, 标签, model, train, learn, federal, machine, participate, node, user, feature, characteristic, forecast, estimate, predict, dimension, label

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 111275491 A (SHENZHEN QIANHAI WEBANK CO., LTD.) 12 June 2020 (2020-06-12) claims 1-11, description paragraph 0083	1-20
X	CN 110633806 A (SHENZHEN QIANHAI WEBANK CO., LTD.) 31 December 2019 (2019-12-31) claims 1-2, 9-10, description paragraphs 0049, 0096-0097	1, 10, 19-20
A	CN 110309923 A (SHENZHEN QIANHAI WEBANK CO., LTD.) 08 October 2019 (2019-10-08) entire document	1-20
A	CN 110598870 A (SHENZHEN QIANHAI WEBANK CO., LTD.) 20 December 2019 (2019-12-20) entire document	1-20
A	WO 2018052987 A1 (OHIO STATE INNOVATION FOUNDATION et al.) 22 March 2018 (2018-03-22) 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

29 January 2021

Date of mailing of the international search report

19 February 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

Authorized officer

Facsimile No. (86-10)62019451

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/129123

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	111275491	A	12 June 2020	None			
CN	110633806	A	31 December 2019	None			
CN	110309923	A	08 October 2019	None			
CN	110598870	A	20 December 2019	None			
WO	2018052987	A1	22 March 2018	US	2019206057	A1	04 July 2019
				EP	3512415	A1	24 July 2019

国际检索报告

国际申请号

PCT/CN2020/129123

A. 主题的分类

G06Q 30/02 (2012.01) 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)

G06Q

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: 模型, 训练, 学习, 联邦, 机器, 参与, 参加, 节点, 用户, 特征, 预设, 预测, 预估, 维度, 标签, model, train, learn, federal, machine, participate, node, user, feature, characteristic, forecast, estimate, predict, dimension, label

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 111275491 A (深圳前海微众银行股份有限公司) 2020年 6月 12日 (2020 - 06 - 12) 权利要求1-11, 说明书第0083段	1-20
X	CN 110633806 A (深圳前海微众银行股份有限公司) 2019年 12月 31日 (2019 - 12 - 31) 权利要求1-2, 9-10, 说明书第0049, 0096-0097段	1, 10, 19-20
A	CN 110309923 A (深圳前海微众银行股份有限公司) 2019年 10月 8日 (2019 - 10 - 08) 全文	1-20
A	CN 110598870 A (深圳前海微众银行股份有限公司) 2019年 12月 20日 (2019 - 12 - 20) 全文	1-20
A	WO 2018052987 A1 (OHIO STATE INNOVATION FOUNDATION 等) 2018年 3月 22日 (2018 - 03 - 22) 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 1月 29日

国际检索报告邮寄日期

2021年 2月 19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

胡妮

电话号码 86-(10)-53961429

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/129123

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	111275491	A	2020年 6月 12日	无			
CN	110633806	A	2019年 12月 31日	无			
CN	110309923	A	2019年 10月 8日	无			
CN	110598870	A	2019年 12月 20日	无			
WO	2018052987	A1	2018年 3月 22日	US	2019206057	A1	2019年 7月 4日
				EP	3512415	A1	2019年 7月 24日