



(12) 发明专利申请

(10) 申请公布号 CN 116805251 A

(43) 申请公布日 2023. 09. 26

(21) 申请号 202210251940.2

(22) 申请日 2022.03.15

(71) 申请人 腾讯科技(深圳)有限公司

地址 518057 广东省深圳市南山区高新区
科技中一路腾讯大厦35层

(72) 发明人 田红

(74) 专利代理机构 华进联合专利商标代理有限
公司 44224

专利代理师 杨欢

(51) Int. Cl.

G06Q 30/0202 (2023.01)

G06F 18/2431 (2023.01)

G06F 18/214 (2023.01)

A63F 13/77 (2014.01)

A63F 13/792 (2014.01)

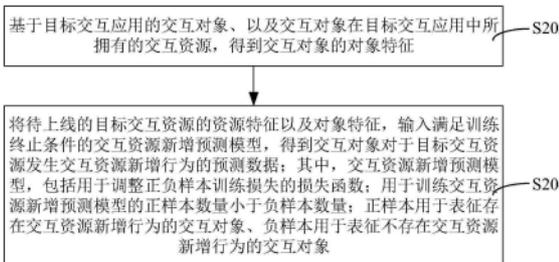
权利要求书2页 说明书20页 附图9页

(54) 发明名称

数据预测方法、装置、计算机设备和存储介
质

(57) 摘要

本申请涉及一种数据预测方法、装置、计算机设备和存储介质。所述方法包括：基于目标交互应用的交互对象、以及交互对象在目标交互应用中所拥有的交互资源，得到交互对象的对象特征；将待上线的目标交互资源的资源特征以及对象特征，输入满足训练终止条件的交互资源新增预测模型，得到交互对象对于目标交互资源发生交互资源新增行为的预测数据；其中，交互资源新增预测模型，包括用于调整正负样本训练权重的损失函数；用于训练交互资源新增预测模型的正样本数量小于负样本数量；正样本用于表征存在交互资源新增行为的交互对象、负样本用于表征不存在交互资源新增行为的交互对象。采用本申请实施例的方法能够提高数据预测的准确性。



1. 一种数据预测方法,其特征在于,所述方法包括:

基于目标交互应用的交互对象、以及所述交互对象在所述目标交互应用中所拥有的交互资源,得到所述交互对象的对象特征;

将待上线的目标交互资源的资源特征以及所述对象特征,输入满足训练终止条件的交互资源新增预测模型,得到所述交互对象对于所述目标交互资源发生交互资源新增行为的预测数据;

其中,所述交互资源新增预测模型,包括用于调整正负样本训练权重的损失函数;用于训练所述交互资源新增预测模型的正样本数量小于负样本数量;所述正样本用于表征存在交互资源新增行为的交互对象、所述负样本用于表征不存在交互资源新增行为的交互对象。

2. 根据权利要求1所述的方法,其特征在于,所述基于目标交互应用的交互对象、以及所述交互对象在所述目标交互应用中所拥有的交互资源,得到所述交互对象的对象特征,包括:

确定待上线的所述目标交互资源在所述目标交互应用的上线时间;

获取所述上线时间之前的预设时长内的交互对象数据,所述交互对象数据包括:交互对象在所述目标交互应用中所拥有的交互资源的资源信息、以及所述交互对象的对象信息;

对所述交互对象数据进行特征提取,得到所述交互对象的对象特征。

3. 根据权利要求2所述的方法,其特征在于,所述交互对象数据为校验通过的数据;所述对象信息包括:交互对象基础信息、交互对象社群信息以及交互对象行为信息;

所述交互对象数据的校验方式,包括:

校验所述交互对象基础信息与所述交互对象在所述目标交互应用中所拥有的交互资源之间的关系,保留符合预设关系的交互对象基础信息;

校验所述交互对象基础信息之间的关系,保留存在交集的交互对象基础信息;

校验并保留表达符合表达筛选条件的交互对象数据;

确定所述交互对象数据与待上线的所述目标交互资源之间的相关性,筛选出符合预设相关性的交互对象数据并保留。

4. 根据权利要求1所述的方法,其特征在于,所述正负样本训练权重,包括:正负样本在损失函数中的权重大小以及训练过程中难易分类样本的损失权重;所述正样本的权重大于所述负样本的权重,难分类样本的损失权重大于易分类样本的损失权重。

5. 根据权利要求1所述的方法,其特征在于,所述交互资源新增预测模型包括:分类预测子模型和回归预测子模型;

所述将待上线的目标交互资源的资源特征以及所述对象特征,输入满足训练终止条件的交互资源新增预测模型,得到所述交互对象对于所述目标交互资源发生交互资源新增行为的预测数据,包括:

根据待上线的所述目标交互资源的资源特征以及所述对象特征,通过所述分类预测子模型,预测出对于所述目标交互资源发生交互资源新增行为的新增交互对象;

从所述对象特征中筛选确定所述新增交互对象的新增对象特征;

基于所述新增对象特征以及所述资源特征,通过所述回归预测子模型,得到所述交互

对象对于所述目标交互资源发生交互资源新增行为的预测数据。

6. 根据权利要求5所述的方法,其特征在于,所述方法还包括:

在所述分类预测子模型满足训练结束条件时,根据所述分类预测子模型中的各叶子节点所对应的增益情况,确定影响所述交互资源新增行为的目标对象特征;

所述目标对象特征用于调整所述目标交互资源所对应的资源信息。

7. 根据权利要求6所述的方法,其特征在于,所述根据所述分类预测子模型中的各叶子节点所对应的增益情况,确定影响所述交互资源新增行为的目标对象特征,包括:

确定所述分类预测子模型中的各叶子节点的分割特征;

统计各所述叶子节点的分割特征,得到各所述分割特征的统计结果;

根据各所述分割特征的统计结果和预设的对象特征个数,确定影响所述交互资源新增行为的目标对象特征。

8. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

在待上线的所述目标交互资源上线后,确定所述交互对象对于所述目标交互资源发生交互资源新增行为的实际数据;

若基于所述预测数据与所述实际数据确定满足模型更新条件,更新所述交互资源新增预测模型,得到更新后的交互资源新增预测模型。

9. 根据权利要求8所述的方法,其特征在于,所述更新后的交互资源新增预测模型包括:更新后的分类预测子模型;所述方法还包括:

根据所述更新后的分类预测子模型中的各叶子节点所对应的增益情况,更新影响所述交互资源新增行为的目标对象特征。

10. 一种数据预测装置,其特征在於,所述装置包括:

获取模块,用于基于目标交互应用的交互对象、以及所述交互对象在所述目标交互应用中所拥有的交互资源,得到所述交互对象的对象特征;

预测模块,用于将待上线的目标交互资源的资源特征以及所述对象特征,输入满足训练终止条件的交互资源新增预测模型,得到所述交互对象对于所述目标交互资源发生交互资源新增行为的预测数据;其中,所述交互资源新增预测模型,包括用于调整正负样本训练权重的损失函数;用于训练所述交互资源新增预测模型的正样本数量小于负样本数量;所述正样本用于表征存在交互资源新增行为的交互对象、所述负样本用于表征不存在交互资源新增行为的交互对象。

11. 一种计算机设备,包括存储器和处理器,所述存储器存储有计算机程序,其特征在於,所述处理器执行所述计算机程序时实现权利要求1至9中任一项所述的方法的步骤。

12. 一种计算机可读存储介质,其上存储有计算机程序,其特征在於,所述计算机程序被处理器执行时实现权利要求1至9中任一项所述的方法的步骤。

13. 一种计算机程序产品,包括计算机程序,其特征在於,该计算机程序被处理器执行时实现权利要求1至9中任一项所述的方法的步骤。

数据预测方法、装置、计算机设备和存储介质

技术领域

[0001] 本申请涉及人工智能技术领域,特别是涉及一种数据预测方法、装置、计算机设备、计算机可读存储介质和计算机程序产品。

背景技术

[0002] 互联网用户之间会不断的产生数据交互,例如,交互对象可以在各种交互应用中进行账户创建、管理、升级、购买等行为。为了提高交互应用收益,策划者通常会在活动节日之前设计交互资源产品内容并发布,以供交互对象购买。传统方式中是根据交互应用的策划者的业务经验和部分外网交互对象的反馈数据进行推测,来设计交互资源产品中包含的道具内容,依赖历史销量预测新的交互资源产品的销量。

[0003] 然而,这种方式非常依赖于交互应用的策划者的业务经验和交互对象反馈数据的覆盖度,导致交互资源产品的设计和销量预测受到不确定因素的干扰,使得数据预测的准确性不高。

发明内容

[0004] 基于此,有必要针对上述技术问题,提供一种能够提高数据预测的准确性的数据预测方法、装置、计算机设备、计算机可读存储介质和计算机程序产品。

[0005] 第一方面,本申请提供了一种数据预测方法。所述方法包括:

[0006] 基于目标交互应用的交互对象、以及所述交互对象在所述目标交互应用中所拥有的交互资源,得到所述交互对象的对象特征;

[0007] 将待上线的目标交互资源的资源特征以及所述对象特征,输入满足训练终止条件的交互资源新增预测模型,得到所述交互对象对于所述目标交互资源发生交互资源新增行为的预测数据;

[0008] 其中,所述交互资源新增预测模型,包括用于调整正负样本训练权重的损失函数;用于训练所述交互资源新增预测模型的正样本数量小于负样本数量;所述正样本用于表征存在交互资源新增行为的交互对象、所述负样本用于表征不存在交互资源新增行为的交互对象。

[0009] 第二方面,本申请还提供了一种数据预测装置。所述装置包括:

[0010] 获取模块,用于基于目标交互应用的交互对象、以及所述交互对象在所述目标交互应用中所拥有的交互资源,得到所述交互对象的对象特征;

[0011] 预测模块,用于将待上线的目标交互资源的资源特征以及所述对象特征,输入满足训练终止条件的交互资源新增预测模型,得到所述交互对象对于所述目标交互资源发生交互资源新增行为的预测数据;其中,所述交互资源新增预测模型,包括用于调整正负样本训练权重的损失函数;用于训练所述交互资源新增预测模型的正样本数量小于负样本数量;所述正样本用于表征存在交互资源新增行为的交互对象、所述负样本用于表征不存在交互资源新增行为的交互对象。

[0012] 第三方面,本申请还提供了一种计算机设备。所述计算机设备包括存储器和处理器,所述存储器存储有计算机程序,所述处理器执行所述计算机程序时实现上述的方法的步骤。

[0013] 第四方面,本申请还提供了一种计算机可读存储介质。所述计算机可读存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现上述的方法的步骤。

[0014] 第五方面,本申请还提供了一种计算机程序产品。所述计算机程序产品,包括计算机程序,该计算机程序被处理器执行时实现上述的方法的步骤。

[0015] 上述数据预测方法、装置、计算机设备、计算机可读存储介质和计算机程序产品,通过基于目标交互应用的交互对象、以及交互对象在目标交互应用中所拥有的交互资源,得到交互对象的对象特征,可以提高获得的对象特征所对应的数据覆盖度。进而,通过将待上线的目标交互资源的资源特征以及对象特征,输入满足训练终止条件的交互资源新增预测模型,得到交互对象对于目标交互资源发生交互资源新增行为的预测数据;其中,交互资源新增预测模型,包括用于调整正负样本训练权重的损失函数;用于训练交互资源新增预测模型的正样本数量小于负样本数量;正样本用于表征存在交互资源新增行为的交互对象、负样本用于表征不存在交互资源新增行为的交互对象。通过设置上述的损失函数,能够平衡正负样本的不均衡情况,而且,可以使得训练时算法更关注于难分类样本,从而提高训练得到的交互资源新增预测模型的预测精度,提高数据预测的准确性。

附图说明

- [0016] 图1为一个实施例中数据预测方法的应用环境图;
- [0017] 图2为一个实施例中数据预测方法的流程示意图;
- [0018] 图3为一个实施例中数据预测方法的架构流程示意图;
- [0019] 图4为一个实施例中数据预测方法的系统流程示意图;
- [0020] 图5为一个实施例中交互资源新增预测模型的训练流程示意图;
- [0021] 图6为一个实施例中影响交互资源新增行为的对象特征的示意图;
- [0022] 图7为一个具体实施例中目标交互资源的示意图;
- [0023] 图8为一个具体实施例中交互对象数据的数据内容示意图;
- [0024] 图9为一个具体实施例中影响交互资源新增行为的目标对象特征的示意图;
- [0025] 图10为一个实施例中数据预测装置的结构框图;
- [0026] 图11为一个实施例中计算机设备的内部结构图;
- [0027] 图12为另一个实施例中计算机设备的内部结构图。

具体实施方式

[0028] 为了使本申请的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本申请进行进一步详细说明。应当理解,此处描述的具体实施例仅仅用以解释本申请,并不用于限定本申请。

[0029] 需要首先说明的是,本申请所涉及的目标交互应用、交互对象、交互资源等相关信息和数据,均为经过各方充分授权的信息和数据,且相关信息与数据的收集、使用和处理需要遵守相关国家和地区的相关法律法规和标准。

[0030] 一个实施例中,本申请提供的数据预测方法,可以应用于如图1所示的应用环境中。该应用环境可以同时涉及终端102和服务器104,在另外一些实施例中,还可以同时涉及终端106。终端102、终端106通过网络与服务器104进行通信。数据存储系统可以存储服务器104需要处理的数据。数据存储系统可以集成在服务器104上,也可以放在云上或其他服务器上。

[0031] 具体地,在终端102和终端106中内置有各交互应用,每个交互对象可以通过其持有的终端102和/或终端106在任意一个交互应用中实现与其他交互对象的数据交互。其中,若同一用户同时持有终端102和终端106、或持有更多终端,该用户在所持有的终端102和终端106中可以是登录相同的账号,也可以是登录不同的账号,以通过该账号与其他用户进行数据交互。

[0032] 具体地,可以是在服务器104中训练得到交互资源新增预测模型,而后,服务器104确定目标交互应用,基于目标交互应用的交互对象、以及交互对象在目标交互应用中所拥有的交互资源,得到交互对象的对象特征;服务器104将待上线的目标交互资源的资源特征以及对象特征,输入满足训练终止条件的交互资源新增预测模型,得到交互对象对于目标交互资源发生交互资源新增行为的预测数据;其中,交互资源新增预测模型,包括用于调整正负样本训练权重的损失函数;用于训练交互资源新增预测模型的正样本数量小于负样本数量;正样本用于表征存在交互资源新增行为的交互对象、负样本用于表征不存在交互资源新增行为的交互对象。

[0033] 一个实施例中,在服务器104中训练得到交互资源新增预测模型之后,还可以是服务器104将交互资源新增预测模型发送至终端102和终端106。在终端102和终端106中,通过基于目标交互应用的交互对象、以及交互对象在目标交互应用中所拥有的交互资源,得到交互对象的对象特征;并将待上线的目标交互资源的资源特征以及对象特征,输入交互资源新增预测模型,得到交互对象对于目标交互资源发生交互资源新增行为的预测数据。

[0034] 一个实施例中,在终端102或终端106的计算能力满足要求的情况下,该应用环境可以仅涉及终端102或终端106。其中,在终端102或终端106中训练得到交互资源新增预测模型,并通过基于目标交互应用的交互对象、以及交互对象在目标交互应用中所拥有的交互资源,得到交互对象的对象特征;将待上线的目标交互资源的资源特征以及对象特征,输入交互资源新增预测模型,得到交互对象对于目标交互资源发生交互资源新增行为的预测数据。

[0035] 其中,终端102、终端106可以但不限于各种台式计算机、笔记本电脑、智能手机、平板电脑、物联网设备和便携式可穿戴设备,物联网设备可为智能电视、智能车载设备等。便携式可穿戴设备可为智能手表、智能手环等。服务器104可以用独立的服务器或者是多个服务器组成的服务器集群来实现。

[0036] 在一个实施例中,如图2所示,提供了一种数据预测方法,以该方法应用于图1中的服务器104为例进行说明,包括以下步骤:

[0037] 步骤S202,基于目标交互应用的交互对象、以及交互对象在目标交互应用中所拥有的交互资源,得到交互对象的对象特征。

[0038] 交互应用是指能够实现数据交互的应用程序或平台。其中,交互是指行为的交互,本质上是数据交互。例如,交互应用可以是电子游戏、购物软件、音乐软件等,数据交互可以

是联机游戏、购物分享、音乐分享等。交互对象是指交互应用的使用者,具体可以是用户,例如,在电子游戏中的交互对象为游戏玩家、在购物软件中的交互对象为购物用户、在音乐软件中的交互对象为音乐用户等,一般地,交互对象可以通过在交互应用中注册并登录账号,以实现使用该交互应用,并在该交互应用中与其他交互对象进行数据交互。交互资源是指交互对象在交互应用中所拥有的资源,该资源可以是虚拟资源或实体衍生产品资源,例如,游戏玩家在电子游戏中拥有的道具和礼包或周边等、购物用户在购物软件中拥有的购物券等、音乐玩家在音乐软件中拥有的电子专辑或实体专辑等。

[0039] 目标交互应用是指本实施例中需要进行数据预测的某一交互应用,例如,若需要对电子游戏中的游戏礼包的销量数据进行预测,则目标交互应用为电子游戏,目标交互应用的交互对象即为游戏玩家,目标对象在目标交互应用中所拥有的交互资源可以是游戏玩家所拥有的游戏礼包。

[0040] 在一个实施例中,确定目标交互应用,以及目标交互应用的交互对象。其中,目标交互应用可以是任意一个交互应用,目标交互应用的交互对象可以通过交互对象在该目标交互应用中所注册和管理的账号确定,一个交互对象可以在该目标交互应用中注册和管理一个或多个账号。

[0041] 基于目标交互应用的交互对象、以及交互对象在目标交互应用中所拥有的交互资源,得到交互对象的对象特征。其中,对象特征用于表征该交互对象的相关信息,以便进行后续处理。具体地,可以通过获取目标交互应用的交互对象的交互对象数据,对交互对象数据进行特征提取,以得到对象特征。

[0042] 需要说明的是,若一个交互对象在目标交互应用中注册和管理多个账号,则确定交互对象在目标交互应用中所拥有的交互资源时,包括该交互对象在所有账号下的所有交互资源。

[0043] 步骤S204,将待上线的目标交互资源的资源特征以及对象特征,输入满足训练终止条件的交互资源新增预测模型,得到交互对象对于目标交互资源发生交互资源新增行为的预测数据;其中,交互资源新增预测模型,包括用于调整正负样本训练权重的损失函数;用于训练交互资源新增预测模型的正样本数量小于负样本数量;正样本用于表征存在交互资源新增行为的交互对象、负样本用于表征不存在交互资源新增行为的交互对象。

[0044] 目标交互资源是指需要进行数据预测的交互资源。其中,目标交互资源是待上线的交互资源,待上线即还未在目标交互应用中上线,也即此时的交互对象还未能拥有该目标交互资源。一般地,目标交互应用的策划者会设置在特定时间上线目标交互资源,特定时间可以是设定的活动节日时间,具体可以根据实际技术需要进行设置。

[0045] 目标交互资源在上线后,目标交互应用的交互对象可以选择是否购买该目标交互资源,交互资源新增行为即指交互对象选择购买该目标交互资源的行为。交互资源新增预测模型是预先训练的模型,用于预测交互对象的交互资源新增行为,得到一系列的预测数据。其中,交互资源新增预测模型的模型类型和模型结构可以根据实际技术需要进行设置。预测数据具体可以包括:预测会发生交互资源新增行为的交互对象的预测对象数目,以及该目标交互资源所对应的交互资源新增行为的预测新增数目。

[0046] 在一个实施例中,结合实际情况来看,交互对象在使用目标交互应用时,可能仅仅是为了体验该目标交互应用,而在使用过程中不会发生任何的交互资源新增行为。因此,在

交互资源新增预测模型的训练过程中,根据目标交互应用的样本交互对象确定训练样本时,用于训练的正样本的数量往往小于或远小于负样本的数量。其中,正样本用于表征存在交互资源新增行为的交互对象、负样本用于表征不存在交互资源新增行为的交互对象。

[0047] 为了平衡正负样本的不均衡情况,交互资源新增预测模型的损失函数选择可以用于调整正负样本训练权重的损失函数。一个实施例中,正负样本训练权重,包括:正负样本在损失函数中的权重大小以及训练过程中难易分类样本的损失权重。其中,难分类样本是指该样本的特征非常接近于正样本的特征,但其并非正样本。在本实施例的损失函数中,正样本的权重大于负样本的权重,难分类样本的损失权重大于易分类样本的损失权重,以使得该模型能够平衡正负样本,并聚焦于难分类样本,提高数据的预测精度。

[0048] 具体地,待上线的目标交互资源的资源特征可以根据目标交互资源的资源信息进行特征提取得到。在确定待上线的目标交互资源的资源特征后,将待上线的目标交互资源的资源特征以及对象特征,输入满足训练终止条件的交互资源新增预测模型,即可得到交互对象对于目标交互资源发生交互资源新增行为的预测数据。

[0049] 上述数据预测方法中,通过基于目标交互应用的交互对象、以及交互对象在目标交互应用中所拥有的交互资源,得到交互对象的对象特征,可以提高获得的对象特征所对应的数据覆盖度。进而,通过将待上线的目标交互资源的资源特征以及对象特征,输入满足训练终止条件的交互资源新增预测模型,得到交互对象对于目标交互资源发生交互资源新增行为的预测数据;其中,交互资源新增预测模型,包括用于调整正负样本训练权重的损失函数;用于训练交互资源新增预测模型的正样本数量小于负样本数量;正样本用于表征存在交互资源新增行为的交互对象、负样本用于表征不存在交互资源新增行为的交互对象。通过设置上述的损失函数,能够平衡正负样本的不均衡情况,而且,可以使得训练时算法更关注于难分类样本,从而提高训练得到的交互资源新增预测模型的预测精度,提高数据预测的准确性。

[0050] 在一个实施例中,基于目标交互应用的交互对象、以及交互对象在目标交互应用中所拥有的交互资源,得到交互对象的对象特征,可以包括以下步骤:

[0051] 步骤S302,确定待上线的目标交互资源在目标交互应用的上线时间。

[0052] 目标交互应用的策划者会设置在特定时间上线目标交互资源,特定时间可以是设定的活动节日时间,具体可以根据实际技术需要进行设置。具体地,确定待上线的目标交互资源在目标交互应用的上线时间。其中,为了提高预测数据的准确性,该上线时间可以具体到小时时间点。

[0053] 步骤S304,获取上线时间之前的预设时长内的交互对象数据,交互对象数据包括:交互对象在目标交互应用中所拥有的交互资源的资源信息、以及交互对象的对象信息。

[0054] 预设时长是指预先设定的、用于确定数据获取时间段的时长,可以根据实际技术需要进行设置,一个实施例中,预设时长可以设置为两个月。具体地,获取上线时间之前的预设时长内的交互对象数据,即获取上线时间之前的两个月内的各交互对象的交互对象数据。具体地,交互对象数据包括:交互对象在目标交互应用中所拥有的交互资源的资源信息、以及交互对象的对象信息。

[0055] 其中,可以通过腾讯分布式数据仓库(Tencent distributed Data Warehouse, TDW)平台中的相关数据库获取上述的交互对象数据。需要说明的是,此处获取的交互对象

数据,是预先经过各交互对象授权、或经过各方充分授权的数据。

[0056] 在一个实施例中,由于交互对象可以在目标交互应用内进行多种类型的操作,为了提高交互对象数据的数据覆盖度,获取的每一交互对象的交互对象的对象信息可以包括以下几种:交互对象基础信息、交互对象社群信息、以及交互对象交易信息。

[0057] 交互对象基础信息可以描述交互对象的基础属性,交互对象社群信息用于描述交互对象的活跃程度,交互对象交易信息用于表征交互对象的购买能力,交互对象在目标交互应用中所拥有的交互资源的资源信息用于体现交互对象所拥有的交互资源。

[0058] 为了方便理解,以下实施例以目标交互应用是游戏、交互对象是游戏玩家、交互资源是游戏礼包为例,对上述的交互对象数据进行说明。

[0059] 一个实施例中,交互对象基础信息包括:游戏玩家所持有的账号、账号下的游戏角色信息,具体包括每个账号下的所有游戏角色所对应的角色类别信息、名望值信息、付费角色信息、参与副本情况信息,以及每个账号下的前预定数目名望值角色的角色信息等。

[0060] 一个实施例中,角色类别信息可以包括预定等级的角色数量、预定等级的角色中的预定类型角色数量、小于预定等级的角色数量。例如,预定等级设置为100级、预定类型设置为主C、34C和奶系,上述的角色类别信息为:100级角色数量、100级角色中主C角色数量、100级角色中34C角色数量、100级角色中奶系角色数量、小于100级角色数量。

[0061] 一个实施例中,名望值用于表征游戏角色的强弱。名望值信息可以包括小于第一名望值区间的角色数量、处于第一名望值区间的角色数量、处于第二名望值区间的角色数量、处于第三名望值区间的角色数量、处于第四名望值区间的角色数量、处于第五名望值区间的角色数量、处于第六名望值区间的角色数量、处于第七名望值区间的角色数量、处于第八名望值区间的角色数量、大于第八名望值区间的角色数量。其中,上述各名望值区间所对应的名望值可以是依次增加。例如,第一名望值区间设置为1-1.1万、第二名望值区间设置为1.1-1.2万、第三名望值区间设置为1.2-1.3万、第四名望值区间设置为1.3-1.4万、第五名望值区间设置为1.4-1.5万、第六名望值区间设置为1.5-1.6万、第七名望值区间设置为1.6-1.7万、第八名望值区间设置为1.7-1.8万。上述的名望值信息为:名望值在1万以下的角色数、名望值在1-1.1万之间的角色数、名望值在1.1-1.2万之间的角色数、名望值在1.2-1.3万之间的角色数、名望值在1.3-1.4万之间的角色数、名望值在1.4-1.5万之间的角色数、名望值在1.5-1.6万之间的角色数、名望值在1.6-1.7万之间的角色数、名望值在1.7-1.8万之间的角色数、名望值在1.8万以上的角色数。

[0062] 一个实施例中,付费角色信息用于描述游戏角色的付费情况,付费角色信息可以包括预定时长内有付费的角色数量。例如,预定时长设置为最近30天,上述的付费角色信息为:最近30天有付费的角色数量。

[0063] 一个实施例中,参与副本情况信息用于描述游戏角色是否参与了游戏副本,参与副本情况可以包括预定时长内分别参与过第一副本、第二副本的角色数量。例如,预定时长设置为最近30天,第一副本设置为“希洛克团本”,第二副本设置为“地下城”。上述的参与副本情况信息为:最近30天参与过“希洛克团本”的角色数量、最近30天参与过“地下城”的角色数量。

[0064] 一个实施例中,每个账号下的前预定数目名望值角色的角色信息,角色信息可以包括该角色的角色名望值、预设时长角色名望值提升速度、角色预设穿着排名(无该预设穿

着排名则为0)、角色穿着装备成套数(至少3件装备为成套)、角色金币数量、预设时长参与预定副本的总次数、角色预设时长消费金额(第一类型券)、角色预设时长消费金额(第二类型券)。例如,预定数目可以设置为5,预定时长设置为最近30天,预设穿着为“神话”,预定副本设置为“地下城”,第一类型券为点券,第二类型券为非点券。分别针对名望值排名前5的角色1至角色5,每一角色各自对应的上述的角色信息为:角色名望值、最近30天角色名望值提升速度、角色穿着“神话”排名(无“神话”排名则为0)、角色穿着装备成套数(至少3件装备为成套)、角色金币数量、最近30天参与“地下城”的总次数、角色最近30天消费金额(点券)、角色最近30天消费金额(非点券)。

[0065] 一个实施例中,交互对象社群信息包括游戏玩家活跃信息,可以用第一预定时长的登录游戏天数、第二预定时长的登录游戏天数和第二预定时长的活跃角色数量进行表征。例如,第一预定时长设置为最近90天或最近60天,第二预定时长设置为最近30天,上述的游戏玩家活跃度信息为:最近90天(60天)的登录游戏天数、最近30天的登录游戏天数和最近30天的活跃角色数量。

[0066] 一个实施例中,交互对象交易信息包括游戏玩家付费信息,可以包括:第一时长的累计消费第一类型券金额、第二时长的累计消费第一类型券金额、第二时长的消费第一类型券次数、第二时长的平均第一类型券余额、第一时长的末次付费距离统计日期天数(付费包含第一类型券与第二类型券)、第一时长的累计消费第二类型券金额、第二时长的累计消费第二类型券金额、第二时长的消费第二类型券次数、第二时长的平均第二类型券余额、第二时长的平均金币余额。例如,第一时长设置为最近1年,第二时长设置为最近30天,第一类型券设置为点券,第二类型券设置为非点券。上述的付费信息包括:最近1年的累计消费点券金额、最近30天的累计消费点券金额、最近30天的消费点券次数、最近30天的平均点券余额、最近1年的末次付费距离统计日期天数(付费包含点券与非点券)、最近1年的累计消费非点券金额、最近30天的累计消费非点券金额、最近30天的消费非点券次数、最近30天的平均非点券余额、最近30天的平均金币余额。

[0067] 一个实施例中,交互对象在目标交互应用中所拥有的交互资源的资源信息,可以包括:游戏玩家所拥有的游戏礼包的优惠券数量信息和礼包属性信息,礼包属性信息可以包括礼包内的各道具及其对应的名望值提升信息。例如,道具可以是装扮套装、光环装扮、称号、宠物和宠物装备。上述的礼包属性信息为:装扮套装-名望值提升信息、光环装扮-名望值提升信息、称号-名望值提升信息、宠物-名望值提升信息、宠物装备-名望值提升信息。

[0068] 步骤S306,对交互对象数据进行特征提取,得到交互对象的对象特征。

[0069] 在获取上线时间之前的预设时长内的各交互对象的交互对象数据后,通过对交互对象数据进行特征提取,即可得到交互对象的对象特征。其中,特征提取的方式可以根据实际技术需要进行设置。

[0070] 需要说明的是,分别针对每一交互对象的交互对象数据进行特征提取,得到各交互对象各自对应的交互对象特征。然后,将各交互对象的交互对象特征进行拼接处理,以便后续输入至交互资源新增预测模型。在拼接处理后,交互对象特征可以表现为矩阵形式,矩阵的行表示交互对象,行数与交互对象的数量相同,列表示不同类型的交互对象数据,列数与交互对象数据的总数目相同。

[0071] 本实施例中,根据待上线的目标交互资源在目标交互应用的上线时间,确定需要

获取的预设时长内的交互对象数据,能够使得获取的交互对象数据的数据覆盖度更大、数据表征更充分,在后续确定对象特征并输入至交互资源新增预测模型时,从而可以提高该模型表征的准确度。

[0072] 在一个实施例中,上述步骤S306,对交互对象数据进行特征提取,得到交互对象的对象特征,可以包括以下步骤:

[0073] 步骤S402,校验交互对象数据,获得校验通过的数据。

[0074] 由于获取的交互对象数据的数据量大,因此还需要校验交互对象数据,获得校验通过的数据,也即交互对象数据需要为校验通过的数据,进而,得到对象特征,以确保交互对象数据和得到的对象特征的准确性。

[0075] 校验交互对象数据的方式可以根据实际技术需要进行设置,一个实施例中,校验方式包括以下一种或者多种的组合。需要说明的是,以下的校验方式之间不存在优先级,根据实际情况选择一种或者多种即可。

[0076] 一个实施例中,校验交互对象基础信息与交互对象在目标交互应用中所拥有的交互资源之间的关系,保留符合预设关系的交互对象基础信息。其中,上述的交互对象基础信息与交互对象在目标交互应用中所拥有的交互资源之间的关系可以是“因果”关系,需要删除的交互对象基础信息所对应的关系可以设置为“果”关系,需要保留的交互对象基础信息所对应的关系可以设置为“因”关系,即预设关系为“因”关系。即,如果交互对象基础信息发生增大,是交互对象在目标交互应用中所拥有的交互资源的交互资源新增行为的结果,也即,在购买了新的交互资源之后才导致交互对象基础信息发生了增大,说明此种交互对象基础信息无法表征交互资源新增行为,需要将此种类型的交互对象基础信息删除,仅保留除该类型之外的数据。

[0077] 具体地,确定交互对象在目标交互应用中拥有任一交互资源的时间点,确定交互对象基础信息对应的增长时间段与该时间点之间的先后顺序,并根据先后顺序确定交互对象基础信息与交互对象在目标交互应用中所拥有的交互资源之间的关系。其中,若交互对象基础信息对应的增长时间段在该时间点之后,则确定交互对象基础信息与交互对象在目标交互应用中所拥有的交互资源之间的关系为“果关系”,则相关数据需要删除,若交互对象基础信息对应的增长时间段在该时间点之前,则确定交互对象基础信息与交互对象在目标交互应用中所拥有的交互资源之间的关系为“因关系”,则相关数据可以保留。

[0078] 一个实施例中,校验交互对象基础信息之间的关系,保留存在交集的交互对象基础信息。即,删除交互对象基础信息之间无交集的相关信息。例如,交互对象基础信息为每个账号下的前预定数目名望值角色的角色信息,角色信息中包括角色预设穿着排名,若部分角色不存在角色预设穿着排名这一参数,则表示部分角色的角色预设穿着排名无交集,则该参数需要删除。

[0079] 一个实施例中,校验并保留表达符合表达筛选条件的交互对象数据。其中,交互对象数据的表达筛选条件,是指该交互对象数据能够充分表达的条件,是否表达充分需要根据实际技术需要,并结合该数据的含义和种类确定。例如,有的数据使用最值(最大值或最小值)表达较为合理,有的数据采用平均数或众数表达较为合理,合理即为表达充分,若存在不合理的表达则需要进行合理替换。

[0080] 一个实施例中,上述获得的交互对象数据中可能存在与待上线的目标交互资源的

交互资源新增行为相关性极低的因素,因此可以将此种交互对象数据进行过滤,以增加模型的可解释性。可以确定交互对象数据与待上线的目标交互资源之间的相关性,筛选出符合预设相关性的交互对象数据并保留。

[0081] 具体地,可以采用任意一种可实现的方式确定相关性,包括但不限于是:判断该交互对象数据对应的对象特征是否发散,删除不发散的交互对象数据,保留发散的交互对象数据。其中,不发散即表示该特征不能正确表达。或者,计算交互对象数据与待上线的目标交互资源之间的相关系数,筛选出大于或等于预设相关系数的交互对象数据并保留。其中,相关系数越高,该数据越应保留。预设相关系数的数值可以根据实际技术需要进行设置。相关系数可以采用任意一种可实现的计算方式,例如,可以是计算皮尔逊相关系数(Pearson correlation coefficient)。

[0082] 步骤S404,对校验通过的数据进行特征提取,得到交互对象的对象特征。

[0083] 在得到校验通过的数据后,可以通过任意一种可能的方式对校验通过的数据进行特征提取,得到交互对象的对象特征。

[0084] 本实施例中,通过对获取的交互对象数据进行校验,并对校验通过的数据进行特征提取,再得到交互对象的对象特征,可以实现数据过滤,提高交互对象数据的准确度,从而提高对象特征的准确性,在后续模型利用对象特征进行处理时,提高输出的预测数据的精度。

[0085] 在一个实施例中,交互资源新增预测模型是预先训练的组合模型,其中,交互资源新增预测模型包括:分类预测子模型和回归预测子模型。分类预测子模型和回归预测子模型的模型类型和模型结构可以根据实际技术需要进行设置,一个实施例中,分别采用的是XGBoost分类预测子模型和XGBoost回归预测子模型。分类预测子模型和回归预测子模型是采用各自的训练集分别训练得到,其训练集分别称为分类数据集和回归数据集。

[0086] 其中,XGBoost分类预测子模型采用的是改进的XGBoost分类算法,并基于分类数据集进行训练得到。改进的XGBoost分类算法是指对损失函数的相关改进,在后续的实施例中会进一步说明。XGBoost回归预测子模型采用的是XGBoost回归算法,并基于回归数据集训练得到。

[0087] XGBoost分类预测子模型的分类是指将交互对象分为会发生交互资源新增行为、以及不会发生交互资源新增行为这两种类型。进而,XGBoost回归预测子模型基于会发生交互资源新增行为的交互对象的对象特征,预测出会发生交互资源新增行为所对应的预测数据。具体地,结合图3所示,本申请实施例方法,即通过获取交互对象数据以及待上线的目标交互资源的相关数据,进一步构建得到对象特征以及资源特征,分别通过分类、预测,最终得到交互对象对于目标交互资源发生交互资源新增行为的预测数据。

[0088] 一个实施例中,结合图4中交互资源新增预测模型的训练部分、以及图5所示的交互资源新增预测模型的训练过程所示,以下实施例对该交互资源新增预测模型的训练过程进行相关说明。

[0089] 对于XGBoost分类预测子模型,采用与上述步骤的实施例类似的方式,获得多个样本对象各自对应的样本对象数据,并对样本对象数据进行校验以及特征提取处理,得到样本对象特征。为了避免量纲对模型训练的影响,将样本对象特征进行归一化处理,也称标签化处理,标签0表示该样本对象未发生过交互资源新增行为,标签1表示该样本对象至少发

生过一次交互资源新增行为。标签1所表示的样本也称为正样本，标签0所表示的样本称为负样本，至此形成分类数据集，作为改进的XGBoost分类预测子模型的输入。

[0090] XGBoost模型是一种提升树模型，其是将许多树模型集成在一起，形成分类器，分类的思想是不断添加决策树，每增加一棵树，实质上是学习一个新的函数来拟合上次预测的误差。假设训练结束后得到了K棵树的模型，一个样本所对应的样本预测分数，实质上取决于该样本的同维度特征所分散的叶子节点，即模型中的每个叶子节点分别对应一个预测分数，所有树的预测分数相加即为该样本的预测分数。样本对应的预测分数可以表示如下：

$$[0091] \quad \hat{y} = \varphi(x_i) = \sum_{k=1}^K f_k(x_i)$$

[0092] where $F = \{f(x) = \omega_{q(x)}\} (q: R^m \rightarrow T, \omega \in R^T)$

[0093] 其中， \hat{y} 为样本对应的预测分数， x_i 表示第i个特征向量， $\omega_{q(x)}$ 为叶子节点q的分数， $f(x)$ 是其中一棵回归树， $q: R^m \rightarrow T$ 表示叶子节点q从集合R映射到集合T，R和T无具体指代意义。

[0094] XGBoost模型算法的训练基于目标函数，目标函数由衡量样本的预测分数和真实分数之间差距的损失函数、以及正则化项组成，目标函数可以表示为：

$$[0095] \quad Obj = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k)$$

[0096] 其中，Obj表示目标函数， \hat{y}_i 表示样本对应的预测分数， y_i 表示样本对应的真实分数， f_k 表示第k个正则化项。

[0097] 将损失函数部分进行泰勒展开，泰勒展开后的损失函数表示为：

$$[0098] \quad l(y_i, \hat{y}_i^{(t-1)}) + g_i f(x_i) + \frac{1}{2} h_i f_t^2(x_i)$$

[0099] 其中， h_i 表示损失函数对 \hat{y}_i 的一阶导数， g_i 表示损失函数对 \hat{y}_i 的二阶导数，(t-1)表示叶子节点的前(t-1)棵树。

[0100] 将前(t-1)棵树常量化后定义一棵树的复杂度表示为：

$$[0101] \quad \Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2$$

[0102] 其中， γ 和 λ 为表示复杂度的常量，训练时根据数据先验知识或训练结果动态设定，T表示叶子节点总数， ω_j 表示第j个叶子节点的预测分数。

[0103] 结合上述公式，将目标函数进一步表示为：

$$Obj = \sum_{j=1}^T \left[\left(\sum_{i \in I} g_i \right) \omega_j + \frac{1}{2} \left(\sum_{i \in I} h_i + \lambda \right) \omega_j^2 \right] + \gamma T$$

[0104]

$$= \sum_{j=1}^T \left[G_j \omega_j + \frac{1}{2} (H_j + \lambda) \omega_j^2 \right] + \gamma T$$

[0105] 根据上述公式,当 $\omega_j^* = -\frac{G_j}{H_j + \lambda}$ 时,最优切分点使得目标函数为:

$$Obj = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T$$

[0106]

[0107] 其中, ω_j^* 表示 ω_j 取特定常数值, G_j 是叶子节点j所包含样本的一阶导数之和, H_j 是叶子节点j所包含样本的二阶导数之和。

[0108] 此时,基于叶子节点的分裂依据,叶子节点的增益情况表示为:

$$Gain = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] - \gamma$$

[0109]

[0110] 其中,Gain表示叶子节点的增益情况,G和H分别表示一阶导数之和与二阶导数之和,下标L和R分别表示取叶子节点左边、叶子节点右边的导数之和。

[0111] 通过确定分类数据集中的正负样本的比例,本实施例中设置的XGBoost分类预测子模型的损失函数表示为:

$$L_{FL-\alpha} = \begin{cases} -\alpha(1 - \hat{y}_i)^\gamma \log \hat{y}_i, & y_i = 1 \\ -(1 - \alpha)\hat{y}_i^\gamma \log(1 - \hat{y}_i), & y_i = 0 \end{cases}$$

[0112]

[0113] 其中,平衡系数 α 用于调整正负样本在损失函数中的权重大小,其取值可以根据实际技术需要进行设置,一个实施例中,平衡系数 α 的大小在(0,1]之间,用于增加正样本的权重,以使正样本的权重大于负样本的权重。聚焦系数 γ 用于调整训练过程中难易分类样本的损失权重,其取值可以根据实际技术需要进行设置,一个实施例中,聚焦系数 γ 大于0。对于一个正样本,若该样本对应的预测分数越接近于1,表示该样本越容易被分类,被视为易分类样本,此时 $(1 - \hat{y}_i)^\gamma$ 越小,即表示减小了易分类样本的损失权重,从而训练过程更关注于难分类样本,以使难分类样本的损失权重大于易分类样本的损失权重。

[0114] 计算上述改进的损失函数的一阶导数与二阶导数,将相应的数值代入至叶子节点的增益情况公式中,计算选取不同特征做叶子节点的分割特征时,叶子节点的增益情况。XGBoost分类预测子模型的训练过程,即学习确定每一叶子节点的最大增益的过程。在XGBoost分类预测子模型达到训练结束条件时,即训练得到XGBoost分类预测子模型。其中,训练结束条件可以根据实际技术需要进行设置,例如,可以设置为模型损失收敛。

[0115] 对于XGBoost回归预测子模型,从样本对象特征中提取出至少发生过一次交互资源新增行为的样本对象的相关特征,形成回归数据集,回归数据集中的样本标签是该样本

对象的交互资源新增行为所对应的新增数目。XGBoost回归预测子模型的损失函数表示为：

$$[0116] \quad l(y_i, \hat{y}_i) = \frac{1}{2} \sum (y_i - \hat{y}_i)^2$$

[0117] 上述的损失函数对应的一阶导数 g_i 表示为：

$$[0118] \quad g_i = \frac{d}{d\hat{y}_i} \frac{1}{2} (y_i - \hat{y}_i)^2 = -(y_i - \hat{y}_i)$$

[0119] 损失函数对应的二阶导数 $h_i = 1$ 。在XGBoost回归预测子模型达到训练结束条件时，即训练得到XGBoost回归预测子模型。其中，训练结束条件可以根据实际技术需要进行设置，例如，可以设置为模型损失收敛。通过训练XGBoost回归预测子模型，即可确定样本对象对应的预测数据。

[0120] 在分类预测子模型与回归预测子模型分别训练结束后，即可得到满足训练终止条件的交互资源新增预测模型。此时，将待上线的目标交互资源的资源特征以及对象特征，输入满足训练终止条件的交互资源新增预测模型，即可得到交互对象对于目标交互资源发生交互资源新增行为的预测数据，该交互资源新增预测模型的具体数据处理包括以下步骤：

[0121] 步骤S502，根据待上线的目标交互资源的资源特征以及对象特征，通过分类预测子模型，预测出对于目标交互资源发生交互资源新增行为的新增交互对象。

[0122] 交互资源新增预测模型包括分类预测子模型。结合图4中交互资源新增预测模型的预测部分所示，将待上线的目标交互资源的资源特征以及对象特征，输入满足训练终止条件的交互资源新增预测模型，即分类预测子模型先根据待上线的目标交互资源的资源特征以及对象特征，通过分类预测子模型，预测出对于目标交互资源发生交互资源新增行为的交互对象，此种对于目标交互资源会发生交互资源的交互对象也称为新增交互对象。

[0123] 步骤S506，从对象特征中筛选确定新增交互对象的新增对象特征。

[0124] 在确定新增交互对象后，需要针对新增交互对象的相关特征进行进一步处理，将新增交互对象所对应的对象特征称为新增对象特征。其中，由于新增交互对象是交互对象中的一部分对象，因此，可以从对象特征中筛选确定新增交互对象的新增对象特征。

[0125] 步骤S508，基于新增对象特征以及资源特征，通过回归预测子模型，得到交互对象对于目标交互资源发生交互资源新增行为的预测数据。

[0126] 交互资源新增预测模型包括回归预测子模型。在确定新增交互对象特征后，基于新增对象特征以及资源特征，通过回归预测子模型，即可得到交互对象对于目标交互资源发生交互资源新增行为的预测数据。其中，预测数据具体包括：预测会发生交互资源新增行为的交互对象的预测对象数目，以及该目标交互资源所对应的交互资源新增行为的预测新增数目。

[0127] 本实施例中，通过分类预测子模型的损失函数进行改进，能够平衡正负样本，改进正样本的数量小于负样本的数量的不均衡情况，而且，可以使得训练时算法更关注于难分类样本。通过分别针对分类预测子模型和回归预测子模型进行训练，进而得到交互对象新增预测模型，能够达到使得在数据预测的过程中，先进行数据分类预测，再进行数据回归预测，得到预测数据，提高数据预测的准确性。

[0128] 在一个实施例中，本实施例的分类预测子模型，还可以确定影响交互资源新增行

为的对象特征,从而,后续目标交互应用的策划者可以对目标交互资源的相关资源信息进行适应性调整。

[0129] 具体地,在分类预测子模型满足训练结束条件时,根据分类预测子模型中的各叶子节点所对应的增益情况,确定影响交互资源新增行为的目标对象特征,目标对象特征用于调整目标交互资源所对应的资源信息。其中,目标对象特征的个数可以是预定数目个,预定数目可以根据实际技术需要设置,例如,可以设置为15。预定数目个目标对象特征,可以是根据增益情况对各对象特征进行排序,选择排序在前的预定数目个目标对象特征,或者,选择排序在前的预定比例个对象特征,作为预定数目个目标对象特征。

[0130] 在一个实施例中,根据分类预测子模型中的各叶子节点所对应的增益情况,确定影响交互资源新增行为的目标对象特征,包括以下步骤:

[0131] 步骤S602,确定分类预测子模型中的各叶子节点的分割特征。

[0132] 在分类预测子模型的实际训练过程中,可以学习确定各叶子节点的最优切分点,通过确定增益最大的特征作为分割特征,用该特征的最佳分割点作为分割位置,在该叶子节点上分割出左右两个新的叶子节点,直至满足训练结束条件为止。由此,在分类预测子模型满足训练结束条件时,即可确定分类预测子模型中的各叶子节点的分割特征。

[0133] 步骤S604,统计各叶子节点的分割特征,得到各分割特征的统计结果。

[0134] 确定分类预测子模型中的各叶子节点的分割特征后,统计各叶子节点的分割特征,得到各分割特征的统计结果,即确定每一分割特征对应的预测分数和出现次数。

[0135] 步骤S606,根据各分割特征的统计结果和预设的对象特征个数,确定影响交互资源新增行为的目标对象特征。

[0136] 根据各分割特征的统计结果和预设的对象特征个数,对各分割特征进行排序,即可确定排序在前的预定数目个对象特征,也即确定了影响交互资源新增行为的预定数目个目标对象特征,从而,后续可以便于目标交互应用的策划者对目标交互资源的资源信息进行调整。结合图6所示,通过统计出每一分割特征及其对应的特征重要性程度得分,其中特征重要性程度得分可以根据该特征的预测分数和出现次数确定,然后,选择前预定数目个特征重要性程度高的特征,作为影响交互资源新增行为的目标对象特征,目标对象特征为预定数目个。

[0137] 本实施例中,通过在分类预测子模型训练结束后,确定影响交互资源新增行为的目标对象特征,能够为目标交互应用的策划者在设计与调整目标交互资源的资源信息时提供参考价值。

[0138] 在一个实施例中,为确保交互资源新增预测模型的预测精度,在待上线的目标交互资源上线后,还可以根据交互对象对于目标交互资源发生交互资源新增行为的实际数据,对交互资源新增预测模型进行调整,具体包括以下步骤:

[0139] 步骤S702,在待上线的目标交互资源上线后,确定交互对象对于目标交互资源发生交互资源新增行为的实际数据。

[0140] 在待上线的目标交互资源上线后,统计并确定预设时长内,交互对象对于目标交互资源发生交互资源新增行为的实际数据。其中,预设时长可以根据实际技术需要进行设置,在此不做限制。实际具体数据可以包括:针对该目标交互资源发生了交互资源新增行为的交互对象的实际对象数目,以及该目标交互资源所对应的交互资源新增行为的实际新增

数目。

[0141] 步骤S704,若基于预测数据与实际数据确定满足模型更新条件,更新交互资源新增预测模型,得到更新后的交互资源新增预测模型。

[0142] 确定预测数据与实际数据之间的数据偏差,数据偏差可以包括:预测对象数目与实际对象数目之间的偏差、以及预测新增数目与实际新增数目之间的偏差。若任一种类型的数据偏差大于或等于其对应的预设偏差,则表明此时的交互资源新增预测模型的预测精度低、预测数据不准确,即确定满足模型更新条件,需要采用上述实施例的训练方式,对该交互资源新增预测模型进行重新训练。具体地,若基于预测数据与实际数据确定满足模型更新条件,更新交互资源新增预测模型,得到更新后的交互资源新增预测模型。

[0143] 在一个实施例中,交互资源新增预测模型包括:分类预测子模型和回归预测子模型,因此,更新后的交互资源新增预测模型包括:更新后的分类预测子模型、以及更新后的回归预测子模型。

[0144] 由于分类预测子模型在满足训练结束条件时,可以基于该子模型确定出影响交互资源新增行为的目标对象特征。因此,在交互资源新增预测模型进行了更新后,确定的上述目标对象特征也需要进行相应的更新。具体地,根据更新后的分类预测子模型中的各叶子节点所对应的增益情况,更新影响交互资源新增行为的目标对象特征。其中,上述的目标对象特征的确定方式可以与步骤S602至步骤S606实施例的方式相同,在此不再赘述。

[0145] 在一个实施例中,在更新影响交互资源新增行为的目标对象特征之后,可以根据更新后的影响交互资源新增行为的目标对象特征,调整目标交互资源的资源信息。例如,以目标交互应用是游戏、交互对象是游戏玩家、交互资源是游戏礼包为例,通过阶段性的更新影响游戏玩家购买游戏礼包的特征,可以让游戏策划者动态的调节玩家数值、副本难度、活动频次、礼包价格等,达到游戏礼包销量和游戏玩家口碑之间的相对平衡。

[0146] 本实施例中,通过对交互资源新增预测模型进行更新,得到更新后的交互资源新增预测模型,能够使得交互资源新增预测模型的预测精度保持在较高精度水平,维持其预测数据的准确性。通过更新影响交互资源新增行为的目标对象特征,能够为目标交互应用的策划者在设计与调整目标交互资源的资源信息时提供参考价值。

[0147] 为了使本申请的目的、技术方案及优点更加清楚明白,以下结合附图及具体实施例,对本申请进行进一步详细说明。应当理解,此处描述的具体实施例仅仅用以解释本申请,并不用于限定本申请。

[0148] 在具体实施例中,以目标交互应用是游戏、交互对象是游戏玩家、交互资源是游戏礼包、待上线的目标交互资源为在该游戏中待上线的游戏礼包为例。如图7所示为目标交互资源的示意图,该目标交互资源即为游戏礼包,该游戏礼包中包含游戏道具,包括但不限于是套装、称号、宠物、宠物装备、光环等内容,游戏玩家通过购买游戏礼包,可以实现游戏账号升级、装备加强等。以下实施例以数据预测方法是对游戏礼包的销量数据进行预测为例进行说明:

[0149] 确定待上线的目标交互资源在目标交互应用的上线时间,获取上线时间之前的预设时长内的交互对象数据,交互对象数据包括:交互对象在目标交互应用中所拥有的交互资源的资源信息、以及交互对象的对象信息。

[0150] 其中,交互对象的对象信息包括:交互对象基础信息、交互对象社群信息以及交互

对象行为信息。具体地,可以通过腾讯分布式数据仓库平台(TDW),获取游戏礼包在游戏中的上线时间之前的、两个月内的游戏玩家数据,游戏玩家数据中的交互对象基础信息即为游戏玩家基础信息、交互对象社群信息即为游戏玩家活跃信息、交互对象交易信息即为游戏玩家付费信息、交互对象在目标交互应用中所拥有的交互资源的资源信息即为游戏玩家所拥有的游戏礼包的优惠券数量信息和礼包属性信息。

[0151] 如图8所示为交互对象数据的数据内容示意图,该示意图包括数据类别、数据定义对应的英文表示以及描述。

[0152] 其中,游戏玩家基础信息具体包括:游戏玩家所持有的账号、账号下的游戏角色信息,具体包括每个账号下的所有游戏角色所对应的角色类别信息、名望值信息、付费角色信息、参与副本情况信息,以及每个账号下的前5个名望值角色的角色信息等。

[0153] 具体地,每个账号下的所有游戏角色所对应的角色类别信息,包括:100级角色数量、100级角色中主C角色数量、100级角色中34C角色数量、100级角色中奶系角色数量、小于100级角色数量。每个账号下的所有游戏角色所对应的名望值信息,包括:名望值在1万以下的角色数、名望值在1-1.1万之间的角色数、名望值在1.1-1.2万之间的角色数、名望值在1.2-1.3万之间的角色数、名望值在1.3-1.4万之间的角色数、名望值在1.4-1.5万之间的角色数、名望值在1.5-1.6万之间的角色数、名望值在1.6-1.7万之间的角色数、名望值在1.7-1.8万之间的角色数、名望值在1.8万以上的角色数。每个账号下的所有游戏角色所对应的付费角色信息,包括:最近30天有付费的角色数量。每个账号下的所有游戏角色所对应的参与副本情况信息,包括:最近30天参与过“希洛克团本”的角色数量、最近30天参与过“地下城”的角色数量。分别针对名望值排名前5的角色1至角色5,每一角色各自对应的角色信息为:角色名望值、最近30天角色名望值提升速度、角色穿着“神话”排名(无“神话”排名则为0)、角色穿着装备成套数(至少3件装备为成套)、角色金币数量、最近30天参与“地下城”的总次数、角色最近30天消费金额(点券)、角色最近30天消费金额(非点券)。

[0154] 其中,游戏玩家活跃信息具体包括:最近90天(60天)的登录游戏天数、最近30天的登录游戏天数和最近30天的活跃角色数量。

[0155] 其中,游戏玩家付费信息具体包括:最近1年的累计消费点券金额、最近30天的累计消费点券金额、最近30天的消费点券次数、最近30天的平均点券余额、最近1年的末次付费距离统计日期天数(付费包含点券与非点券)、最近1年的累计消费非点券金额、最近30天的累计消费非点券金额、最近30天的消费非点券次数、最近30天的平均非点券余额、最近30天的平均金币余额。

[0156] 其中,游戏玩家所拥有的游戏礼包的优惠券数量信息和礼包属性信息具体包括:账号下所拥有的游戏礼包的优惠券数量信息,礼包属性信息为:装扮套装-名望值提升信息、光环装扮-名望值提升信息、称号-名望值提升信息、宠物-名望值提升信息、宠物装备-名望值提升信息。

[0157] 在获得上述的交互对象数据后,校验交互对象数据,获得校验通过的数据。具体地,校验交互对象基础信息与交互对象在目标交互应用中所拥有的交互资源之间的关系,保留符合预设关系的交互对象基础信息。例如,游戏玩家的某一数据值是在购买了新的游戏礼包之后才发生了增大,则需要将此种类型的游戏玩家数据删除,保留其他数据。校验交互对象基础信息之间的关系,仅保留交互对象基础信息之间存在交集的相关信息。例如,角

色信息中包括角色穿着“神话”排名,若部分角色不存在角色穿着“神话”排名这一参数,则表示部分角色的角色预设穿着排名无交集,则该参数需要删除,保留其他数据。校验并保留表达符合表达筛选条件的交互对象数据。例如,有的数据使用最值(最大值或最小值)表达较为合理,有的数据采用平均数或众数表达较为合理,若存在不合理的表达则需要进行合理替换。计算交互对象数据与待上线的目标交互资源之间的相关系数,筛选出大于或等于预设相关系数的交互对象数据并删除,相关系数可以是皮尔逊相关系数。

[0158] 需要注意的是,目前主要设置有7个检查项,包括:游戏玩家最近30天、最近一年消费记录,金币、点券、非点券余额,角色数量、登录、等级信息,角色名望值信息,副本参与信息,礼包属性值信息,以及上述所有数据信息在TDW数据库存储的有效期等等。

[0159] 对校验通过的数据进行特征提取,得到交互对象的对象特征。具体地,通过对上述的校验通过的游戏对象数据进行特征提取,针对每一个游戏玩家得到一个92维的向量,多个游戏玩家的特征共同组成对象特征。

[0160] 将待上线的目标交互资源的资源特征以及对象特征,输入满足训练终止条件的交互资源新增预测模型,得到交互对象对于目标交互资源发生交互资源新增行为的预测数据;其中,交互资源新增预测模型,包括用于调整正负样本训练权重的损失函数;用于训练交互资源新增预测模型的正样本数量小于负样本数量;正样本用于表征存在交互资源新增行为的交互对象、负样本用于表征不存在交互资源新增行为的交互对象。

[0161] 其中,预先训练的交互资源新增预测模型包括:预先训练的XGBoost分类预测子模型、预先训练的XGBoost回归预测子模型。XGBoost分类预测子模型的损失函数即为上述的可以调整正负样本训练权重的损失函数,表示如下:

$$[0162] \quad L_{FL-\alpha} = \begin{cases} -\alpha(1 - \hat{y}_i)^\gamma \log \hat{y}_i, & y_i = 1 \\ -(1 - \alpha)\hat{y}_i^\gamma \log(1 - \hat{y}_i), & y_i = 0 \end{cases}$$

[0163] 其中,平衡系数 α 用于调整正负样本在损失函数中的权重大小,其取值为0.2,聚焦系数 γ 用于调整训练过程中难易分类样本的损失权重,其取值为2。需要说明的是,当负样本过多时,大量的负样本的损失会主导损失函数。为了减少这种主导,可以给负样本增加的损失也相应增加一个衰减。理论上,平衡系数 α 的取值越大越好,但是由于聚焦系数 γ 的存在,负样本的损失已经呈指数下降,所以平衡系数 α 可以取较小值,以相对较小幅度地减少负样本的损失。

[0164] 具体地,将待上线的游戏礼包的礼包特征以及上述得到的对象特征,输入至预先训练的交互资源新增预测模型中,其中,通过分类预测子模型,预测出对于游戏礼包会发生该游戏礼包购买行为的购买游戏玩家;从对象特征中筛选确定购买游戏玩家的购买对象特征;基于购买对象特征以及礼包特征,通过回归预测子模型,得到游戏玩家对于该游戏礼包会发生该游戏礼包购买行为的预测销量数据,预测销量数据包括:预测的会购买的游戏玩家数目以及该游戏礼包的销量。

[0165] 其中,在上述的分类预测子模型满足训练结束条件时,还可以根据分类预测子模型中的各叶子节点所对应的增益情况,确定影响交互资源新增行为的目标对象特征;其中,目标对象特征用于调整目标交互资源所对应的资源信息。具体地,可以确定分类预测子模型中的各叶子节点的分割特征;统计各叶子节点的分割特征,得到各分割特征的统计结果;根据各分割特征的统计结果和预设的对象特征个数,确定影响交互资源新增行为的目标对

象特征。

[0166] 具体地,通过统计确定影响发生游戏礼包购买行为的前15个对象特征,作为目标对象特征。图9所示为一个实施例中确定的影响交互资源新增行为的目标对象特征的示意图。其中,统计出每一叶子节点的分割特征及其对应的特征重要性程度得分,选择前15个特征重要性程度高的特征,作为影响发生游戏礼包购买行为的目标对象特征,目标对象特征具体包括:角色1最近30天参与“地下城”的总次数、最近一年的累计消费点券金额、角色1金币数量、最近90天登录游戏天数、最近30天的平均金币余额、最近一年末次付费距统计日期天数(含非点券)、角色2最近30天参与“地下城”的总次数、最近一年的累计消费非点券金额、角色1等级提升、角色3最近30天参加“地下城”的总次数、角色2金币数量、最近30天的平均非点券金额、角色3金币数量、最近90天登录游戏天数、以及角色4金币数量。

[0167] 在待上线的目标交互资源上线后,确定交互对象对于目标交互资源发生交互资源新增行为的实际数据;若基于预测数据与实际数据确定满足模型更新条件,更新交互资源新增预测模型,得到更新后的交互资源新增预测模型。

[0168] 具体地,在待上线的游戏礼包上线后,确定游戏玩家对于该游戏礼包发生了游戏礼包购买行为的实际销量数据,若实际销量数据与预测销量数据之间的偏差大于预设偏差,则确定满足模型更新条件,此时需要通过新的训练数据集,对交互资源新增预测模型进行重新训练,得到更新后的交互资源新增预测模型。

[0169] 根据更新后的分类预测子模型中的各叶子节点所对应的增益情况,更新影响交互资源新增行为的目标对象特征;根据更新后的影响交互资源新增行为的目标对象特征,调整目标交互资源的资源信息。具体地,调整目标交互资源的资源信息,可以包括调节玩家数值、副本难度、活动频次、礼包价格等内容。

[0170] 应该理解的是,虽然如上所述的各实施例所涉及的流程图中的各个步骤按照箭头的指示依次显示,但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明,这些步骤的执行并没有严格的顺序限制,这些步骤可以以其它的顺序执行。而且,如上所述的各实施例所涉及的流程图中的至少一部分步骤可以包括多个步骤或者多个阶段,这些步骤或者阶段并不必然是在同一时刻执行完成,而是可以在不同的时刻执行,这些步骤或者阶段的执行顺序也不必然是依次进行,而是可以与其它步骤或者其它步骤中的步骤或者阶段的至少一部分轮流或者交替地执行。

[0171] 基于同样的发明构思,本申请还提供了一种用于实现上述所涉及的数据预测方法的数据预测装置。该装置所提供的解决问题的实现方案与上述方法中所记载的实现方案相似,故下面所提供的一个或多个数据预测装置实施例中的具体限定可以参见上文中对于数据预测方法的限定,在此不再赘述。

[0172] 在一个实施例中,如图10所示,提供了一种数据预测装置,包括:获取模块11和预测模块12,其中:

[0173] 获取模块11,用于基于目标交互应用的交互对象、以及所述交互对象在所述目标交互应用中所拥有的交互资源,得到所述交互对象的对象特征。

[0174] 预测模块12,用于将待上线的目标交互资源的资源特征以及所述对象特征,输入满足训练终止条件的交互资源新增预测模型,得到所述交互对象对于所述目标交互资源发生交互资源新增行为的预测数据;其中,所述交互资源新增预测模型,包括用于调整正负样

本训练权重的损失函数；用于训练所述交互资源新增预测模型的正样本数量小于负样本数量；所述正样本用于表征存在交互资源新增行为的交互对象、所述负样本用于表征不存在交互资源新增行为的交互对象。

[0175] 在一个实施例中，所述获取模块11，用于确定待上线的所述目标交互资源在所述目标交互应用的上线时间；获取所述上线时间之前的预设时长内的交互对象数据，所述交互对象数据包括：交互对象在所述目标交互应用中所拥有的交互资源的资源信息、以及所述交互对象的对象信息；对所述交互对象数据进行特征提取，得到所述交互对象的对象特征。

[0176] 在一个实施例中，所述交互对象数据为校验通过的数据；所述对象信息包括：交互对象基础信息、交互对象社群信息以及交互对象行为信息；所述获取模块11，还用于校验所述交互对象数据；所述交互对象数据的校验方式，包括：校验所述交互对象基础信息与所述交互对象在所述目标交互应用中所拥有的交互资源之间的关系，保留符合预设关系的交互对象基础信息；校验所述交互对象基础信息之间的关系，保留存在交集的交互对象基础信息；校验并保留表达符合表达筛选条件的交互对象数据；确定所述交互对象数据与待上线的所述目标交互资源之间的相关性，筛选出符合预设相关性的交互对象数据并保留。

[0177] 在一个实施例中，所述预测模块12中交互资源新增预测模型的所述正负样本训练权重，包括：正负样本在损失函数中的权重大小以及训练过程中难易分类样本的损失权重；所述正样本的权重大于所述负样本的权重，难分类样本的损失权重大于易分类样本的损失权重。

[0178] 在一个实施例中，所述交互资源新增预测模型包括：分类预测子模型和回归预测子模型；所述预测模块12，用于根据待上线的所述目标交互资源的资源特征以及所述对象特征，通过所述分类预测子模型，预测出对于所述目标交互资源发生交互资源新增行为的新增交互对象；从所述对象特征中筛选确定所述新增交互对象的新增对象特征；基于所述新增对象特征以及所述资源特征，通过所述回归预测子模型，得到所述交互对象对于所述目标交互资源发生交互资源新增行为的预测数据。

[0179] 在一个实施例中，所述装置还包括：排序模块。

[0180] 在一个实施例中，所述排序模块，用于在所述分类预测子模型满足训练结束条件时，根据所述分类预测子模型中的各叶子节点所对应的增益情况，确定影响所述交互资源新增行为的目标对象特征；所述目标对象特征用于调整所述目标交互资源所对应的资源信息。

[0181] 在一个实施例中，所述排序模块，用于确定所述分类预测子模型中的各叶子节点的分割特征；统计各所述叶子节点的分割特征，得到各所述分割特征的统计结果；根据各所述分割特征的统计结果和预设的对象特征个数，确定影响所述交互资源新增行为的目标对象特征。

[0182] 在一个实施例中，所述装置还包括：更新模块。

[0183] 在一个实施例中，所述更新模块，用于在待上线的所述目标交互资源上线后，确定所述交互对象对于所述目标交互资源发生交互资源新增行为的实际数据；若基于所述预测数据与所述实际数据确定满足模型更新条件，更新所述交互资源新增预测模型，得到更新后的交互资源新增预测模型。

[0184] 在一个实施例中,所述更新后的交互资源新增预测模型包括:更新后的分类预测子模型;所述更新模块,用于根据所述更新后的分类预测子模型中的各叶子节点所对应的增益情况,更新影响所述交互资源新增行为的目标对象特征。

[0185] 上述数据预测装置中的各个模块可全部或部分通过软件、硬件及其组合来实现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中,也可以以软件形式存储于计算机设备中的存储器中,以便于处理器调用执行以上各个模块对应的操作。

[0186] 在一个实施例中,提供了一种计算机设备,该计算机设备可以是服务器,其内部结构图可以如图11所示。该计算机设备包括处理器、存储器、输入/输出接口(Input/Output,简称I/O)和通信接口。其中,处理器、存储器和输入/输出接口通过系统总线连接,通信接口通过输入/输出接口连接到系统总线。其中,该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质和内存。该非易失性存储介质存储有操作系统、计算机程序和数据库。该内存为非易失性存储介质中的操作系统和计算机程序的运行提供环境。该计算机设备的数据库用于存储数据预测数据。该计算机设备的输入/输出接口用于处理器与外部设备之间交换信息。该计算机设备的通信接口用于与外部的终端通过网络连接通信。该计算机程序被处理器执行时以实现一种数据预测方法。

[0187] 在一个实施例中,提供了一种计算机设备,该计算机设备可以是终端,其内部结构图可以如图12所示。该计算机设备包括处理器、存储器、输入/输出接口、通信接口、显示单元和输入装置。其中,处理器、存储器和输入/输出接口通过系统总线连接,通信接口、显示单元和输入装置通过输入/输出接口连接到系统总线。其中,该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存。该非易失性存储介质存储有操作系统和计算机程序。该内存为非易失性存储介质中的操作系统和计算机程序的运行提供环境。该计算机设备的输入/输出接口用于处理器与外部设备之间交换信息。该计算机设备的通信接口用于与外部的终端进行有线或无线方式的通信,无线方式可通过WIFI、移动蜂窝网络、NFC(近场通信)或其他技术实现。该计算机程序被处理器执行时以实现一种数据预测方法。该计算机设备的显示单元用于形成视觉可见的画面,可以是显示屏、投影装置或虚拟现实成像装置,显示屏可以是液晶显示屏或电子墨水显示屏,该计算机设备的输入装置可以是显示屏上覆盖的触摸层,也可以是计算机设备外壳上设置的按键、轨迹球或触控板,还可以是外接的键盘、触控板或鼠标等。

[0188] 本领域技术人员可以理解,图11和图12中示出的结构,仅仅是与本申请方案相关的部分结构的框图,并不构成对本申请方案所应用于其上的计算机设备的限定,具体的计算机设备可以包括比图中所示更多或更少的部件,或者组合某些部件,或者具有不同的部件布置。

[0189] 在一个实施例中,提供了一种计算机设备,包括存储器和处理器,存储器中存储有计算机程序,该处理器执行计算机程序时实现上述的方法的步骤。

[0190] 在一个实施例中,提供了一种计算机可读存储介质,其上存储有计算机程序,计算机程序被处理器执行时实现上述的方法的步骤。

[0191] 在一个实施例中,提供了一种计算机程序产品,包括计算机程序,该计算机程序被处理器执行时实现上述的方法的步骤。

[0192] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,是可以

通过计算机程序来指令相关的硬件来完成,所述的计算机程序可存储于一非易失性计算机可读取存储介质中,该计算机程序在执行时,可包括如上述各方法的实施例的流程。其中,本申请所提供的各实施例中所使用的对存储器、数据库或其它介质的任何引用,均可包括非易失性和易失性存储器中的至少一种。非易失性存储器可包括只读存储器(Read-Only Memory,ROM)、磁带、软盘、闪存、光存储器、高密度嵌入式非易失性存储器、阻变存储器(ReRAM)、磁变存储器(Magnetoresistive Random Access Memory,MRAM)、铁电存储器(Ferroelectric Random Access Memory,FRAM)、相变存储器(Phase Change Memory,PCM)、石墨烯存储器等。易失性存储器可包括随机存取存储器(Random Access Memory,RAM)或外部高速缓冲存储器等。作为说明而非局限,RAM可以是多种形式,比如静态随机存取存储器(Static Random Access Memory,SRAM)或动态随机存取存储器(Dynamic Random Access Memory,DRAM)等。本申请所提供的各实施例中涉及的数据库可包括关系型数据库和非关系型数据库中至少一种。非关系型数据库可包括基于区块链的分布式数据库等,不限于此。本申请所提供的各实施例中涉及的处理器可为通用处理器、中央处理器、图形处理器、数字信号处理器、可编程逻辑器、基于量子计算的数据处理逻辑器等,不限于此。

[0193] 以上所述实施例的各技术特征可以进行任意的组合,为使描述简洁,未对上述实施例中的各个技术特征所有可能的组合都进行描述,然而,只要这些技术特征的组合不存在矛盾,都应当认为是本说明书记载的范围。

[0194] 以上所述实施例仅表达了本申请的几种实施方式,其描述较为具体和详细,但不能因此而理解为对本申请专利范围的限制。应当指出的是,对于本领域的普通技术人员来说,在不脱离本申请构思的前提下,还可以做出若干变形和改进,这些都属于本申请的保护范围。因此,本申请的保护范围应以所附权利要求为准。

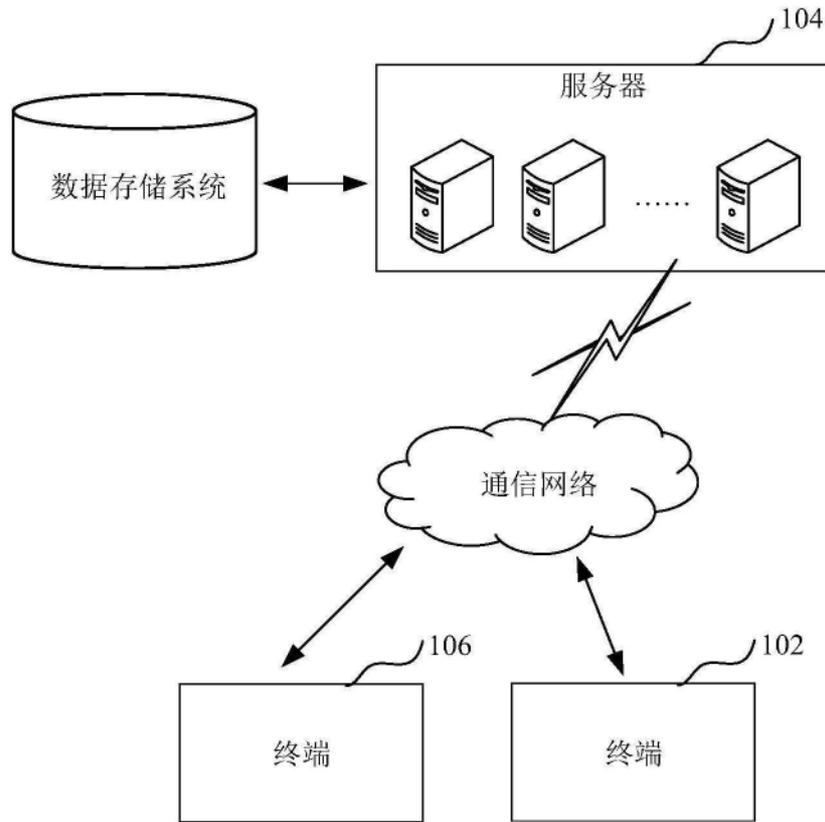


图1

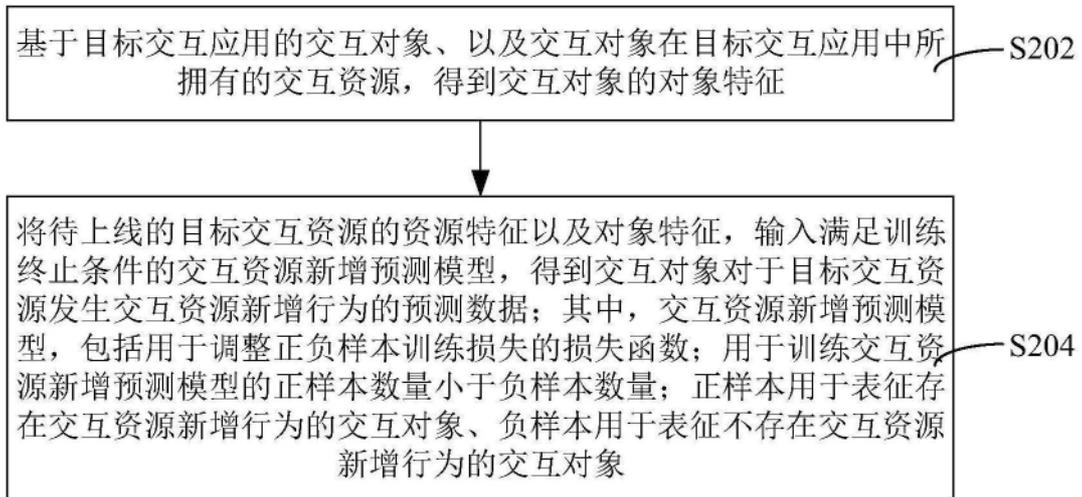


图2

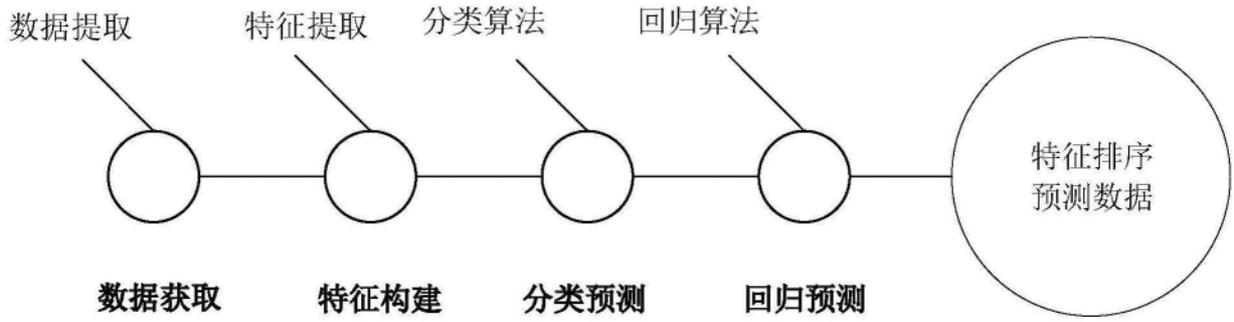


图3

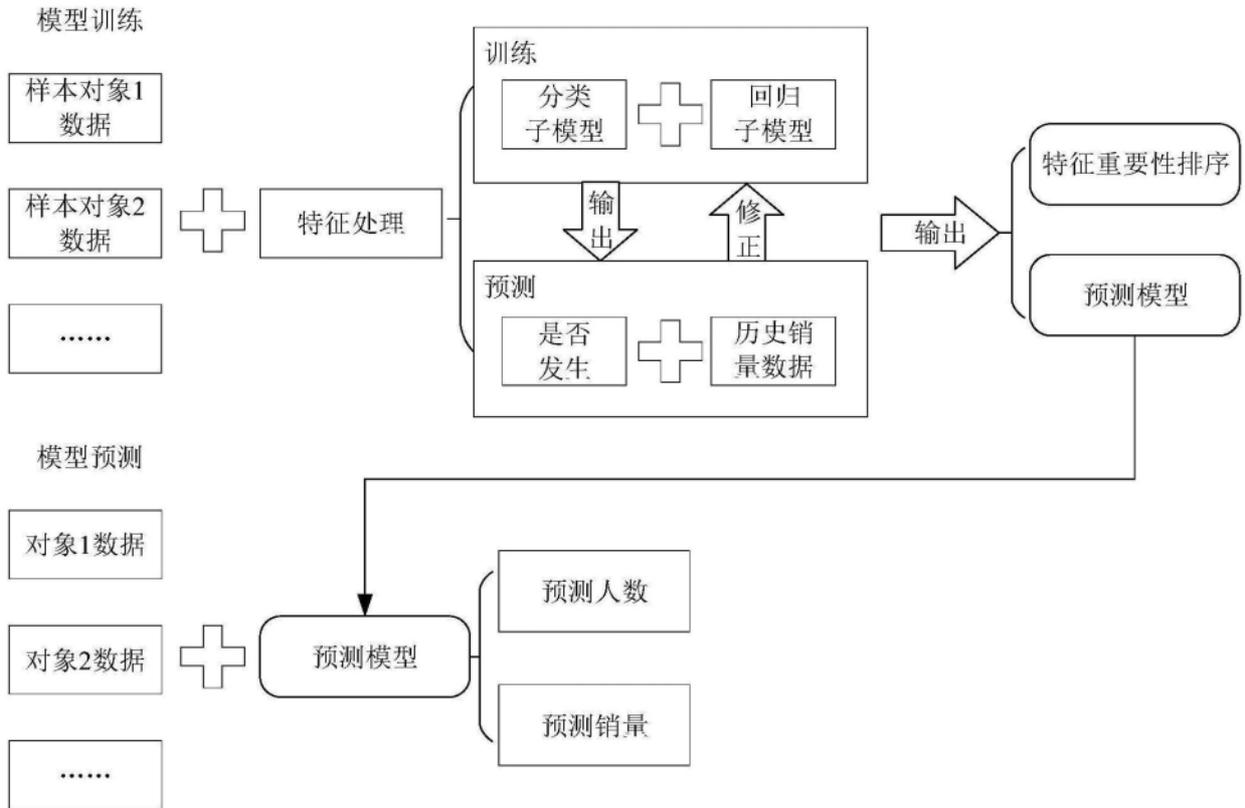


图4

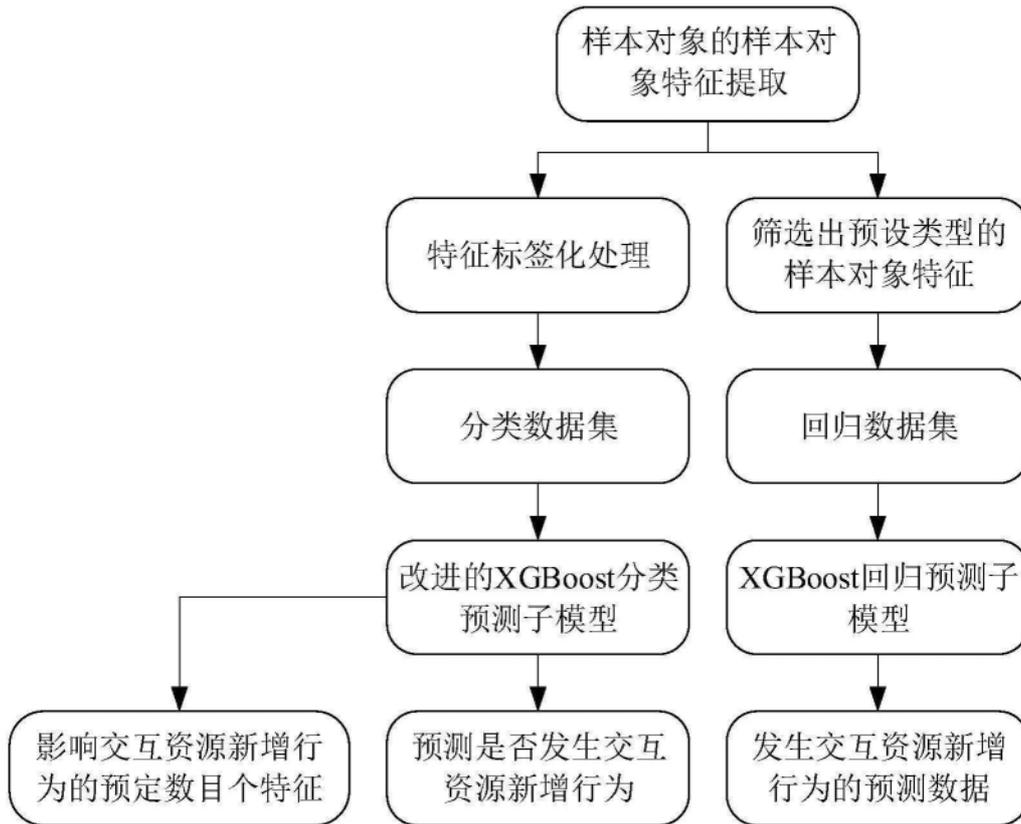


图5

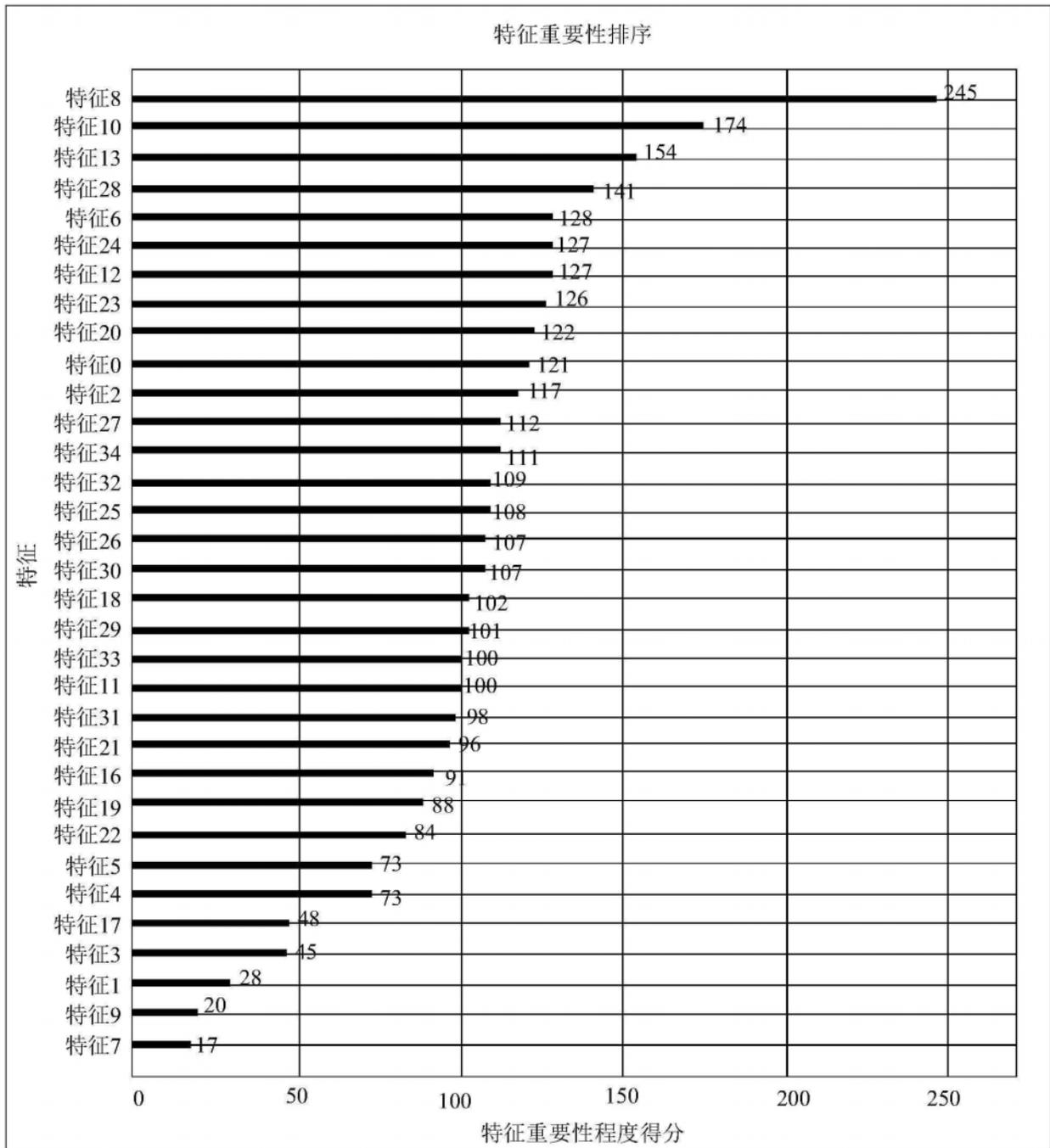


图6

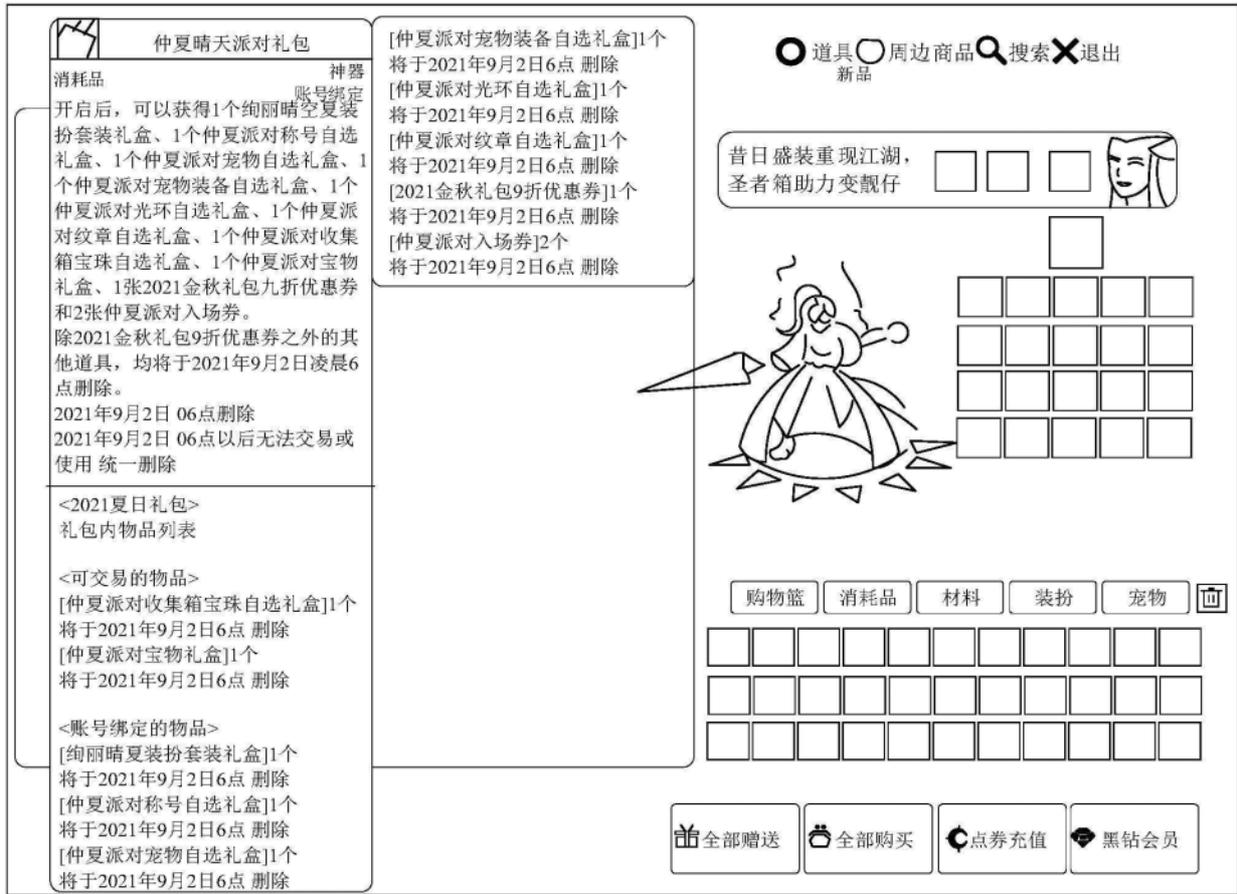


图7

序号	数据类别	数据定义	描述
1	玩家付费能力	Payment_365day_antry_sum	最近一年的累计消费点券金额
2		Payment_30day_antry_sum	最近30天的累计消费点券金额
3		times_30day_antry_sum	最近30天的消费点券次数
4		Money_remaining_entry_sum	最近30天的平均点券余额（20210307之后才开始按天存储）
5		Lastpay_day_num	最近一年，末次付费据统计日期天数（含非点券）
6		Payment_365day_cerapoint_sum	最近一年的累计消费非点券金额
7		Payment_30day_cerapoint_sum	最近30天的累计消费非点券金额
8		times_30day_cerapoint_sum	最近30天的消费非点券次数
9		Avg_money_remaining_cerapoint	最近30天的平均非点券余额
10		Avg_gold	最近30天的平均金币余额
11	玩家活跃情况	Logindays_90day_sum	最近90天（60天）登录游戏天数
12		Logindays_30day_sum	最近30天登录游戏天数
13		Activerote_30day_sum	最近30天活跃角色数量
14	账号下角色信息	Account_lv100_sum	账号下100级角色数量
15	角色类别	Account_lv100_C_sum	账号下100级角色中主C角色数量
16		Account_lv100_34_sum	账号下100级角色中34C角色数量
17		Account_lv100_nai_sum	账号下100级角色中奶系角色数量
18		Account_lesslv100_sum	账号下小于100级角色数量
19	名望值	Account_fame_10000_sum	名望值在1万以下的角色数
20		Account_fame_10000_11000_sum	名望值在1万~1.1万之间的角色数
21		Account_fame_11000_12000_sum	名望值在1.1万~1.2万之间的角色数
22		Account_fame_12000_13000_sum	名望值在1.2万~1.3万之间的角色数
23		Account_fame_13000_14000_sum	名望值在1.3万~1.4万之间的角色数
24		Account_fame_14000_15000_sum	名望值在1.4万~1.5万之间的角色数
25		Account_fame_15000_16000_sum	名望值在1.5万~1.6万之间的角色数
26		Account_fame_16000_17000_sum	名望值在1.6万~1.7万之间的角色数
27		Account_fame_17000_18000_sum	名望值在1.7万~1.8万之间的角色数
28		Account_fame_18000_sum	名望值在1.8万以上的角色数
29	付费情况	Account_pay_30day_sum	最近30天有付费的角色数
30	参与副本情况	Account_hot_gamescript_30day_sum	最近30天参与过“希洛克团本”的角色数量
31		Account_total_30day_sum	最近30天参加过地下城的角色数量
32	账号下名望值前5的角色信息-角色1	Role1_fame	角色名望值
33		Role1_fame_improve_speed	最近30天角色名望值提升速度
34		Role1_myth	角色穿着神话排名（无神话则为0）
35		Role1_equipment_sum	角色穿着装备成套数（至少3件才算成套）
36		Role1_gold_number	角色金币数量
37		Role1_gamescript_30day_number	最近30天参加地下城总次数
38		Role1_entry_30day_sum	角色最近30天消费金额（点券）
39		Role1_carepoint_30day_sum	角色最近30天消费金额（非点券）
40	优惠券	Coupon_number	账号下对应礼包的优惠券数量
41	账号下名望值前5的角色信息-角色2	Role2_fame	角色名望值
42		Role2_fame_improve_speed	最近30天角色名望值提升速度
43		Role2_myth	角色穿着神话排名（无神话则为0）
44		Role2_equipment_sum	角色穿着装备成套数（至少3件才算成套）
45		Role2_gold_number	角色金币数量
46		Role2_gamescript_30day_number	最近30天参加地下城总次数
47		Role2_entry_30day_sum	角色最近30天消费金额（点券）
48		Role2_carepoint_30day_sum	角色最近30天消费金额（非点券）
...
...
...	礼包属性	avatar_fame_improve	装扮套装-名望值提升
...		Halo_fame_improve	光环装扮-名望值提升
...		title_fame_improve	称号-名望值提升
...		pet_improve	宠物-名望值提升
...		pet_zhuangbei_improve	宠物装扮-名望值提升

图8

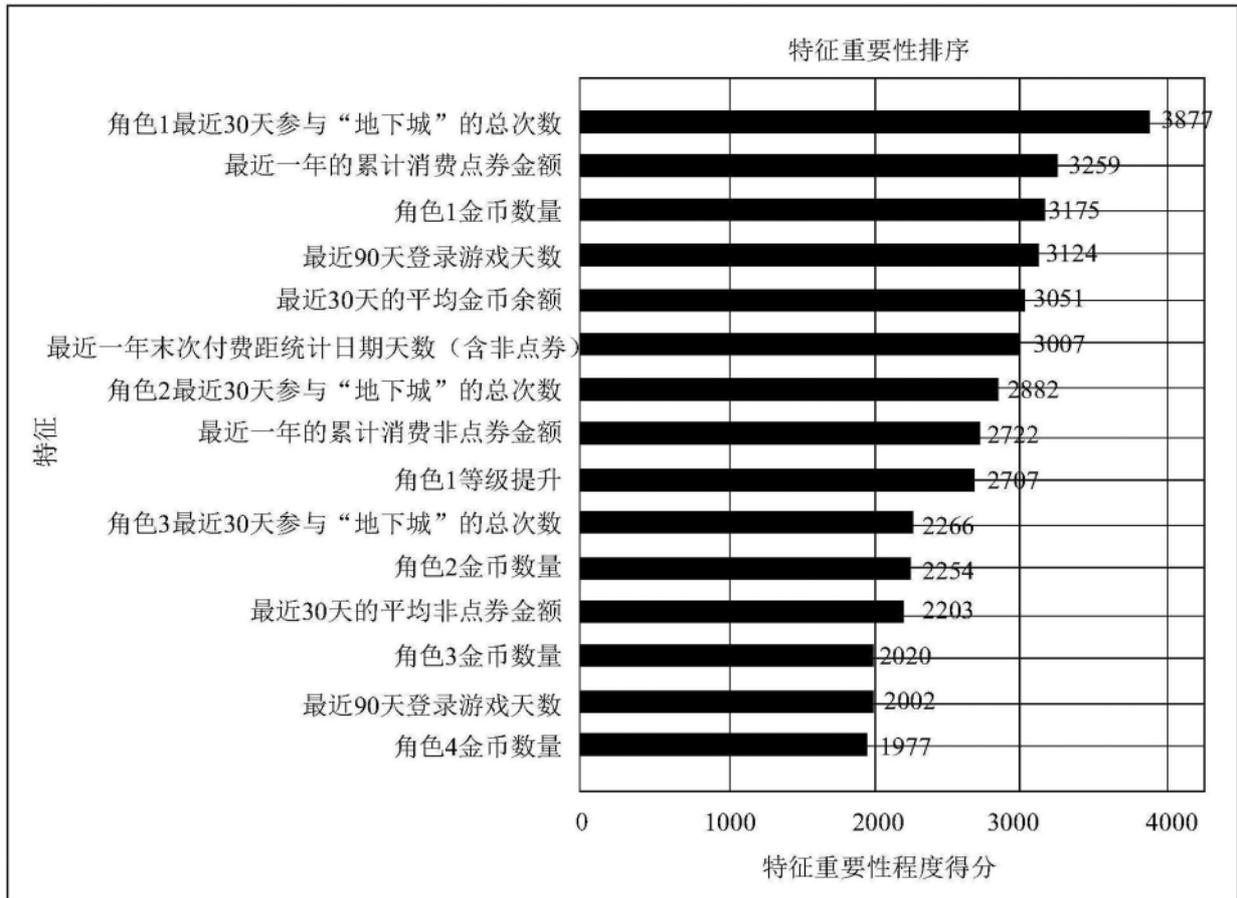


图9



图10

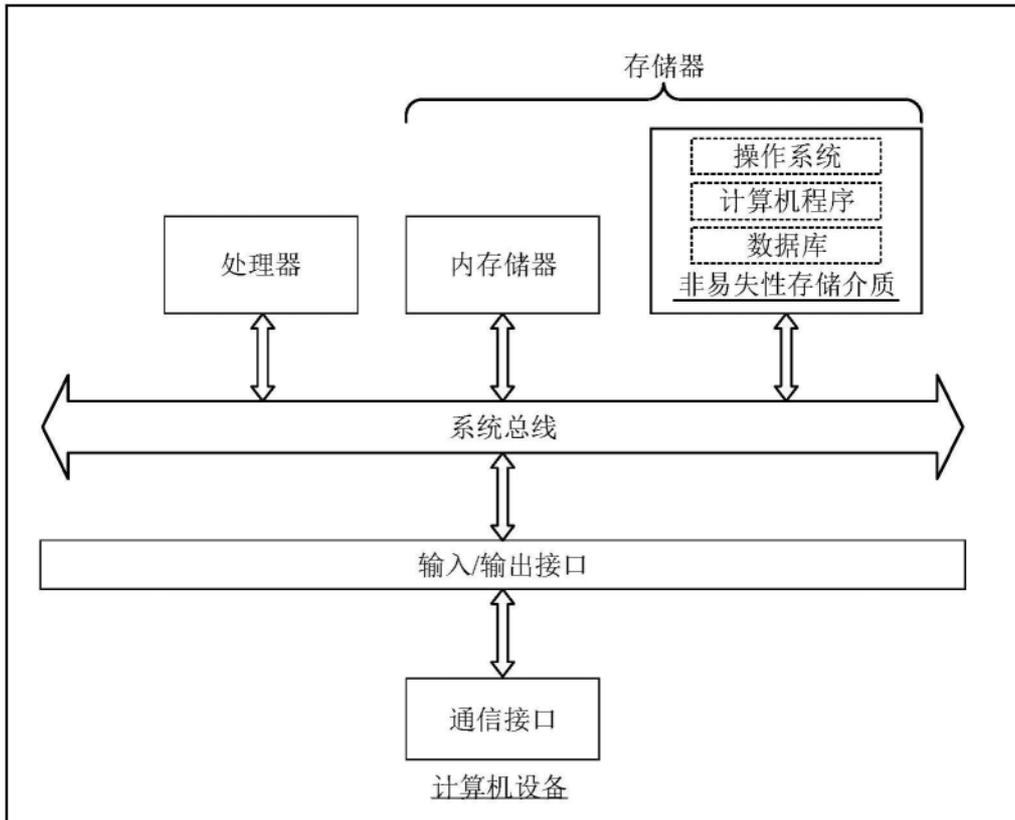


图11

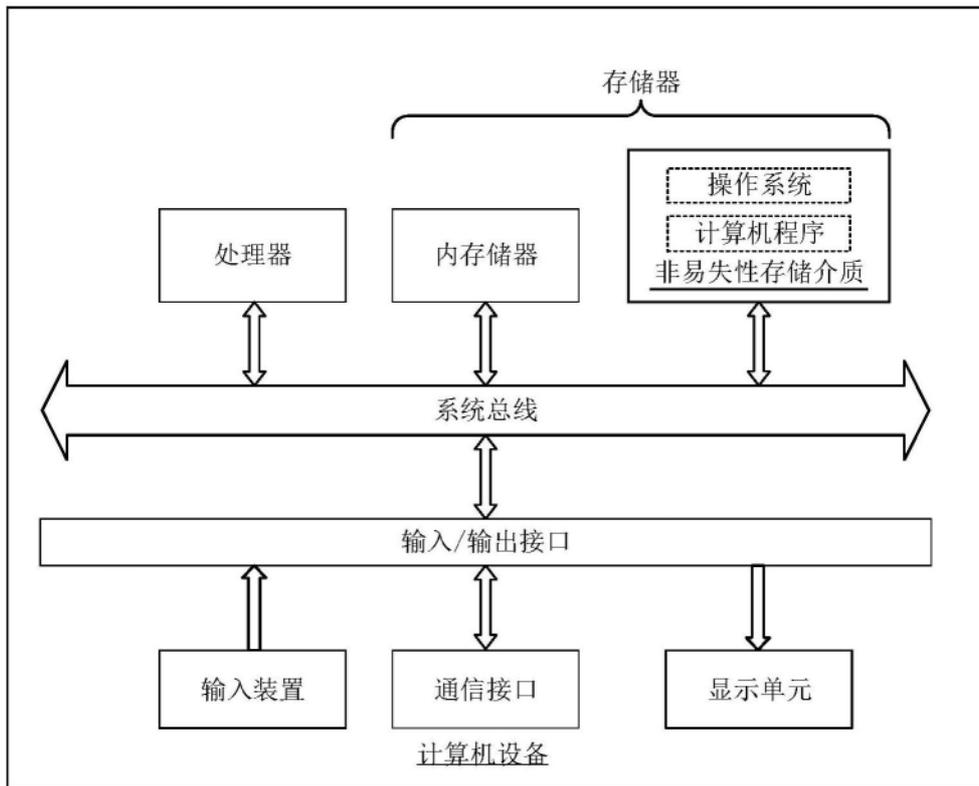


图12