

说明书

模型的训练方法、设备及存储介质

5 技术领域

本说明书涉及互联网技术领域，尤其涉及一种模型的训练方法、设备及存储介质。

背景技术

随着互联网技术的快速发展，越来越多的人使用各种服务平台享受便利服务，
10 与此同时，会有一些不法分子在服务平台中投机，给服务平台或者正常用户带来风险。
在此形势下，服务平台需要训练用来识别风险用户的风控模型，以保障服务平台或者
正常用户的安全性。因此服务平台如何得到风控效果较好的风控模型是目前亟需解决
的问题。

背景技术部分的内容仅仅是发明人个人所知晓的信息，并不代表上述信息在本公
15 开申请日之前已经进入公共领域，也不代表其可以成为本公开的现有技术。

发明内容

本说明书提供一种模型的训练方法、设备及存储介质，通过多个参与方联合训练
得到风控模型，在训练过程中充分利用了多个参与方本地的数据，打破了数据孤岛，
20 提升了风控模型的风控效果。

第一方面，本说明书提供一种模型的训练方法，用于通过联邦系统训练得到具有
识别风险用户的能力的风控模型，所述联邦系统包括控制设备和位于不同区域的 M 个
参与设备，所述 M 为大于 1 的整数，每个参与设备存储有本地数据集，所述本地数据
集至少包括所述参与设备所在区域内的样本用户的用户特征，所述方法应用于所述控
25 制设备。所述方法包括：初始化所述风控模型的参数；进而，控制所述 M 个参与设备
对所述风控模型进行多次迭代的联合训练，直至达到预设的训练结束条件时，得到所
述风控模型的目标参数，其中，每次迭代过程包括：向所述 M 个参与设备分发所述风
控模型的当前模型参数，以便所述 M 个参与设备基于所述当前模型参数和所述 M 个参
与设备各自对应的所述本地数据集确定中间训练信息，所述中间训练信息表征所述风

控模型对所述本地数据集中的样本用户的风险预测情况，以及从所述 M 个参与设备分别获得所述中间训练信息，并基于获得的 M 个中间训练信息，对所述风控模型的当前模型参数进行更新。

第二方面，本说明书提供一种模型的训练方法，用于通过联邦系统训练得到具有识别风险用户的能力的风控模型，所述联邦系统包括控制设备和位于不同区域的 M 个参与设备，所述 M 为大于 1 的整数，每个参与设备存储有本地数据集，所述本地数据集至少包括所述参与设备所在区域内的样本用户的用户特征，所述方法应用于所述参与设备。所述方法包括：从所述控制设备接收所述风控模型的当前模型参数；基于所述当前模型参数和所述参与设备对应的所述本地数据集确定中间训练信息，所述中间训练信息表征所述风控模型对所述本地数据集中的样本用户的风险预测情况；以及，向所述控制设备发送所述中间训练信息，以便所述控制设备基于所述 M 个参与设备各自对应的所述中间训练信息对所述风控模型的当前模型参数进行更新。

第三方面，本说明书还提供一种电子设备，位于包括控制设备和 M 个参与设备的联邦系统中，所述联邦系统用于训练得到具有识别风险用户的能力的风控模型，所述 M 个参与设备位于不同区域，每个参与设备存储有本地数据集，所述本地数据集至少包括所述参与设备所在区域内的样本用户的用户特征，所述电子设备作为所述控制设备，包括：至少一个存储介质，存储有至少一个指令集；以及至少一个处理器，同所述至少一个存储介质通信连接，其中，所述至少一个处理器运行时读取所述至少一个指令集，并且根据所述至少一个指令集的指示：初始化所述风控模型的参数，控制所述 M 个参与设备对所述风控模型进行多次迭代的联合训练，直至达到预设的训练结束条件时，得到所述风控模型的目标参数，其中，每次迭代过程包括：向所述 M 个参与设备分发所述风控模型的当前模型参数，以便所述 M 个参与设备基于所述当前模型参数和所述 M 个参与设备各自对应的所述本地数据集确定中间训练信息，所述中间训练信息表征所述风控模型对所述本地数据集中的样本用户的风险预测情况，以及从所述 M 个参与设备分别获得所述中间训练信息，并基于获得的 M 个中间训练信息，对所述风控模型的当前模型参数进行更新。

第四方面，本说明书还提供一种电子设备，位于包括控制设备和 M 个参与设备的联邦系统中，所述联邦系统用于训练得到具有识别风险用户的能力的风控模型，所述 M 个参与设备位于不同区域，每个参与设备存储有本地数据集，所述本地数据集至少包括所述参与设备所在区域内的样本用户的用户特征，所述电子设备作为所述 M 个参

与设备中的其中一个参与设备，包括：至少一个存储介质，存储有至少一个指令集；以及至少一个处理器，同所述至少一个存储介质通信连接，其中，所述至少一个处理器运行时读取所述至少一个指令集，并且根据所述至少一个指令集的指示：从所述控制设备接收所述风控模型的当前模型参数，基于所述当前模型参数和所述参与设备对应的所述本地数据集确定中间训练信息，所述中间训练信息表征所述风控模型对所述本地数据集中的样本用户的风险预测情况，以及向所述控制设备发送所述中间训练信息，以便所述控制设备基于所述 M 个参与设备各自对应的所述中间训练信息对所述风控模型的当前模型参数进行更新。

第五方面，本说明书提供一种风险用户的识别方法，包括：获得目标用户的用户特征，将所述目标用户的用户特征输入到已训练的风控模型中得到所述风控模型输出的风险概率，基于所述风险概率确定所述目标用户是否为风险用户，其中，所述风控模型是通过联邦系统训练得到的，所述联邦系统包括控制设备和 M 个参与设备，所述 M 为大于 1 的整数，在所述训练过程中所述控制设备执行如第一方面所述的方法、所述参与设备执行如第二方面所述的方法。

第六方面，本说明书提供一种电子设备，包括：至少一个存储介质和至少一个处理器，其中，所述至少一个存储介质存储有至少一个指令集，所述至少一个处理器同所述至少一个指令集通信连接，在所述至少一个处理器运行时读取所述至少一个指令集，并且根据所述至少一个指令集的指示：获得目标用户的用户特征，将所述目标用户的用户特征输入到已训练的风控模型中得到所述风控模型输出的风险概率，基于所述风险概率确定所述目标用户是否为风险用户，其中，所述风控模型是通过联邦系统训练得到的，所述联邦系统包括控制设备和 M 个参与设备，所述 M 为大于 1 的整数，在所述训练过程中所述控制设备执行如第一方面所述的方法、所述参与设备执行如第二方面所述的方法。

第七方面，本说明书提供一种计算机可读非暂时性存储介质，其中，所述计算机可读非暂时性存储介质中存储有至少一个指令集，所述至少一个指令集被至少一个处理器执行时，实现如第一方面或第二方面所述的训练方法，或者实现如第五方面所述的识别方法。

本说明书提供的模型的训练方法、设备及存储介质的其他功能将在以下说明中部分列出。根据描述，以下数字和示例介绍的内容将对那些本领域的普通技术人员显而易见。本说明书提供的模型的训练方法、设备及存储介质的创造性方面可以通过实践

或使用下面详细示例中所述的方法、装置和组合得到充分解释。

附图说明

为了更清楚地说明本说明书实施例中的技术方案，下面将对实施例描述中所需要
5 使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本说明书的一些实施例，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 示出了根据本说明书的实施例提供的一种模型训练的应用场景示意图；

图 2 示出了根据本说明书的实施例提供的一种电子设备的硬件结构图；

10 图 3 示出了根据本说明书的实施例提供的一种模型的训练方法的流程图；以及

图 4 示出了根据本说明书的实施例提供的模型的训练方法的详细流程图。

具体实施方式

以下描述提供了本说明书的特定应用场景和要求，目的是使本领域技术人员能够
15 制造和使用本说明书中的内容。对于本领域技术人员来说，对所公开的实施例的各种局部修改是显而易见的，并且在不脱离本说明书的精神和范围的情况下，可以将这里定义的一般原理应用于其他实施例和应用。因此，本说明书不限于所示的实施例，而是与权利要求一致的最宽范围。

这里使用的术语仅用于描述特定示例实施例的目的，而不是限制性的。比如，除
20 非上下文另有明确说明，这里所使用的，单数形式“一”，“一个”和“该”也可以包括复数形式。当在本说明书中使用时，术语“包括”、“包含”和/或“含有”意思是指所关联的整数，步骤、操作、元素和/或组件存在，但不排除一个或多个其他特征、整数、步骤、操作、元素、组件和/或组的存在或在该系统/方法中可以添加其他特征、整数、步骤、操作、元素、组件和/或组。

25 考虑到以下描述，本说明书的这些特征和其他特征、以及结构的相关元件的操作和功能、以及部件的组合和制造的经济性可以得到明显提高。参考附图，所有这些形成本说明书的一部分。然而，应该清楚地理解，附图仅用于说明和描述的目的，并不旨在限制本说明书的范围。还应理解，附图未按比例绘制。

本说明书中使用的流程图示出了根据本说明书中的一些实施例的系统实现的操作。

应该清楚地理解，流程图的操作可以不按顺序实现。相反，操作可以以反转顺序或同时实现。此外，可以向流程图添加一个或多个其他操作。可以从流程图中移除一个或多个操作。

5 服务平台可以基于自身平台的数据训练得到风控模型。例如，服务平台需要训练用于在营销场景中识别风险用户的风控模型。此时，服务平台可以从自身平台中的数据清洗出多个样本用户的用户特征形成训练样本，并基于训练样本训练得到所述风控模型。

但是，对于新开展的服务平台而言，自身平台中的数据较少，风控经验有限，基于平台自身的数据训练得到风控模型的难度较大。

10 例如，在发展初期，服务平台会推出很多营销活动来吸引用户，与此同时，也会吸引不法分子通过作弊行为套取服务平台的营销资金，或者吸引很多薅羊毛用户仅在活动的时候使用服务平台。因此，服务平台需要有效的风控模型来识别营销场景中的风险用户，以减少平台的营销资产损失，避免由于平台的营销资金被不法分子非法套取导致达不到预想的营销效果的问题。

15 在此情况下，所述服务平台可以借鉴一些开展较早的、拥有大量历史数据的其他服务平台的风控模型，这样可以规避一些风险。但是由于不同服务平台面向的用户群体以及开展的服务类型有差异，因此不同服务平台防控的风险也不同。借鉴其他服务平台的风控模型很难准确、全面的防控所述服务平台中存在的风险。

除此之外，所述服务平台可以委托风控经验丰富的其他服务平台为其训练所述风控模型，这种情况下，训练过程中还是需要结合所述服务平台中的数据。但是由于数据隐私问题，不同服务平台间的数据不能共享。尤其比如所述其他服务平台和所述服务平台分属于不同国家或地区的情况下，由于用户数据出境的限制，所述其他服务平台无法获得所述服务平台真实的数据，进而无法通过 SaaS（Software as a service，软件即服务）的方式为所述服务平台训练所述风控模型。

25 鉴于此，本说明书提供一种模型的训练方法、设备及存储介质，通过包括控制设备和多个参与设备的联邦系统进行多次迭代的联合训练，从而得到可以识别风险用户的风控模型。在每次迭代训练过程中，通过控制设备的协调，可以在参与设备的本地数据集不出本地的前提下完成风控模型的训练，有利于数据隐私保护。并且在训练过程中使用来自多个参与设备的本地数据集，打破数据孤岛，使得训练数据更加多样化，
30 从而可以提高训练得到的风控模型的泛化能力和适应性，并提升风控模型的风控效果。

下面结合图 1 对本说明书提供的训练风控模型的应用场景进行介绍。

图 1 示出了根据本说明书的实施例提供的一种模型训练的应用场景示意图。如图 1 所示，场景 001 示例了训练风控模型的联邦系统，联邦系统包括：控制设备 100 和多个参与设备 200。

5 控制设备 100 作为联合训练的控制方，可以协调控制每个参与设备 200 参与多次迭代训练，直到风控模型的性能满足预设条件。控制设备 100 和参与设备 200 之间可以建立通信连接，用以互相传输数据。例如，在每一次迭代训练中，控制设备 100 可以将风控模型的当前模型参数分发给每个参与设备 200，以使得参与设备 200 基于当前模型参数训练风控模型得到中间训练信息，控制设备 100 还可以从多个参与设备 200
10 处分别获得中间训练信息，控制设备 100 可以基于中间训练信息更新风控模型的当前模型参数。直至多次迭代训练后，控制设备 100 得到满足预设条件的风控模型。

 控制设备 100 是具有一定数据处理能力的设备。控制设备 100 可以执行本说明书描述的模型的训练方法的数据和指令，并可以执行或用于执行所述数据和指令。在一些实施例中，控制设备 100 可以包括具有数据信息处理功能的硬件设备和驱动该硬件
15 设备工作所需必要的程序。

 不同参与设备 200 可以位于不同区域（例如国家或地区），每个参与设备 200 存储有本地数据集，不同参与设备 200 可以是同一类型的服务平台在不同区域内的服务设备。以一个参与设备 200 为例，参与设备 200 的本地数据集包括参与设备 200 所在区域内的用户的用户特征。比如，参与设备 200 是 A 国的服务平台的服务设备，可以获
20 取的平台数据包括但不限于：A 国的用户在服务平台中注册的个人基础信息（例如用户名、所在城市、联系方式等）和 A 国的用户在服务平台中执行的事件的事件信息（例如营销场景中的权益发放事件、支付事件等）等。参与设备 200 可以基于所述平台数据清洗出用于训练风控模型的用户特征，用户特征例如消费频次、支付成功率以及参加营销活动的次数等。进而，参与设备 200 可以基于用户特征生成本地数据集，
25 后续在每一次迭代训练中，参与设备 200 可以基于本地数据集和控制设备 100 发送的当前模型参数训练风控模型得到中间训练信息，并将中间训练信息以及控制设备 100 所需要的其他信息（例如样本用户的数量等）发送给控制设备 100。

 风控模型训练完成后，控制设备 100 可以将风控模型的目标参数发送给参与设备 200，参与设备 200 可以将风控模型安装于风控执行设备（图 1 中未示出）。例如，参
30 与设备 200 可以将风控模型安装于所述参与设备 200 所在区域的营销服务器中，用来

识别营销场景中的风险用户。识别营销场景中的风险用户可以理解为：识别参加营销活动的用户是否是风险用户。营销活动例如服务平台向所述服务平台的用户发放权益。在本说明书中，权益可以是指用户在服务平台上的权利和利益。例如用户在服务平台购买商品或服务时享有的价格减免的权利，用户在满足一定条件(例如服务平台的新用户等)情况下领取权益的权利等。权益可以包括但不限于：现金、折扣券、满减券、单品券、优惠券、礼品、积分等。具体的，当风控模型识别出当前参加营销活动的用户为风险用户，则可以拒绝向该用户发送权益；当风控模型识别出当前参加营销活动的用户不是风险用户，则可以正常向该用户发送权益；当风控模型不能确定当前参加营销活动的用户是否是风险用户时，服务平台还可以基于其他的风控策略进一步判断该用户是否是风险用户。

参与设备 200 是具有一定数据处理能力的设备。参与设备 200 可以执行本说明书描述的模型的训练方法的数据和指令，并可以执行或用于执行所述数据和指令。在一些实施例中，参与设备 200 可以包括具有数据信息处理功能的硬件设备和驱动该硬件设备工作所需必要的程序。

应该理解，图 1 中的参与设备 200 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参与设备 200。

需要说明的是，所述模型的训练方法部分在控制设备 100 上执行，部分在参与设备 200 上执行。

图 2 示出了根据本说明书的实施例提供的一种电子设备 300 的硬件结构图。电子设备 300 可以作为图 1 中的控制设备 100，也可以作为图 1 中的任一参与设备 200。

如图 2 所示，电子设备 300 可以包括至少一个存储介质 330 和至少一个处理器 320。在一些实施例中，电子设备 300 还可以包括通信端口 350 和内部通信总线 310。此外，电子设备 300 还可以包括 I/O 组件 360。

内部通信总线 310 可以连接不同的系统组件。例如，内部通信总线 310 可以连接存储介质 330、处理器 320、通信端口 350 和 I/O 组件 360。

I/O 组件 360 支持电子设备 300 和其他组件之间的输入/输出。

通信端口 350 用于电子设备 300 同外界的数据通信。比如，通信端口 350 可以用于电子设备 300 同网络 400 之间的数据通信。通信端口 350 可以是有线通信端口也可以是无线通信端口。

网络 400 用以在控制设备 100 和参与设备 200 之间提供通信连接的介质。网络 400 可以促进信息或数据的交换。在一些实施例中，网络 400 可以是任何类型的有线或无线网络，也可以是其组合。比如，网络 400 可以包括电缆网络，有线网络、光纤网络、电信通信网络、内联网、互联网、局域网（LAN）、广域网（WAN）、无线局域网（WLAN）、大都市市区网（MAN）、公用电话交换网（PSTN）、蓝牙网络、ZigBee 网络、近场通信（NFC）网络或类似网络。在一些实施例中，网络 400 可以包括一个或多个网络接入点。例如，网络 400 可以包括有线或无线网络接入点，如基站或互联网交换点，通过该接入点，控制设备 100、参与设备 200 的一个或多个组件可以连接到网络 400 以交换数据或信息。

10 存储介质 330 可以包括数据存储装置。所述数据存储装置可以是非暂时性存储介质，也可以是暂时性存储介质。比如，所述数据存储装置可以包括磁盘 332、只读存储介质（ROM）334 或随机存取存储介质（RAM）336 中的一种或多种。存储介质 330 还包括存储在所述数据存储装置中的至少一个指令集。所述指令集可以包括计算机程序代码。所述计算机程序代码可以包括执行本说明书提供的模型的训练方法的程序、
15 例程、对象、组件、数据结构、过程、模块等等。

处理器 320 可以同存储介质 330 通信连接。处理器 320 用以执行上述至少一个指令集。当电子设备 300 作为控制设备 100 或参与设备 200 运行时，处理器 320 读取所述至少一个指令集，并且根据所述至少一个指令集的指示，执行本说明书提供的模型的训练方法。

20 处理器 320 可以是一个或多个处理器的形式，在一些实施例中，处理器 320 可以包括一个或多个硬件处理器，例如微控制器，微处理器，精简指令集计算机（RISC），专用集成电路（ASIC），特定于应用的指令集处理器（ASIP），中心处理单元（CPU），图形处理单元（GPU），物理处理单元（PPU），微控制器单元，数字信号处理器（DSP），现场可编程门阵列（FPGA），高级 RISC 机器（ARM），可编程逻辑器件
25 （PLD），能够执行一个或多个功能的任何电路或处理器等，或其任何组合。

仅仅为了说明问题，在附图中电子设备 300 中仅示出一个处理器 320。然而，应当注意，本说明书中电子设备 300 还可以包括多个处理器，因此，本说明书中披露的操作和/或方法步骤可以由一个处理器执行，也可以由多个处理器联合执行。例如，如果在本说明书中电子设备 300 的处理器 320 执行步骤 A 和步骤 B，则应该理解，步骤 A

和步骤 B 也可以由两个不同处理器 320 联合或分开执行（例如，第一处理器执行步骤 A，第二处理器执行步骤 B，或者第一和第二处理器共同执行步骤 A 和 B）。

图 3 示出了根据本说明书的实施例提供一种模型的训练方法 P100 的流程图。如前，控制设备 100 和参与设备 200 可以协同执行本说明书的方法 P100。具体地，控制设备 100 中的处理器可以读取存储在其本地存储介质中的指令集，然后根据指令集的规定，执行本说明书的方法 P100 中由控制设备 100 执行的步骤或操作；参与设备 200 中的处理器可以读取存储在其本地存储介质中的指令集，然后根据指令集的规定，执行本说明书的方法 P100 中由参与设备 200 执行的步骤或操作。

方法 P100 可应用于支付平台、交易平台、网络安全平台、金融平台等各种需要进行风控的服务平台。为清楚、完整的描述方法 P100，后文以支付平台为例描述，方法 P100 在其他服务平台的应用可以此类推。

为了描述方便，后文中以联邦系统包括控制设备和 M 个参与设备为例进行描述，所述 M 为大于 1 的整数。其中，每个参与设备存储有本地数据集，参与设备的本地数据集至少包括该参与设备所在区域内的样本用户的用户特征。如图 3 所示，模型的训练方法 P100 可以包括如下步骤。

S110：控制设备初始化风控模型的参数。

其中，控制设备可以采用例如随机初始化、Xavier（Glorot）初始化、He 初始化或者正交初始化等方式得到所述风控模型的初始化参数。

初始化完成后，控制设备可以控制所述 M 个参与设备对所述风控模型进行多次迭代的联合训练，直至所述风控模型达到预设的训练结束条件时，控制设备得到所述风控模型的目标参数。

示例性的，所述 M 个参与设备中，一部分参与设备是具有较多风控经验的服务平台对应的服务设备，还有另一部分参与设备是具有较少风控经验的服务平台（比如新开展服务的服务平台）对应的服务设备。

在每次迭代过程中：控制设备向所述 M 个参与设备分发所述风控模型的当前模型参数，所述 M 个参与设备基于所述当前模型参数和所述 M 个参与设备各自对应的所述本地数据集确定中间训练信息。所述中间训练信息表征所述风控模型对本地数据集中的样本用户的风险预测情况。控制设备可以从所述 M 个参与设备分别获得每个参与设备确定的中间训练信息，基于获得的 M 个中间训练信息对所述风控模型的当前模型参数进行更新。

如图 3 所示，在第一次迭代过程中：

S120：控制设备向所述 M 个参与设备分发所述风控模型的当前模型参数#1，当前模型参数#1 即 S110 中得到的风控模型的初始化参数。

例如，图 4 所示，假设 M 个参与设备包括第一参与设备和第二参与设备，S120 包
5 括 S1201 和 S1202，可以理解，S1201 和 S1202 可以同时执行，也可以先后依次执行，
本说明书对此不做限制。

S1201：控制设备向第一参与设备发送当前模型参数#1，对应的，第一参与设备接收当前模型参数#1。

S1202：控制设备向第二参与设备发送当前模型参数#1，对应的，第二参与设备接
10 收当前模型参数#1。

S130： M 个参与设备中的每个参与设备基于当前模型参数#1 和本地数据集确定中间训练信息#1。

可以理解，在执行 S130 之前，每个参与设备需要生成各自的本地数据集。所述本地数据集可以被存储在参与设备所在区域内的存储设备中，或者，所述本地数据集也
15 可以被存储在所述参与设备中。

在一些实施例中，控制设备分别向 M 个参与设备发送第一指示信息，第一指示信息用于指示 N 个目标特征维度， M 个参与设备可以基于所述 N 个目标特征维度生成本地数据集， N 为大于或等于 1 的整数。

示例性的，对于营销场景而言， N 个目标特征维度包括但不限于：用户的基础信息、用户的交易频率、交易次数、交易时间间隔，支付成功率、充值次数、用户历史
20 参加营销活动的次数、用户历史参加营销活动的频率和用户历史参加的营销活动类型等。

可以理解，控制设备可以基于多个场景中用于风控的特征维度中筛选得到适用于特定场景（比如营销场景）的所述 N 个目标特征维度。除此之外，控制设备还可以基
25 于其他方式得到所述 N 个目标特征维度。

在一些实施例中，控制设备从所述 M 个参与设备分别接收第二指示信息，所述第二指示信息用于指示多个候选特征维度，并基于接收到的 M 个第二指示信息，确定所述 N 个目标特征维度。

示例性的，所述 M 个参与设备包括第一参与设备和第二参与设备。具体的，第一
30 参与设备可以得到第一服务平台中的第一平台数据，从所述第一平台数据中确定出 $N1$

个候选特征维度；第二参与设备可以得到第二服务平台中的第二平台数据，从所述第二平台数据中确定出 $N2$ 个候选特征维度。第一参与设备通过第二指示信息#1 向控制设备指示所述 $N1$ 个候选特征维度，第二参与设备通过第二指示信息#2 向控制设备指示所述 $N2$ 个候选特征维度。控制设备确定所述 $N1$ 个候选特征维度和所述 $N2$ 个候选特征维度中相同的候选特征维度，将相同的候选特征维度作为所述 N 个目标特征维度。

在一些实施例中，为了数据隐私，第一参与设备向控制设备发送的第二指示信息#1 为候选特征维度密文#1，即第一参与设备确定 $N1$ 个候选特征维度后，将 $N1$ 个候选特征维度通过公钥进行加密得到候选特征维度密文#1。控制设备接收到候选特征维度密文#1 后可以通过所述公钥对应的私钥将其解密得到所述 $N1$ 个候选特征维度。

同理，第二参与设备向控制设备发送的第二指示信息#2 为候选特征维度密文#2，即第二参与设备确定 $N2$ 个候选特征维度后，将 $N2$ 个候选特征维度通过所述公钥进行加密得到候选特征维度密文#2。控制设备接收到候选特征维度密文#2 后可以通过所述私钥将其解密得到所述 $N2$ 个候选特征维度。

可以理解，在联合训练之前，控制设备可以生成所述公钥和所述私钥，将所述私钥存在控制设备本地，将所述公钥发送给每个参与设备，每个参与设备将公钥存在本地。各参与设备与控制设备之间通过密文形式传递候选特征维度，能够提高信息传递过程中的安全性。

控制设备确定所述 N 个目标特征维度后，通过所述第一指示信息将其指示给第一参与设备和第二参与设备。第一参与设备基于所述 N 个目标特征维度生成第一本地数据集，并将其存储在所述第一参与设备本地，以备训练使用。所述第一本地数据集中包括所述第一参与设备所在区域内的第一样本用户的用户特征、以及所述第一样本用户的风险标签。第二参与设备基于所述 N 个目标特征维度生成第二本地数据集，并将其存储在第二参与设备本地，以备训练使用。所述第二本地数据集中包括所述第二参与设备所在区域内的第二样本用户的用户特征。

具体的，第一参与设备将使用第一服务平台的用户作为第一样本用户，基于所述 N 个目标特征维度从所述第一平台数据中清洗出第一样本用户的用户特征和风险标签，所述用户特征包括上文示例的所述 N 个目标特征维度对应的用户特征，所述风险标签包括正常用户或者风险用户。并且，将一个第一样本用户对应的所有用户特征和风险标签作为第一本地数据集中的训练样本，第一参与设备基于这种方式生成包括多个训练样本的第一本地数据集。

第二参与设备将使用第二服务平台的用户作为第二样本用户，基于所述 N 个目标特征维度从所述第二平台数据中清洗出第二样本用户的用户特征，所述用户特征包括上文示例的所述 N 个目标特征维度对应的用户特征。并且，将一个第二样本用户对应的所有用户特征作为第二本地数据集中的训练样本，第二参与设备基于这种方式生成包括多个训练样本的第二本地数据集。

可以理解，第一平台数据和第二平台数据可以为同一时间段的数据，比如都是同一年的平台数据。这样可以避免由于平台数据所属时间段不同，导致双方的候选特征维度中相同的候选特征维度较少，从而导致训练得到的风控模型效果不好。

在每个参与设备生成本地数据集之后，可以基于当前模型参数#1 和本地数据集确定中间训练信息#1。其中，当前模型参数#1 可以理解为第一次迭代过程中所述风控模型的当前模型参数。中间训练信息#1 可以理解为第一次迭代过程中，参与设备得到的中间训练信息。每个参与设备在第一次迭代过程中得到的中间训练信息#1 不同。

例如，图 4 所示，S130 包括 S1301 和 S1302。

S1301：第一参与设备基于当前模型参数#1 和本地数据集#1a 确定中间训练信息#1a。

第一参与设备生成第一本地数据集后，可以将第一本地数据集分为 n 份， n 为大于 1 的整数。例如，第一本地数据集包括本地数据集#1a、本地数据集#2a、...、以及本地数据集#na。其中，每一份本地数据集参与一次迭代训练，每一份本地数据集包括的样本用户的数量可以相同，也可以不同。

在一些实施例中，每个参与设备在每轮迭代中确定的中间训练信息可以包括梯度信息。比如，中间训练信息#1a 包括梯度信息#1a，第一参与设备基于所述当前模型参数#1 和本地数据集#1a 确定梯度信息#1a，所述梯度信息#1a 表征损失函数 1 关于所述当前模型参数#1 的梯度，所述损失函数 1 用于确定第一参与设备基于所述当前模型参数#1 对所述本地数据集#1a 中的样本用户进行风险预测的损失情况。梯度信息#1a 可以通过以下公式计算得到。

$$\nabla L_1(\theta, D_1) = \frac{1}{|D_1|} \sum_{x_1 \in D_1} \nabla l_1(\theta, x_1)$$

其中， L_1 为损失函数 1， $\nabla L_1(\theta, D_1)$ 为梯度信息#1a， D_1 为本地数据集#1a 中样本用户的数量， $\nabla l_1(\theta, x_1)$ 为本地数据集#1a 中的单个样本 x_1 的损失 l_1 关于参数 θ （即当前模型参数#1）的梯度。

在一些实施例中，每个参与设备在每轮迭代中确定的中间训练信息可以包括模型参数的更新方式。比如，中间训练信息#1a 包括更新后的参数与当前模型参数 θ 之间的差异。假设所述更新后的参数记为 θ_1 ，则中间训练信息#1a 可以包括 $\Delta\theta_1$ ：

$$\Delta\theta_1 = \theta_1 - \theta = -\eta_1 \cdot \nabla L_1$$

5 其中， η_1 表示第一参与设备训练所述风控模型的学习率。

在一些实施例中，每个参与设备在每轮迭代中确定的中间训练信息可以包括损失函数。比如，中间训练信息#1a 可以包括本地数据集#1a 中所有样本的损失之和。

S1302：第二参与设备基于当前模型参数#1 和本地数据集#1b 确定中间训练信息#1b。

10 第二参与设备生成第二本地数据集后，可以将第二本地数据集分为 n 份，例如，第二本地数据集包括本地数据集#1b、本地数据集#2b、...、以及本地数据集#nb。其中，每一份本地数据集参与一次迭代训练，每一份本地数据集包括的样本用户的数量可以相同，也可以不同。

15 在一些实施例中，每个参与设备在每轮迭代中确定的中间训练信息可以包括梯度信息。比如，中间训练信息#1b 包括梯度信息#1b，第二参与设备基于所述当前模型参数#1 和本地数据集#1b 确定梯度信息#1b，所述梯度信息#1b 表征损失函数 2 关于所述当前模型参数#1 的梯度，所述损失函数 2 用于确定第二参与设备基于所述当前模型参数#1 对所述本地数据集#1b 中的样本用户进行风险预测的损失情况。梯度信息#1b 可以通过以下公式计算得到。

$$20 \quad \nabla L_2(\theta, D_2) = \frac{1}{|D_2|} \sum_{x_2 \in D_2} \nabla l_2(\theta, x_2)$$

其中， L_2 为损失函数 2， $\nabla L_2(\theta, D_2)$ 为梯度信息#1b， D_2 为本地数据集#1b 中样本用户的数量， $\nabla l_2(\theta, x_2)$ 为本地数据集#1b 中的单个样本 x_2 的损失 l_2 关于参数 θ （即当前模型参数#1）的梯度。

25 在一些实施例中，每个参与设备在每轮迭代中确定的中间训练信息可以包括模型参数的更新方式。比如，中间训练信息#1b 包括更新后的参数与当前模型参数 θ 之间的差异，假设所述更新后的参数记为 θ_2 ，则中间训练信息#1b 可以包括 $\Delta\theta_2$ ：

$$\Delta\theta_2 = \theta_2 - \theta = -\eta_2 \cdot \nabla L_2$$

其中， η_2 表示第二参与设备训练所述风控模型的学习率。

在一些实施例，每个参与设备在每轮迭代中确定的中间训练信息可以包括损失函数。比如，中间训练信息#1b可以包括本地数据集#1b中所有样本的损失之和。

可以理解，第一参与设备生成的第一本地数据集中的训练样本包括风险标签，因此第一参与设备在模型训练过程中，进行有监督学习；而第二参与设备生成的第二本地数据集有部分训练样本不包括风险标签，因此第二参与设备在模型训练过程中，进行无监督学习。由此可见，第一参与设备和第二参与设备在模型训练过程中，使用的损失函数不同。有监督学习的损失函数例如交叉熵损失函数或者均方误差损失函数等；无监督学习的损失函数例如重构误差函数或者距离度量函数等。

S140: M个参与设备将各自得到的中间训练信息#1发送给控制设备。

10 例如，图4所示，S140包括S1401和S1402。

S1401: 第一参与设备向控制设备发送所述中间训练信息#1a，对应的，控制设备接收中间训练信息#1a。

示例性的，所述中间训练信息#1a为中间训练信息密文#1a，第一参与设备确定中间训练信息#1a后，通过所述公钥将中间训练信息#1a加密后得到中间训练信息密文#1a，并向控制设备发送所述中间训练信息密文#1a，控制设备可以通过所述私钥将所述中间训练信息密文#1a解密得到所述中间训练信息#1a。

S1402: 第二参与设备向控制设备发送中间训练信息#1b，对应的，控制设备接收中间训练信息#1b。

示例性的，所述中间训练信息#1b为中间训练信息密文#1b，第二参与设备确定中间训练信息#1b后，通过所述公钥将中间训练信息#1b加密后得到中间训练信息密文#1b，并向控制设备发送所述中间训练信息密文#1b，控制设备可以通过所述私钥将所述中间训练信息密文#1b解密得到所述中间训练信息#1b。

各参与设备与控制设备之间通过密文形式传递中间训练信息，能够提高信息传递过程中的安全性，避免信息泄露。

25 S150: 控制设备基于M个中间训练信息#1对所述风控模型的当前模型参数#1进行更新得到当前模型参数#2。

示例性的，控制设备得到的中间训练信息#1a包括梯度信息#1a，控制设备得到的中间训练信息#1b包括梯度信息#1b，则控制设备可以向第一参与设备获取本地数据集#1a中样本用户的数量 D_1 ，向第二参与设备获取本地数据集#1b中样本用户的数量 D_2 。

30 控制设备可以基于 D_1 和 D_2 确定第一参与设备对应的训练权重#1和第二参与设备对应

的训练权重#2。比如训练权重#1 为 $\frac{|D_1|}{|D_1|+|D_2|}$ ，训练权重#2 为 $\frac{|D_2|}{|D_1|+|D_2|}$ 。控制设备得到梯度信息#1a、梯度信息#1b、训练权重#1 和训练权重#2 之后，可以基于以下公式得到聚合梯度信息 ΔL 。

$$\Delta L = \frac{1}{|D_1|+|D_2|} (|D_1| \cdot \nabla L_1 + |D_2| \cdot \nabla L_2)$$

5

进而控制设备可以基于聚合梯度信息确定当前模型参数#2。比如当前模型参数#2 基于以下公式得到：

$$\theta' = \theta - \eta \cdot \Delta L$$

其中， θ' 表示更新后的当前模型参数#2， η 表示控制设备训练所述风控模型的学习率， ΔL 表示聚合梯度。 η_1 和 η_2 可以相同，也可以不同。

10

示例性的，控制设备得到的中间训练信息#1a 包括 $\Delta\theta_1$ ，控制设备得到的中间训练信息#1b 包括 $\Delta\theta_2$ ，则控制设备可以向第一参与设备和第二参与设备获取 η_1 和 η_2 ，并根据 $\Delta\theta_1$ 和 η_1 确定梯度信息#1a，根据 $\Delta\theta_2$ 和 η_2 确定梯度信息#1b。后续控制设备可以通过上文示例的方式确定当前模型参数#2。

15

示例性的，控制设备得到的中间训练信息#1a 包括本地数据集#1a 中所有样本的损失之和#1a，则控制设备可以基于所述损失之和#1a、 D_1 确定梯度信息#1a；控制设备得到的中间训练信息#1b 包括本地数据集#1b 中所有样本的损失之和#1b，则控制设备可以基于所述损失之和#1b、 D_2 确定梯度信息#1b。后续控制设备可以通过上文示例的方式确定当前模型参数#2。

20

至此，第一次迭代训练结束。依次类推，经过 $n-1$ 次迭代训练后，控制设备可以基于预设的训练结束条件判断当前模型参数#n 对应的所述风控模型是否还需要继续训练，若确定不需要继续训练，则将当前模型参数#n 作为所述风控模型的目标参数；若确定需要继续训练，则进入第 n 次迭代训练。当前模型参数#n 为控制设备在第 $n-1$ 次迭代训练更新得到的所述风控模型的当前模型参数。

25

训练结束条件可以包括预设迭代次数，即迭代次数达到预设迭代次数后，可以停止模型训练。或者所述训练结束条件可以包括损失函数 1 和损失函数 2 都收敛。或者训练结束条件可以包括与验证集性能相关的条件。比如第一参与设备可以在第 $n-1$ 次迭代训练过程中，基于第一平台数据生成的验证集确定当前模型参数#n-1 对应的第一准

准确率，并将第一准确率发送至控制设备；第二参与设备可以在第 $n-1$ 次迭代训练过程中，基于第二平台数据生成的验证集确定当前模型参数 $\#n-1$ 对应的第二准确率，并将第二准确率发送至控制设备。控制设备确定第一准确率和第二准确率都达到预设准确率时，可以确定训练结束，如果第一准确率或第二准确率有一个没有达到预设准确率，则确定还需要进行第 n 次迭代训练。或者，所述训练结束条件还可以是其他条件，不再一一列举。

在第 n 次迭代过程中：

S12x：控制设备向所述 M 个参与设备分发所述风控模型的当前模型参数 $\#n$ ，当前模型参数 $\#n$ 即 $n-1$ 次迭代过程中控制设备更新得到的模型参数。

例如，图 4 所示，S12x 包括 S12x1 和 S12x2，可以理解，S12x1 和 S12x2 可以同时执行，也可以先后依次执行，本说明书对此不做限制。

S12x1：控制设备向第一参与设备发送当前模型参数 $\#n$ ，对应的，第一参与设备接收当前模型参数 $\#n$ 。

S12x2：控制设备向第二参与设备发送当前模型参数 $\#n$ ，对应的，第二参与设备接收当前模型参数 $\#n$ 。

S13x： M 个参与设备中的每个参与设备基于当前模型参数 $\#n$ 和本地数据集确定中间训练信息 $\#n$ 。

当前模型参数 $\#n$ 可以理解为第 n 次迭代过程中所述风控模型的当前模型参数。中间训练信息 $\#n$ 理解为第 n 次迭代过程中，参与设备得到的中间训练信息，每个参与设备在第 n 次迭代过程中得到的中间训练信息 $\#n$ 不同。

例如，图 4 所示，S13x 包括 S13x1 和 S13x2。

S13x1：第一参与设备基于当前模型参数 $\#n$ 和本地数据集 $\#na$ 确定中间训练信息 $\#na$ 。

第一参与设备确定中间训练信息 $\#na$ 与确定中间训练信息 $\#1a$ 的方式类似，可参考上文描述，不再赘述。

S13x2：第二参与设备基于当前模型参数 $\#n$ 和本地数据集 $\#nb$ 确定中间训练信息 $\#nb$ 。

第二参与设备确定中间训练信息 $\#nb$ 与确定中间训练信息 $\#1b$ 的方式类似，可参考上文描述，不再赘述。

S14x： M 个参与设备将各自得到的中间训练信息 $\#n$ 发送给控制设备。

例如，图 4 所示，S14x 包括 S14x1 和 S14x2。

S14x1：第一参与设备向控制设备发送所述中间训练信息#na，对应的，控制设备接收中间训练信息#na。

可以理解，S14x1 和 S1401 类似，相关内容和实现方式可参考 S1401 中的描述，不再赘述。

S14x2：第二参与设备向控制设备发送所述中间训练信息#nb，对应的，控制设备接收中间训练信息#nb。

可以理解，S14x2 和 S1402 类似，相关内容和实现方式可参考 S1402 中的描述，不再赘述。

S15x：控制设备基于 M 个中间训练信息#n 对所述风控模型的当前模型参数#n 进行更新得到当前模型参数#n+1。

可以理解，S15x 和 S150 类似，相关内容和实现方式可参考 S150 中的描述，不再赘述。

控制设备得到当前模型参数#n+1 之后，基于上文示例的训练结束条件判断当前模型参数#n+1 对应的所述风控模型是否达到要求，若确定已达到要求，则不需要继续训练，将所述当前模型参数#n+1 作为所述风控模型的目标参数。

方法 P100 还包括：

S160：控制设备分别向所述 M 个参与设备发送所述风控模型的目标参数。

例如，图 4 所示，S160 包括 S1601 和 S1602，可以理解，S1601 和 S1602 可以同时执行，也可以先后依次执行，本说明书对此不做限制。

S1601：控制设备向第一参与设备发送所述目标参数，对应的，第一参与设备接收所述目标参数。

S1602：控制设备向第二参与设备发送所述目标参数，对应的，第二参与设备接收所述目标参数。

至此，联合训练结束，所述目标参数即联合训练结束后，得到的能够识别风险用户的所述风控模型。

需要说明的是，在一些应用场景下，上述 S1601 和 S1602 也可以择一执行。例如，当，第一参与设备为开展服务时间较长的第一服务平台的服务设备，第二参与设备为新开展服务的第二服务平台的服务设备，该情况下，方法 P100 训练后的风控模型可用于识别第二参与设备所在区域内的用户是否为风险用户。这样，可以避免第二参与设

备在服务前期由于防控历史数据较少导致的冷启动问题。或者，在一些应用场景下，S1601 和 S1602 也可以都不执行。此时，第一参与设备或者第二参与设备可以将自身在第 n 次迭代过程中更新后的参数作为所述风控模型的目标参数。

5 具体来说，第二参与设备得到所述风控模型后，可以将所述风控模型部署在所述第二服务平台，用于防控所述第二服务平台中的风险用户。

例如，第二参与设备将所述风控模型部署在所述第二服务平台的营销场景。当用户甲在第二服务平台发起一笔营销场景下的支付事件（比如该支付事件使用优惠券等权益）时，第二服务平台可以基于所述 N 个目标特征维度挖掘所述支付事件关联的用户甲的用户特征，并将获得的用户特征输入所述风控模型，得到预测结果。第二服务平台可以基于预测结果确定用户甲是否是风险用户。例如，所述预设结果表征用户甲为风险用户的概率，当所述概率小于或等于第一阈值时，第二服务平台确定用户甲不是风险用户；当所述概率大于所述第一阈值且小于或等于第二阈值时，第二阈值大于所述第一阈值，第二服务平台确定用户甲可能是风险用户，即第二服务平台还需要结合其他条件进一步判断用户甲是否是风险用户；当所述概率大于所述第二阈值时，第二服务平台确定用户甲是风险用户。若第二服务平台确定用户甲不是风险用户，则不用拦截所述支付事件，即可以让用户甲完成所述支付事件。若第二服务平台确定用户甲是风险用户，则拦截所述支付事件，即不让用户甲完成所述支付事件，或者，第二服务平台增加额外的验证流程，在用户甲完成验证流程后，让用户甲完成所述支付事件。

20 可以理解，图 3 或者图 4 示例的是识别风险用户的风控模型的训练过程，以此类推，还可以基于方法 P100 展示的训练逻辑训练其他风控模型。比如基于方法 P100 训练用于其他场景的识别风险用户的风控模型，在此情况下，训练过程中的本地数据集包括的用户特征需要结合所述其他场景。或者基于方法 P100 训练用于识别其他风险的风控模型，在此情况下，训练过程中的本地数据集包括的风险特征与识别的具体风险关联。

还可以理解，对于营销场景，可以基于 P100 展示的训练逻辑得到识别不同风险维度的多个风控模型，其中，多个风控模型包括用户风险维度对应的风控模型（即方法 P100 中示例的用于识别风险用户的风控模型）。上述多个风控模型可以被联合使用以识别营销场景下的目标事件是否为风险事件。

30 例如，用户甲在第二服务平台发起一笔营销场景下的支付事件，第二服务平台可

以确定所述多个风控模型中每个风控模型对应的风险特征集。对于每个风控模型而言，第二服务平台将所述风控模型对应的风险特征集输入所述风控模型，得到支付事件的风险子值，依次类推，得到多个风控模型对应的多个风险子值。进而第二服务平台基于所述多个风险子值确定所述支付事件的风险值，当所述风险值大于或等于风险阈值时，确定所述支付事件为风险事件，不让用户甲完成所述支付事件；当所述风险值小于所述风险阈值时，确定所述支付事件不是风险事件，让用户甲继续完成所述支付事件。

需要说明的是，上述描述方法 P100 的过程中，以支付平台的营销场景举例，因此具体描述中采用的是该营销场景中会涉及的术语，例如本地数据集包括的用户特征和风险标签等。如果将方法 P100 应用到前文提及的其他平台，所涉及的术语可以适应性的调整，此处不一一枚举。

综上所述，本说明书提供的模型的训练方法、设备及存储介质，通过包括控制设备和多个参与设备的联邦系统进行多次迭代的联合训练，得到可以识别风险用户的风控模型。在每次迭代训练过程中，通过控制设备的协调，可以在参与设备的本地数据集不出本地的前提下完成风控模型的训练，有利于数据隐私保护。并且联合训练的训练数据来自多个参与设备，打破了数据孤岛，训练数据更加丰富、全面。因此采用本方案得到的风控模型的泛化能力、适应性和风控的准确性都更好一些。

本说明书另一方面提供一种非暂时性存储介质，存储有至少一组用来进行模型训练的可执行指令。当所述可执行指令被处理器执行时，所述可执行指令指导所述处理器实施本说明书所述的方法 P100 的步骤。在一些可能的实施方式中，本说明书的各个方面还可以实现为一种程序产品的形式，其包括程序代码。当所述程序产品在控制设备 100 上运行时，所述程序代码用于使控制设备 100 执行本说明书描述的方法 P100 中由控制设备 100 执行的步骤。当所述程序产品在参与设备 200 上运行时，所述程序代码用于使参与设备 200 执行本说明书描述的方法 P100 中由参与设备 200 执行的步骤。当所述程序产品在风控执行设备上运行时，所述程序代码用于使风控执行设备执行本说明书描述的风险用户的识别方法的步骤。用于实现上述方法的程序产品可以采用便携式紧凑盘只读存储器（CD-ROM）包括程序代码，并可以在控制设备 100 或参与设备 200 或风控执行设备上运行。然而，本说明书的程序产品不限于此，在本说明书中，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使用或者与其结合使用。所述程序产品可以采用一个或多个可读介质的任意组合。可读

介质可以是可读信号介质或者可读存储介质。可读存储介质例如可以为但不限于电、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。可读存储介质的更具体的例子包括：具有一个或多个导线的电连接、便携式盘、硬盘、随机存取存储器（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。所述计算机可读存储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，其中承载了可读程序代码。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。可读存储介质还可以是可读存储介质以外的任何可读介质，该可读介质可以发送、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。可读存储介质上包含的程序代码可以用任何适当的介质传输，包括但不限于无线、有线、光缆、RF 等等，或者上述的任意合适的组合。可以以一种或多种程序设计语言的任意组合来编写用于执行本说明书操作的程序代码，所述程序设计语言包括面向对象的程序设计语言—诸如 Java、C++等，还包括常规的过程式程序设计语言—诸如“C”语言或类似的设计语言。程序代码可以完全地在控制设备 100 和参与设备 200 上执行，部分地在控制设备 100 和参与设备 200 上执行，作为一个独立的软件包执行，部分在控制设备 100 和参与设备 200 上、部分在远程计算设备上执行、或者完全在远程计算设备上执行。程序代码可以完全地在风控执行设备上执行，部分地在风控执行设备上执行，作为一个独立的软件包执行，部分在风控执行设备上、部分在远程计算设备上执行、或者完全在远程计算设备上执行。

本说明书提供的技术方案中，所涉及的平台数据的采集，以及控制设备和参与设备之间信息的传输，均符合相关法律法规的规定，且不违背公序良俗。

本说明书提及的“多个”，如果没有特殊的声明，应当理解为两个或者两个以上。

上述对本说明书特定实施例进行了描述。其他实施例在所附权利要求书的范围内。在一些情况下，在权利要求书中记载的动作或步骤可以按照不同于实施例中的顺序来执行并且仍然可以实现期望的结果。另外，在附图中描绘的过程不一定要求示出特定顺序或者连续顺序才能实现期望的结果。在某些实施方式中，多任务处理和并行处理也是可以的或者是可能有利的。

综上所述，在阅读本详细公开内容之后，本领域技术人员可以明白，前述详细公开内容可以仅以示例的方式呈现，并且可以不是限制性的。尽管这里没有明确说明，

本领域技术人员可以理解本说明书需求囊括对实施例的各种合理改变，改进和修改。这些改变，改进和修改旨在由本说明书提出，并且在本说明书的示例性实施例的精神和范围内。

权 利 要 求 书

1. 一种模型的训练方法，用于通过联邦系统训练得到具有识别风险用户的能力的风控模型，所述联邦系统包括控制设备和位于不同区域的 M 个参与设备，所述 M 为大
5 于 1 的整数，每个参与设备存储有本地数据集，所述本地数据集至少包括所述参与设备所在区域内的样本用户的用户特征，所述方法应用于所述控制设备，包括：
- 初始化所述风控模型的参数；以及
- 控制所述 M 个参与设备对所述风控模型进行多次迭代的联合训练，直至达到预设的训练结束条件时，得到所述风控模型的目标参数，其中，每次迭代过程包括：
- 10 向所述 M 个参与设备分发所述风控模型的当前模型参数，以便所述 M 个参与设备基于所述当前模型参数和所述 M 个参与设备各自对应的所述本地数据集确定中间训练信息，所述中间训练信息表征所述风控模型对所述本地数据集中的样本用户的风险预测情况，以及
- 从所述 M 个参与设备分别获得所述中间训练信息，并基于获得的 M 个中
15 间训练信息，对所述风控模型的当前模型参数进行更新。
2. 根据权利要求 1 所述的方法，其中，所述 M 个参与设备至少包括第一参与设备和第二参与设备，
- 所述第一参与设备存储有第一本地数据集，所述第一本地数据集中包括所述第一
20 参与设备所在区域内的第一样本用户的用户特征、以及所述第一样本用户的风险标签，
- 所述第二参与设备存储有第二本地数据集，所述第二本地数据集中包括所述第二参与设备所在区域内的第二样本用户的用户特征，以及
- 训练后的所述风控模型用于识别所述第二参与设备所在区域内的用户是否为风险
25 用户。
3. 根据权利要求 1 所述的方法，其中，所述方法还包括：
- 分别向所述 M 个参与设备发送第一指示信息，所述第一指示信息用于指示 N 个目标特征维度，以便所述 M 个参与设备基于所述 N 个目标特征维度生成所述本地数据集，所述 N 为大于或等于 1 的整数。

4. 根据权利要求 3 所述的方法, 其中, 所述方法还包括:

从所述 M 个参与设备分别接收第二指示信息, 所述第二指示信息用于指示多个候选特征维度; 以及

5 基于接收到的 M 个第二指示信息, 确定所述 N 个目标特征维度。

5. 根据权利要求 1 所述的方法, 其中, 所述中间训练信息包括梯度信息, 所述梯度信息表征损失函数关于所述当前模型参数的梯度, 所述损失函数用于确定所述参与设备基于所述当前模型参数对所述样本用户进行风险预测的损失情况;

10 所述基于获得的 M 个中间训练信息, 对所述风控模型的当前模型参数进行更新, 包括:

对所述 M 个中间训练信息中的所述梯度信息进行聚合, 得到聚合梯度信息; 以及基于所述聚合梯度信息对所述风控模型的当前模型参数进行更新。

15 6. 根据权利要求 5 所述的方法, 其中, 所述对所述 M 个中间训练信息中的所述梯度信息进行聚合, 得到聚合梯度信息, 包括:

获得所述 M 个参与设备各自对应的所述本地数据集中的所述样本用户的数量;

基于所述数量, 确定所述 M 个参与设备各自对应的训练权重; 以及

20 基于所述训练权重对所述 M 个中间训练信息中的 M 个梯度信息进行聚合得到所述聚合梯度信息。

7. 根据权利要求 1 所述的方法, 其中, 从所述 M 个参与设备分别获得所述中间训练信息, 包括:

从所述 M 个参与设备分别获得密文; 以及

25 通过私钥对所述密文进行解密得到所述中间训练信息。

8. 根据权利要求 1 所述的方法, 其中, 所述得到所述风控模型的目标参数之后, 所述方法还包括:

分别向所述 M 个参与设备发送所述风控模型的目标参数。

9. 一种模型的训练方法，用于通过联邦系统训练得到具有识别风险用户的能力的风控模型，所述联邦系统包括控制设备和位于不同区域的 M 个参与设备，所述 M 为大于 1 的整数，每个参与设备存储有本地数据集，所述本地数据集至少包括所述参与设备所在区域内的样本用户的用户特征，所述方法应用于所述参与设备，包括：

5 从所述控制设备接收所述风控模型的当前模型参数；

 基于所述当前模型参数和所述参与设备对应的所述本地数据集确定中间训练信息，所述中间训练信息表征所述风控模型对所述本地数据集中的样本用户的风险预测情况；以及

 向所述控制设备发送所述中间训练信息，以便所述控制设备基于所述 M 个参与设备各自对应的所述中间训练信息对所述风控模型的当前模型参数进行更新。

10

10. 根据权利要求 9 所述的方法，其中，所述 M 个参与设备至少包括第一参与设备和第二参与设备，

 所述第一参与设备存储有第一本地数据集，所述第一本地数据集中包括所述第一参与设备所在区域内的第一样本用户的用户特征、以及所述第一样本用户的风险标签，

15

 所述第二参与设备存储有第二本地数据集，所述第二本地数据集中包括所述第二参与设备所在区域内的第二样本用户的用户特征，以及

 训练后的所述风控模型用于识别所述第二参与设备所在区域内的用户是否为风险用户。

20

11. 根据权利要求 9 所述的方法，其中，所述方法还包括：

 从所述控制设备接收第一指示信息，所述第一指示信息用于指示 N 个目标特征维度，所述 N 为大于或等于 1 的整数；以及

 基于所述 N 个目标特征维度生成所述本地数据集。

25

12. 根据权利要求 11 所述的方法，其中，所述方法还包括：

 向所述控制设备发送第二指示信息，所述第二指示信息用于指示多个候选特征维度，以便所述控制设备基于所述多个候选特征维度确定所述 N 个目标特征维度。

13. 根据权利要求 9 所述的方法, 其中, 所述基于所述当前模型参数和所述参与设备对应的所述本地数据集确定中间训练信息, 包括:

基于所述当前模型参数和所述参与设备对应的所述本地数据集确定梯度信息, 所述梯度信息表征损失函数关于所述当前模型参数的梯度, 所述损失函数用于确定所述参与设备基于所述当前模型参数对所述样本用户进行风险预测的损失情况;

所述向所述控制设备发送所述中间训练信息, 包括:

向所述控制设备发送梯度信息, 以便所述控制设备基于所述梯度信息得到用于更新所述当前模型参数的聚合梯度信息。

14. 根据权利要求 13 所述的方法, 其中, 所述方法还包括:

向所述控制设备发送所述本地数据集中的所述样本用户的数量, 以便所述控制设备基于所述数量和所述梯度信息得到所述聚合梯度信息。

15. 根据权利要求 9 所述的方法, 其中, 所述向所述控制设备发送所述中间训练信息, 包括:

通过公钥对所述中间训练信息加密生成密文; 以及
向所述控制设备发送所述密文。

16. 根据权利要求 9 所述的方法, 其中, 所述方法还包括:

从所述控制设备接收所述风控模型的目标参数, 所述目标参数为训练完成后的所述风控模型的模型参数。

17. 一种电子设备, 位于包括控制设备和 M 个参与设备的联邦系统中, 所述联邦系统用于训练得到具有识别风险用户的能力的风控模型, 所述 M 个参与设备位于不同区域, 每个参与设备存储有本地数据集, 所述本地数据集至少包括所述参与设备所在区域内的样本用户的用户特征, 所述电子设备作为所述控制设备, 包括:

至少一个存储介质, 存储有至少一个指令集; 以及

至少一个处理器, 同所述至少一个存储介质通信连接, 其中, 所述至少一个处理器运行时读取所述至少一个指令集, 并且根据所述至少一个指令集的指示:

初始化所述风控模型的参数,

控制所述 M 个参与设备对所述风控模型进行多次迭代的联合训练，直至达到预设的训练结束条件时，得到所述风控模型的目标参数，其中，每次迭代过程包括：

5 向所述 M 个参与设备分发所述风控模型的当前模型参数，以便所述 M 个参与设备基于所述当前模型参数和所述 M 个参与设备各自对应的所述本地数据集确定中间训练信息，所述中间训练信息表征所述风控模型对所述本地数据集中的样本用户的风险预测情况，以及
从所述 M 个参与设备分别获得所述中间训练信息，并基于获得的 M 个中间训练信息，对所述风控模型的当前模型参数进行更新。

10 18. 根据权利要求 17 所述的电子设备，其中，所述至少一个处理器还：

分别向所述 M 个参与设备发送第一指示信息，所述第一指示信息用于指示 N 个目标特征维度，以便所述 M 个参与设备基于所述 N 个目标特征维度生成所述本地数据集，所述 N 为大于或等于 1 的整数。

15 19. 一种电子设备，位于包括控制设备和 M 个参与设备的联邦系统中，所述联邦系统用于训练得到具有识别风险用户的能力的风控模型，所述 M 个参与设备位于不同区域，每个参与设备存储有本地数据集，所述本地数据集至少包括所述参与设备所在区域内的样本用户的用户特征，所述电子设备作为所述 M 个参与设备中的其中一个参与设备，包括：

20 至少一个存储介质，存储有至少一个指令集；以及

至少一个处理器，同所述至少一个存储介质通信连接，其中，所述至少一个处理器运行时读取所述至少一个指令集，并且根据所述至少一个指令集的指示：

从所述控制设备接收所述风控模型的当前模型参数，

25 基于所述当前模型参数和所述参与设备对应的所述本地数据集确定中间训练信息，所述中间训练信息表征所述风控模型对所述本地数据集中的样本用户的风险预测情况，以及

向所述控制设备发送所述中间训练信息，以便所述控制设备基于所述 M 个参与设备各自对应的所述中间训练信息对所述风控模型的当前模型参数进行更新。

20. 一种计算机可读非暂时性存储介质，其中，所述计算机可读非暂时性存储介质中存储有至少一个指令集，所述至少一个指令集被至少一个处理器执行时，实现如权利要求 1-9 中任一项所述的方法，或者实现如权利要求 10-16 中任一项所述的方法。