

虚拟资源分配、模型建立、数据预测方法及装置

技术领域

本说明书涉及计算机应用领域，尤其涉及一种虚拟资源分配、模型的建立、数据预测方法及装置。

5 背景技术

随着互联网技术的飞速发展，用户的个人数据的网络化和透明化已经成为不可阻挡的大趋势。对于一些面向用户提供互联网服务的服务平台而言，可以通过采集用户日常产生的服务数据，收集到海量的用户数据。而这些用户数据对于服务平台的运营方来说，是非常珍贵的“资源”，服务平台的运营
10 方可以通过数据挖掘和机器学习，基于这些“资源”来构建用户评估模型，并利用该用户评估模型对用户进行评估决策。

例如，在信贷发放的场景下，可以从海量的用户数据中提取出若干个维度的数据特征，并基于提取出的这些特征构建训练样本，通过特定的机器学习算法训练来创建用户风险评估模型，然后使用该用户风险评估模型，对用
15 户进行风险评估，并基于风险评估结果来决策该用户是否为风险用户，进而决定是否需要向用户发放贷款。

发明内容

本说明书提出一种虚拟资源分配方法，包括：

接收多个数据提供方上传的若干用户的评估结果；其中，所述评估结果
20 为各数据提供方基于自己的评估模型对用户分别进行评估后得到；

将各数据提供方上传的评估结果作为训练数据，构建出若干训练样本，每一条训练样本包含同一个用户在所述各数据提供方的评估结果；其中，所述训练样本根据用户对业务的实际执行情况被标定了标签；

基于所述若干训练样本以及各个训练样本的标签对模型进行训练，将训练

好的模型中的各变量的系数作为各数据提供方贡献度，基于各数据提供方的贡献度为各数据提供方分配虚拟资源。

可选的，所述训练好的模型为线性模型。

可选的，为各数据提供方分配的所述虚拟资源的数量，与各数据提供方的贡献度成正比。

可选的，还包括：

接收多个数据提供方上传的针对某一个用户的评估结果，将所述评估结果输入所述训练好的模型中，得到该用户的最终的评估结果。

可选的，所述虚拟资源为向各数据提供方发放的用户数据使用资金。

可选的，所述评估模型为用户风险评估模型；所述评估结果为风险评分；所述标签指示用户是否为风险用户。

本说明书还提出一种虚拟资源分配装置，包括：

接收模块，接收多个数据提供方上传的若干用户的评估结果；其中，所述评估结果为各数据提供方基于自己的评估模型对用户分别进行评估后得到；

训练模块，将各数据提供方上传的评估结果作为训练数据，构建出若干训练样本，每一条训练样本包含同一个用户在所述各数据提供方的评估结果；其中，所述训练样本根据用户对业务的实际执行情况被标定了标签；

分配模块，基于所述若干训练样本以及各个训练样本的标签对模型进行训练，将训练好的模型中的各变量的系数作为各数据提供方贡献度，基于各数据提供方的贡献度为各数据提供方分配虚拟资源。

可选的，所述训练好的模型为线性模型。

可选的，为各数据提供方分配的所述虚拟资源的数量，与各数据提供方的贡献度成正比。

可选的，还包括：

评估模块，接收多个数据提供方上传的针对某一个用户的评估结果，将所述评估结果输入所述训练好的模型中，得到该用户的最终的评估结果。

可选的，所述虚拟资源为向各数据提供方发放的用户数据使用资金。

可选的,所述评估模型为用户风险评估模型;所述评估结果为风险评分;所述标签指示用户是否为风险用户。

可选的,本说明书还提出一种模型建立方法,包括:

- 接收多个数据提供方上传的若干用户的评估结果;其中,所述评估结果
- 5 为各数据提供方基于自己的评估模型对用户分别进行评估后得到;

将各数据提供方上传的评估结果作为训练数据,构建出若干训练样本,每一条训练样本包含同一个用户在所述各数据提供方的评估结果;其中,所述训练样本根据用户对业务的实际执行情况被标定了标签;

- 基于所述若干训练样本以及各个训练样本的标签对模型进行训练,得到训
- 10 练好的模型。

可选的,其中,所述训练好的模型为线性模型。

可选的,所述评估模型为用户风险评估模型;所述评估结果为风险评分;所述标签指示用户是否为风险用户。

本说明书还提出一种数据预测的方法,包括:

- 15 接收多个数据提供方上传的若干用户的评估结果;其中,所述评估结果
- 为各数据提供方基于自己的评估模型对用户分别进行评估后得到;

将各数据提供方上传的评估结果作为训练数据,构建出若干训练样本,每一条训练样本包含同一个用户在所述各数据提供方的评估结果;其中,所述训练样本根据用户对业务的实际执行情况被标定了标签;

- 20 基于所述若干训练样本以及各个训练样本的标签对模型进行训练,得到训
- 练好的模型;

接收多个数据提供方上传的针对某一个用户的评估结果,将所述评估结果输入所述训练好的模型中,得到该用户的最终的评估结果。

本说明书还提出一种虚拟资源分配系统,包括:

- 25 多个数据提供方服务端,向风险评估方服务端上传若干用户的评估结果;
- 其中,所述评估结果为各数据提供方基于自己的评估模型对用户分别进行评估后得到;

风险评估方服务端，将各数据提供方上传的评估结果作为训练数据，构建出若干训练样本，每一条训练样本包含同一个用户在所述各数据提供方的评估结果；其中，所述训练样本根据用户对业务的实际执行情况被标定了标签；以及，基于所述若干训练样本以及各个训练样本的标签对模型进行训练，

5 将训练好的模型中的各变量的系数作为各数据提供方贡献度，基于各数据提供方的贡献度为各数据提供方分配虚拟资源。

本说明书还提出一种电子设备，包括：

处理器；

用于存储机器可执行指令的存储器；

10 其中，通过读取并执行所述存储器存储的与虚拟资源分配的控制逻辑对应的机器可执行指令，所述处理器被促使：

接收多个数据提供方上传的若干用户的评估结果；其中，所述评估结果为各数据提供方基于自己的评估模型对用户分别进行评估后得到；

将各数据提供方上传的评估结果作为训练数据，构建出若干训练样本，

15 每一条训练样本包含同一个用户在所述各数据提供方的评估结果；其中，所述训练样本根据用户对业务的实际执行情况被标定了标签；

基于所述若干训练样本以及各个训练样本的标签对模型进行训练，将训练好的模型中的各变量的系数作为各数据提供方贡献度，基于各数据提供方的贡献度为各数据提供方分配虚拟资源。

20 本说明书中，多个数据提供方可以基于自己的评估模型对若干用户分别进行评估后得到的评估结果上传给风险评估方；而风险评估方可以将各数据提供方上传的评估结果作为训练数据，构建出若干训练样本，来训练模型，并将训练好的模型中的与各变量对应的系数作为各数据提供方的贡献度，然后可以基于该贡献度为各数据提供方分配虚拟资源：

25 一方面，由于风险评估方在基于各数据提供方维护的用户数据训练模型时，数据提供方仅需要向风险评估方传输对若干用户进行初步评估后得到的评估结果，因此对于数据提供方而言，不再需要将本地维护的原始用户数据

传输至风险评估方，可以显著降低用户隐私泄露的发生风险；

另一方面，由于训练好的模型中的各变量的系数能够真实的反映各数据提供方对训练越好的模型的贡献度，因此基于该贡献度向各数据提供方分配虚拟资源，能够做到虚拟资源的合理分配。

5 附图说明

图 1 是本说明书一实施例示出的一种虚拟资源分配方法的流程图；

图 2 是本说明书一实施例示出的一种风险评估方基于多个数据提供方上传的评估结果训练模型的示意图；

图 3 是本说明书一实施例示出的一种模型建立方法的流程图；

10 图 4 是本说明书一实施例示出的一种数据预测方法的流程图；

图 5 是本说明书一实施例提供的承载虚拟资源分配装置的电子设备所涉及的硬件结构图；

图 6 是本说明书一实施例提供的所述虚拟资源分配装置的逻辑框图。

具体实施方式

15 在大数据时代，通过对海量数据进行挖掘，可以获得各种形式的有用信息，因此数据的重要性不言而喻。不同的机构都拥有各自的数据，但是任何一家机构的数据挖掘效果，都会受限于其自身拥有的数据数量和种类。针对该问题，一种直接的解决思路是：多家机构相互合作，将数据进行共享，从而实现更好的数据挖掘效果，实现共赢。

20 然而对于数据拥有方而言，数据本身是一种具有很大价值的资产，而且出于保护隐私、防止泄露等需求，数据拥有方往往并不愿意直接把数据提供出来，这种状况导致“数据共享”在现实中很难实际运作。因此，如何在充分保证数据安全的前提下实现数据共享，已经成为行业内备受关注的问題。

本说明书中，则旨在提出一种风险评估方在“共享”多个数据提供方维护
25 的用户数据来训练模型时，各个数据提供方不再需要将原始的用户数据传输至

风险评估方，就可以完成“数据共享”的技术方案。

在实现时，各个数据提供方可以分别基于机器学习算法对其本地维护的用户数据进行训练，构建用户评估模型，并使用该用户评估模型对若干样本用户进行评估，然后将评估结果上传给风险评估方。

- 5 而风险评估方可以将各数据提供方上传的评估结果作为训练数据，来构建出若干训练样本；其中，每一条训练样本包含同一个用户在各数据提供方的评估结果。例如，可以将各数据提供方上传的对某一个用户的评估结果分别作为建模特征，来构建一个特征向量作为训练样本。

- 10 并且，构建完成的训练样本可以根据用户对业务的实际执行情况对训练样本标定对应的标签；例如，在信贷发放的业务场景下，为各训练样本标定的标签，具体可以是基于用户真实的还款情况，为用户标定的能够指示该用户是否为风险用户的用户标签。

- 15 最后，风险评估方可以基于构建完成的训练样本以及与各个训练样本对模型进行训练，并将训练好的模型中的各变量的系数作为各数据提供方对模型的贡献度，然后基于各数据提供方的贡献度为各数据提供方分配虚拟资源。

- 一方面，由于风险评估方在基于各数据提供方维护的用户数据训练模型时，数据提供方仅需要向风险评估方传输对若干用户进行初步评估后得到的评估结果，因此对于数据提供方而言，不再需要将本地维护的原始用户数据传输至风险评估方，可以显著降低用户隐私泄露的发生风险；

- 20 另一方面，由于训练好的模型中的各变量的系数能够真实的反映各数据提供方对训练越好的模型的贡献度，因此基于该贡献度向各数据提供方分配虚拟资源，能够做到虚拟资源的合理分配。

- 例如，以信贷发放的业务场景为例，上述用户评估模型具体可以是一个用于决策用户是否为风险用户的用户风险评估模型；而上述评估结果则可以是利用上述用户风险评估模型对用户进行风险评估后输出的一个风险评分。

在这种场景下，各个数据提供方可以基于自有的用户数据构建用户风险评估模型；当风险评估方（比如，可以是贷款的发放一方）需要共享各数据提供

方的用户数据来训练用户风险评估模型时，可以将各数据提供方上传的评估结果作为训练数据，来构建出若干训练样本，并基于用户真实的还款情况，为各训练样本标定能够指示该用户是否为风险用户的标签，然后可以基于构建完成的训练样本以及与各个训练样本对模型进行训练，并将训练好的模型中的各变量的系数作为各数据提供方对模型的贡献度，基于各数据提供方的贡献度为各数据提供方分配虚拟资源。从而，在整个过程中，各个数据提供方并不需要向风险评估方提供原始的用户数据，就可以完成“数据共享”。

下面通过具体实施例并结合具体的应用场景进行详细描述。

请参考图 1，图 1 是本说明书一实施例提供的一种虚拟资源分配方法，应用于风险评估方服务端，执行以下步骤：

步骤 102，接收多个数据提供方上传的若干用户的评估结果；其中，所述评估结果为各数据提供方基于自己的评估模型对用户分别进行评估后得到；

步骤 104，将各数据提供方上传的评估结果作为训练数据，构建出若干训练样本，每一条训练样本包含同一个用户在所述各数据提供方的评估结果；其中，所述训练样本根据用户对业务的实际执行情况被标定了标签；

步骤 106，基于所述若干训练样本以及各个训练样本的标签对模型进行训练，将训练好的模型中的各变量的系数作为各数据提供方贡献度，基于各数据提供方的贡献度为各数据提供方分配虚拟资源。

上述数据提供方，具体可以包括与上述风险评估方具有合作关系的一方。在实际应用中，上述数据提供方和风险评估方可以分别对应不同的运营方；例如，上述建模方可以是 A 公司的数据运营平台，而上述数据提供方可以是与 A 公司的数据运营平台对接的诸如电商平台、第三方银行、快递公司、其它金融机构、电信运营商等服务平台。

上述所述用户评估模型，具体可以包括任意类型的用于对用户进行评估的机器学习模型；

例如，在示出的一种实施方式中，上述用户评估模型，具体可以是基于特定的机器学习算法训练出的用户风险评估模型（比如，用于对用户进行风

险评估的线性逻辑回归模型或者评分卡模型)；相应的，通过该用户评估模型对用户进行评估后输出的评估结果，则可以是表征该用户的风险水平的风险评分；其中，在实际应用中，该风险评分通常是一个 0-1 之间的浮点数值（比如，上述风险评分具体可以是一个表征用户风险水平的概率值）；或者，

5 上述评估结果也可以是风险评分以外其它形式的评分，比如，信用评分等。

在本说明书中，为了降低将原始的用户数据传输至风险评估方进行建模，而造成的用户隐私泄露的风险，各个数据提供方可以不再需要将本地维护的原始用户数据，传输至风险评估方，而是各自利用本地维护的原始用户数据分别进行建模。

10 在实现时，各个数据提供方的服务端可以分别在后台收集用户日常产生的用户数据，从收集到的这些用户数据中采集若干条用户数据作为数据样本，并基于采集到的这些数据样本生成一个初始化的数据样本集合。

其中，采集到的上述数据样本的具体数量，在本说明书中不进行特别限定，本领域技术人员可以基于实际的需求进行设置。

15 上述用户数据的具体形态，取决于具体的业务场景以及建模需求，具体可以涵盖任意类型的，从中可以提取出用于训练用户评估模型的建模特征的用户数据，在本说明书中也不进行特别限定；

例如，在实际应用中，如果希望创建出一个用于对用户发起的贷款申请，或者支付交易进行风险评估的评分卡模型，上述用户数据则可以包括诸如用

20 户的交易数据、购物记录、还款记录、消费记录、理财产品购买记录等，能够从中提取出用于训练风险评估模型的建模特征的用户数据。

当基于采集到的数据样本生成上述数据样本集合后，上述数据提供方服务端还可以对该数据样本集合中的数据样本进行预处理。

其中，对上述数据样本集合中的数据样本进行预处理，通常包括对上述

25 数据样本集合中的数据样本进行数据清洗、补充缺省值、归一化处理或者其它形式的预处理过程。通过对数据样本集合中的数据样本进行预处理，可以将采集到的数据样本转换成适宜进行模型训练的标准化的数据样本。

当对上述数据样本集合中的数据样本预处理完成后，上述数据提供方服务端可以从上述数据样本集合中的各数据样本中，分别提取出若干个维度的数据特征（即最终参与建模的建模特征）。其中，提取出的上述若干个维度的数据特征的数量，在本说明书中不进行特别限定，本领域技术人员可以基
5 于实际的建模需求进行选择。

另外，提取出的数据特征的具体类型，在本说明书中也不进行特别限定，本领域技术人员可以基于实际的建模需求，从上述数据样本中实际所包含的信息中来人工选定。

当从数据样本中分别提取出若干个维度的数据特征后，上述数据提供方
10 服务端可以基于提取出的这些维度的数据特征对应的数据特征取值，为各数据样本分别生成一个数据特征向量，然后基于各数据样本的数据特征向量，构建出一个目标矩阵；比如，以从 N 条数据样本中分别提取出 M 个维度的数据特征为例，上述目标矩阵则可以是一个 $N*M$ 维的矩阵。

此时，构建出的上述目标矩阵，即最终进行模型训练的训练样本集，各
15 个数据提供方服务端，可以基于特定的机器学习算法，将上述目标矩阵作为原始的样本训练集进行机器学习，分别训练出一个用户评估模型。

其中，需要说明的是，各个数据提供方在训练用户评估模型时所采用的机器学习算法，可以相同也可以不同，在本说明书中不进行特别限定。

在本说明书中，上述机器学习模型具体可以是有监督的机器学习模型；
20 例如，上述机器学习模型具体可以是 LR（Logistic Regression，逻辑回归）模型。

在这种情况下，上述训练样本集中的每一条数据样本，可以分别携带一个预先标定的样本标签。其中，该样本标签的具体形态，通常也取决于具体的业务场景以及建模需求，在本说明书中也不进行特别限定；

25 例如，在实际应用中，如果希望创建出一个是否可以给用户发放贷款的模型，那么上述样本标签则具体可以是一个用于指示该用户是否为风险用户的用户标签；其中，该用户标签具体可以由风险评估方来标定和提供。在这

种情况下，上述目标矩阵中的每一个数据特征向量，可以分别对应一个样本标签。

其中，各个数据提供方基于有监督的机器学习算法训练用户评估模型的具体过程，在本说明书中不再进行详述，本领域技术人员在将本说明书记载
5 的技术方案付诸实现时，可以参考相关技术中的记载。

例如，以上述有监督的机器学习算法为 LR 算法为例，在基于 LR 算法训练逻辑回归模型时，通常可以采用损失函数（Loss Function）来评估训练样本和对应的样本标签之间的拟合误差。在实现时，可以将训练样本和对应的样本标签作为输入值输入至损失函数中，并采用梯度下降法进行反复迭代
10 计算，直至收敛，进而可以求解出模型参数（即训练样本中各个建模特征的最优权重值，该权重值可以表征各个建模特征对模型输出结果的贡献度）的取值，然后将求解出的该模型参数的取值作为最优参数，来构建上述逻辑回归模型。

请参见图 2，图 2 为本说明书示出的一种风险评估方基于多个数据提供方上传的评估结果训练模型的示意图。
15

在初始状态下，上述风险评估方可以预先准备若干样本用户，并将这些样本用户的用户 ID 通知给各个数据提供方；例如，在实现时可以将各样本用户的用户 ID 以列表的形式下发至各个数据提供方。

各数据提供方在收到上述若干样本用户的用户 ID 后，可以使用自己的
20 用户评估模型对各样本用户分别进行用户评估，然后将评估结果上传给风险评估方，由上述风险评估方进行建模。

当然，如果各用户提供方发给风险评估方的评估结果中对于相同的用户本身就具有相同的 ID，那么风险评估方无需将样本用户的用户 ID 通知给各个数据提供方。

25 可见，通过这种方式，对于各个数据提供方而言，不再需要将本地维护的原始的用户数据“共享”给风险评估方，只需要向风险评估方“共享”一个对用户的初步的评估结果即可。

一方面，数据提供方“共享”给风险评估方的初步的评估结果，可以理解为对本地维护的用户数据的一个降维；即各个数据提供方“共享”的初步的评估结果，可以看作是一个将本地维护的用户数据，降低到维度为 1 的一个数据特征。

5 另一方面，由于上述初步的评估结果，是由各个数据提供方基于对本地维护的用户数据进行机器学习建模得出的，因此将初步的评估结果“共享”给风险评估方，相当于是将基于机器学习从本地维护的用户数据中学习分析出的数据价值，共享给风险评估方。虽然各个数据提供方没有将原始的用户数据“共享”给风险评估方，但通过数据价值“共享”，仍然能够达到数据共享的
10 目的。

在本说明书中，风险评估方在收到各数据提供方上传的对应于这些样本用户的评估结果后，可以将各数据提供方上传的评估结果作为训练数据，为各样本用户分别创建出一条对应的训练样本。

此时，构建完成的每一个训练样本中，将包含各个数据提供方基于训练
15 完成上述用户评估模型，对该训练样本对应的样本用户进行初步的评估后，得到的评估结果。而对于每一个数据提供方的评估结果而言，将分别对应上述训练样本中的一个特征变量。

其中，上述特征变量具体是指构成上述训练样本的特征字段，在本说明书中，每一个训练样本中将包含若干个特征字段，而每个特征字段将分别对
20 应一个数据提供方上传的评估结果。

当为各样本用户分别创建出对应的训练样本后，还可以基于创建的训练样本生成一个训练样本集，并基于各样本用户对业务的实际执行情况对训练样本标定对应的标签；例如，在信贷发放的业务场景下，为各训练样本标定的标签，具体可以是基于用户真实的还款情况，为用户标定的能够指示该用
25 户是否为风险用户的用户标签。在这种场景下，风险评估方可以基于各样本用户最终是否违约还款，对各样本用户标定用户标签；比如，假设最终向某一样本用户发放贷款后，该用户出现了违约还款，那么最终在训练样本集中，

与该样本用户对应的训练样本将被标记一个用于指示该用户为风险用户的标签。

当为训练样本集中的训练样本分别标定了用户标签后，上述风险评估方服务端，可以基于构建的训练样本集以及与各训练样本对应的标签，对预设
5 的机器学习模型进行训练。

在示出的一种实施方式中，由于各数据提供方上传的对同一个用户的评估结果，与该用户的用户标签（即最终的用户评估结果）之间，可能会保持着一定的线性关系；

例如，对于风险评估方而言，可以通过将各个数据提供方上传的对同一
10 个用户的评估结果乘以相应的系数进行相加计算，然后将计算结果作为针对该用户的最终评估结果。

因此，在本实施例，风险评估方一侧训练的机器学习模型，具体可以是线性模型；例如，在实际应用中，风险评估方一侧训练的机器学习模型，可以是线性逻辑回归模型。

其中，风险评估方在基于构建的训练样本集以及与各训练样本对应的标签，对线性模型进行训练的过程，即为将各数据提供方上传的评估结果作为自变量，将对应用户标签作为因变量，代入到线性模型的表达式中进行线性拟合，求解出与各自变量对应的系数的过程，其具体的实施过程，在本说明书不再进行详述，本领域技术人员在将本说明书中的技术方案付诸实现时，
15 20 可以参考相关技术中的记载。

在本说明书中，当风险评估方通过以上训练过程，训练出与训练样本中的各变量（即各数据提供方上传的评估结果）对应的系数后，此时模型训练完毕。

当风险评估方通过使用各个数据提供方上传的对目标用户的初步评估结果，完成模型的训练后，还可以基于各个数据提供方对训练好的模型的贡献度为各个数据提供方分配一定数额的虚拟资源；其中，为各数据提供方分配的虚拟资源的数量，可以与各个数据提供方的权重值（即系数）成正比；
25

在示出的一种实施方式中，为各数据提供方分配的虚拟资源，具体可以是风险评估方向各数据提供方发放的用户数据使用资金。在这种情况下，上述风险评估方可以基于各个数据提供方对训练好的模型的贡献度，对能够向各个数据提供方发放的用户数据使用资金，进行利益分配。

- 5 其中，在示出的一种实现方式中，各个数据提供方对训练好的模型的贡献度，具体可以用训练出的与训练样本中的各变量对应的系统来表征。在这种情况下，当风险评估方通过以上示出的模型训练过程，训练出了与训练样本中的各变量对应的系数后，可以将训练出的各变量对应的系数作为各个数据提供方的贡献度，然后基于各变量对应的系数大小为各数据提供方进行利
10 益分配；

例如，风险评估方可以将各变量的系数作为对模型的贡献度，换算出对应的分配比例，然后按照换算出的分配比例，将能够用于向各数据提供方分配的用户数据使用资金的总额，分别分配给各个数据提供方。在这种情况下，对模型的贡献较高的数据提供方，将可以拿到更多的数据使用资金。通过这
15 种方式，可以使高质量的数据提供方能够获益更多，从而迫使各个数据提供方能够不断去提高各自的数据质量。

当然，在实际应用中，在初始状态下，如果风险评估方无法收集到足够的训练样本完成模型的训练，那么可以为该模型中的各变量设置一个初始的系数，并利用该初始的系数来表征各数据提供方对模型的初始贡献度。

- 20 其中，上述初始贡献度的设置策略，在本说明书中不进行特别限定，本领域技术人员在将本说明书的技术方案付诸实现时，可以基于实际的需求来进行设置；

例如，在一种实现方式中，当风险评估方的服务端在设备冷启动（即设备首次开机运行）时，可以采用加权平均的方式，为模型中的各变量设置一个相同的初始的系数，并基于该初始的系数作为各数据提供方的初始贡献度，
25 来为各数据提供方平均分配虚拟资源。

例如，以风险评估方为各数据提供方分配的虚拟资源，为风险评估方向

各数据提供方发放的用户数据使用资金为例，在这种情况下，上述风险评估方可以基于各数据提供方的初始贡献度，对能够用于向各数据提供方分配的用户数据使用资金的总额，平均分配给各个数据提供方。

可见，通过这种方式，由于训练好的模型中的各变量的系数能够真实的
5 反映各数据提供方对训练越好的模型的贡献度，因此基于该贡献度向各数据提供方分配虚拟资源，能够做到虚拟资源的合理分配。

请继续参见图 2，当风险评估方完成模型的训练后，后续可以使用训练好的模型对某一个目标用户来进行风险评估。

其中，上述目标用户，具体可以包括上述风险评估方需要进行风险评估
10 的用户；例如，以信贷发放的业务场景为例，上述风险评估方，具体可以是贷款发放的一方；而上述目标用户具体可以是指发起了一笔贷款申请，需要由风险评估方进行风险评估和决策是否发放贷款的用户。

当该多个数据提供方在收到该目标用户的用户 ID 后，可以基于该用户 ID 查找已经利用自己的用户评估模型评估完成的评估结果，然后将评估结果
15 上传给风险评估方。

而风险评估方在收到各数据提供方针对该目标用户的评估结果后，可以将各数据提供方上传的评估结果作为训练数据，为该目标用户创建出一条对应的预测样本，然后将该预测样本输入至训练好的模型中进行预测计算，得到该用户的最终的评估结果，并基于该最终的评估结果来进行相应的业务决策。
20

例如，仍以信贷发放的业务场景为例，在这种场景下，上述最终的评估结果仍然可以是风险评分；风险评估方在基于该风险评分来决策是否向该用户发放贷款时，可以通过将该风险评分与预设的风险阈值进行比较来实现；一方面，如果风险评分高于或者等于上述风险阈值，则表明上述目标用户为
25 风险用户，此时可以为该用户标定一个用于指示该用户为风险用户的用户标签，并终止该用户发起的贷款申请。

另一方面，如果上述风险评分低于上述风险阈值，则表明上述目标用户

为低风险用户，可以为该用户标定一个用于指示该用户为低风险用户的用户标签，并正常响应该用户发起的贷款申请，向该用户发放贷款。

当然，为该目标用户发放贷款后，后续还可以基于该目标用户最终是否违约还款，对已经为该用户标定的用户标签进行维护和更新；比如，假设目
5 标用户被标定为非风险用户，最终向该用户发放贷款后，如果该用户出现违约还款，那么可以立即对已经标定的用户标签进行更新，将该用户重新标定为风险用户。

最后，需要补充说明的是，在本说明书中，与上述风险评估方具有合作关系的数据提供方，可以是动态变化的。

10 作为数据建模方，可以支持任何一个数据提供方随时退出“数据共享”，也可以支持任何一个数据提供方随时加入“数据共享”。即对于风险评估方而言，可以不需要关注与其具有合作关系的数据提供方的数量以及类型，仅需要对当前与自身保持合作关系的各个数据提供方上传的对上述目标用户的初步的评估结果进行加权计算即可。可见，在本说明书中，上述风险评估方，
15 可以灵活的与不同类型的数据提供方进行对接。

通过以上各实施例可知，一方面，由于风险评估方在基于各数据提供方维护的用户数据训练模型时，数据提供方仅需要向风险评估方风险评估方传输对该用户进行初步评估后得到的评估结果，因此对于数据提供方而言，不再需要将本地维护的原始用户数据传输至风险评估方，可以显著降低用户隐
20 私泄露的发生风险；

另一方面，由于训练好的模型中的各变量的系数能够真实的反映各数据提供方对训练越好的模型的贡献度，因此基于该贡献度向各数据提供方分配虚拟资源，能够做到虚拟资源的合理分配。

请参见图 3，与上述方法实施例相对应，本说明书还提供了一种模型建
25 立方法，应用于风险评估方服务端，执行以下步骤：

步骤 302，接收多个数据提供方上传的若干用户的评估结果；其中，所述评估结果为各数据提供方基于自己的评估模型对用户分别进行评估后得到；

步骤 304, 将各数据提供方上传的评估结果作为训练数据, 构建出若干训练样本, 每一条训练样本包含同一个用户在所述各数据提供方的评估结果; 其中, 所述训练样本根据用户对业务的实际执行情况被标定了标签;

5 步骤 306, 基于所述若干训练样本以及各个训练样本的标签对模型进行训练, 得到训练好的模型。

在本实施例中, 上述训练好的模型可以为线性模型。例如, 在实际应用中, 可以是线性逻辑回归模型。上述评估模型可以为用户风险评估模型; 上述评估结果可以为风险评分(或信用评分); 上述标签指示用户是否为风险用户。

10 其中, 以上各步骤的具体实施细节, 在本实施例中不再进行赘述, 本领域技术人员可以参考之前实施例的记载。

请参见图 4, 与上述方法实施例相对应, 本说明书还提供了一种数据预测的方法, 应用于风险评估方服务端, 执行以下步骤:

步骤 402, 接收多个数据提供方上传的若干用户的评估结果; 其中, 所述评估结果为各数据提供方基于自己的评估模型对用户分别进行评估后得到;

步骤 404, 将各数据提供方上传的评估结果作为训练数据, 构建出若干训练样本, 每一条训练样本包含同一个用户在所述各数据提供方的评估结果; 其中, 所述训练样本根据用户对业务的实际执行情况被标定了标签;

20 步骤 406, 基于所述若干训练样本以及各个训练样本的标签对模型进行训练, 得到训练好的模型;

步骤 408, 接收多个数据提供方上传的针对某一个用户的评估结果, 将所述评估结果输入所述训练好的模型中, 得到该用户的最终的评估结果。

其中, 以上各步骤的具体实施细节, 在本实施例中不再进行赘述, 本领域技术人员可以参考之前实施例的记载。与上述方法实施例相对应, 本说明书还提供了一种虚拟资源分配装置的实施例。

本说明书的虚拟资源分配装置的实施例可以应用在电子设备上。装置实施例可以通过软件实现, 也可以通过硬件或者软硬件结合的方式实现。以软

件实现为例，作为一个逻辑意义上的装置，是通过其所在电子设备的处理器将非易失性存储器中对应的计算机程序指令读取到内存中运行形成的。从硬件层面而言，如图 5 所示，为本说明书的虚拟资源分配装置所在电子设备的一种硬件结构图，除了图 5 所示的处理器、内存、网络接口、以及非易失性存储器之外，实施例 5 中装置所在的电子设备通常根据该电子设备的实际功能，还可以包括其他硬件，对此不再赘述。

图 6 是本说明书一示例性实施例示出的一种虚拟资源分配装置的框图。

请参考图 6，所述虚拟资源分配装置 60 可以应用在前述图 6 所示的电子设备中，包括有：接收模块 601、训练模块 602、分配模块 603。

10 其中，接收模块 601，接收多个数据提供方上传的若干用户的评估结果；其中，所述评估结果为各数据提供方基于自己的评估模型对用户分别进行评估后得到；

训练模块 602，将各数据提供方上传的评估结果作为训练数据，构建出若干训练样本，每一条训练样本包含同一个用户在所述各数据提供方的评估结果；其中，所述训练样本根据用户对业务的实际执行情况被标定了标签；

分配模块 603，基于所述若干训练样本以及各个训练样本的标签对模型进行训练，将训练好的模型中的各变量的系数作为各数据提供方贡献度，基于各数据提供方的贡献度为各数据提供方分配虚拟资源。

在本实施例中，所述训练好的模型为线性模型。

20 在本实施例中，为各数据提供方分配的所述虚拟资源的数量，与各数据提供方的贡献度成正比。

在本实施例中，所述装置还包括：

评估模块 604（图 6 中未示出），接收多个数据提供方上传的针对某一个用户的评估结果，将所述评估结果输入所述训练好的模型中，得到该用户 25 的最终评估结果。

在本实施例中，所述虚拟资源为向各数据提供方发放的用户数据使用资金。

在本实施例中，所述评估模型为用户风险评估模型；所述评估结果为风险评分；所述标签指示用户是否为风险用户。

上述装置中各个模块的功能和作用的实现过程具体详见上述方法中对应步骤的实现过程，在此不再赘述。

- 5 对于装置实施例而言，由于其基本对应于方法实施例，所以相关之处参见方法实施例的部分说明即可。以上所描述的装置实施例仅仅是示意性的，其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或
- 10 者全部模块来实现本说明书方案的目的。本领域普通技术人员在不付出创造性劳动的情况下，即可以理解并实施。

- 上述实施例阐明的系统、装置、模块或单元，具体可以由计算机芯片或实体实现，或者由具有某种功能的产品来实现。一种典型的实现设备为计算机，计算机的具体形式可以是个人计算机、膝上型计算机、蜂窝电话、相机
- 15 电话、智能电话、个人数字助理、媒体播放器、导航设备、电子邮件收发设备、游戏控制台、平板计算机、可穿戴设备或者这些设备中的任意几种设备的组合。

与上述方法实施例相对应，本说明书还提供一种虚拟资源分配系统的实施例。

- 20 该虚拟资源分配系统，可以包括多个数据提供方服务端和风险评估方服务端。

 其中，多个数据提供方服务端，向风险评估方服务端上传若干用户的评估结果；其中，所述评估结果为各数据提供方基于自己的评估模型对用户分别进行评估后得到；

- 25 风险评估方服务端，将各数据提供方上传的评估结果作为训练数据，构建出若干训练样本，每一条训练样本包含同一个用户在所述各数据提供方的评估结果；其中，所述训练样本根据用户对业务的实际执行情况被标定了标

签;以及,基于所述若干训练样本以及各个训练样本的标签对模型进行训练,将训练好的模型中的各变量的系数作为各数据提供方贡献度,基于各数据提供方的贡献度为各数据提供方分配虚拟资源。

与上述方法实施例相对应,本说明书还提供了一种电子设备的实施例。

- 5 该电子设备包括:处理器以及用于存储机器可执行指令的存储器;其中,处理器和存储器通常通过内部总线相互连接。在其他可能的实现方式中,所述设备还可能包括外部接口,以能够与其他设备或者部件进行通信。

在本实施例中,通过读取并执行所述存储器存储的与上述虚拟资源分配的控制逻辑对应的机器可执行指令,所述处理器被促使:

- 10 接收多个数据提供方上传的若干用户的评估结果;其中,所述评估结果为各数据提供方基于自己的评估模型对用户分别进行评估后得到;

将各数据提供方上传的评估结果作为训练数据,构建出若干训练样本,每一条训练样本包含同一个用户在所述各数据提供方的评估结果;其中,所述训练样本根据用户对业务的实际执行情况被标定了标签;

- 15 基于所述若干训练样本以及各个训练样本的标签对模型进行训练,将训练好的模型中的各变量的系数作为各数据提供方贡献度,基于各数据提供方的贡献度为各数据提供方分配虚拟资源。

在本实施例中,所述训练好的模型为线性模型。

- 20 在本实施例中,为各数据提供方分配的所述虚拟资源的数量,与各数据提供方的贡献度成正比。

在本实施例中,通过读取并执行所述存储器存储的与上述虚拟资源分配的控制逻辑对应的机器可执行指令,所述处理器被促使:

接收多个数据提供方上传的针对某一个用户的评估结果,将所述评估结果输入所述训练好的模型中,得到该用户的最终的评估结果。

- 25 在本实施例中,所述虚拟资源为向各数据提供方发放的用户数据使用资金。

在本实施例中,所述评估模型为用户风险评估模型;所述评估结果为风

险评分；所述标签指示用户是否为风险用户。

本领域技术人员在考虑说明书及实践这里公开的发明后，将容易想到本说明书的其它实施方案。本说明书旨在涵盖本说明书的任何变型、用途或者适应性变化，这些变型、用途或者适应性变化遵循本说明书的一般性原理并包括本说明书未公开的本技术领域中的公知常识或惯用技术手段。说明书和
5 实施例仅被视为示例性的，本说明书的真正范围和精神由下面的权利要求指出。

应当理解的是，本说明书并不局限于上面已经描述并在附图中示出的精确结构，并且可以在不脱离其范围进行各种修改和改变。本说明书的范围仅
10 由所附的权利要求来限制。

上述对本说明书特定实施例进行了描述。其它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的特定顺序或者连续顺序才能实现期望的结果。在
15 某些实施方式中，多任务处理和并行处理也是可以的或者可能是有利的。

以上所述仅为本申请的较佳实施例而已，并不用以限制本申请，凡在本申请的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本申请保护的范围之内。