

用于评估联合训练模型的方法和装置

5 相关申请的交叉引用

本申请要求于 2020 年 04 月 08 日提交的、申请号为 202010272019.7、发明名称为“用于评估联合训练模型的方法和装置”的中国专利申请的优先权，该申请的全文通过引用结合在本申请中。

技术领域

- 10 本申请实施例涉及计算机技术领域，具体涉及用于评估联合训练模型的方法和装置。

背景技术

- 随着人工智能技术的发展，机器学习取得越来越广泛的应用。近
- 15 年来，为了解决数据孤岛和数据隐私的问题，相关的方式通常采用联合训练模型来实现在不暴露原始数据的情况下完成机器学习模型的构建。上述联合训练模型的训练往往通过多个模型训练参与方之间传递的模型梯度对模型参数进行调整来实现。

20 发明内容

本申请实施例提出了用于评估联合训练模型的方法和装置。

- 第一方面，本申请实施例提供了一种用于评估联合训练模型的方法，该方法包括：接收目标设备发送的模型评价数据请求，其中，目标设备包括联合训练模型的参与方；获取与模型评价数据请求相匹配
- 25 的样本集合，其中，匹配的样本集合属于与联合训练模型相关联的有标签数据；根据匹配的样本集合，生成联合训练模型的模型评价数据。

- 在一些实施例中，上述模型评价数据请求中包括数据划分信息，上述数据划分信息用于指示样本的划分方法；以及上述获取与模型评价数据请求相匹配的样本集合，包括：获取用于训练联合训练模型的
- 30 样本集合；根据数据划分信息，从样本集合中选取样本，组成与模型

评价数据请求相匹配的样本集合。

在一些实施例中，上述根据匹配的样本集合，生成联合训练模型的模型评价数据，包括：将匹配的样本集合划分为训练样本集合和测试样本集合；利用匹配的样本集合，生成联合训练模型的指标变化信息，其中，联合训练模型的指标变化信息用于指示联合训练模型的模型评价数据随时间维度的变化。

在一些实施例中，上述模型评价数据请求中还包括数据查询权限信息，数据查询权限信息用于指示指标变化信息中的至少一个维度；以及该方法还包括：根据数据查询权限信息，从指标变化信息中提取与数据查询权限信息匹配的模型评价数据；向目标设备发送匹配的模型评价数据。

第二方面，本申请实施例提供了一种用于评估联合训练模型的装置，该装置包括：接收单元，被配置成接收目标设备发送的模型评价数据请求，其中，目标设备包括联合训练模型的参与方；获取单元，被配置成获取与模型评价数据请求相匹配的样本集合，其中，匹配的样本集合属于与联合训练模型相关联的有标签数据；生成单元，被配置成根据匹配的样本集合，生成联合训练模型的模型评价数据。

在一些实施例中，上述模型评价数据请求中包括数据划分信息，数据划分信息用于指示样本的划分方法；上述获取单元包括：获取模块，被配置成获取用于训练联合训练模型的样本集合；第一生成模块，被配置成根据数据划分信息，从样本集合中选取样本，组成与模型评价数据请求相匹配的样本集合。

在一些实施例中，上述生成单元包括：划分模块，被配置成将匹配的样本集合划分为训练样本集合和测试样本集合；第二生成模块，被配置成利用匹配的样本集合，生成联合训练模型的指标变化信息，其中，联合训练模型的指标变化信息用于指示联合训练模型的模型评价数据随时间维度的变化。

在一些实施例中，上述模型评价数据请求中还包括数据查询权限信息，数据查询权限信息用于指示指标变化信息中的至少一个维度；该装置还包括：提取单元，被配置成根据数据查询权限信息，从指标

变化信息中提取与数据查询权限信息匹配的模型评价数据；发送单元，被配置成向目标设备发送匹配的模型评价数据。

第三方面，本申请实施例提供了一种用于评估联合训练模型的系统，该系统包括：第一参与端，被配置成获取本地样本集合，其中，
5 本地样本集合中的本地样本属于用于训练联合训练模型的无标签数据；生成与本地样本集合相关联的联合训练模型的模型评价数据请求；向联合训练模型的第二参与端发送模型评价数据请求，其中，第二参与端与第一参与端共同训练联合训练模型；第二参与端，被配置成接收第一参与端发送的模型评价数据请求；获取与模型评价数据请求相
10 匹配的样本集合，其中，匹配的样本集合属于与联合训练模型相关联的有标签数据；根据匹配的样本集合，生成联合训练模型的模型评价数据。

在一些实施例中，上述模型评价数据请求中包括数据划分信息；上述第一参与端进一步被配置成：根据本地样本集合中样本的特征维
15 度，生成数据划分信息，其中，数据划分信息用于指示样本的划分方法。

在一些实施例中，上述第一参与端还被配置成：从第二参与端获取联合训练模型的模型评价数据；基于模型评价数据，调整联合训练模型。

第四方面，本申请实施例提供了一种服务器，该服务器包括：一个或多个处理器；存储装置，其上存储有一个或多个程序；当一个或多个程序被一个或多个处理器执行，使得一个或多个处理器实现如第一方面中任一实现方式描述的方法。

第五方面，本申请实施例提供了一种计算机可读介质，其上存储
25 有计算机程序，该程序被处理器执行时实现如第一方面中任一实现方式描述的方法。

本申请实施例提供的用于评估联合训练模型的方法和装置，通过首先接收目标设备发送的模型评价数据请求。其中，目标设备包括联合训练模型的参与方。而后，获取与模型评价数据请求相匹配的样本
30 集合。其中，匹配的样本集合属于与联合训练模型相关联的有标签数

据。最后，根据匹配的样本集合，生成联合训练模型的模型评价数据。从而实现了在不暴露原始样本数据的前提下共享联合训练模型的效果指标。从而为联合训练模型的优化和改进提供及时、有效的数据参考基础。

5

附图说明

通过阅读参照以下附图所作的对非限制性实施例所作的详细描述，本申请的其它特征、目的和优点将会变得更明显：

图 1 是本申请的一个实施例可以应用于其中的示例性系统架构图；
10 图；

图 2 是根据本申请的用于评估联合训练模型的方法的一个实施例的流程图；

图 3 是根据本申请的实施例的用于评估联合训练模型的方法的一个应用场景的示意图；

15 图 4 是根据本申请的用于评估联合训练模型的方法的又一个实施例的流程图；

图 5 是根据本申请的用于评估联合训练模型的装置的一个实施例的结构示意图；

图 6 是根据本申请的用于评估联合训练模型的系统的一个实施例中各个设备之间交互的时序图。
20

图 7 是适于用来实现本申请的实施例的电子设备的结构示意图。

具体实施方式

下面结合附图和实施例对本申请作进一步的详细说明。可以理解
25 的是，此处所描述的具体实施例仅仅用于解释相关发明，而非对该发明的限定。另外还需要说明的是，为了便于描述，附图中仅示出了与有关发明相关的部分。

需要说明的是，在不冲突的情况下，本申请中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本
30 申请。

图 1 示出了可以应用本申请的用于评估联合训练模型的方法或用于评估联合训练模型的装置的示例性架构 100。

如图 1 所示,系统架构 100 可以包括终端设备 101、102,网络 103、104 和服务器 105、服务器 106。网络 103 用以在终端设备 101、102 和服务器 105、106 之间提供通信链路的介质。网络 104 用以在服务器 105 和服务器 106 之间提供通信链路的介质。网络 103、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。

终端设备 101、102 通过网络 103 与服务器 105、106 交互,以接收或发送消息等。终端设备 101、102 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、文本编辑类应用等。

终端设备 101、102 可以是硬件,也可以是软件。当终端设备 101、102 为硬件时,可以是具有显示屏并且支持数据传输的各种电子设备,包括但不限于智能手机、平板电脑、电子书阅读器、膝上型便携计算机和台式计算机等等。当终端设备 101、102 为软件时,可以安装在上述所列举的电子设备中。其可以实现成多个软件或软件模块(例如用来提供分布式服务的软件或软件模块),也可以实现成单个软件或软件模块。在此不做具体限定。

服务器 105、106 可以是提供各种服务的服务器,例如为终端设备 101、102 上安装的各种通讯客户端应用提供支持的后台服务器。后台服务器可以对接收到的数据请求进行分析等处理,并将处理结果(如推荐的网页内容)反馈给终端设备。服务器 105、106 可以通过网络 104 进行中间数据的传输以共同训练机器学习模型。

需要说明的是,服务器可以是硬件,也可以是软件。当服务器为硬件时,可以实现成多个服务器组成的分布式服务器集群,也可以实现成单个服务器。当服务器为软件时,可以实现成多个软件或软件模块(例如用来提供分布式服务的软件或软件模块),也可以实现成单个软件或软件模块。在此不做具体限定。

需要说明的是,本申请实施例所提供的用于评估联合训练模型的方法一般由服务器 105 或服务器 106 执行,相应地,用于评估联合训

练模型的装置一般设置于服务器 105 或服务器 106 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。

5 继续参考图 2，示出了根据本申请的用于评估联合训练模型的方法的一个实施例的流程 200。该用于评估联合训练模型的方法包括以下步骤：

步骤 201，接收目标设备发送的模型评价数据请求。

在本实施例中，用于评估联合训练模型的方法的执行主体（如图
10 1 所示的服务器 105 或服务器 106）可以通过有线连接方式或者无线连接方式接收目标设备发送的模型评价数据请求。其中，上述目标设备可以包括联合训练模型的参与方。上述联合训练模型可以包括各种能够实现在不暴露原始数据的前提下由多方共同参与训练的模型。其中，参与上述联合训练模型的参与方之间可以传输加密后的特征和模型中
15 间结果（例如梯度），以调整各参与方训练的子模型的模型参数。上述模型评价数据请求可以用于获取上述执行主体和上述目标设备共同训练的联合训练模型的模型评价数据。上述模型评价数据可以包括但不限于以下至少一项：准确率，召回率，AUC(area under curve, ROC 曲线下的面积)，PSI(Population Stability Index, 模型稳定度指标),对数损失（Log Loss）。
20

需要说明的是，作为示例，在两个参与方的联合训练模型的应用场景下，各参与方分别存储有样本集。两方样本集通常具有一定的交集。通常，两个参与方中仅有一方的样本集中的样本具有标签，而另一方的样本集中的样本仅包括特征。而且，在训练前，通常需要两个
25 参与方进行样本对齐。

在本实施例的一些可选的实现方式中，上述模型评价数据请求中还可以包括数据查询权限信息。其中，上述数据查询权限信息可以用于指示上述模型评价数据中的至少一个维度。上述维度可以包括但不限于以下至少一项：模型评价数据的标识（key），采样粒度，预设时间
30 段。

步骤 202, 获取与模型评价数据请求相匹配的样本集合。

在本实施例中, 上述执行主体可以通过有线连接方式或者无线连接方式获取上述与模型评价数据请求相匹配的样本集合。其中, 上述匹配的样本集合通常属于与上述联合训练模型相关联的有标签数据, 5 即上述样本集合中的样本可以包括特征和与特征对应的标签。

作为示例, 上述执行主体通常可以获取预先存储于本地的、用于训练上述模型评价数据请求所指示的联合训练模型的样本集合。作为又一示例, 上述执行主体还可以从通信连接的电子设备(例如存储服务)获取上述匹配的样本集合。需要说明的是, 为了保证数据安全性, 10 上述通信连接的电子设备与上述联合训练模型的其他参与方之间并不直接共享原始的样本数据。

在本实施例的一些可选的实现方式中, 上述模型评价数据请求中可以包括数据划分信息。其中, 上述数据划分信息可以用于指示样本的划分方法。上述样本标识例如可以包括网络连接中的请求标识 15 (Request ID)。上述执行主体还可以按照如下步骤获取与模型评价数据请求相匹配的样本集合:

第一步, 获取用于训练联合训练模型的样本集合。

在这些实现方式中, 上述执行主体可以从本地或通信连接的电子设备获取用于训练上述联合训练模型的样本集合。通常, 上述用于训练 20 上述联合训练模型的样本集合可以包括经过样本对齐的样本。需要说明的是, 为了保证数据安全性, 上述通信连接的电子设备与上述联合训练模型的其他参与方之间并不直接共享原始的样本数据。

第二步, 根据数据划分信息, 从样本集合中选取样本, 组成与模型评价数据请求相匹配的样本集合。

25 在这些实现方式中, 根据数据划分信息所指示的样本划分方法, 上述执行主体可以从上述第一步所获取的样本集合中选取样本, 组成与上述模型评价数据请求相匹配的样本集合。作为示例, 上述数据划分信息可以包括白名单形式的样本标识列表。上述执行主体可以从上述第一步所获取的样本集合中选取与上述样本标识列表中的标识相匹配 30 的样本。其中, 上述样本标识可以包括各种能够唯一标识样本的字

字符串，例如序号或样本特征。

基于上述可选的实现方式，上述执行主体可以选取与上述数据划分信息相匹配的特定样本集合，为后续联合训练模型的训练和模型评价数据的生成提供数据基础。

- 5 步骤 203，根据匹配的样本集合，生成联合训练模型的模型评价数据。

在本实施例中，根据上述步骤 202 所获取的匹配的样本集合，上述执行主体可以通过各种方式生成上述联合训练模型的模型评价数据。作为示例，上述执行主体可以利用上述用于训练上述模型评价数据请求所指示的联合训练模型的样本集合训练模型，从而生成与上述模型评价数据请求相匹配的模型评价数据。例如上述模型评价数据请求用于指示获取模型准确率的数据，则上述执行主体可以根据上述模型训练完成后的测试结果，生成上述模型的准确率指标。可选地，上述执行主体也可以在上述模型训练完成后，生成上述模型的各项模型评价数据集。而后，上述执行主体可以从上述模型评价数据集中选取与上述模型评价数据请求所指示的数据相匹配的模型评价数据。

10

15

在本实施例的一些可选的实现方式中，上述执行主体还可以向上述目标设备发送所生成的联合训练模型的模型评价数据作为上述模型评价数据请求的响应信息。

20

继续参见图 3，图 3 是根据本申请实施例的用于评估联合训练模型的方法的应用场景的一个示意图。在图 3 的应用场景中，资讯类应用的后台服务器 301 和购物类应用的后台服务器 302 可以分别利用样本集合 A303 和样本集合 B304 共同训练联合训练模型 305。上述样本集合 A303 中的样本可以包括各种用户属性信息（例如性别、年龄、偏好等）。上述样本集合 B304 中的样本可以包括用户标识和用户是否有购物行为（例如将商品加入购物车）。通常，后台服务器 301 和后台服务器 302 不能直接获取对方样本数据。当用户通过终端 306 使用咨询类应用时，后台服务器 301 可以利用上述训练好的联合训练模型 305 获取相应的商品链接并发送至终端 306，以供终端 306 呈现给用户。

25

30

后台服务器 301 可以向后台服务器 302 发送模型评价数据请求 307 以获取上述联合训练模型 305 的效果指标（例如准确率）。后台服务器 302 可以获取与上述模型评价数据请求 307 相匹配的样本集合 B304。根据上述样本集合 B304，后台服务器 302 可以生成利用上述样本集合 B304 训练完成的联合训练模型 305 的效果指标（例如准确率）作为上述模型评价数据。可选地，后台服务器 302 还可以进一步将包括上述模型评价数据的响应信息 308 发送至后台服务器 301。

目前，现有技术之一通常是采用分布式训练的方式来提升大规模模型训练的效率，导致不具备标签样本的一方难以在保证数据安全的前提下获知模型的训练效果。而本申请的上述实施例提供的方法，通过根据不具备标签样本的一方发送的模型评价数据请求生成相应的联合训练模型的模型评价数据，实现了在不暴露原始样本数据的前提下共享联合训练模型的效果指标。从而为联合训练模型的优化和改进提供及时、有效的数据参考基础。

15

进一步参考图 4，其示出了用于评估联合训练模型的方法的又一个实施例的流程 400。该用于评估联合训练模型的方法的流程 400，包括以下步骤：

步骤 401，接收目标设备发送的模型评价数据请求。

在本实施例中，上述模型评价数据请求中可以包括数据划分信息。上述数据划分信息可以用于指示样本的划分方法。

在本实施例的一些可选的实现方式中，上述模型评价数据请求中还可以包括数据查询权限信息。其中，上述数据查询权限信息可以用于指示上述指标变化信息中的至少一个维度。

步骤 402，获取用于训练联合训练模型的样本集合。

步骤 403，根据数据划分信息，从样本集合中选取样本，组成与模型评价数据请求相匹配的样本集合。

上述步骤 401、步骤 402、步骤 403 分别与前述实施例中的步骤 201 和步骤 202 中可选的实现方式一致，上文针对步骤 201 和步骤 202 中可选的实现方式的描述也适用于步骤 401、步骤 402 和步骤 403，此

处不再赘述。

步骤 404, 将匹配的样本集合划分为训练样本集合和测试样本集合。

在本实施例中, 用于评估联合训练模型的方法的执行主体 (例如
5 图 1 所示的服务器 105 或服务器 106) 可以通过各种方式将上述步骤
403 所获取的匹配的样本集合划分为训练样本集合和测试样本集合。
其中, 上述样本集合划分的方式可以包括但不限于以下至少一项: 留出法(hold-out), 交叉验证法(cross validation), 自助法(bootstrapping)。

步骤 405, 利用匹配的样本集合, 生成联合训练模型的指标变化
10 信息。

在本实施例中, 利用匹配的样本集合, 上述执行主体可以通过各种
方式生成联合训练模型的指标变化信息。其中, 上述联合训练模型的
指标变化信息可以用于指示上述联合训练模型的模型评价数据随时
间维度的变化。其中, 上述时间维度可以包括但不限于以下至少一项:
15 时间, 全局步数(global step)。作为示例, 上述执行主体可以根据上
述时间维度, 利用上述匹配的样本集合中的训练样本集合对联合训练
模型的参数进行调整, 而后利用上述匹配的样本集合中的测试样本集
合生成上述调整后的联合训练模型的各项模型评价数据。其中, 上述
模型评价数据可以包括但不限于以下至少一项: 准确率, 召回率, AUC。
20 从而, 上述执行主体可以生成上述各项模型评价数据随时间变化的曲
线作为联合训练模型的指标变化信息。

在本实施例的一些可选的实现方式中, 基于上述模型评价数据请
求中包括的数据查询权限信息, 上述执行主体可以继续执行以下步骤:

第一步, 根据数据查询权限信息, 从指标变化信息中提取与数据
25 查询权限信息匹配的模型评价数据。

在这些实现方式中, 基于上述模型评价数据请求中包括的数据查
询权限信息, 上述执行主体可以从所生成的指标变化信息中提取与上
述数据查询权限信息匹配的模型评价数据。其中, 上述数据查询权限
信息可以用于指示上述指标变化信息中的至少一个维度。上述维度可
30 以包括但不限于以下至少一项: 时间, 模型评价数据的标识 (例如名

称), 采样粒度。

第二步, 向目标设备发送匹配的模型评价数据。

在这些实现方式中, 上述执行主体可以向上述目标设备发送上述第一步所提取的匹配的模型评价数据。

- 5 基于上述可选的实现方式, 上述目标设备可以利用数据查询权限信息实现按照时间、模型评价数据的名称(例如准确率、AUC)等维度获取所需要的模型评价数据。

从图4中可以看出, 本实施例中的用于评估联合训练模型的方法的流程400体现了利用划分后的匹配的样本集合, 生成联合训练模型的指标变化信息的步骤。由此, 本实施例描述的方案可以利用划分的
10 样本集合对联合训练模型进行训练并生成模型训练过程中的模型评价数据随时间维度的变化信息, 从而在保证数据安全的前提下实现了联合训练模型之间共享模型评价指标。

- 15 进一步参考图5, 作为对上述各图所示方法的实现, 本申请提供了用于评估联合训练模型的装置的一个实施例, 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5所示, 本实施例提供的用于评估联合训练模型的装置500包括接收单元501、获取单元502和生成单元503。其中, 接收单元
20 501, 被配置成接收目标设备发送的模型评价数据请求, 目标设备包括联合训练模型的参与方; 获取单元502, 被配置成获取与模型评价数据请求相匹配的样本集合, 匹配的样本集合属于与联合训练模型相关联的有标签数据; 生成单元503, 被配置成根据匹配的样本集合, 生成联合训练模型的模型评价数据。

- 25 在本实施例中, 用于评估联合训练模型的装置500中: 接收单元501、获取单元502和生成单元503的具体处理及其所带来的技术效果可分别参考图2对应实施例中的步骤201、步骤202、步骤203或图4对应的实施例中的步骤401-步骤405的相关说明, 在此不再赘述。

在本实施例的一些可选的实现方式中, 上述模型评价数据请求中
30 可以包括数据划分信息。其中, 上述数据划分信息可以用于指示样本

的划分方法。上述获取单元 502 可以包括：获取模块（图中未示出）、
第一生成模块（图中未示出）。其中，上述获取模块，可以被配置成获
取用于训练联合训练模型的样本集合。上述第一生成模块，可以被配
置成根据数据划分信息，从样本集合中选取样本，组成与模型评价数
据请求相匹配的样本集合。

在本实施例的一些可选的实现方式中，上述生成单元 503 可以包
括划分模块（图中未示出）、第二生成模块（图中未示出）。其中，上
述划分模块可以被配置成将匹配的样本集合划分为训练样本集合和测
试样本集合。上述第二生成模块可以被配置成利用匹配的样本集合，
生成联合训练模型的指标变化信息。其中，上述联合训练模型的指标
变化信息可以用于指示联合训练模型的模型评价数据随时间维度的变
化。

在本实施例的一些可选的实现方式中，上述模型评价数据请求中
还可以包括数据查询权限信息。上述数据查询权限信息可以用于指示
指标变化信息中的至少一个维度。该用于评估联合训练模型的装置
500 还可以包括：提取单元（图中未示出）、发送单元（图中未示出）。
其中，上述提取单元，可以被配置成根据数据查询权限信息，从指标
变化信息中提取与数据查询权限信息匹配的模型评价数据。上述发送
单元，可以被配置成向目标设备发送匹配的模型评价数据。

本申请的上述实施例提供的装置，通过接收单元 501 接收目标设
备发送的模型评价数据请求。其中，目标设备包括联合训练模型的参
与方。而后，获取单元 502 获取与模型评价数据请求相匹配的样本集
合。其中，匹配的样本集合属于与联合训练模型相关联的有标签数据。
最后，生成单元 503 根据匹配的样本集合，生成联合训练模型的模型
评价数据。从而实现了在不暴露原始样本数据的前提下共享联合训练
模型的效果指标。从而为联合训练模型的优化和改进提供及时、有效
的数据参考基础。

进一步参考图 6，其示出了用于评估联合训练模型的方法的一个
实施例中各个设备之间交互的时序 600。该用于评估联合训练模型的

系统可以包括：第一参与端（例如图 1 所示的服务器 105），第二参与端（例如图 1 所示的服务器 106）。其中，上述第一参与端，可以被配置成获取本地样本集合，其中，上述本地样本集合中的本地样本属于用于训练联合训练模型的无标签数据；生成与上述本地样本集合相关的联合训练模型的模型评价数据请求；向上述联合训练模型的第二参与端发送上述模型评价数据请求，其中，上述第二参与端与上述第一参与端共同训练上述联合训练模型。上述第二参与端，可以被配置成接收上述第一参与端发送的模型评价数据请求；获取与上述模型评价数据请求相匹配的样本集合，其中，上述匹配的样本集合属于与上述联合训练模型相关联的有标签数据；根据上述匹配的样本集合，生成上述联合训练模型的模型评价数据。

在本实施例的一些可选的实现方式中，上述模型评价数据请求中可以包括数据划分信息。上述第一参与端可以进一步被配置成：根据上述本地样本集合中样本的特征维度，生成数据划分信息，其中，上述数据划分信息用于指示样本的划分方法。

在本实施例的一些可选的实现方式中，上述第一参与端还可以被配置成：从上述第二参与端获取上述联合训练模型的模型评价数据；基于上述模型评价数据，调整上述联合训练模型。

如图 6 所示，在步骤 601 中，第一参与端获取本地样本集合。

在本实施例中，上述第一参与端可以获取本地样本集合。其中，上述本地样本集合中的本地样本通常属于可用于训练联合训练模型的无标签数据，即上述样本集合中的样本通常仅包括特征。而且，上述本地样本集合中的本地样本中所包括的特征通常多于对应的有标签数据的样本中所包括的特征。作为示例，上述无标签数据的样本中包括的特征可以包括性别、年龄、地区、偏好、历史浏览信息；上述无标签数据的样本中包括的特征可以包括账号。

在步骤 602 中，第一参与端生成与本地样本集合相关联的联合训练模型的模型评价数据请求。

在本实施例中，上述第一参与端可以通过各种方式生成与上述步骤 601 所获取的本地样本集合相关联的联合训练模型的模型评价数据

请求。其中，上述模型评价数据请求可以用于获取上述联合训练模型的模型评价数据。上述第一参与端可以根据预设的模型指标列表生成上述模型评价数据请求。其中，上述模型指标列表可以包括一项至少一项：准确率，召回率，AUC，PSI,对数损失。作为示例，上述第一参与端可以从上述模型指标列表中随机选取至少一项指标。作为又一示例，上述第一参与端也可以根据技术人员使用的终端所发送的用于指示模型指标的信息，生成用于获取该模型指标的模型评价数据请求。

在本实施例的一些可选的实现方式中，根据上述步骤 601 所获取的本地样本集合中样本的特征维度，上述第一参与端可以生成数据划分信息。其中，上述数据划分信息可以用于指示样本的划分方法。上述特征维度例如可以包括用户维度（例如 ID）和用户特征维度（例如在线时长等）。作为示例，上述数据划分信息例如可以包括白名单形式的网络连接的请求标识列表。可选地，上述数据划分信息也可以是预设的散列函数。其中，上述预设的散列函数可以是利用各种加密方法所生成的安全策略函数，上述安全策略函数无法反推出原始数据，从而进一步保证了联合训练模型的数据安全性。

在步骤 603 中，第一参与端向联合训练模型的第二参与端发送模型评价数据请求。

在本实施例中，上述第二参与端与上述第一参与端共同训练上述联合训练模型。

在步骤 604 中，第二参与端接收第一参与端发送的模型评价数据请求。

在步骤 605 中，第二参与端获取与模型评价数据请求相匹配的样本集合。

在步骤 606 中，第二参与端根据匹配的样本集合，生成联合训练模型的模型评价数据。

上述步骤 604、步骤 605 和步骤 606 分别与前述实施例中的步骤 201、步骤 202、步骤 203 或步骤 401-步骤 405 及其可选的实现方式一致，上文针对步骤 201、步骤 202、步骤 203 或步骤 401-步骤 405 及其可选的实现方式的描述也适用于步骤 604、步骤 605 和步骤 606，此处

不再赘述。

在本实施例的一些可选的实现方式中，上述第一参与端还可以继续执行以下步骤：

步骤 607，从第二参与端获取联合训练模型的模型评价数据。

5 步骤 608，基于模型评价数据，调整联合训练模型。

在这些可选的实现方式中，基于模型评价数据，上述第一参与端可以通过各种方式调整上述联合训练模型。作为示例，上述第一参与端可以首先确定上述模型评价数据是否符合预设的模型调整条件。上述预设的模型调整条件例如可以包括模型准确率小于预设阈值。而后，
10 响应于确定上述模型评价数据符合预设的模型调整条件，上述执行主体可以通过各种方式调整上述联合训练模型。例如，上述执行主体可以利用未用于训练的样本继续训练上述联合训练模型。再例如，上述执行主体可以调整样本中对模型输出不敏感的特征（例如以年龄替换地区）。

15 本申请的上述实施例提供的用于评估联合训练模型的系统，首先，上述第一参与端获取本地样本集合。其中，上述本地样本集合中的本地样本属于用于训练联合训练模型的无标签数据。而后，上述第一参与端生成与上述本地样本集合相关联的联合训练模型的模型评价数据请求。之后，上述第一参与端向上述联合训练模型的第二参与端发送
20 上述模型评价数据请求。其中，上述第二参与端与上述第一参与端共同训练上述联合训练模型。接下来，上述第二参与端接收上述第一参与端发送的模型评价数据请求。而后，上述第二参与端获取与上述模型评价数据请求相匹配的样本集合。其中，上述匹配的样本集合属于与上述联合训练模型相关联的有标签数据。最后，上述第二参与端根据
25 上述匹配的样本集合，生成上述联合训练模型的模型评价数据。从而通过根据不具备标签样本的联合训练模型参与方生成并发送模型评价数据请求，具备标签样本的另一参与方生成相应的联合训练模型的模型评价数据，实现了在不暴露原始样本数据的前提下共享联合训练模型的效果指标，进而为联合训练模型的优化和改进提供及时、有效的
30 的数据参考基础。

下面参考图 7，其示出了适于用来实现本申请的实施例的电子设
备（例如图 1 中的服务器 105）700 的结构示意图。图 7 示出的服务器
仅仅是一个示例，不应对本申请的实施例的功能和使用范围带来任何
5 限制。

如图 7 所示，电子设备 700 可以包括处理装置（例如中央处理器、
图形处理器等）701，其可以根据存储在只读存储器（ROM）702 中的
程序或者从存储装置 708 加载到随机访问存储器（RAM）703 中的程
序而执行各种适当的动作和处理。在 RAM 703 中，还存储有电子设备
10 700 操作所需的各种程序和数据。处理装置 701、ROM 702 以及 RAM
703 通过总线 704 彼此相连。输入/输出（I/O）接口 705 也连接至总线
704。

通常，以下装置可以连接至 I/O 接口 705：包括例如触摸屏、触摸
板、键盘、鼠标、等的输入装置 706；包括例如液晶显示器（LCD，
15 Liquid Crystal Display）、扬声器、振动器等输出装置 707；包括例如
磁带、硬盘等的存储装置 708；以及通信装置 709。通信装置 709 可以
允许电子设备 700 与其他设备进行无线或有线通信以交换数据。虽然
图 7 示出了具有各种装置的电子设备 700，但是应理解的是，并不要
求实施或具备所有示出的装置。可以替代地实施或具备更多或更少的
20 装置。图 7 中示出的每个方框可以代表一个装置，也可以根据需要代
表多个装置。

特别地，根据本申请的实施例，上文参考流程图描述的过程可以
被实现为计算机软件程序。例如，本申请的实施例包括一种计算机程
序产品，其包括承载在计算机可读介质上的计算机程序，该计算机程
序包含用于执行流程图所示的方法的程序代码。在这样的实施例中，
25 该计算机程序可以通过通信装置 709 从网络上被下载和安装，或者从
存储装置 708 被安装，或者从 ROM 702 被安装。在该计算机程序被处
理装置 701 执行时，执行本申请的实施例的方法中限定的上述功能。

需要说明的是，本申请的实施例所述的计算机可读介质可以是计
30 算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组

合。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、光纤、便携式紧凑磁盘只读存储器（CD-ROM）、光存储器件、磁存储器件、或者上述的任意合适的组合。在本申请的实施例中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本申请的实施例中，计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是计算机可读存储介质以外的任何计算机可读介质，该计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：电线、光缆、RF（Radio Frequency，射频）等等，或者上述的任意合适的组合。

上述计算机可读介质可以是上述电子设备中所包含的；也可以是单独存在，而未装配入该服务器中。上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序被该服务器执行时，使得该服务器：接收目标设备发送的模型评价数据请求，其中，目标设备包括联合训练模型的参与方；获取与模型评价数据请求相匹配的样本集合，其中，匹配的样本集合属于与联合训练模型相关联的有标签数据；根据匹配的样本集合，生成联合训练模型的模型评价数据。

可以以一种或多种程序设计语言或其组合来编写用于执行本申请实施例的操作的计算机程序代码，所述程序设计语言包括面向对象的程序设计语言——诸如 Java、Smalltalk、C++，还包括常规的过程式程序设计语言——诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立

的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)——连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。

附图中的流程图和框图，图示了按照本申请的各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

描述于本申请实施例中所涉及到的单元可以通过软件的方式实现，也可以通过硬件的方式来实现。所描述的单元也可以设置在处理器中，例如，可以描述为：一种处理器，包括接收单元、获取单元、生成单元。其中，这些单元的名称在某种情况下并不构成对该单元本身的限定，例如，接收单元还可以被描述为“接收目标设备发送的模型评价数据请求的单元，其中，目标设备包括联合训练模型的参与方”。

以上描述仅为本申请的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解，本申请的实施例中所涉及的发明范围，并不限于上述技术特征的特定组合而成的技术方案，同时也应涵盖在不脱离上述发明构思的情况下，由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本申请实施例中公开的(但不限于)具有类似功能的技术特征进行互相替换而形成的技术方案。