



(12) 发明专利申请

(10) 申请公布号 CN 106295832 A

(43) 申请公布日 2017. 01. 04

(21) 申请号 201510239620. 5

(22) 申请日 2015. 05. 12

(71) 申请人 阿里巴巴集团控股有限公司

地址 英属开曼群岛大开曼资本大厦一座四
层 847 号邮箱

(72) 发明人 王义

(74) 专利代理机构 北京国昊天诚知识产权代理
有限公司 11315

代理人 许志勇

(51) Int. Cl.

G06Q 10/04(2012. 01)

G06Q 30/02(2012. 01)

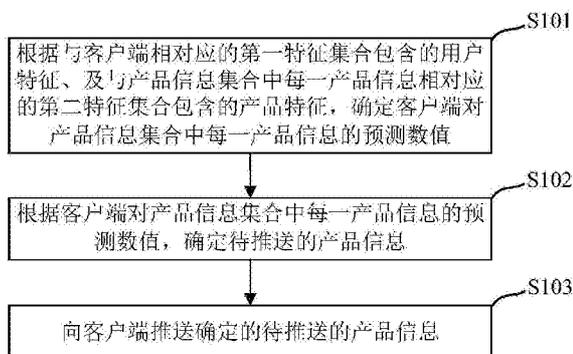
权利要求书3页 说明书13页 附图2页

(54) 发明名称

产品信息推送方法及装置

(57) 摘要

本申请实施例公开了一种产品信息推送方法及装置,该方法包括:根据与客户端相对应的第一特征集合包含的用户特征、及与产品信息集合中每一产品信息相对应的第二特征集合包含的产品特征,确定客户端对产品信息集合中每一产品信息的预测数值;其中,所述预测数值表征该客户端通过行为操作获取某产品信息的可能性;根据所述客户端对产品信息集合中每一产品信息的预测数值,确定待推送的产品信息;向所述客户端推送确定的待推送的产品信息。本申请实施例可以确保向用户推送的产品信息与用户的真实兴趣相契合,从而提升产品信息推送的准确性。



1. 一种产品信息推送方法,其特征在于,包括:

根据与客户端相对应的第一特征集合包含的用户特征、及与产品信息集合中每一产品信息相对应的第二特征集合包含的产品特征,确定所述客户端对产品信息集合中每一产品信息的预测数值;其中,所述预测数值表征该客户端通过行为操作获取某产品信息的可能性;

根据所述客户端对产品信息集合中每一产品信息的预测数值,确定待推送的产品信息;

向所述客户端推送确定的待推送的产品信息。

2. 如权利要求 1 所述的方法,其特征在于,所述确定所述客户端对产品信息集合中每一产品信息的预测数值,包括:

利用预先获得的预测模型计算所述客户端对产品信息集合中每一产品信息的预测数值;其中,所述预测模型是通过机器训练获得的。

3. 如权利要求 2 所述的方法,其特征在于,所述方法还包括:

获取用以机器训练的包含用户特征、产品特征及点击率的样本数据;

基于所述样本数据进行机器训练,获得所述预测模型。

4. 如权利要求 3 所述的方法,其特征在于,基于所述样本数据进行机器训练,获得所述预测模型,包括:

以与所述用户特征相对应的特征值、及与所述产品特征相对应的特征值为所述预测模型的自变量,并以与点击率相对应的变换数值为因变量进行机器训练,以获得所述预测模型;或,

以与所述用户特征相对应的特征值、与所述产品特征相对应的特征值、及与组合特征相对应的特征值为所述预测模型的自变量,并以与点击率相对应的变换数值为因变量进行机器训练,以获得所述预测模型;其中,所述组合特征是所述用户特征与所述产品特征的组合。

5. 如权利要求 4 所述的方法,其特征在于,所述确定所述客户端对产品信息集合中每一产品信息的预测数值,包括:

根据与所述客户端相对应的第一特征集合包含的用户特征、及与所述产品信息相对应的第二特征集合包含的产品特征,获得与每一用户特征相对应的特征值及与每一产品特征相对应的特征值;

将与每一用户特征相对应的特征值及与每一产品特征相对应的特征值作为所述预测模型的自变量,计算所述预测模型的因变量的值;

对计算得到的因变量的值作相应处理,以得到该客户端对该产品信息的预测数值。

6. 如权利要求 4 所述的方法,其特征在于,所述确定所述客户端对产品信息集合中每一产品信息的预测数值,包括:

根据与所述客户端相对应的第一特征集合包含的用户特征、及与所述产品信息相对应的第二特征集合包含的产品特征,获得与每一用户特征相对应的特征值及与每一产品特征相对应的特征值;

根据所述第一特征集合包含的用户特征及第二特征集合包含的产品特征,获得与每一组合特征相对应的特征值;

将与每一用户特征相对应的特征值、与每一产品特征相对应的特征值、及每一组合特征的特征值作为所述预测模型的自变量,计算所述预测模型的因变量的值;

对计算得到的因变量的值作相应处理,以得到该客户端对该产品信息的预测数值。

7. 如权利要求 1 所述的方法,其特征在于,所述确定所述客户端对产品信息集合中每一产品信息的预测数值之前,所述方法还包括:

根据所述客户端发送的搜索请求所携带的关键词,确定与该客户端相匹配的产品信息集合;或,

根据与客户端相对应的第一特征集合包含的用户特征,确定与该客户端相匹配的产品信息集合。

8. 如权利要求 1 所述的方法,其特征在于,根据所述客户端对产品信息集合中每一产品信息的预测数值,确定待推送的产品信息,包括:

根据所述客户端对产品信息集合中每一产品信息的预测数值,从所述产品信息集合中选取至少一个待推送的产品信息;则,

向所述客户端推送确定的待推送的产品信息,包括:

向所述客户端推送选取的至少一个待推送的产品信息。

9. 如权利要求 1 所述的方法,其特征在于,根据所述客户端对产品信息集合中每一产品信息的预测数值,确定待推送的产品信息,包括:

将所述产品信息集合中的产品信息按照所述预测数值从大到小进行排序;则,

向所述客户端推送确定的待推送的产品信息,包括:

向所述客户端推送排序后的所述产品信息。

10. 一种产品信息推送装置,其特征在于,包括:

计算模块,用于根据与客户端相对应的第一特征集合包含的用户特征、及与产品信息集合中每一产品信息相对应的第二特征集合包含的产品特征,确定所述客户端对产品信息集合中每一产品信息的预测数值;其中,所述预测数值表征该客户端通过行为操作获取某产品信息的可能性;

确定模块,用于根据所述客户端对产品信息集合中每一产品信息的预测数值,确定待推送的产品信息;

推送模块,用于向所述客户端推送确定的待推送的产品信息。

11. 如权利要求 10 所述的装置,其特征在于,所述计算模块用于:

利用预先获得的预测模型计算所述客户端对产品信息集合中每一产品信息的预测数值;其中,所述预测模型是通过机器训练获得的。

12. 如权利要求 11 所述的装置,其特征在于,所述装置还包括:

样本数据获取单元,用于获取用以机器训练的包含用户特征、产品特征及点击率的样本数据;

机器训练单元,用于基于所述样本数据进行机器训练,获得所述预测模型。

13. 如权利要求 12 所述的装置,其特征在于,所述机器训练单元用于:

以与所述用户特征相对应的特征值、及与所述产品特征相对应的特征值为所述预测模型的自变量,并以与点击率相对应的变换数值为因变量进行机器训练,以获得所述预测模型;或,

以与所述用户特征相对应的特征值、与所述产品特征相对应的特征值、及与组合特征相对应的特征值为所述预测模型的自变量,并以与点击率相对应的变换数值为因变量进行机器训练,以获得所述预测模型;其中,所述组合特征是所述用户特征与所述产品特征的组合。

14. 如权利要求 13 所述的装置,其特征在于,所述计算模块包括:

特征值获取单元,用于根据与所述客户端相对应的第一特征集合包含的用户特征、及与所述产品信息相对应的第二特征集合包含的产品特征,获得与每一用户特征相对应的特征值及与每一产品特征相对应的特征值;

模型计算单元,用于将与每一用户特征相对应的特征值及与每一产品特征相对应的特征值作为所述预测模型的自变量,计算所述预测模型的因变量的值;

预测数值计算单元,用于对计算得到的因变量的值作相应处理,以得到该客户端对该产品信息的预测数值。

15. 如权利要求 13 所述的装置,其特征在于,所述计算模块包括:

第一特征值获取单元,用于根据与所述客户端相对应的第一特征集合包含的用户特征、及与所述产品信息相对应的第二特征集合包含的产品特征,获得与每一用户特征相对应的特征值及与每一产品特征相对应的特征值;

第二特征值获取单元,用于根据所述第一特征集合包含的用户特征及第二特征集合包含的产品特征,获得与每一组合特征相对应的特征值;

模型计算单元,用于将与每一用户特征相对应的特征值、与每一产品特征相对应的特征值、及每一组合特征的特征值作为所述预测模型的自变量,计算所述预测模型的因变量的值;

预测数值计算单元,用于对计算得到的因变量的值作相应处理,以得到该客户端对该产品信息的预测数值。

16. 如权利要求 10 所述的装置,其特征在于,所述装置还包括产品信息集合确定单元,用于:

根据所述客户端发送的搜索请求所携带的关键词,确定与该客户端相匹配的产品信息集合;或,

根据与客户端相对应的第一特征集合包含的用户特征,确定与该客户端相匹配的产品信息集合。

17. 如权利要求 10 所述的装置,其特征在于,所述确定模块用于:

根据所述客户端对产品信息集合中每一产品信息的预测数值,从所述产品信息集合中选取至少一个待推送的产品信息;则,

所述推送模块用于:

向所述客户端推送选取的至少一个待推送的产品信息。

18. 如权利要求 10 所述的装置,其特征在于,所述确定模块用于:

将所述产品信息集合中的产品信息按照所述预测数值从大到小进行排序;则,

所述推送模块用于:

向所述客户端推送排序后的所述产品信息。

产品信息推送方法及装置

技术领域

[0001] 本申请涉及互联网技术领域,特别涉及一种产品信息推送方法及装置。

背景技术

[0002] 互联网技术中,通过互联网向用户推送产品信息是当前热点技术之一。用户可以通过浏览器或应用等各种方式访问网站服务商或应用服务商提供的各种页面,而网站服务商或应用服务商则可以通过将待推送的产品信息推送到用户终端所访问的各种页面上,来实现产品信息推广功能。

[0003] 例如,向用户推送的是产品信息是金融产品信息,网站服务商或应用服务商一般会预先生成所需推送的金融产品信息,在用户访问该网站服务商或应用服务商提供的各种页面时,将上述金融产品信息推送到上述各种页面上,以供用户查看。

[0004] 现有技术中,网站服务商或应用服务商通常会预先生成固定的产品信息集合,并向每一个访问该网站服务商或应用服务商提供的各种页面的用户推送上述固定的产品信息集合,来供每个用户查看上述产品信息集合中的各个产品信息。例如生成的产品信息集合是:{A1,A2,A3,A4},则向访问上述页面的各个用户逐条推送该产品信息集合中的各个产品信息 A1、A2、A3、A4。

[0005] 在实现本申请的过程中,发明人发现现有技术至少存在以下问题:

[0006] 对于某个产品信息而言,可能只是部分用户人群对该产品信息较为感兴趣,或对该产品信息有购买需求。若向每一个访问上述页面的用户推送上述固定的产品信息集合包含的各产品信息,难以确保向用户推送的产品信息与该用户的真实兴趣相契合,从而造成产品信息推送的精确性不高。

发明内容

[0007] 本申请实施例的目的是提供一种产品信息推送方法及装置,以解决现有技术存在的产品信息推送精确性不高的问题。

[0008] 为解决上述技术问题,本申请实施例提供的产品信息推送方法及装置是这样实现的:

[0009] 一种产品信息推送方法,包括:

[0010] 根据与客户端相对应的第一特征集合包含的用户特征、及与产品信息集合中每一产品信息相对应的第二特征集合包含的产品特征,确定所述客户端对产品信息集合中每一产品信息的预测数值;其中,所述预测数值表征该客户端通过行为操作获取某产品信息的可能性;

[0011] 根据所述客户端对产品信息集合中每一产品信息的预测数值,确定待推送的产品信息;

[0012] 向所述客户端推送确定的待推送的产品信息。

[0013] 一种产品信息推送装置,包括:

[0014] 计算模块,用于根据与客户端相对应的第一特征集合包含的用户特征、及与产品信息集合中每一产品信息相对应的第二特征集合包含的产品特征,确定所述客户端对产品信息集合中每一产品信息的预测数值;其中,所述预测数值表征该客户端通过行为操作获取某产品信息的可能性;

[0015] 确定模块,用于根据所述客户端对产品信息集合中每一产品信息的预测数值,确定待推送的产品信息;

[0016] 推送模块,用于向所述客户端推送确定的待推送的产品信息。

[0017] 由以上本申请实施例提供的技术方案可见,根据与每个客户端相对应的用户特征以及各个产品信息相对应的产品特征,确定该客户端对各个产品信息的预测数值。由于该预测数值用以表征某客户端通过用户的行为操作获取某产品信息的可能性,若根据该预测数值来确定应该向客户端推送怎样的产品信息,可以确保向客户端所推送的产品信息与该客户端的用户的真实兴趣相契合,从而提升产品信息推送的准确性。

附图说明

[0018] 为了更清楚地说明本申请实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要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申请中记载的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图。

[0019] 图 1 为本申请实施例中产品信息推送方法的流程图;

[0020] 图 2 为本申请实施例中产品信息推送方法中离线训练的流程图;

[0021] 图 3 为本申请第一实施例中产品信息推送装置的模块图;

[0022] 图 4 为本申请第二实施例中产品信息推送装置的模块图;

[0023] 图 5 为本申请第三实施例中产品信息推送装置的模块图。

具体实施方式

[0024] 为了使本技术领域的人员更好地理解本申请中的技术方案,下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都应当属于本申请保护的范围。

[0025] 图 1 为本申请实施例中产品信息推送方法的流程图。值得说明的是,本实施例的方法的执行主体可以是服务器,该服务器可以是单个服务器或多个服务器组成的服务器群组。则上述产品信息推送方法包括如下步骤:

[0026] S101:根据与客户端相对应的第一特征集合包含的用户特征、及与产品信息集合中每一产品信息相对应的第二特征集合包含的产品特征,确定所述客户端对产品信息集合中每一产品信息的预测数值;其中,所述预测数值表征该客户端通过行为操作获取某产品信息的可能性。

[0027] 上述客户端可以是电脑、手机、个人数字助理(PDA)、智能手表等。上述用户特征也称为用户画像,是与使用客户端的用户相对应的特征数据,上述产品特征是与产品营销商

所提供的各个产品信息相对应的特征数据。每个客户端的用户均对应于一个包括至少一个用户特征的第一特征集合,每个产品信息也对应于一个包括至少一个产品特征的第二特征集合,上述第一特征集合和第二特征集合被存放于相应的数据库中。上述产品信息包括但不限于:金融产品信息、或保险产品信息、或楼盘信息、或健身产品信息等。

[0028] 在实际使用过程中,可以通过客户端的 IP 地址、MAC 地址、或用户在该客户端上登陆的账户信息来获取与该用户相对应的第一特征集合,账户信息包括在某网页上登陆的账户信息或在某应用(Application, APP)上登陆的账户信息。

[0029] 本申请实施例中,上述确定得到的预测数值可以是所述客户端对产品信息集合中每一产品信息的点击率,该点击率是通过特定的计算方法来预测到的,通常,某个客户端对某个产品信息的点击率越高,表明该客户端对该产品信息的关注度越高,或者说该产品信息与该客户端的用户的真实兴趣相契合;反之,则表明客户端对该产品信息的关注都越低。该预测得到的点击率可以表征该客户端通过行为操作获取某产品信息的可能性,所述行为操作包括浏览、点击、购买等行为。值得一提的是,上述预测数值还可以是成交率(购买数量/浏览量)。

[0030] 为了能够全面地构建用户的用户特征(用户画像),本申请实施例构建用户画像数据的过程如下:

[0031] 从静态和动态两个维度,可以刻画用户的静态特性数据和动态特征数据。上述静态特征数据是指用户在某网站或客户端应用上的注册信息,例如,构建的静态特征数据是:{用户性别、年龄、受教育年限、最高学历、家庭住址、毕业院校}。上述动态特征数据是根据用户在某个时间段的互联网行为数据所构建的一类特征数据,例如,上述互联网行为数据是各个用户在某个金融产品网站上浏览、点击或购买等行为数据,则构建的动态特征数据是:{用户购买力、用户品牌偏好、用户风险偏好、用户投资经验值、用户偏好理财类型}。该类动态特征数据会根据用户在某个时间段的互联网行为数据作统计分析,从而得到该用户在该时间段所对应的上述动态特征数据。以一个月的时间为例,采集各个用户在每个月的互联网行为数据作统计分析,以获得各个用户在每个月对应的动态特征数据。

[0032] 以某个金融产品网站为例,若定义该金融产品网站为站内,该金融产品网站之外的其他网站为站外,则从站内和站外两个维度来构建上述用户画像数据,可以得到更为全面的用户画像数据。具体地,一般获取用户在站内的注册数据,以及用户在站内对相关产品的浏览、点击、购买等行为数据,通过上述注册数据及站内的行为数据,可以比较准确地获得各个用户的背景信息以及用户的消费习惯信息。同时,用户在站外的其他网站上的购买商品、点击广告、浏览视频等行为数据也可以作为用户画像数据的一个有效补充。此外,通过站外方式还可以获得其他用户画像数据,例如:用户的房产信息,用户家庭状况等。同时用户的移动终端可以作为获取用户画像数据的一个重要手段,例如:手机品牌可以衡量用户的购买力,用户经常出现的消费场所以及工作场所可以作为用户消费潜力,个人背景的衡量手段,等等。

[0033] 本申请实施例中,上述步骤 S101 具体包括:通过机器训练获得的预测模型,计算所述客户端对产品信息集合中每一产品信息的预测数值。

[0034] 上述预测模型可以是逻辑回归(Logistic Regression, LR)模型、线性回归模型、一般线性回归模型等。本实施例中,若上述预测模型是逻辑回归模型,则该逻辑回归模型

是：

$$[0035] \quad g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \beta_{n+1} y_1 + \beta_{n+2} y_2 + \dots + \beta_{n+m} y_m;$$

[0036] 其中, $\{x_1, x_2, \dots, x_n\}$ 是与第一特征集中各个用户特征相对应的特征值; $\{y_1, y_2, \dots, y_m\}$ 是与第二特征集中各个产品特征相对应的特征值。 $\{x_1, x_2, \dots, x_n\}$ 、 $\{y_1, y_2, \dots, y_m\}$ 是该逻辑回归模型的自变量, $\{\beta_1, \beta_2, \dots, \beta_{n+m}\}$ 是该逻辑回归模型的回归系数, β_0 是常数项, $g(x)$ 是该逻辑回归模型的因变量。

[0037] 若定义上述客户端对产品信息集中每一产品信息的预测数值为 $f(x)$, 该预测数值 $f(x)$ 越大, 表示客户端对某产品信息的关注度越高 (或购买某产品的可能性越大), 该预测数值 $f(x)$ 越小, 表示客户端对某产品信息的关注度越低 (购买某产品的可能性越小)。由于该预测数值 $f(x)$ 的取值介于 0 到 1 之间, 不太适合作为逻辑回归模型的因变量, 故对上述 $f(x)$ 作 Logistic 变换得到变换数值, 并将该变换数值作为逻辑回归模型的因变量 $g(x)$, 则：

$$[0038] \quad g(x) = \text{logit}(f(x)) = \ln\left(\frac{f(x)}{1-f(x)}\right);$$

$$[0039] \quad f(x) = \frac{e^{g(x)}}{1+e^{g(x)}};$$

[0040] 上述逻辑回归模型是通过特定的机器训练来获得的, 关于具体怎样机器训练的过程将在下文进行详细叙述。基于上述逻辑回归模型, 对于任意一个客户端和任意一个产品信息, 在获取到该客户端的用户的第二特征集中包含的若干用户特征, 及与该产品信息对应的第二特征集中包含的若干产品特征后, 即可以通过上述逻辑回归模型, 计算出该客户端对该产品信息的预测数值 $f(x)$, 以通过该 $f(x)$ 表征该客户端的用户购买该产品信息的可能性。

[0041] 值得一提的是, 上述变换数值是一个大于 1 的数值, 是通过 Logistic 函数进行 Logistic 变换来获得的, 当然, 本领域普通技术人员能够想到的是, 本申请实施例获得该变换数值的方式并不限于上述 Logistic 变换。

[0042] 本申请一实施例中, 上述步骤 S101 具体包括：

[0043] a) 根据与所述客户端相对应的第一特征集合包含的用户特征、及与所述产品信息相对应的第二特征集合包含的产品特征, 获得与每一用户特征相对应的特征值及与每一产品特征相对应的特征值。

[0044] 举例而言, 假设第一特征集合包括三个用户特征, 如下：

[0045] { 用户性别、用户购买力、用户风险偏好 }；

[0046] 假设第二特征集合包括三个产品特征, 如下：

[0047] { 产品最低额度、产品收益水平、产品风险指数 }；

[0048] 基于上述用户特征及产品特征, 用户性别可以分为“男性”及“女性”, 假设用户购买力分为三个等级: ¥0 ~ 50000、¥50001 ~ 200000、¥200001 以上, 假设用户风险偏好分为两个等级: 高、低; 假设产品最低额度分为三个等级: ¥0 ~ 1000、¥1001 ~ 10000、¥10000 以上, 假设产品收益水平 (以年收益为例) 分为三个等级: 3% ~ 5%、5% ~ 7%、7% 以上, 假设产品风险指数分为三个等级: 高、中、低。

[0049] 则上述第一特征集合的用户特征及第二特征集合中的产品特征与特征值的对应

关系如表一所示（特征值的取值是 0 或 1）：

[0050] 表一：

[0051]

用户特征/产品特征		特征值
用户性别	男性	$x_1=1$
	女性	$x_1=0$
用户购买力	¥0~50000	$x_2=1, x_3=0$
	¥50001~200000	$x_2=0, x_3=1$
	¥200001 以上	$x_2=0, x_3=0$
用户风险偏好	高	$x_4=1$
	低	$x_4=0$
产品最低额度	¥0~1000	$y_1=1, y_2=0$
	¥1001~10000	$y_1=0, y_2=1$
	¥10000 以上	$y_1=0, y_2=0$
产品收益水平	3%~5%	$y_3=1, y_4=0$
	5%~7%	$y_3=0, y_4=1$
	7%以上	$y_3=0, y_4=0$
产品风险指数	高	$y_5=1, y_6=0$
	中	$y_5=0, y_6=1$
	低	$y_5=0, y_6=0$

[0052] 如表一所示,对于第一特征集合中的每一个用户特征及第二特征集合中的每一个产品特征,均可以将上述特征转化成相应的特征值来表示。值得一提的是,上述用户特征或产品特征均是以离散的区间或值进行划分,在本申请其他实施例中,若上述用户特征或产品特征是连续的数值,则也可以依据以上方式获得与各个数值相对应的特征值,在此不再予以赘述。

[0053] b) 将与每一用户特征相对应的特征值及与每一产品特征相对应的特征值作为所述预测模型的自变量,计算所述预测模型的因变量的值。

[0054] 以表一为例,对于任意一个客户端的第一特征集合及任意一个产品信息的第二特征集合,均可以得到表一中的各个特征值 $\{x_1, x_2, x_3, x_4, y_1, y_2, y_3, y_4, y_5, y_6\}$,将得到的各个特征值 $\{x_1, x_2, x_3, x_4, y_1, y_2, y_3, y_4, y_5, y_6\}$ 作为自变量代入上文提及的预测模型:

$$[0055] \quad g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \beta_{n+1} y_1 + \beta_{n+2} y_2 + \dots + \beta_{n+m} y_m; (n = 4, m = 6)$$

[0056] 至此,可以通过计算得到因变量 $g(x)$ 的值。

[0057] c) 将计算得到的因变量的值作 Logistic 变换,以得到该客户端对该产品信息的预测数值。

[0058] 具体地,将上述步骤 b 得到的因变量 $g(x)$ 的值代入以上 Logistic 变换公式:

$$[0059] \quad f(x) = \frac{e^{g(x)}}{1 + e^{g(x)}}$$

[0060] 则通过计算可以得出一个 0 到 1 之间的 $f(x)$ 的值,至此,该 $f(x)$ 的值即是该客户

端对该产品信息的预测数值（购买某产品的概率）。

[0061] 在本申请实施例中,为了进一步提升上述预测模型的泛化能力（即误差更小）,还引入组合特征的概念。组合特征是指第一特征集合中的用户特征与第二特征集合中的产品特征的组合,该组合的形式可以是一个用户特征与一个产品特征进行交叉组合,或一个用户特征与多个产品特征进行交叉组合,或多个用户特征与一个产品特征进行交叉组合等。

[0062] 继续以表一为例,若分别将第一特征集合中的任意一个用户特征与第二特征集合中的任意一个产品特征进行交叉组合,则可以得到若干个组合特征,如下:

[0063] {用户性别&产品最低额度、用户性别&产品收益水平、用户性别&产品风险指数、用户购买力&产品最低额度、用户购买力&产品收益水平、用户购买力&产品风险指数、用户风险偏好&产品最低额度、用户风险偏好&产品收益水平、用户风险偏好&产品风险指数};

[0064] 基于上述组合特征,上述步骤 S101 具体包括:

[0065] d) 根据与所述客户端相对应的第一特征集合包含的用户特征、及与所述产品信息相对应的第二特征集合包含的产品特征,获得与每一用户特征相对应的特征值及与每一产品特征相对应的特征值。该步骤 d 与上述步骤 a 一致。

[0066] e) 根据所述第一特征集合包含的用户特征及第二特征集合包含的产品特征,获得与每一组合特征相对应的特征值。

[0067] 与每一组合特征相对应的特征值如下表二（未一一列出）所示:

[0068] 表二:

[0069]

用户特征&产品特征		特征值
用户性别&产品最低额度	男性 & ¥0~1000	$z_1=0, z_2=0, z_3=0$
	男性 & ¥1001~10000	$z_1=0, z_2=0, z_3=1$
	男性 & ¥10000 以上	$z_1=0, z_2=1, z_3=0$
	女性 & ¥0~1000	$z_1=1, z_2=0, z_3=0$
	女性 & ¥1001~10000	$z_1=1, z_2=1, z_3=0$
	女性 & ¥10000 以上	$z_1=0, z_2=1, z_3=1$
.....

[0070] f) 将与每一用户特征相对应的特征值、与每一产品特征相对应的特征值、及每一组合特征的特征值作为所述预测模型的自变量,计算所述预测模型的因变量的值。

[0071] 若引入组合特征的概率,上文提及的预测模型如下:

[0072]
$$g(x) = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n + \beta_{n+1} y_1 + \dots + \beta_{n+m} y_m + \beta_{n+m+1} z_1 + \dots + \beta_{n+m+p} z_p;$$

[0073] 则依次代入与每个用户特征相对应的特征值 $\{x_1, x_2, \dots, x_n\}$,与每个产品特征相对应的特征值 $\{y_1, y_2, \dots, y_m\}$,以及与各个组合特征相对应的特征值 $\{z_1, z_2, \dots, z_p\}$,可以计算出上述预测模型的因变量 $g(x)$ 的值。

[0074] h) 将计算得到的因变量的值作 Logistic 变换,以得到该客户端对该产品信息的

预测数值。

[0075] 具体地,将上述步骤 f 得到的因变量 $g(x)$ 的值代入以上 Logistic 变换公式:

$$[0076] \quad f(x) = \frac{e^{g(x)}}{1+e^{g(x)}}$$

[0077] 则通过计算可以得出一个 0 到 1 之间的 $f(x)$ 的值,至此,该 $f(x)$ 的值即是该客户端对该产品信息的预测数值(购买某产品的概率)。

[0078] S102:根据所述客户端对产品信息集合中每一产品信息的预测数值,确定待推送的产品信息。

[0079] 基于一个产品信息集合中每一产品信息的预测数值,服务器可以确定待推送到相应的客户端上的产品信息。

[0080] 本申请实施例中,对于每个客户端的用户而言,当产品信息库中的产品信息量特别大时,为减小机器运算负担,均会确定一个该客户端的用户可能感兴趣的产品信息集合,则确定产品信息集合的步骤包括:

[0081] 根据所述客户端发送的搜索请求所携带的关键词,确定与该客户端相匹配的产品信息集合。

[0082] 如果用户在某个客户端上通过网页的方式或应用 APP 的方式输入一个关键词,随后该客户端将包含该关键词的搜索请求发送至服务器端,服务器端便根据用户在客户端输入的关键词匹配到相应的产品信息集合。例如:用户输入的关键词是“基金”,则服务器通过搜索引擎技术即可得到相应的产品信息集合是“A 基金、B 基金、C 基金、...”。

[0083] 在本申请另一实施例中,确定产品信息集合的步骤包括:

[0084] 根据与客户端相对应的第一特征集合包含的用户特征,确定与该客户端相匹配的产品信息集合。

[0085] 通过预先构建的用户特征(用户画像)数据库,可以得到可以反映用户较为感兴趣的产品类型。例如:第一特征集合中包含的用户特征“用户偏好理财类型”是:基金、互联网金融,则产品信息集合是“A 基金、B 基金、C 基金、...互联网金融 D、互联网金融 F、...”。

[0086] 值得一提的是,确定上述产品信息集合的方式并不受限于此。

[0087] 本申请实施例中,确定待推送的产品信息的步骤 S102 具体包括:

[0088] 将所述产品信息集合中的产品信息按照所述预测数值从大到小进行排序。

[0089] 例如:

[0090] 产品信息集合是:{A 基金、B 基金、C 基金},当前的客户端对上述各个产品信息的预测数值分别是:0.22、0.67、0.48。则按照预测数值从大到小进行排序后,上述产品信息集合中的产品信息的次序是:B 基金、C 基金、A 基金。

[0091] 当然,在本申请另一实施例中,确定待推送的产品信息的步骤 S102 具体包括:

[0092] 根据所述客户端对产品信息集合中每一产品信息的预测数值,从产品信息集合中选取至少一个待推送的产品信息。

[0093] 例如:

[0094] 产品信息集合是:{A 基金、B 基金、C 基金},当前的客户端对上述各个产品信息的预测数值分别是:0.22、0.67、0.48。若预先设定的规则是预测数值大于 0.5 的产品信息是所需推送的产品信息,则按照该规则可选取到的待推送的产品信息是:B 基金、C 基金。

[0095] S103 :向所述客户端推送确定的待推送的产品信息。

[0096] 基于上述步骤 S102,服务器向当前的客户端推送所确定到的待推送的产品信息。
例如 :

[0097] 上述产品信息集中的产品信息的次序是 :B 基金、C 基金、A 基金。则将产品信息集中的产品信息按照上述次序推送到客户端上,通常将预测数值最大的产品信息显示于搜索结果列表的最前端,以方便用户快速查看到该产品信息。又如 :按照该规则可选取到的待推送的产品信息是 :B 基金、C 基金。则服务器只会将 B 基金、C 基金推送到客户端上。

[0098] 值得说明的是,上述产品信息可以是推送到用户在客户端所浏览的网页上、或以邮件方式推送到用户的特定邮箱中、或推送到用户在客户端(如手机)上安装的某个应用 APP(如金融产品应用 APP)上。服务器可以在获知用户登陆某个应用 APP、网页的个人账户后触发该信息推送动作、或在接收到用户在某个应用 APP、网页上输入的搜索关键词后触发该信息推送动作。

[0099] 接下来,介绍通过机器训练得到上述预测模型的过程。图 2 为本申请实施例中产品信息推送方法中离线训练的流程图,离线训练过程包括 :

[0100] S201 :获取用以机器训练的包含用户特征、产品特征及点击率的样本数据。

[0101] 基于预先构建的用户特征(用户画像)库及产品特征库,以及用户在某网站(如金融产品网站)上的产品的浏览、点击、购买等互联网行为数据,可以整理得到原始数据,该原始数据的形式是用户+产品+浏览量(page view, PV)+点击量Click+购买数量Trans :
例如,得到的原始数据如表三所示 :

[0102] 表三 :

[0103]

用户	产品	浏览量	点击量	购买数量
用户 A	产品 D	34	6	1
用户 B	产品 D	60	15	2
用户 C	产品 D	46	9	0
用户 A	产品 E	12	1	1
用户 B	产品 E	5	0	0
用户 C	产品 E	33	5	2
...

[0104] 假设用户的第一特征集合是 $\{X1, X2, X3\}$, 产品的第二特征集合是 $\{Y1, Y2, Y3\}$, 则对于每一条原始数据均可以得到最终的样本数据 :

[0105] $\{X1, X2, X3, Y1, Y2, Y3, X1\&Y1, X1\&Y2, X1\&Y3, X2\&Y1, X2\&Y2, X2\&Y3, X3\&Y1, X3\&Y2, X3\&Y3, PV, Click, Trans\}$ 。

[0106] S202 :基于样本数据进行机器训练,获得上述预测模型。

[0107] 若采取逻辑回归模型,则以与用户特征相对应的特征值、及与产品特征相对应的特征值为所述预测模型的自变量,并以与点击率相对应的变换数值为因变量进行机器训练,以获得所述逻辑回归模型。

[0108] 值得说明的是,由于一般成交数据比较稀疏,上述逻辑回归模型的目标是用户到特定产品的点击率,在此过程中,将购买数量折算成点击数量,从而保证在点击率相同的情况下,购买数量大的产品的预测数值越高。则,

[0109] 点击率 = (点击量 + 购买数量 * N) / 浏览量。通常,上述 N 可以取值为 10 (经验值)。例如:在表三中,用户 A 对产品 D 的浏览量是 34,点击量是 6,购买数量是 1,则用户 A 对产品 D 的点击率 = $(6+1*10)/34 = 0.47$ 。

[0110] 通过样本数据的机器训练,可以计算得到逻辑回归模型(如极大似然):

[0111] $g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \beta_{n+1} y_1 + \beta_{n+2} y_2 + \dots + \beta_{n+m} y_m$ 中各个各个回归系数 $\{\beta_1, \beta_2, \dots, \beta_{n+m}\}$ 及常数项 β_0 的值。

[0112] 至此,完成预测模型的线下训练,并将预测模型文件加载到服务器端。

[0113] 综上所述,本申请实施例的方法根据与每个客户端相对应的用户特征以及与各个产品信息相对应的产品特征,确定该客户端对各个产品信息的预测数值。由于该预测数值表征客户端的用户购买某个产品的可能性,若根据上述可以表征用户购买各个产品的可能性的预测数值,来确定应该向客户端推送怎样的产品信息,可以确保向用户推送的产品信息与该用户的真实兴趣相契合,从而提升产品信息推送的准确性。此外,通过以上实施例的技术方案,可以解决现有技术中向用户盲目推送产品信息所带来的营销效率低的问题,并且现有技术盲目推送产品信息还会造成服务器端的资源耗用大,网络数据量大的问题,本申请实施例也可以改善这一现象。

[0114] 值得述及的是,在上述实施例中,所预测到的预测数值是一个与点击率相关的值,当某个客户端的用户对某个产品信息的预测点击率越高时,则表明该客户端的用户会购买该产品信息的可能性会越高,至少也可以表明该客户端的用户对该产品更为感兴趣。当然,在其他实施例中,也可以通过离线训练获得的预测模型来预测某客户端的用户对某产品的成交率(购买数量/浏览量)。

[0115] 接下来,将基于同样的原理,介绍本申请实施例提供的一种产品信息推送装置。

[0116] 图 3 为本申请第一实施例中产品信息推送装置的模块图。本实施例中,所述产品信息推送装置包括:

[0117] 计算模块 101,用于根据与客户端相对应的第一特征集合包含的用户特征、及与产品信息集合中每一产品信息相对应的第二特征集合包含的产品特征,确定所述客户端对产品信息集合中每一产品信息的预测数值;其中,所述预测数值表征该客户端通过行为操作获取某产品信息的可能性;

[0118] 确定模块 102,用于根据所述客户端对产品信息集合中每一产品信息的预测数值,确定待推送的产品信息;

[0119] 推送模块 103,用于向所述客户端推送确定的待推送的产品信息。

[0120] 本申请实施例中,所述计算模块 101 用于:

[0121] 通过机器训练获得的预测模型,计算所述客户端对产品信息集合中每一产品信息的预测数值。

[0122] 图 4 为本申请第二实施例中产品信息推送装置的模块图。该产品信息推送装置包括线下训练装置及线上推送装置,其中,线上推送装置包括计算模块 101、确定模块 102 及推送模块 103,线下训练装置包括:

[0123] 样本数据获取单元 201,用于获取用以机器训练的包含用户特征、产品特征及点击率的样本数据;

[0124] 机器训练单元 202,用于基于样本数据进行机器训练,获得预测模型。

[0125] 本申请实施例中,所述机器训练单元 202 用于:

[0126] 以与用户特征相对应的特征值、及与产品特征相对应的特征值为预测模型的自变量,并以与点击率相对应的变换数值为因变量进行机器训练,以获得预测模型;或,

[0127] 以与用户特征相对应的特征值、与产品特征相对应的特征值、及与组合特征相对应的特征值为预测模型的自变量,并以与点击率相对应的变换数值为因变量进行机器训练,以获得预测模型;其中,所述组合特征是用户特征与产品特征的组合。

[0128] 图 5 为本申请第三实施例中产品信息推送装置的模块图。本实施例中,上述计算模块 101 具体包括:

[0129] 特征值获取单元 1011,用于根据与客户端相对应的第一特征集合包含的用户特征、及与产品信息相对应的第二特征集合包含的产品特征,获得与每一用户特征相对应的特征值及与每一产品特征相对应的特征值;

[0130] 模型计算单元 1012,用于将与每一用户特征相对应的特征值及与每一产品特征相对应的特征值作为所述预测模型的自变量,计算所述预测模型的因变量的值;

[0131] 预测数值计算单元 1013,用于对计算得到的因变量的值作相应处理,以得到该客户端对该产品信息的预测数值。其中,所述处理可以是采取 Logistic 函数进行处理。

[0132] 本申请其他实施例中,所述计算模块 101 包括:

[0133] 第一特征值获取单元,用于根据与所述客户端相对应的第一特征集合包含的用户特征、及与所述产品信息相对应的第二特征集合包含的产品特征,获得与每一用户特征相对应的特征值及与每一产品特征相对应的特征值;

[0134] 第二特征值获取单元,用于根据所述第一特征集合包含的用户特征及第二特征集合包含的产品特征,获得与每一组合特征相对应的特征值;

[0135] 模型计算单元,用于将与每一用户特征相对应的特征值、与每一产品特征相对应的特征值、及每一组合特征的特征值作为所述预测模型的自变量,计算所述预测模型的因变量的值;

[0136] 预测数值计算单元,用于对计算得到的因变量的值作相应处理,以得到该客户端对该产品信息的预测数值。其中,所述处理可以是采取 Logistic 函数进行处理。

[0137] 本申请其他实施例中,所述装置还包括产品信息集合确定单元,用于:

[0138] 根据客户端发送的搜索请求所携带的关键词,确定与该客户端相匹配的产品信息集合;或,

[0139] 根据与客户端相对应的第一特征集合包含的用户特征,确定与该客户端相匹配的产品信息集合。

[0140] 本申请其他实施例中,所述确定模块 102 用于:

[0141] 根据所述客户端对产品信息集合中每一产品信息的预测数值,从所述产品信息集

合中选取至少一个待推送的产品信息;则,

[0142] 所述推送模块 103 用于:向所述客户端推送选取的至少一个待推送的产品信息。

[0143] 本申请其他实施例中,所述确定模块 102 用于:

[0144] 将所述产品信息集中的产品信息按照所述预测数值从大到小进行排序;则,

[0145] 所述推送模块 103 用于:向所述客户端推送排序后的所述产品信息。

[0146] 综上所述,本申请实施例的装置根据与每个客户端相对应的用户特征以及与各个产品信息相对应的产品特征,确定该客户端对各个产品信息的预测数值。由于该预测数值表征客户端的用户购买某个产品的可能性,若根据上述可以表征用户购买各个产品的可能性的预测数值,来确定应该向客户端推送怎样的产品信息,可以确保向用户推送的产品信息与该用户的真实兴趣相契合,从而提升产品信息推送的准确性。此外,通过以上实施例的技术方案,可以解决现有技术中向用户盲目推送产品信息所带来的营销效率低的问题,并且现有技术盲目推送产品信息还会造成服务器端的资源耗用大,网络数据量大的问题,本申请实施例也可以改善这一现象。

[0147] 在 20 世纪 90 年代,对于一个技术的改进可以很明显地区分是硬件上的改进(例如,对二极管、晶体管、开关等电路结构的改进)还是软件上的改进(对于方法流程的改进)。然而,随着技术的发展,当今的很多方法流程的改进已经可以视为硬件电路结构的直接改进。设计人员几乎都通过将改进的方法流程编程到硬件电路中来得到相应的硬件电路结构。因此,不能说一个方法流程的改进就不能用硬件实体模块来实现。例如,可编程逻辑器件(Programmable Logic Device,PLD)(例如现场可编程门阵列(Field Programmable Gate Array,FPGA))就是这样一种集成电路,其逻辑功能由用户对器件编程来确定。由设计人员自行编程来把一个数字系统“集成”在一片 PLD 上,而不需要请芯片制造厂商来设计和制作专用的集成电路芯片。而且,如今,取代手工地制作集成电路芯片,这种编程也多半改用“逻辑编译器(logic compiler)”软件来实现,它与程序开发撰写时所用的软件编译器相类似,而要编译之前的原始代码也得用特定的编程语言来撰写,此称之为硬件描述语言(Hardware Description Language,HDL),而 HDL 也并非仅有一种,而是有许多种,如 ABEL(Advanced Boolean Expression Language)、AHDL(Altera Hardware Description Language)、Confluence、CUPL(Cornell University Programming Language)、HDCal、JHDL(Java Hardware Description Language)、Lava、Lola、MyHDL、PALASM、RHDL(Ruby Hardware Description Language)等,目前最普遍使用的是 VHDL(Very-High-Speed Integrated Circuit Hardware Description Language)与 Verilog。本领域技术人员也应该清楚,只需要将方法流程用上述几种硬件描述语言稍作逻辑编程并编程到集成电路中,就可以很容易得到实现该逻辑方法流程的硬件电路。

[0148] 控制器可以按任何适当的方式实现,例如,控制器可以采取例如微处理器或处理器以及存储可由该(微)处理器执行的计算机可读程序代码(例如软件或固件)的计算机可读介质、逻辑门、开关、专用集成电路(Application Specific Integrated Circuit,ASIC)、可编程逻辑控制器和嵌入微控制器的形式,控制器的例子包括但不限于以下微控制器:ARC 625D、Atmel AT91SAM、Microchip PIC18F26K20 以及 Silicone Labs C8051F320,存储器控制器还可以被实现为存储器的控制逻辑的一部分。本领域技术人员也知道,除了以纯计算机可读程序代码方式实现控制器以外,完全可以通过将方法步骤进行逻辑编程来

使得控制器以逻辑门、开关、专用集成电路、可编程逻辑控制器和嵌入微控制器等的形式来实现相同功能。因此这种控制器可以被认为是一种硬件部件，而对其内包括的用于实现各种功能的装置也可以视为硬件部件内的结构。或者甚至，可以将用于实现各种功能的装置视为既可以是实现方法的软件模块又可以是硬件部件内的结构。

[0149] 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。

[0150] 为了描述的方便，描述以上装置时以功能分为各种单元分别描述。当然，在实施本申请时可以把各单元的功能在同一个或多个软件和 / 或硬件中实现。

[0151] 本领域内的技术人员应明白，本发明的实施例可提供为方法、系统、或计算机程序产品。因此，本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

[0152] 本发明是参照根据本发明实施例的方法、设备（系统）、和计算机程序产品的流程图和 / 或方框图来描述的。应理解可由计算机程序指令实现流程图和 / 或方框图中的每一流程和 / 或方框、以及流程图和 / 或方框图中的流程和 / 或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的装置。

[0153] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制品，该指令装置实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能。

[0154] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的步骤。

[0155] 在一个典型的配置中，计算设备包括一个或多个处理器（CPU）、输入 / 输出接口、网络接口和内存。

[0156] 内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器（RAM）和 / 或非易失性内存等形式，如只读存储器（ROM）或闪存（flash RAM）。内存是计算机可读介质的示例。

[0157] 计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存（PRAM）、静态随机存取存储器（SRAM）、动态随机存取存储器（DRAM）、其他类型的随机存取存储器（RAM）、只读存储器（ROM）、电可擦除可编程只读存储器（EEPROM）、快闪记忆体或其他内存技术、只读光盘只读存储器（CD-ROM）、数字多功能光盘（DVD）或其他光学存储、磁盒式磁带，磁带磁盘存储或其他磁

性存储设备或任何其他非传输介质,可用于存储可以被计算设备访问的信息。按照本文中的界定,计算机可读介质不包括暂存电脑可读媒体(transitory media),如调制的数据信号和载波。

[0158] 还需要说明的是,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、产品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、产品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、产品或者设备中还存在另外的相同要素。

[0159] 本领域技术人员应明白,本申请的实施例可提供为方法、系统或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0160] 本申请可以在由计算机执行的计算机可执行指令的一般上下文中描述,例如程序模块。一般地,程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等等。也可以在分布式计算环境中实践本申请,在这些分布式计算环境中,通过通信网络而被连接的远程处理设备来执行任务。在分布式计算环境中,程序模块可以位于包括存储设备在内的本地和远程计算机存储介质中。

[0161] 本说明书中的各个实施例均采用递进的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。尤其,对于系统实施例而言,由于其基本相似于方法实施例,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。

[0162] 以上所述仅为本申请的实施例而已,并不用于限制本申请。对于本领域技术人员来说,本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等,均应包含在本申请的权利要求范围之内。

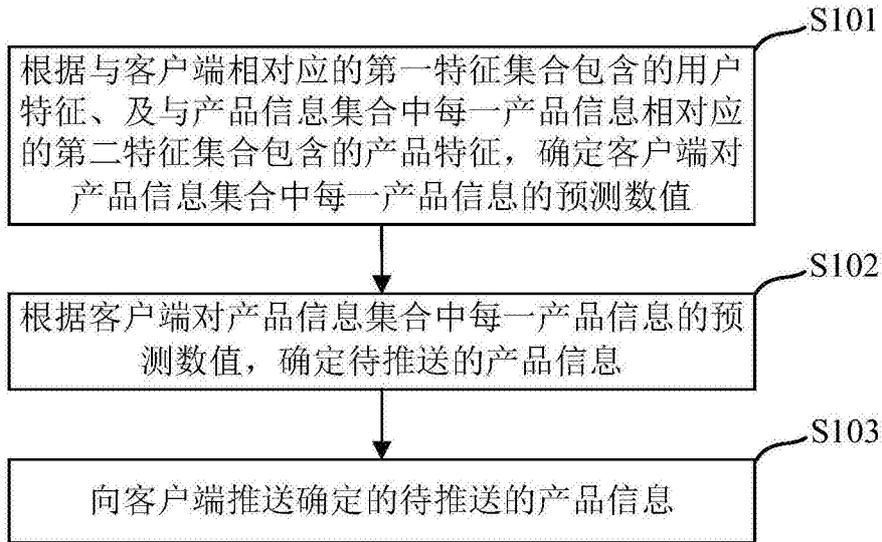


图 1

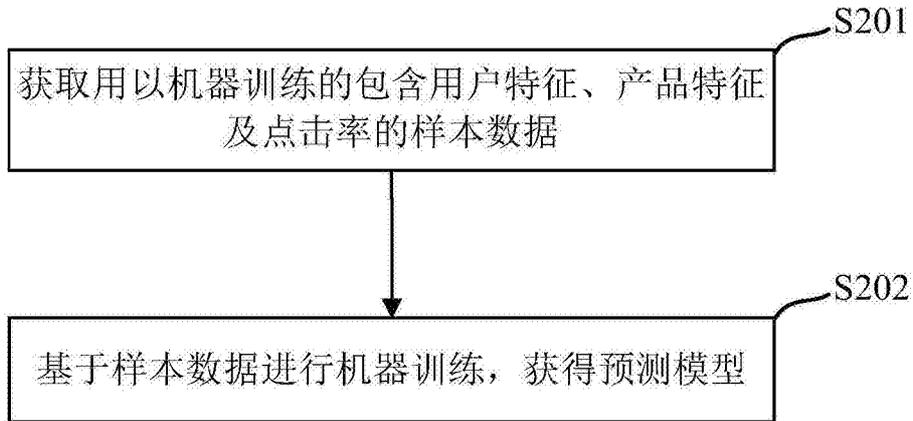


图 2

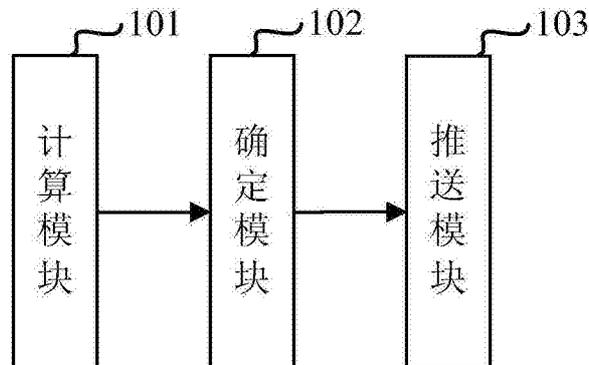


图 3

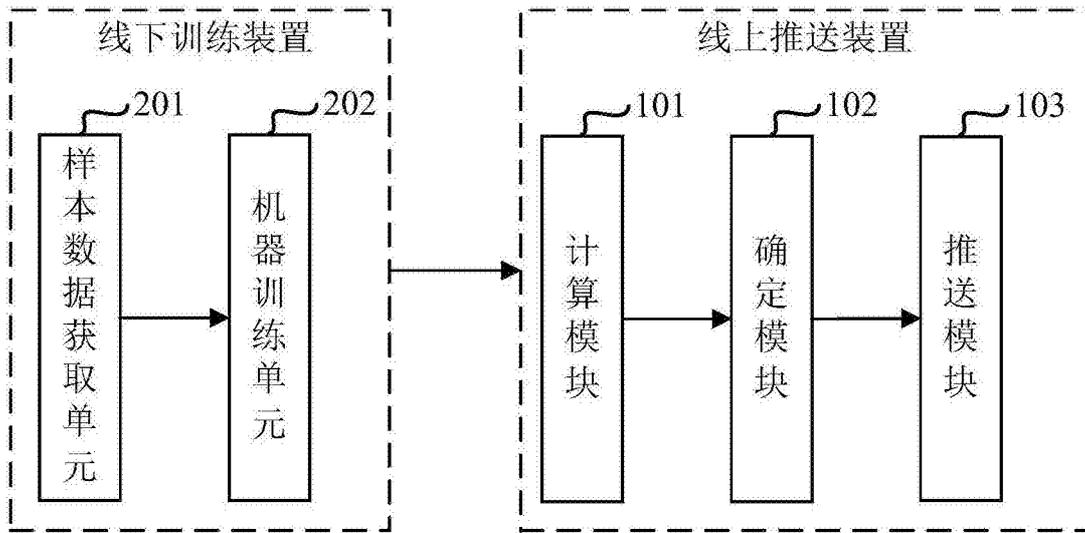


图 4

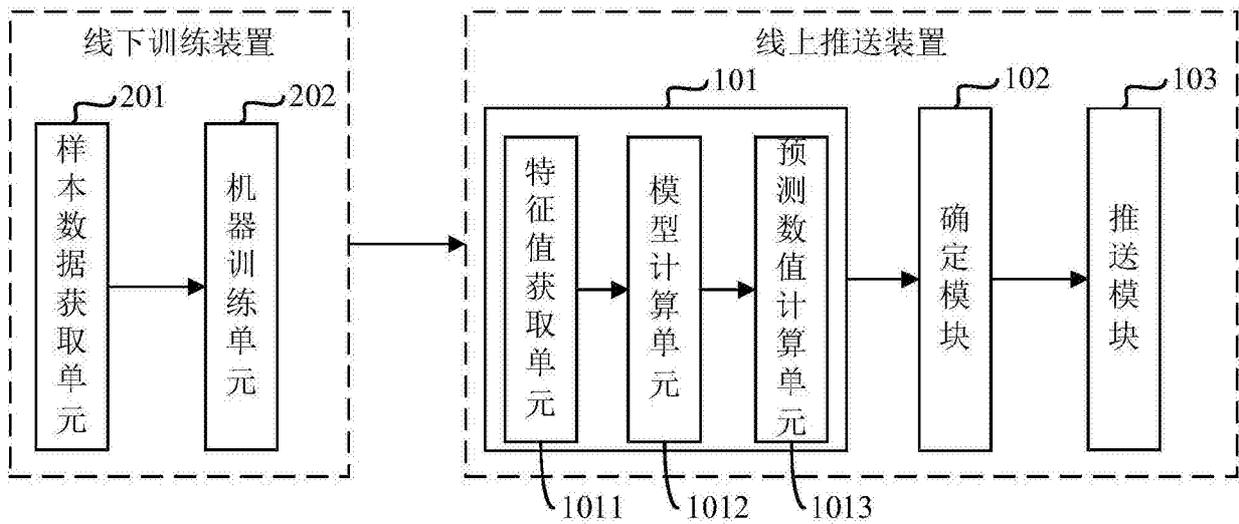


图 5