



(12) 发明专利

(10) 授权公告号 CN 111494964 B

(45) 授权公告日 2020. 11. 20

(21) 申请号 202010613086.0

(22) 申请日 2020.06.30

(65) 同一申请的已公布的文献号
申请公布号 CN 111494964 A

(43) 申请公布日 2020.08.07

(73) 专利权人 腾讯科技(深圳)有限公司
地址 518000 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 林文清

(74) 专利代理机构 广州嘉权专利商标事务所有
限公司 44205

代理人 梁嘉琦

(51) Int. Cl.
G06N 20/20 (2019.01)

(56) 对比文件

CN 108665007 A, 2018.10.16

CN 107705183 A, 2018.02.16

CN 108428001 A, 2018.08.21

CN 108268934 A, 2018.07.10

CN 107292713 A, 2017.10.24

CN 110555717 A, 2019.12.10

CN 110070104 A, 2019.07.30

US 2009281884 A1, 2009.11.12

徐逸扬. 基于用户数据分析的移动游戏品质优化.《中国优秀硕士学位论文全文数据库信息科技辑》.2019, (第8期), 第I138-572页.

徐逸扬. 基于用户数据分析的移动游戏品质优化.《中国优秀硕士学位论文全文数据库信息科技辑》.2019, (第8期), 第I138-572页.

审查员 何诚

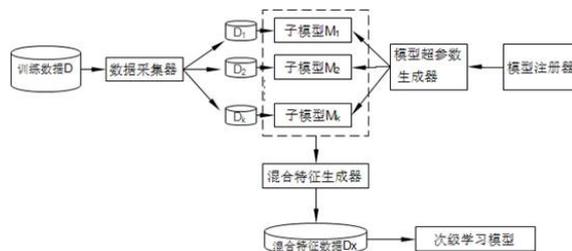
权利要求书2页 说明书15页 附图8页

(54) 发明名称

虚拟物品的推荐方法、模型的训练方法、装置及存储介质

(57) 摘要

本申请公开了虚拟物品的推荐方法、模型的训练方法、装置及存储介质,通过集成学习模型对目标用户所需的虚拟物品进行推荐,其中集成学习模型包括预先注册的多个预设模型,由于预设模型及其对应的配置信息被预先设置好,只需要在预设模型集合中选择预设模型,并根据对应的配置信息对预设模型的超参数进行配置,即可方便生成集成学习模型。由于通过本申请提供的训练方法训练的集成学习模型包括多个不同类型的预设模型,因此能够充分地突出每种学习算法类型的优点,并且通过其他类型的预设模型弥补该预设模型的缺点,通过集成学习模型能够向目标用户准确地推荐所需的虚拟物品。



1. 一种虚拟物品的推荐方法,其特征在于,包括:

获取目标用户的特征数据;

将所述特征数据分别输入集成学习模型中的多个预设模型,对所述目标用户的特征数据进行虚拟物品推荐预测,得到每个所述预设模型的预测值,所述预设模型为所述集成学习模型预先注册的机器学习模型且已配置好超参数;

对多个所述预测值进行统计运算,得到混合特征,所述混合特征包括各个预设模型输出的预测值和根据各个所述预测值的分布特征得到的统计值;

将所述混合特征输入经过训练的次级学习模型,得到所述目标用户的所述特征数据的虚拟物品推荐结果;

其中,所述集成学习模型中包括多种学习算法类型的预设模型,所述集成学习模型中的多个所述预设模型通过在所述集成学习模型预先注册的预设模型集合中随机选择得到,其中,所述预设模型集合包含至少两种学习算法类型的预设模型,所述预设模型集合中注册有由用户创作的学习算法模型,所述预设模型的超参数根据配置信息生成,所述配置信息包括超参数的取值范围,所述预设模型的超参数在所述取值范围内随机生成或通过网格搜索生成,其中,网格搜索指将超参数数值的取值范围做一个划分,然后取划分界线上的值。

2. 一种集成学习模型的训练方法,其特征在于,包括:

在预先注册的预设模型集合中随机选择至少两个预设模型,所述预设模型集合包含至少两种学习算法类型的预设模型,其中,所述预设模型集合中注册有由用户创作的学习算法模型;

获取各个所述预设模型的配置信息;

根据各个所述预设模型的配置信息,在超参数的取值范围内随机生成或通过网格搜索生成各个所述预设模型的超参数数值组合,其中,网格搜索指将超参数数值的取值范围做一个划分,然后取划分界线上的值;

根据各个所述超参数数值组合配置各个所述预设模型的超参数;

获取训练数据,根据所述训练数据对各个配置好超参数的所述预设模型进行训练,所述训练数据包括用户的特征样本数据和所述用户的虚拟物品选择结果;

配置次级学习模型,使训练好的所述至少两个预设模型的预测值输入至所述次级学习模型;

获取各个所述预设模型在训练过程中形成的预测值,对所述预测值进行统计生成一种类型以上的统计数据,将所述统计数据和所述预测值作为第二训练集;

根据所述第二训练集对所述次级学习模型进行训练,完成集成学习模型的训练。

3. 一种计算机设备,其特征在于,包括:

至少一个存储器;

至少一个处理器;

至少一个程序;

所述程序被存储在所述存储器中,所述处理器执行所述至少一个所述程序以实现如权利要求1所述的虚拟物品的推荐方法,或者,实现如权利要求2所述的集成学习模型的训练方法。

4. 一种计算机可读存储介质, 存储有计算机可执行指令, 其特征在于: 所述计算机可执行指令用于执行权利要求1所述的虚拟物品的推荐方法, 或者, 用于执行权利要求2所述的集成学习模型的训练方法。

虚拟物品的推荐方法、模型的训练方法、装置及存储介质

技术领域

[0001] 本发明涉及人工智能技术领域,尤其涉及虚拟物品的推荐方法、模型的训练方法、装置及存储介质。

背景技术

[0002] 在游戏互动过程中,存在一些场景需要向用户推荐虚拟物品,以使用户选择或者购买。例如用户进入虚拟物品购买界面后,在购买界面中向用户推荐虚拟物品。又例如,在一些游戏场景中,当本场游戏结束后,有一定的几率可以让用户购买指定的折扣商品,这样能够提高用户的游戏体验。然而,由于游戏场景中的虚拟物品种类较多,在向用户推荐虚拟物品时无法将所有种类的虚拟物品向用户展示,通常的做法是固定或随机向用户展示供选择或购买的虚拟物品,然而上述的固定推荐方式只是依赖游戏设计者或者维护人员的主观经验进行预测和设置的,预测结果并不理想。

[0003] 为了提高虚拟物品预测的准确性,从而向用户推荐更加感兴趣的虚拟物品,可以通过机器学习的方式对用户的喜好进行预测,例如通过构建机器学习模型,收集用户特征数据和过往的购买行为对机器学习模型进行训练,通过训练好的学习模型预测目标用户对虚拟物品的喜好和购买可能性。例如通过GDBT梯度提升树模型或者Random Forest随机森林模型实现虚拟物品的推荐。然而,上述单一算法模型方案的预测结果并不理想。

发明内容

[0004] 以下是对本文详细描述的主题的概述。本概述并非是为了限制权利要求的保护范围。

[0005] 本申请提供了一种虚拟物品的推荐方法、模型的训练方法、装置及存储介质,能够方便地建立并使用集成学习模型,提高虚拟物品推荐的准确性。

[0006] 根据本申请的第一方面,提供了一种虚拟物品的推荐方法,包括:

[0007] 获取目标用户的特征数据;

[0008] 将所述特征数据分别输入集成学习模型中的多个预设模型,对所述目标用户的特征数据进行虚拟物品推荐预测,得到每个所述预设模型的预测值,所述预设模型为所述集成学习模型预先注册的机器学习模型且已配置好超参数;

[0009] 对多个所述预测值进行统计处理得到所述预测值的分布特征,根据所述分布特征得到所述目标用户的虚拟物品推荐结果;或者,将多个所述预测值输入经过训练的次级学习模型进行计算,得到所述目标用户的虚拟物品推荐结果;

[0010] 其中,所述集成学习模型中包括多种学习算法类型的预设模型,所述预设模型的超参数根据配置信息生成。

[0011] 根据本申请的第二方面,提供了一种集成学习模型的训练方法,包括:

[0012] 在预先注册的预设模型集合中选择至少两个预设模型,所述预设模型集合包含至少两种学习算法类型的预设模型;

- [0013] 获取各个所述预设模型的配置信息；
- [0014] 根据各个所述预设模型的配置信息配置各个所述预设模型的超参数；
- [0015] 获取训练数据，根据所述训练数据对各个配置好超参数的所述预设模型进行训练，所述训练数据包括用户的特征样本数据和所述用户的虚拟物品选择结果；
- [0016] 将各个被训练的所述预设模型进行集成，完成集成学习模型的训练。
- [0017] 根据本申请的第三方面，提供了一种虚拟物品的推荐装置，包括：
- [0018] 数据获取装置，获取目标用户的特征数据；
- [0019] 预测装置，所述特征数据分别输入集成学习模型中的多个预设模型，对所述目标用户的特征数据进行虚拟物品推荐预测，得到每个所述预设模型的预测值，所述预设模型为所述集成学习模型预先注册的机器学习模型且已配置好超参数；
- [0020] 集成装置，对多个所述预测值进行统计处理得到所述预测值的分布特征，根据所述分布特征得到所述目标用户的虚拟物品推荐结果；或者，对多个所述预测值输入经过训练的次级学习模型进行计算，得到所述目标用户的虚拟物品推荐结果；
- [0021] 其中，所述集成学习模型中包括多种学习算法类型的预设模型，所述预设模型的超参数根据配置信息生成。
- [0022] 根据本申请的第四方面，提供了一种集成学习模型的生成装置，包括：
- [0023] 选择装置，在预先注册的预设模型集合中选择至少两个预设模型；
- [0024] 配置信息获取装置，获取各个所述预设模型的配置信息；
- [0025] 超参数配置装置，根据各个所述预设模型的配置信息配置各个所述预设模型的超参数；
- [0026] 训练装置，获取训练数据，根据所述训练数据对各个配置好超参数的所述预设模型进行训练，所述训练数据包括用户的特征样本数据和所述用户的虚拟物品选择结果；
- [0027] 集成装置，将各个被训练的所述预设模型进行集成，完成集成学习模型的训练。
- [0028] 根据本申请的第五方面，提供了一种计算机设备，包括：
- [0029] 至少一个存储器；
- [0030] 至少一个处理器；
- [0031] 至少一个程序；
- [0032] 所述程序被存储在所述存储器中，所述处理器执行所述至少一个所述程序以实现本申请第一方面所述的虚拟物品的推荐方法，或者，以实现本申请第二方面所述的集成学习模型的训练方法。
- [0033] 根据本申请的第六方面，提供了一种计算机可读存储介质，所述计算机可执行指令用于执行本申请第一方面所述的虚拟物品的推荐方法，或者，以实现本申请第二方面所述的集成学习模型的训练方法。
- [0034] 本申请所提供的技术方案，通过集成学习模型对目标用户所需的虚拟物品进行推荐，其中集成学习模型包括预先注册的多个预设模型，预设模型的超参数根据配置信息生成，由于预设模型及其对应的配置信息被预先设置好，因此只需要一次注册好预设模型集合，即可用于生成不同预设模型组合的集成学习模型，用户在构建集成学习模型时无需对集成学习模型的算法的集成进行深入研究及逐一进行配置集成，即可方便生成并训练集成学习模型。另外，即使集成学习模型的训练和预测结果不理想，也可以十分方便地重新生成

集成学习模型,大大节省了构建集成学习模型的时间。由于通过本申请实施例提供的集成学习模型包括多个不同类型的预设模型,因此能够充分地突出每种学习算法类型的优点,并且通过其他类型的预设模型弥补该预设模型的缺点,使得通过集成学习模型能够向目标用户准确地推荐所需的虚拟物品。

附图说明

[0035] 附图用来提供对本申请技术方案的进一步理解,并且构成说明书的一部分,与本申请的实施例一起用于解释本申请的技术方案,并不构成对本申请技术方案的限制。

[0036] 图1是本申请一个示例性的实施例提供的集成学习模型的训练方法的流程图;

[0037] 图2是图1中步骤130的具体方法流程图;

[0038] 图3是图1中步骤140的具体方法流程图;

[0039] 图4是图3中步骤320的具体方法流程图;

[0040] 图5是图1中步骤150的具体方法流程图;

[0041] 图6是图5中步骤520的具体方法流程图;

[0042] 图7是本申请一个示例性的实施例提供的预设模型集合的注册方法流程图;

[0043] 图8是本申请一个示例性的实施例提供的集成学习模型的模块框架图;

[0044] 图9是本申请一个示例性的实施例提供的虚拟物品的推荐方法的流程图;

[0045] 图10是本申请一个示例性的实施例提供的集成学习模型的模块框架图;

[0046] 图11是本申请虚拟物品的推荐方法的一个应用场景的示意图;

[0047] 图12是本申请一个示例性的实施例提供的虚拟物品的推荐装置的结构示图;

[0048] 图13是本申请一个示例性的实施例提供的集成学习模型的生成装置的结构示图;

[0049] 图14是本申请一个示例性的实施例提供的集成学习模型的计算机设备的结构示图;

[0050] 图15是本申请一个示例性的实施例提供的服务器的结构示图。

具体实施方式

[0051] 为了使本发明的目的、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,此处所描述的具体实施例仅用以解释本申请,并不用于限定本申请。

[0052] 需要说明的是,虽然在装置示意图中进行了功能模块划分,在流程图中示出了逻辑顺序,但是在某些情况下,可以以不同于装置中的模块划分,或流程图中的顺序执行所示出或描述的步骤。说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。

[0053] 除非另有定义,本文所使用的所有的技术和科学术语与属于本发明的技术领域的技术人员通常理解的含义相同。本文中所使用的术语只是为了描述本发明实施例的目的,不是旨在限制本发明。

[0054] 首先,对本申请中涉及的若干名词进行解析:

[0055] 虚拟物品:虚拟物品指虚拟账户中的虚拟资源,例如游戏中的游戏道具。虚拟物品可以由系统发放到用户的虚拟账户,也可以被用户所使用,例如在游戏场景中使用游戏道

具。

[0056] 集成学习:集成学习(ensemble learning)指通过构建并结合多个学习器来完成学习任务,有时也被称为多分类器系统(multi-classifier system)、基于委员会的学习(committee-based learning)等。集成学习的一般先产生一组“个体学习器”(individual learner),再用某种策略将它们结合起来。个体学习器通常由一个现有的学习算法从训练数据产生,例如C4.5决策树算法、BP神经网络算法等,此时集成中只包含同种类型的个体学习器,例如“决策树集成”中全是决策树,“神经网络集成”中全是神经网络,这样的集成是“同质”的(homogeneous)。同质集成中的个体学习器亦称“基学习器”(base learner)相应的学习算法称为“基学习算法”(base learning algorithm),集成也可包含不同类型的个体学习器,例如同时包含决策树和神经网络,这样的集成是“异质”的(heterogenous)。

[0057] 超参数:超参数是指模型训练之前提前设定的参数,例如人工神经网络中的权重、支持向量机中的支持向量、线性回归或逻辑回归中的系数和决策树中的决策树个数和决策树最大深度等。

[0058] 梯度提升树模型:梯度提升树模型(Gradient Boosting Decision Tree,GBDT)是通过采用加法模型(即基函数的线性组合),以及不断减小训练过程产生的残差来达到将数据分类或者回归的算法。

[0059] 随机森林模型:随机森林模型是一个包含多个决策树的分类器,并且其输出的类别是由个别树输出的类别的众数而定,主要应用于回归和分类,随机森林生成每棵树时每个节点变量都仅在随机选出的少数变量中产生,因此不但样本是随机的,连每个节点变量(Features)的产生都是随机的。

[0060] 决策树模型:决策树模型(Decision Tree)是一种基本的分类与回归方法,当决策树用于分类时称为分类树,用于回归时称为回归树。

[0061] 逻辑回归模型:逻辑回归(Logistic Regression)是一种用于解决二分类问题的机器学习方法,用于估计某种事物的可能性,逻辑回归是以线性回归为理论支持的,但是逻辑回归通过Sigmoid函数引入了非线性因素,因此可以轻松处理0/1分类问题。

[0062] 支持向量机模型:支持向量机(Support Vector Machine, SVM)是一类按监督学习(supervised learning)方式对数据进行二元分类的广义线性分类器(generalized linear classifier),其决策边界是对学习样本求解的最大边距超平面(maximum-margin hyperplane)。

[0063] 贝叶斯模型:贝叶斯模型是运用贝叶斯统计进行的一种预测。贝叶斯统计不同于一般的统计方法,其不仅利用模型信息和数据信息,而且充分利用先验信息,通过实证分析的方法。

[0064] 本申请实施例中,所提供的虚拟物品的推荐方法、模型的训练方法均可以应用于人工智能之中。人工智能(Artificial Intelligence, AI)是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能,感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。换句话说,人工智能是计算机科学的一个综合技术,它企图了解智能的实质,并生产出一种新的能以人类智能相似的方式做出反应的智能机器。人工智能也就是研究各种智能机器的设计原理与实现方法,使机器具有感知、推理与决策的功能。人工智能技术是一门综合学科,涉及领域广泛,既有硬件层面的技术也有软件层面的技

术。人工智能基础技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理技术、操作/交互系统、机电一体化等技术。人工智能软件技术主要包括计算机视觉技术、语音处理技术、自然语言处理技术以及机器学习/深度学习等几大方向。

[0065] 具体地,本申请实施例中提供的虚拟物品的推荐方法、模型的训练方法可以应用于游戏奖励、游戏道具购买、玩家互动、直播间等应用场景中,在这些应用场景中,人工智能系统可以利用虚拟物品的预设模型预测用户所需的虚拟物品,以便根据预测结果作进一步的应用,也就是说,通过本申请实施例方法所得到的用户的虚拟物品预测结果可以是这些人工智能应用的中间结果。

[0066] 本申请实施例中,主要涉及的人工智能技术是机器学习。

[0067] 机器学习(Machine Learning, ML)属于人工智能的一个分支,是一门多领域交叉学科,涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。专门研究计算机怎样模拟或实现人类的学习行为,以获取新的知识或技能,重新组织已有的知识结构使之不断改善自身的性能。机器学习是人工智能的核心,是使计算机具有智能的根本途径,其应用遍及人工智能的各个领域。机器学习和深度学习通常包括人工神经网络、置信网络、强化学习、迁移学习、归纳学习、式教学习等技术。

[0068] 本申请实施例提供的方案涉及虚拟物品的推荐方法、模型的训练方法、装置及存储介质,具体通过如下实施例进行说明,首先描述本申请实施例中的集成学习模型的训练方法。

[0069] 本申请实施例提供了一种集成学习模型的训练方法,可应用于终端中,也可应用于服务器中,还可以是运行于终端或服务器中的软件。在一些实施例中,可以是应用在Spark:分布式大数据处理平台中,例如通过分布式大数据处理平台对各个预设模型对集成学习模型进行训练。在一些实施例中,终端可以是智能手机、平板电脑、笔记本电脑、台式计算机或者智能手表等;服务器可以配置成独立的物理服务器,也可以配置成多个物理服务器构成的服务器集群或者分布式系统,还可以配置成提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN以及大数据和人工智能平台等基础云计算服务的云服务器;软件可以是训练集成学习模型的应用程序等,但并不局限于以上形式。

[0070] 图1是本申请实施例提供的一种集成学习模型的训练方法的一个可选的流程图,图1中的方法包括步骤110至步骤150。

[0071] 步骤110,在预先注册的预设模型集合中选择至少两个预设模型,所述预设模型集合包含至少两种学习算法类型的预设模型;

[0072] 本步骤中,预设模型集合中预先注册有至少两种学习算法类型的预设模型。预设模型集合可应用于不同的集成学习模型,例如两个不同的集成学习模型可以共用同一个预设模型集合。用户在预设模型集合中预先注册至少两种学习算法类型的预设模型及用于配置预设模型超参数的配置信息。通过配置信息对预设模型的超参数进行配置后,预设模型就可以执行训练任务。其中,预设模型集合中可以预先注册现有经典的学习算法模型,例如梯度提升树模型、随机森林模型、决策树模型、逻辑回归模型、支持向量机模型和贝叶斯模型等。另外预设模型集合中也可以注册由用户创作的学习算法模型或未来出现的新学习算法模型,例如当出现新的学习算法模型,只需要将其注册到预设模型集合中即可进行集成

使用。

[0073] 在预先注册的预设模型集合中选择至少两个预设模型可以是在预设模型集合中随机选择至少两个预设模型,例如在梯度提升树模型、随机森林模型、决策树模型、逻辑回归模型、支持向量机模型和贝叶斯模型中随机选择至少两种,通过随机在预设模型集合中选择多个预设模型,可以提高集成学习模型中学习算法类型种类多样性,即集成学习模型中的预设模型的类型是随机组合的,这样就不用对样本的特征数据进行分析思考集成学习算法的组合。另外,随机选择的方式也节省了用户配置集成算法的操作,用户只需要一个命令即可生成集成学习模型,用户无需知晓或学习各类学习算法类型,降低用户使用集成学习算法的成本,扩大了适用人群。

[0074] 另外,一实施例中,还可以由用户指定选择所需要类型的预设模型,例如指定决策树模型、逻辑回归模型作为预设模型。其中,选择预设模型的数量可以是预设或由用户选择或由用户设定。

[0075] 步骤120,获取各个预设模型的配置信息。

[0076] 本步骤中,配置信息用于配置预设模型的超参数,超参数是模型训练之前设定的参数,预设模型需要配置其超参数才能开始训练,而步骤110中的不同学习算法类型的超参数不同,例如对于随机森林模型来说,其超参数包括决策树个数和决策树最大深度,对于神经网络算法而言其超参数为神经单元的权重。因此,本步骤中,为了对所选择的预设模型的超参数进行配置,需要获取对应各个预设模型的配置信息。一实施例中,在注册上述预设模型集合中的预设模型的同时设置对应预设模型的配置信息。其中,配置信息包括预设模型的每个超参数的类型及其取值范围。超参数的类型包括离散型和连续型,其中离散型表示超参数的取值在多个预设值之间进行选择,例如神经网络中的权重在2、4、6中选择,连续型表示超参数在取值范围中的取值是连续值。例如可以设定随机森林模型的决策树最大深度在最小值2至最大值20之间进行选择。通过配置信息可以配置预设模型的超参数,当预设模型有多个超参数需要配置时,通过配置信息可以生成用于配置各个超参数的超参数组合,例如随机森林模型的决策树个数10个和决策树最大深度5个可以组成一个超参数组合($n=10, w=5$)。

[0077] 步骤130,根据各个预设模型的配置信息配置各个预设模型的超参数。

[0078] 由于配置信息包括超参数的类型和取值范围,因此,可以根据各个预设模型的配置信息配置各个预设模型的超参数,当预设模型的超参数有一个时,根据配置信息在超参数的取值范围内选择一个数值作为预设模型的超参数,该超参数构成一个超参数数值组合,当预设模型的超参数有多个时,根据配置信息获取各个超参数的类型及其取值范围,分别在各个取值范围中选择一个数值作为超参数数值组合,根据超参数组合配置预设模型中的各个超参数,例如,配置信息包括决策树的个数 n 的取值范围为 $5 \leq n \leq 15$,决策树最大深度 m 的取值范围为 $2 \leq m \leq 20$,则根据配置信息中的各类超参数的范围选定 $n=10, w=5$ 作为超参数组合配置该预设模型的超参数 $n=10, w=5$ 。通过配置各个预设模型的超参数数值组合,为后续各个预设模型的训练做好准备。

[0079] 其中,超参数的取值可以由用户在取值范围中选择,又或者自动生成。对于自动生成超参数组合,可以是随机的方式在取值范围中进行选择,又或者通过网络搜索的方式在取值范围中进行选择,网格搜索(GridSearchCV)指将超参数数值的取值范围做一个划分,

然后取划分界线上的值。比如,决策树个数的取值范围是0到120,将这个数值划分为5份,则分界线上的值是 0、24、48、72、96、120,一共6个值,将这6个值分别作为各个预设模型中同一种超参数的取值。

[0080] 步骤140,获取训练数据,根据所述训练数据对各个配置好超参数的所述预设模型进行训练。

[0081] 本步骤中,训练数据包括用户的特征样本数据和所述用户的虚拟物品选择结果。训练数据可以在实际的使用场景中采集,例如,使用场景中各个用户的特征样本数据及对应虚拟物品选择结果,其中特征样本数据包括用户的标签特征和行为特征,用户的标签特征指不随着用户的行为而变化的特征,例如用户的性别、年龄等特征,用户的行为特征指随着用户的行为而变化的特征,例如用户的游戏等级、游戏时间、已拥有的游戏装备等特征。用户的虚拟物品选择结果包括用户曾经选择或购买的虚拟物品列表。获取的训练数据组成第一训练数据集,将第一训练数据集中的训练数据输入至各个配置好超参数的所述预设模型进行训练,输入至各个预设模型的训练数据可以相同或者不同。为了增加各个预设模型的差异性,输入至各个预设模型的训练数据可以在第一训练数据集中随机选择。各个预设模型根据训练数据各自进行训练。

[0082] 步骤150,将各个被训练的所述预设模型进行集成,完成集成学习模型的训练。

[0083] 由于各个预设模型是独立训练的,训练好的预设模型也独立产生预测值的,该预测值可以是数值0至1表示的概率,概率表示用户选择对应虚拟物品的概率,概率越高表示用户选择对应虚拟物品的可能性越大,例如一个预测模型中,衣服道具的预测值为概率0.9,鞋子道具的预测值为概率0.6,则用户选择衣服道具的概率更高,衣服道具更符合用户的需求,另外预测值也可以用分数表示,例如分数越高表示用户选择该虚拟物品的可能性越大。对于相同目标用户的特征数据,各个预设模型的输出的预测值可能是不同的,因此需要将各个预设模型的预测值进行整合,最后输出目标用户的虚拟物品推荐结果。其中,可以采用集成计算的方式进行整合,集成计算指对各个预设模型输出的预测值集中起来进行数学运算,例如,对各个预设模型的预测值进行统计处理得到所述预测值的分布特征,根据分布特征得到目标用户的虚拟物品推荐结果,例如可以采用投票的方式进行决定,即采用各个预测值中预测数量最多的虚拟物品作为虚拟物品推荐结果。又或者计算各个预测值的平均值或加权平均值等统计值作为虚拟物品推荐结果。另外,也可以采用二级学习的方式,通过次级学习模型对各个预设模型的预测值进行计算得出虚拟物品推荐结果。

[0084] 本申请实施例提供的一种集成学习模型的训练方法,通过在预先注册的预设模型集合中选择至少两个预设模型,这样能够构建多个不同类型不同超参数组合的预设模型,由于预设模型集合中的预设模型是预先注册的,因此只需要一次注册好预设模型集合,即可用于生成不同预设模型组合的集成学习模型,用户在构建集成学习模型时无需对集成学习模型的算法的集成进行深入研究及逐一进行配置集成,即可方便生成并训练集成学习模型。另外,即使集成学习模型的训练和预测结果不理想,也可以十分方便地重新生成集成学习模型,大大节省了构建集成学习模型的时间。由于通过本申请实施例提供的训练方法训练的集成学习模型包括多个不同类型的预设模型,因此能够充分地突出每种学习算法类型的优点,并且通过其他类型的预设模型弥补该预设模型的缺点,所训练出来的集成学习模型能够较好应用于虚拟物品推荐的场景,具有较高的准确性。

[0085] 本申请一实施例中,通过在预设模型集合中随机选择至少两个预设模型,并根据各个所述预设模型的配置信息随机生成或通过网格搜索自动生成各个预设模型的超参数数值组合。上述网格搜索(GridSearchCV)指将超参数数值的取值范围做一个划分,然后取划分界线上的值。比如,决策树个数的取值范围是0到120,将这个数值划分为5份,则分界线上的值是 0、24、48、72、96、120,一共6个值,将这6个值分别作为各个预设模型中同一种超参数的取值。本实施例中集成学习模型中的预设模型的种类及各个预设模型的超参数均是自动生成的,无论是预设模型的组合类型和超参数都无需人工进行配置,这样,因此只需要一次注册好预设模型集合,即可直接生成不同预设模型组合的集成学习模型。传统用户需要学习各种学习模型的构建方法及了解各个模型的超参数设置规则和原理,需要花费大量的学习成本,而且训练和预测结果不理想时,用户还需要思考学习模型的设置是否存在问题,从而花费大量的时间在模型的调试和检查算法上,而且重新设置集成学习模型也要花费大量的时间,可以预测的是,当需要更换学习模型类型时,需要学习新的算法模型,所花费的时间更多。本申请实施例中,用户无需理解集成学习模型中的模型类型和超参数组合,只需要输入一个生成命令即可直接生成集成学习模型,进而对集成学习模型进行训练即可。用户无需输入复杂的函数和参数调节,就可以十分方便地生成集成学习模型,即使集成学习模型的训练和预测结果不理想,用户也不用思考是否由于输入的学习模型函数或者超参数配置不正确,只需要考虑训练样本即可。另外,当集成学习模型的训练或者预测结果不理想时,也可以很方便地重新生成集成学习模型。由于预设模型和超参数的配置均是随机的,因此生成的集成学习模型中的各个预设模型具有较好的差异性,使集成学习模型的预测结果更加准确。

[0086] 本申请一实施例中,参照图2所述,依次对集成学习模型中的各个预设模型的超参数进行配置,还包括用于记录各个预设模型超参数组合的超参数组合记录,步骤130具体包括以下步骤:

[0087] 步骤210,获取第一预设模型的第一配置信息;

[0088] 本步骤中,第一配置信息为第一预设模型在预设模型集合中注册时写入,第一配置信息记录了第一预设模型的各个超参数的类型及其取值范围。

[0089] 步骤220,根据配置信息,在超参数的取值范围内随机生成或通过网格搜索生成第一预设模型的第一超参数数值组合,第一超参数数值组合包括第一预设模型中各个超参数的取值。

[0090] 步骤230,将第一超参数数值组合与超参数组合记录中的数据进行对比,若第一超参数数值组合与超参数组合记录中已存在的数据重复,则返回步骤220,若第一超参数数值组合与超参数组合记录中已存在的数据没有重复,则执行步骤240。

[0091] 步骤240,根据第一超参数数值组合配置第一预设模型的超参数,并将第一超参数数值组合记录在超参数组合记录中。

[0092] 通过上述步骤210至步骤240,能够避免出现超参数一样的预设模型,使集成学习模型中的各个预设模型不相同,使集成学习模型对不同类型特征数据的覆盖范围更好。

[0093] 其中,对于超参数的取值范围内随机生成超参数数值组合的方式,如果超参数是离散值,则以 $1/p$ 的概率选取这个超参数值域里的一个值,其中 p 是值域里离散值的数量。例如,超参数为决策树的个数 n ,决策树的个数 n 的取值范围的数值域为 $\{5, 10, 15\}$, $P=3$,则决

策树的个数 n 在5,10,15中随机选择一个,即以1/3的概率选取这个超参数值域里的一个值。一方面,由于超参数值域是离散的,因此能够适配超参数离散值的要求,例如上述决策树的个数 n 不会出现1.5个的情况,另一方面,可以设置超参数值域的值,使其在模型设计者的预设优选数值中选择,即在设置配置信息时,可以根据需要设置允许随机的值,以便使当前预设模型的性能较佳,例如,对于一个决策树模型,模型设计者了解决策树数量为6时预设模型效果欠佳,在设置决策树的个数 n 的取值范围时,只在{1,2,3,4,5,7,8,9,10}中选择,这样就可以控制随机选择的结果,灵活性更好。

[0094] 如果超参数是连续值,以超参数的取值 v 为例,配置信息中 v 的取值范围最大值为 V_{\max} ,最小值为 V_{\min} ,为了操作方便,先将这个超参数数值的取值范围映射到0到1之间的范围。通过将 V 映射为

$V' = \frac{V - V_{\min}}{V_{\max} - V_{\min}}$,这样 V' 是一个0至1之间的数值。在随机选择超参数

时,首先在0至1之间随机生成一个数值 V' ,然后通过公式 $V' = \frac{V - V_{\min}}{V_{\max} - V_{\min}}$ 映射回原来的

尺度 V ,将 V 作为对应超参数的取值。本申请实施例中,由于将连续的取值范围映射到一个预设的区间,因此无需针对不同类型的超参数分别设置对应的随机算法,例如对于取值范围为0至10和取值范围为50.25至100.04的两个超参数,只用同样的随机数生成方法,均可获得对应的取值范围的随机数,这样用户在预设模型集合中注册预设模型时,无需反复额外对随机数的生成方法进行配置,只需要输入超参数的类型和取值范围即可。

[0095] 本申请一实施例中,参照图3所示,步骤140具体包括以下步骤:

[0096] 步骤310,获取第一训练数据集。

[0097] 第一训练数据集包括多个训练数据,训练可以在在实际的使用场景中采集,例如,使用场景中各个用户的特征样本数据及对应虚拟物品选择结果,其中特征样本数据包括用户的标签特征和行为特征,用户的标签特征指不随着用户的行为而变化的特征,例如用户的性别、年龄等特征,用户的行为特征指随着用户的行为而变化的特征,例如用户的游戏等级、游戏时间、已拥有的游戏装备等特征。

[0098] 步骤320,在所述第一训练数据集中随机抽取训练数据分配至各个所述预设模型进行训练。

[0099] 通过上述步骤310至步骤320,通过在所述第一训练数据集中随机抽取训练数据分配至各个所述预设模型进行训练,能为了增加各个预设模型的差异性,使训练出来的集成学习模型的预测结果更好。

[0100] 本申请在一实施例中,参照图4所示,步骤320具体包括以下步骤:

[0101] 步骤410,确定各个预设模型的训练数据量为第二数据量;

[0102] 本步骤中,预设模型的训练数据量为预设数据量或者为执行模型训练机器最大处理能力的数据量。

[0103] 例如,对一个机器,假设它能处理的子模型训练数据量是 m ,而总的训练数据量是 n ,其中, $m \leq n$,则训练数据里的每个数据被选入作为这个预设模型的训练数据的概率是 m/n 。因此,我们根据这个采样概率,为每个子模型生成训练数据。

[0104] 步骤420,计算各个所述预设模型对应的所述第二数据量与所述第一数据量的比

例,根据各个预设模型对应所述比例分别在所述第一训练数据集中随机抽取训练数据分配至对应的所述预设模型进行训练。

[0105] 例如,假设要生成 k 个预设模型。在进行训练数据采样时,对每个训练数据对应生成 k 个0到1之间的随机数,记为 r_i ,其中 $1 \leq i \leq k$ 。因为这 k 都是独立随机产生的,所以都是不同的随机数。如果随机数 $r_i \leq m/n$,则将这个数据作为当前预设模型的训练数据。

[0106] 通过上述步骤410和步骤420,能使各个预设模型的输入数据产生差异,这样能增加训练后各个预设模型之间的差异性,使训练出来的集成学习模型的预测结果更好。

[0107] 本申请一实施例中,参照图5所示,步骤150具体包括以下步骤:

[0108] 步骤510,配置次级学习模型,使训练好的所述至少两个预设模型的预测值输入至所述次级学习模型;

[0109] 其中,所述次级学习模型可以预先配置或者从所述预设模型集合中选择形成。或者,通过另一测试训练集对上述已经训练好的预设模型进行测试,选择预测结果较好的预设模型作为次级学习模型,需要说明的是上述训练好的预设模型已经配置好超参数,因此无需再对次级学习模型的超参数再次进行设置,由于各个预设模型是根据原始的训练样本数据中进行训练的,因此预测结果较好的预设模型表示对于训练样本数据的适配性更好,因此将该预设模型作为次级学习模型能进一步提高集成学习模型的预测准确性。

[0110] 步骤520,获取各个所述预设模型在训练过程中形成的预测值,根据所述预测值生成第二训练集;

[0111] 本步骤中,由于次级学习模型用于对预设模型的结果进行集成,因此需要将各个预设模型的训练结果作为第二训练集对次级学习模型进行训练。

[0112] 步骤530,根据所述第二训练集对所述次级学习模型进行训练。

[0113] 通过第二训练集对次级学习模型进行训练,训练完成后的次级学习模型能够处理各个预设模型输出的预测值。

[0114] 本申请一实施例中,参照图6所示,步骤520具体包括以下步骤:

[0115] 步骤610,对所述训练预测结果进行统计生成一种类型以上的统计数据,将所述统计数据和所述训练预测结果作为所述第二训练集。

[0116] 由于各个预设模块输出的预测值是相互独立的,各个预测值之间相互没有关联,而通过对各个预测模块的预测值进行统计。

[0117] 统计数据可以将各个预设模块输出的预测值进行数据统计,则可产生与各个预设模型关联的统计值,这样,在不增加预测模块个数的基础上能够增加第二训练集中训练数据的数量,能够提高集成学习模型的预测准确性。其中,所述统计数据的类型可以包括以下至少之一:平均值、方差、最小值、25%位数、中位数、75%位数和最大值。

[0118] 例如,集成学习模型中有 k 个预设模型,在对各个预设模型训练时,会产生 k 个预测值,对这 k 个预测值进行统计分析,可以得到这 k 个预测值之间的平均值、方差、最小值、25%位数、中位数、75%位数和最大值一共7个统计值,将 k 个预测值和7个统计值组合作为第二训练集,这样第二训练集中具有 $k+7$ 个特征值,通过这 $k+7$ 个特征值对次级学习模型进行训练,能够提高次级学习模型的结合能力和预测能力,进而提高整个集成学习模型的预测准确性。

[0119] 本申请一实施例中,预设模型集合可以集成任意类型的预设模型,当需要将一个

预设模型加入预设模型集合时,需要先进行预设模型的注册,以便后续可以正确地使用该模型。参照图7所示,注册一个预设模型,包括以下步骤:

[0120] 步骤710,输入训练函数`train()`;

[0121] 其中输入参数是由特征向量和标签值组成的训练数据,通常特征向量的每个元素是一个浮点数值,而标签是一个浮点数值,使训练函数输出一个预设模型。

[0122] 步骤720,输入预测函数 `predict()`:输入参数是由特征向量组成的测试数据,输出是一个浮点数值预测值。

[0123] 步骤730,输入配置信息,即超参数集合 `hyperparameters`,配置信息用于配置预设模型的超参数,超参数是模型训练之前设定的参数,配置信息包括预设模型的每个超参数的类型及其取值范围。超参数的类型包括离散型和连续型,其中离散型表示超参数的取值在多个预设值之间进行选择,例如神经网络中的权重在2、4、6中选择,连续型表示超参数在取值范围中的取值是连续值。

[0124] 可以通过一个数据结构将预设模型的上述3个部分存储在预设模型集合中,完成一个预设模型的注册。那么,后续操作中,集成学习模型可以根据预设模型的注册信息来生成一个预设模型的超参数,从而可以进行预设模型的训练和预测。

[0125] 参照图8所示,为本申请实施例提供的集成学习模型的模块框架图。以下结合图8简单说明本申请集成学习模型的训练方法的训练过程。如图8所示,集成学习模型包括模型注册器和模型超参数生成器,其中模型注册器用于形成或更新预设模型集合,模型超参数生成器用于生成预设模型根据配置信息配置各个预设模型的超参数。

[0126] 首先,用户通过模型注册器注册预设模型,形成或更新预设模型集合,预设模型集合中包括多种不同学习算法类型的预设模型。然后模型超参数生成器在预设模型集合中随机选择预设数量或随机数量的预设模型作为集成学习模型的子模型,模型超参数生成器根据各个预设模型的配置信息在预设范围内生成各个预设模型的超参数数值组合,模型超参数生成器根据各个超参数数值组合配置各个预设模型的超参数,这样,就形成了集成学习模型中的各个子模型 M_1, M_2, \dots, M_k 。然后采集第一训练数据集 D ,其中第一训练数据集 D 包括用户的特征样本数据及用户选择或购买虚拟物品的记录的标签值。通过数据采集器在第一训练数据集 D 中随机抽取对应各个子模型的训练数据 D_1, D_2, \dots, D_k 。通过训练数据 D_1, D_2, \dots, D_k 分别对各个子模型 M_1, M_2, \dots, M_k 进行训练。子模型 M_1, M_2, \dots, M_k 训练完成后,利用测试数据集对子模型 M_1, M_2, \dots, M_k 进行测试。将测试结果最好的子模块作为次级学习模型。通过混合特征生产器获取各个子模型 M_1, M_2, \dots, M_k 的预测值,混合特征生成器将对各个子模块的预测值进行统计,得到多个统计值,混合特征生成器将各个子模块的预测值和统计值混合生成混合特征数据 D_x ,将混合特征数据 D_x 作为第二训练集对次级学习模型进行训练,完成了对集成学习模型的训练。

[0127] 参照以下附图,本申请实施例提供一种虚拟物品的推荐方法可应用于终端中,也可应用于服务器中,还可以是运行于终端或服务器中的软件。在一些实施例中,可以是应用在Spark:分布式大数据处理平台中。在一些实施例中,终端可以是智能手机、平板电脑、笔记本电脑、台式计算机或者智能手表等;服务器可以配置成独立的物理服务器,也可以配置成多个物理服务器构成的服务器集群或者分布式系统,还可以配置成提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、CDN以及

大数据和人工智能平台等基础云计算服务的云服务器;软件可以是训练集成学习模型的应用程序等,但并不局限于以上形式。

[0128] 参照图9所示,本申请实施例提供了一种虚拟物品的推荐方法,包括以下步骤:

[0129] 步骤910,获取目标用户的特征数据;

[0130] 本步骤中,目标用户为当前场景下虚拟物品的推荐对象,例如在游戏结束后的场景,针对目标用户向其推荐选择或购买的虚拟物品,在利用本申请实施例提供的推荐方法进行推荐时,需要获取目标用户的特征数据,进而才能给出合适的推荐。其中特征数据包括用户的标签特征和行为特征,用户的标签特征指不随着用户的行为而变化的特征,例如用户的性别、年龄等特征,用户的行为特征指随着用户的行为而变化的特征,例如用户的游戏等级、游戏时间、已拥有的游戏装备等特征。

[0131] 步骤920,将所述特征数据分别输入集成学习模型中的多个预设模型,对所述目标用户的特征数据进行虚拟物品推荐预测,得到每个所述预设模型的预测值;

[0132] 本步骤中,所述预设模型为所述集成学习模型预先注册的机器学习模型且已配置好超参数。获取了目标用户的特征数据之后,将用户的特征数据输入至集成学习模型中进行推荐预测,其中,集成学习模型为根据图1至图7中任意一个实施例的集成学习模型的训练方法所生成的。目标用户的特征数据分别输入至集成学习模型中的多个预设模型并行地进行虚拟物品推荐预测,每个预设模型均给出一个预测值,因此,能够得到多个预测值。预测值可以是数值0至1表示的概率,概率表示用户选择对应虚拟物品的概率,概率越高表示用户选择对应虚拟物品的可能性越大,例如一个预测模型中,衣服道具的预测值为概率0.9,鞋子道具的预测值为概率0.6,则用户选择衣服道具的概率更高,衣服道具更符合用户的需求,另外预测值也可以用分数表示,例如分数越高表示用户选择该虚拟物品的可能性越大。

[0133] 步骤930,对多个所述预测值进行统计处理得到所述预测值的分布特征,根据所述分布特征得到所述目标用户的虚拟物品推荐结果;或者,将多个所述预测值输入经过训练的次级学习模型进行计算,得到所述目标用户的虚拟物品推荐结果;

[0134] 多个预设模型给出的预测值构成一个预测数据集,一实施例中,对预测数据集中的预测值进行整合得到虚拟物品推荐结果,其中,可以采用集成计算的方式进行整合,集成计算指对各个预设模型输出的预测值集中起来进行数学运算,例如,对各个预设模型的预测值进行统计处理得到所述预测值的分布特征,根据分布特征得到目标用户的虚拟物品推荐结果,例如,可以采用投票的方式进行整合,即采用各个预测值中预测数量最多的虚拟物品作为虚拟物品推荐结果。又或者计算各个预测值的平均值或加权平均值等统计值作为作为虚拟物品推荐结果。

[0135] 另一实施例中,也可以采用二级学习的方式,通过次级学习模型对各个预设模型的预测值进行计算得出虚拟物品推荐结果。其中次级学习模型采用如图5所示的步骤510至560或图6所示步骤610的训练方法进行训练。

[0136] 本申请的一个实施例中,通过对预测数据集中的多个预测值进行统计运算,得到混合特征,混合特征包括各个预设模型输出的预测值和根据各个预测值的分布特征得到的统计值,将预测值和统计值的集合作为混合特征。将混合特征输入次级学习模型进行计算,得到所述目标用户的虚拟物品推荐结果。其中,推荐的虚拟物品可以是一个,也可以是多

个,当推荐结果需要多个时,只需要选择预测值较准确的多个预测值对应的多个虚拟物品进行推荐即可。

[0137] 参照图10所示,为训练好的集成学习模型的模块框图,以下结合图10来说明本申请实施例提供的虚拟物品的推荐方法。参照图10所示,集成学习模型包括多个子模型 M_1, M_2, \dots, M_k ,子模型为图8所示的模型超参数生成器在预设模型集合中随机选择预设模型而生成的。当需要向目标用户推荐虚拟物品时,通过数据采集器采集目标用户的特征数据A,将所述特征数据A分别发送至各个子模型 M_1, M_2, \dots, M_k 进行推荐预测,每个子模型 M_1, M_2, \dots, M_k 各自得出一个预测值,这时,通过混合特征生成器对各个子模型的预测值进行统计计算,得到混合特征 A_x ,其中混合特征包括各个子模块推荐的预测值和对这些预测值进行统计计算的统计值,统计值包括以下至少之一:平均值、方差、最小值、25%位数、中位数、75%位数和最大值。将混合特征输入 A_x 到已经训练好的次级学习模型进行计算,得到目标用户的虚拟物品推荐结果。

[0138] 由于集成学习模型包括多个不同类型不同超参数配置的预设模型,因此能够覆盖目标用户的各种类型特征数据,集成学习模型能够较好应用于虚拟物品推荐的场景,具有较高的准确性。同时由于通过本申请提供的训练方法训练的集成学习模型包括多个不同类型的预设模型,因此能够充分地突出每种学习算法类型的优点,并且通过其他类型的预设模型弥补该预设模型的缺点,通过集成学习模型能够向目标用户准确地推荐所需的虚拟物品。通过本申请实施例提供的虚拟物品的推荐方法能够向目标用户准确地推荐所需的虚拟物品。另外,由于预设模型集合中的预设模型是预先注册的,因此只需要一次注册好预设模型集合,即可用于生成不同预设模型组合的集成学习模型,用户在构建集成学习模型时无需对集成学习模型的算法的集成进行深入研究及逐一进行配置集成,即可方便生成并训练集成学习模型。另外,即使集成学习模型的训练和预测结果不理想,也可以十分方便地重新生成集成学习模型,大大节省了构建集成学习模型的时间。

[0139] 以下结合图11说明本申请实施例提供的虚拟物品的推荐方法的使用场景,图11所示,为游戏应用的一个场景,该场景为游戏结束后向目标用户提供可折扣购买虚拟物品的界面1110,包括虚拟物品清单及其价格,在本游戏场景中,利用本申请实施例提供的虚拟物品的推荐方法,采集目标用户的特征数据(例如性别、年龄,在线时长、等级、对局数量、胜率等),并计算得到推荐的虚拟物品,根据计算结果,得出最准确的多个预测值并找出各个预测值对应的虚拟物品,按照准确程度的优先级排序在虚拟物品界面中显示,如图11所示,计算得到3个预测值并找到对应虚拟物品在虚拟物品的界面1110中显示,方便用户选购。

[0140] 通过将本申请实施例提供的虚拟物品的推荐方法与现有的单一学习算法模型的预测结果在法在历史数据上进行训练和测试,其中训练和测试数据的比例是8:2。参见表1所示,通过利用MSE (Mean Squared Error) 来评测这些模型的表现。

模型	MSE
梯度提升树模型 (GDBT)	0.063553
随机森林模型 (Random Forest)	0.060109
本申请实施例的推荐 算法	0.041923

[0141]

[0142] 表1:本申请实施例的推荐方法在与其他算法的预测对比表

[0143] 表1中,MSE越小,表示模型预测值离真实值的差别越小,也就是模型的表现越好。表1展示了这3个推荐算法的测试效果,可以得出本申请实施例提供的虚拟物品的推荐方法表现都好于其他两个模型,其中本申请实施例提供的虚拟物品的推荐方法比最好的对比方法随机森林模型(Random Forest)相对提升了30.26%。

[0144] 另外,在游戏场景中采用A/B Test的方式对以上的三个模型进行测试。其中,A/B Test采用随机的方式将每个用户以1/3的概率分配给一个模型,因此每个模型得到1/3的用户量。我们通过购买率来衡量这些模型的线上表现,其中购买率的计算为购买的用户量除以曝光的用户量。表2展示了这3个推荐算法的线上表现,其中本申请实施例提供的虚拟物品的推荐方法的购买率最高,并且随机森林模型(Random Forest)的购买率相对提升了30.27%。

模型	购买率
梯度提升树模型 (GDBT)	0.0035945
随机森林模型 (Random Forest)	0.0036555
本申请实施例的推荐 算法	0.0047621

[0145] 表 2:本申请实施例的推荐方法的测试结果表

[0146] 参照图12所示,本申请实施例提供了一种虚拟物品的推荐装置,可以实现上述虚拟物品的推荐方法的实施例,该装置包括:

[0147] 数据获取装置1210,获取目标用户的特征数据;

[0148] 预测装置1220,将所述特征数据分别输入集成学习模型中的多个预设模型,对所述目标用户的特征数据进行虚拟物品推荐预测,得到每个所述预设模型的预测值,所述预设模型为所述集成学习模型预先注册的机器学习模型且已配置好超参数;

[0149] 集成装置1230,对多个所述预测值进行统计处理得到所述预测值的分布特征,根据所述分布特征得到所述目标用户的虚拟物品推荐结果;或者,将多个所述预测值输入经过训练的次级学习模型进行计算,得到所述目标用户的虚拟物品推荐结果;

[0150] 其中,所述集成学习模型中包括多种学习算法类型的预设模型,所述预设模型的超参数根据配置信息生成。

[0151] 参照图13所示,本申请实施例提供了一种集成学习模型的生成装置,可以实现上述集成学习模型的训练方法的实施例,该装置包括:

[0152] 选择装置1310,在预先注册的预设模型集合中选择至少两个预设模型;

[0153] 配置信息获取装置1320,获取各个所述预设模型的配置信息;

[0154] 超参数配置装置1330,根据各个所述预设模型的配置信息配置各个所述预设模型的超参数;

[0155] 训练装置1340,获取训练数据,根据所述训练数据对各个配置好超参数的所述预设模型进行训练,所述训练数据包括用户的特征样本数据和所述用户的虚拟物品选择结果;

[0156] 集成装置1350,将各个被训练的所述预设模型进行集成,完成集成学习模型的训

练。

[0158] 参照图14所示,本申请实施例中提供一种计算机设备,包括:

[0159] 至少一个存储器;

[0160] 至少一个处理器;

[0161] 至少一个程序;

[0162] 所述程序被存储在所述存储器中,所述处理器执行所述至少一个所述程序以执行上述各个实施例的虚拟物品的推荐方法、集成学习模型的训练方法。

[0163] 参照图15所示,本申请实施例中的装置可以被配置为服务器的形式,如图15所示,其为本申请实施例提供的服务器1500的结构图,服务器1500可因配置或性能不同而产生比较大的差异,可以包括一个或一个以上中央处理器(Central Processing Units,CPU) 1522(例如,一个或一个以上处理器)和存储器1532,一个或一个以上存储应用程序1542或数据1544的存储介质1530(例如一个或一个以上海量存储装置)。其中,存储器1532和存储介质1530可以是短暂存储或持久存储。存储在存储介质1530的程序可以包括一个或一个以上模块(图示没标出),每个模块可以包括对服务器中的一系列指令操作。更进一步地,中央处理器1522可以设置为与存储介质1530通信,在服务器1500上执行存储介质1530中的一系列指令操作。

[0164] 服务器1500还可以包括一个或一个以上电源1526,一个或一个以上有线或无线网络接口1550,一个或一个以上输入输出接口1558,和/或,一个或一个以上操作系统1541,例如Windows Server™,Mac OS X™,Unix™,Linux™,FreeBSD™等等。

[0165] 服务器中的处理器可以用于执行上述各个实施例的虚拟物品的推荐方法、集成学习模型的训练方法。

[0166] 本申请实施例还提供一种计算机可读存储介质,所述计算机可读存储介质用于存储程序代码,所述程序代码用于执行上述各个实施例的虚拟物品的推荐方法、集成学习模型的训练方法。

[0167] 本领域普通技术人员可以理解,上文中所公开方法中的全部或某些步骤、系统可以被实施为软件、固件、硬件及其适当的组合。某些物理组件或所有物理组件可以被实施为由处理器,如中央处理器、数字信号处理器或微处理器执行的软件,或者被实施为硬件,或者被实施为集成电路,如专用集成电路。这样的软件可以分布在计算机可读介质上,计算机可读介质可以包括计算机存储介质(或非暂时性介质)和通信介质(或暂时性介质)。如本领域普通技术人员公知的,术语计算机存储介质包括在用于存储信息(诸如计算机可读指令、数据结构、程序模块或其他数据)的任何方法或技术中实施的易失性和非易失性、可移除和不可移除介质。计算机存储介质包括但不限于RAM、ROM、EEPROM、闪存或其他存储器技术、CD-ROM、数字多功能盘(DVD)或其他光盘存储、磁盒、磁带、磁盘存储或其他磁存储装置,或者可以用于存储期望的信息并且可以被计算机访问的任何其他的介质。此外,本领域普通技术人员公知的是,通信介质通常包含计算机可读指令、数据结构、程序模块或者诸如载波或其他传输机制之类的调制数据信号中的其他数据,并且可包括任何信息递送介质。

[0168] 以上是对本申请的较佳实施进行了具体说明,但本申请并不局限于上述实施方式,熟悉本领域的技术人员在不违背本申请精神的前提下还可作出种种的等同变形或替换,这些等同的变形或替换均包含在本申请权利要求所限定的范围内。

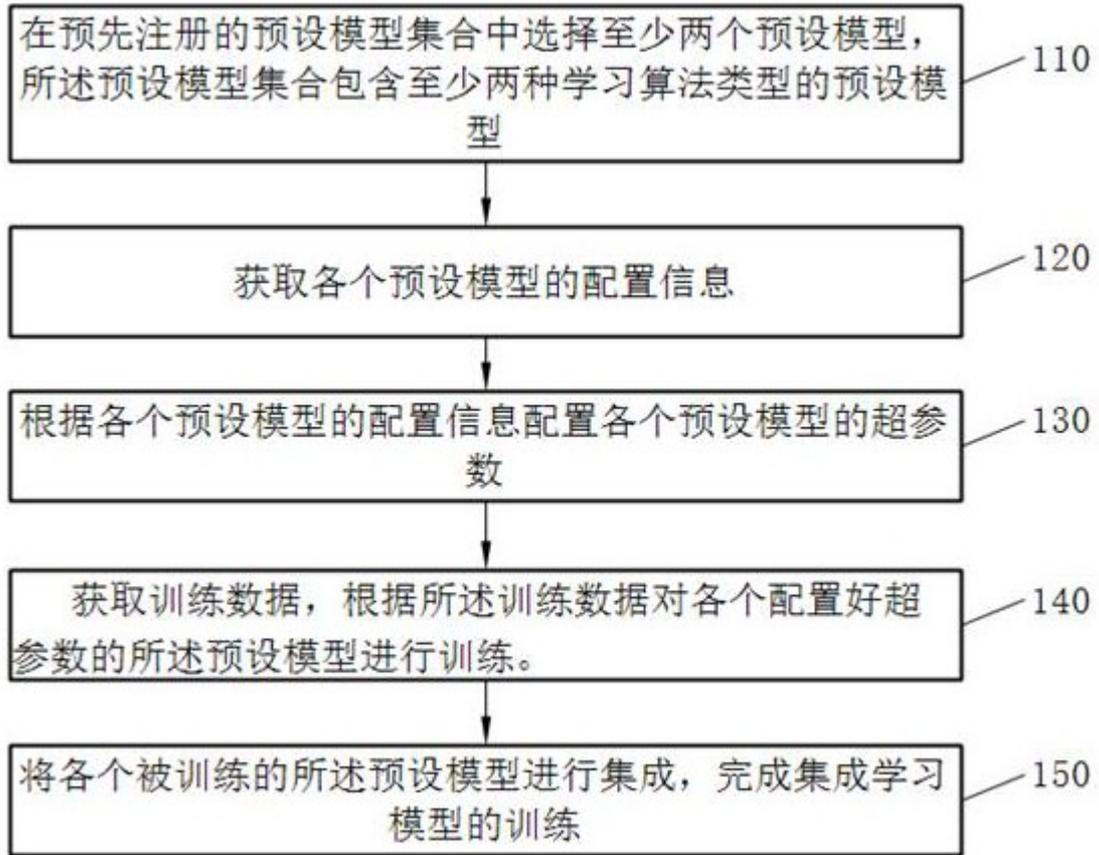


图1

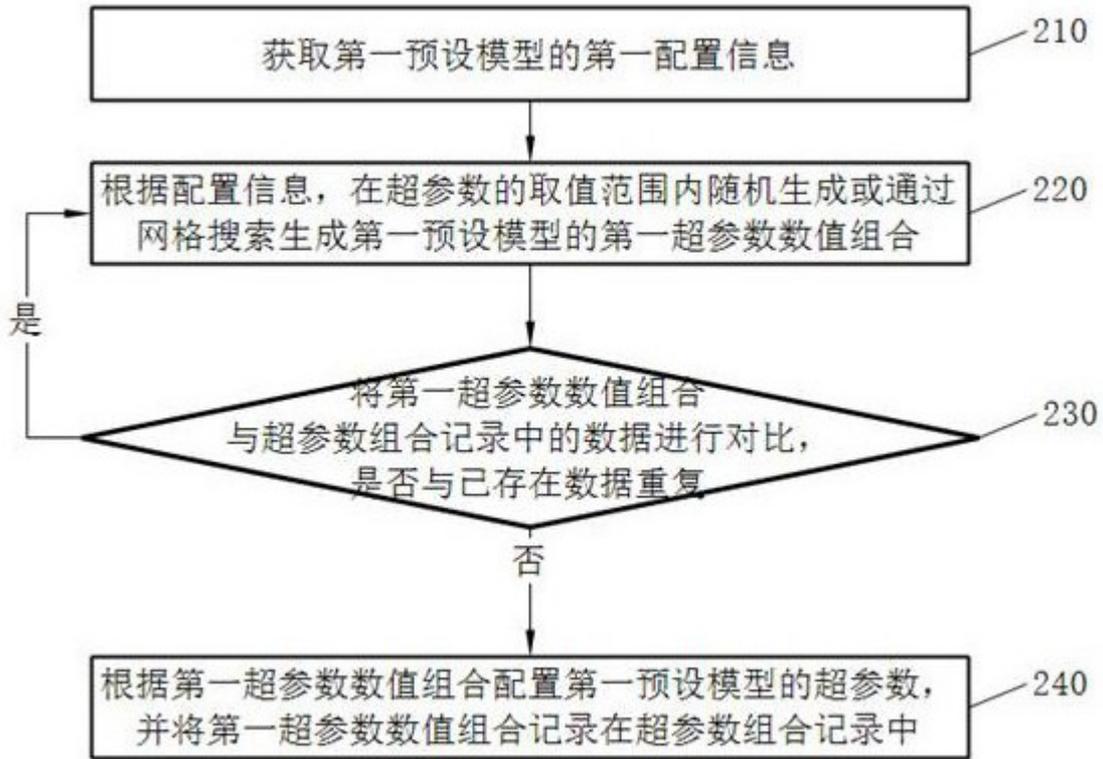


图2



图3

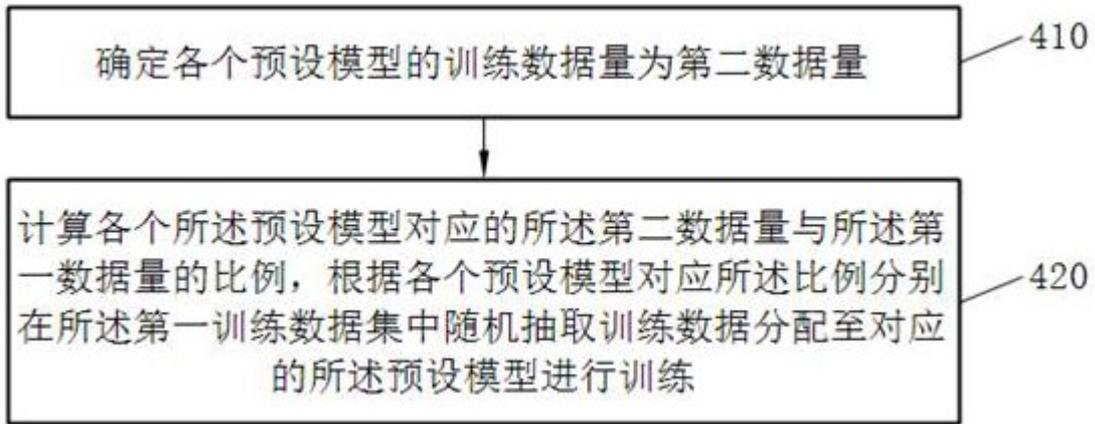


图4

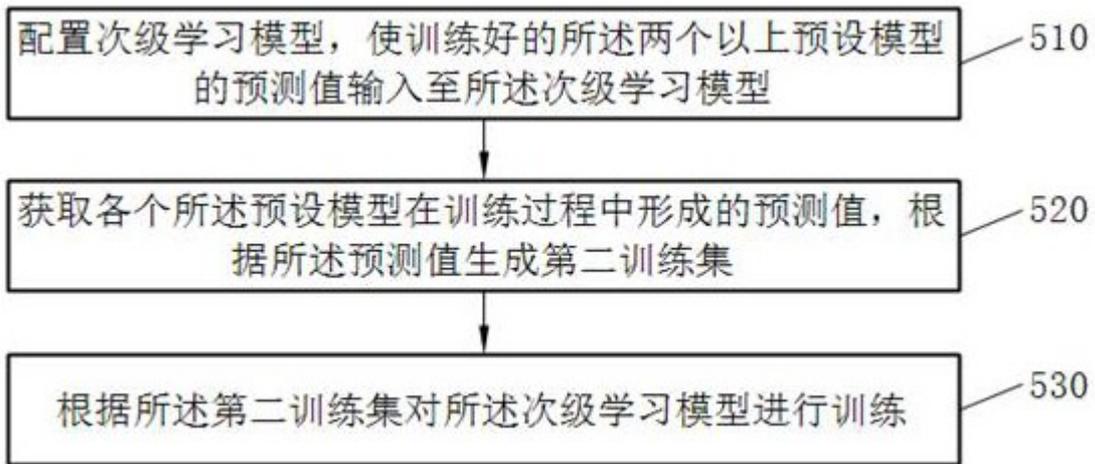


图5

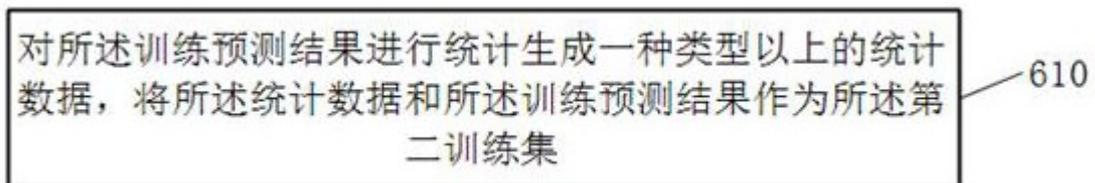


图6

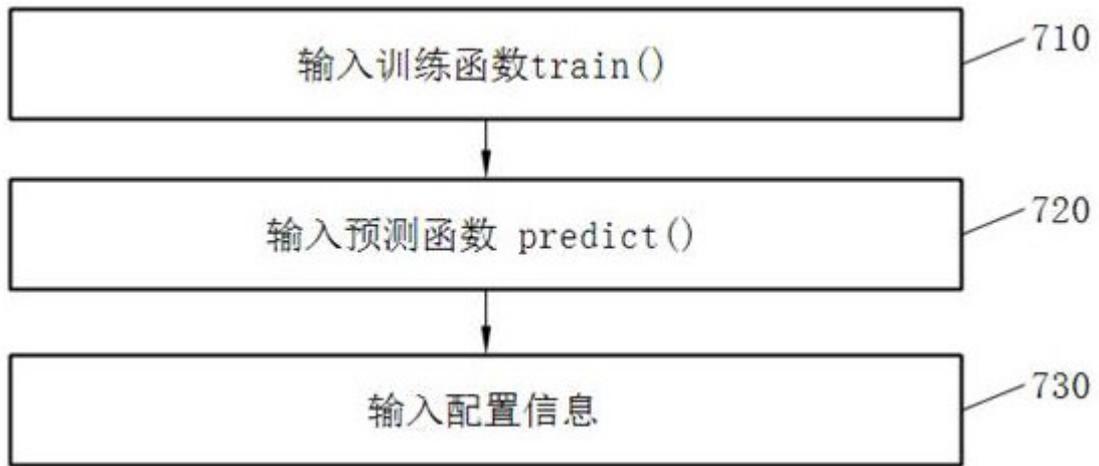


图7

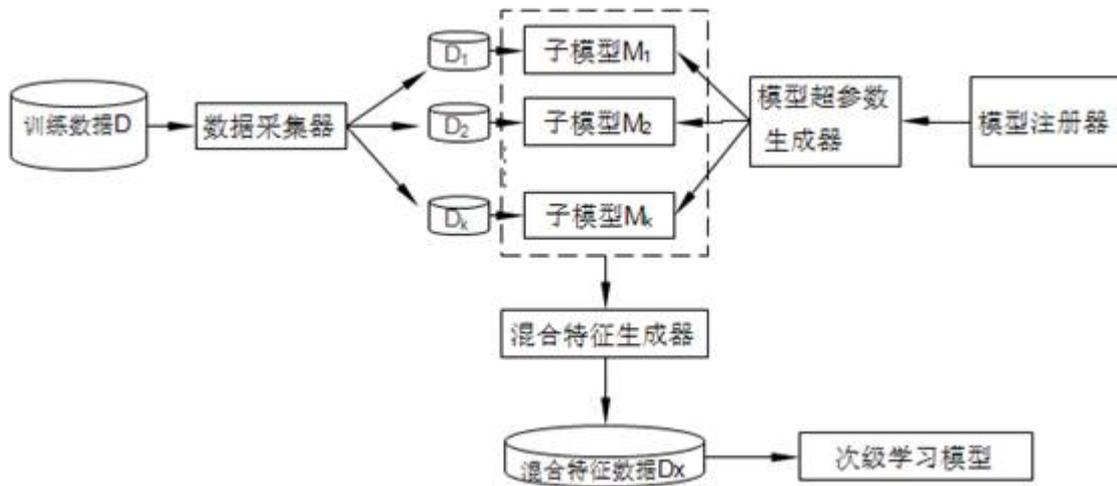


图8

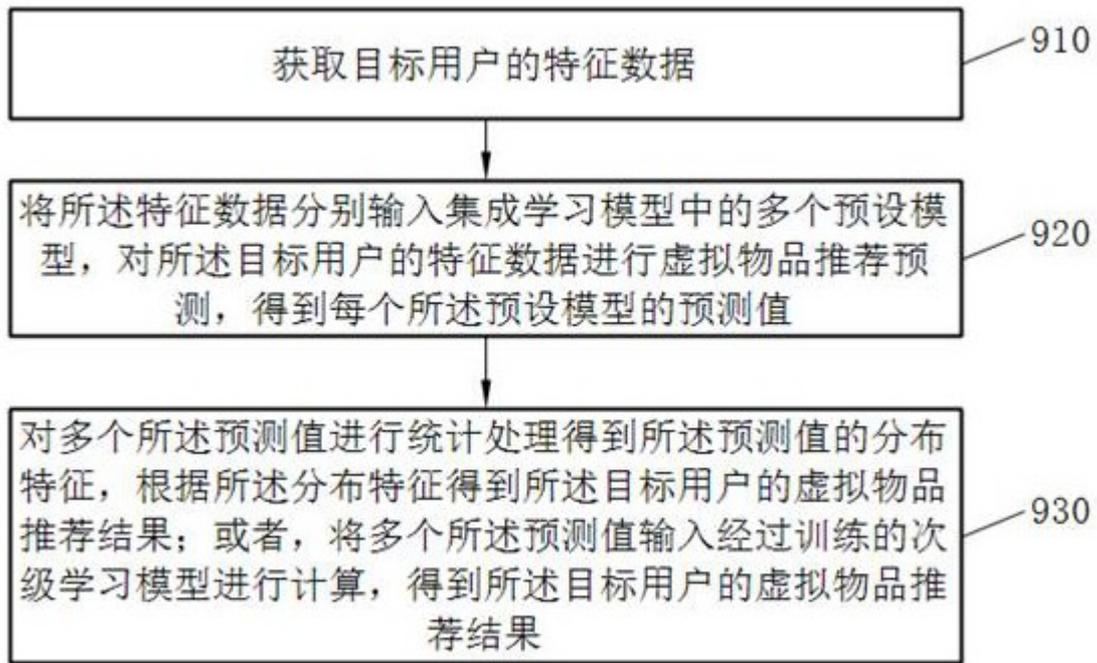


图9

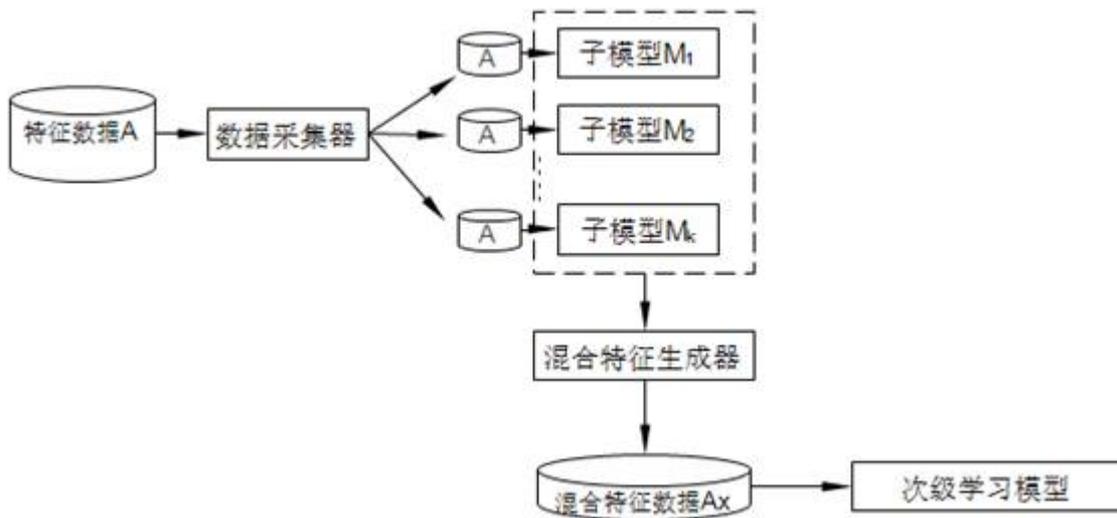


图10

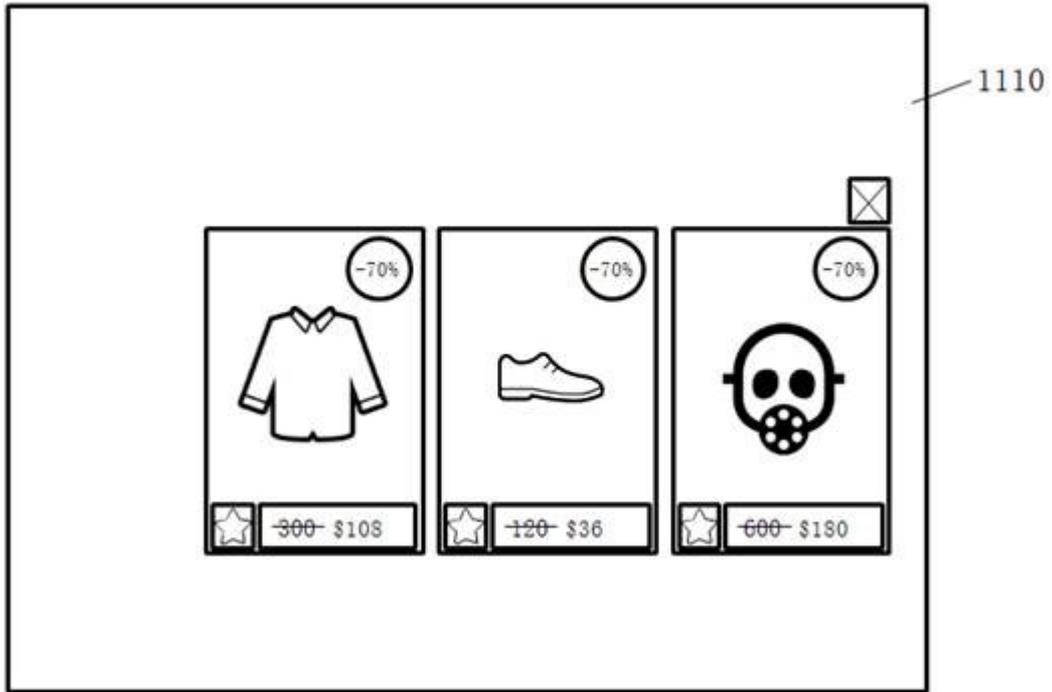


图11

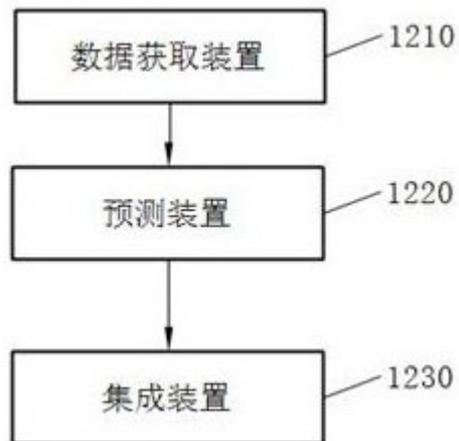


图12

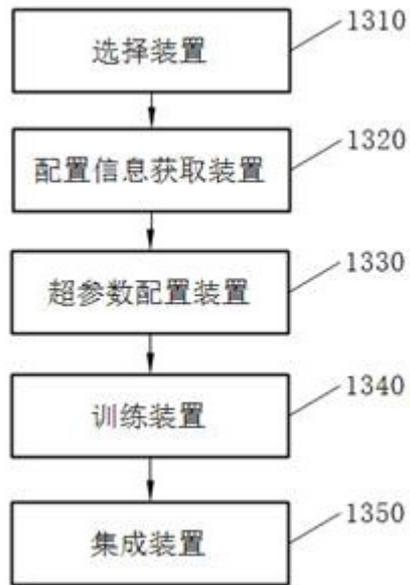


图13

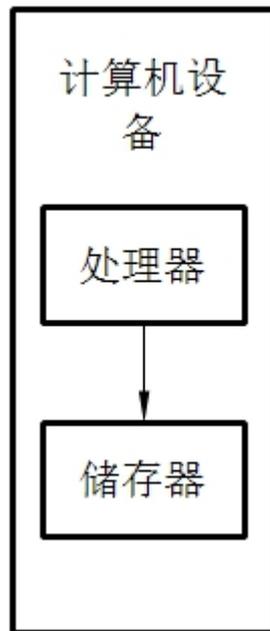


图14

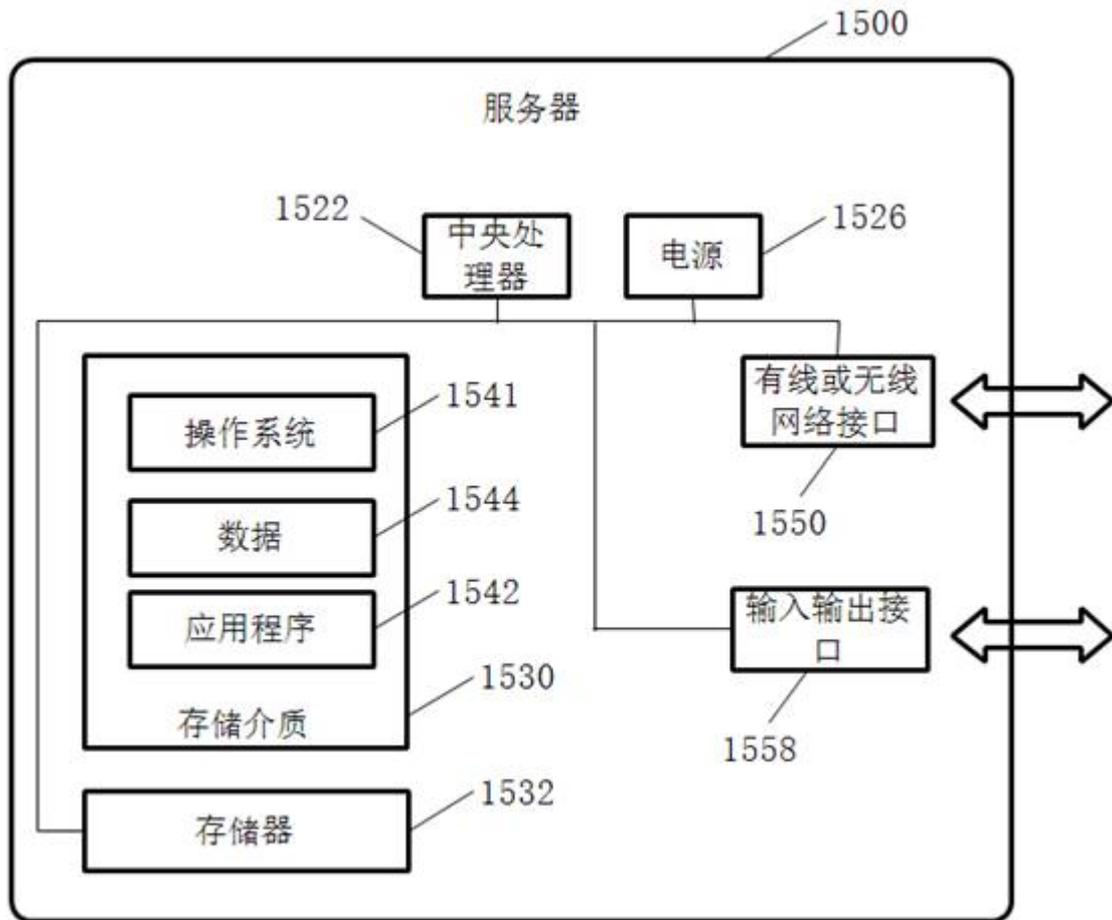


图15