

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织

国 际 局

(43) 国际公布日

2022 年 7 月 7 日 (07.07.2022)



WIPO | PCT



(10) 国际公布号

WO 2022/142013 A1

(51) 国际专利分类号:

G06Q 10/06 (2012.01) G06F 11/36 (2006.01)

(21) 国际申请号:

PCT/CN2021/090460

(22) 国际申请日:

2021 年 4 月 28 日 (28.04.2021)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

202011614892.6 2020年12月30日 (30.12.2020) CN

(71) 申请人: 平安科技(深圳)有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(72) 发明人: 陆琦川 (LU, Qichuan); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市赛恩倍吉知识产权代理有限公司 (SHENZHEN SCIENBIZIP INTELLECTUAL PROPERTY AGENCY CO., LTD.); 中国广东省深圳市龙华新区龙观东路 83 号荣群大厦 9 楼, Guangdong 518109 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,

(54) Title: ARTIFICIAL INTELLIGENCE-BASED AB TESTING METHOD AND APPARATUS, COMPUTER DEVICE AND MEDIUM

(54) 发明名称: 基于人工智能的AB测试方法、装置、计算机设备及介质

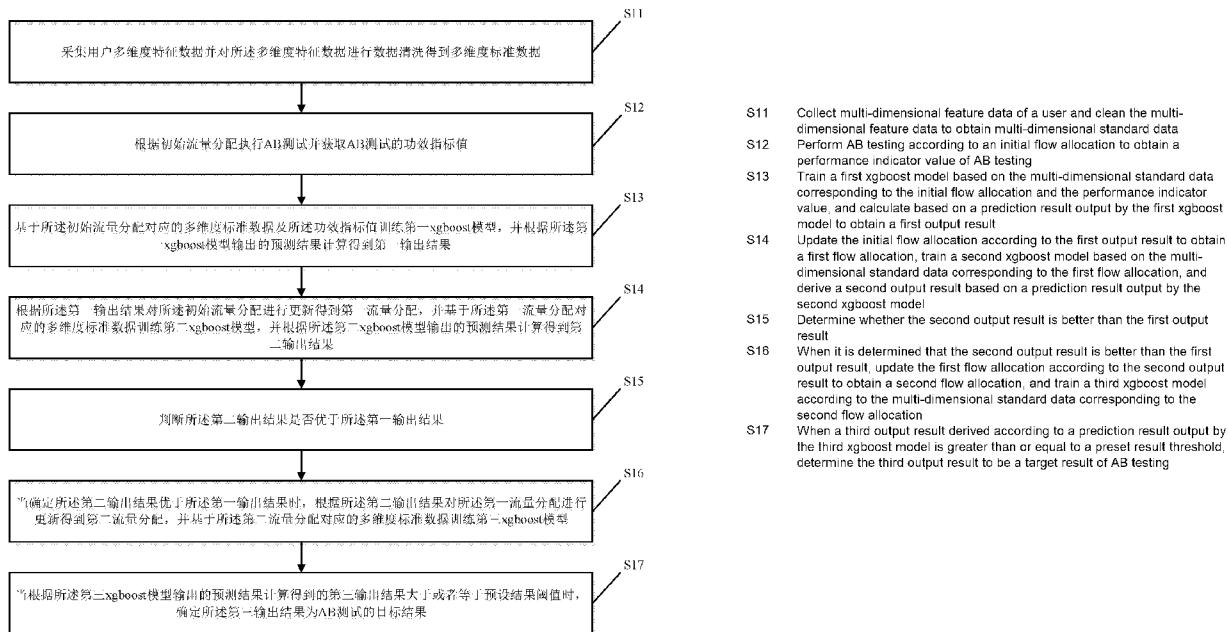


图 1

(57) Abstract: Provided are an artificial intelligence-based AB testing method and apparatus, a computer device, and a medium, comprising: cleaning multi-dimensional feature data to obtain multi-dimensional standard data (S11); performing AB testing according to an initial flow allocation to obtain a performance indicator value (S12); training a first xgboost model based on the multi-dimensional standard data corresponding to the initial flow allocation and the performance indicator value (S13); updating the first flow allocation according to a first output result of the first xgboost model, and training a second xgboost model based on the multi-dimensional standard data corresponding to the first flow allocation (S14); when a second output result of the second xgboost model is better than a first



MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

output result, updating the second flow allocation according to the second output result and training a third xgboost model (S16); when a third output result of a third xgboost model is greater than or equal to a preset result threshold, determining the third output result to be a target result of AB testing (S17). The present application can improve the efficiency of AB testing and ensure the consistency of the long-term effect and the short-term effect of the AB testing result.

(57) 摘要: 提供一种基于人工智能的AB测试方法、装置、计算机设备及介质, 包括: 清洗多维度特征数据为多维度标准数据 (S11); 根据初始流量分配执行AB测试得到功效指标值 (S12); 基于初始流量分配对应的多维度标准数据及功效指标值训练第一xgboost模型 (S13); 根据第一xgboost模型的第一输出结果更新第一流量分配, 基于第一流量分配对应的多维度标准数据训练第二xgboost模型 (S14); 当第二xgboost模型的第二输出结果优于第一输出结果, 根据第二输出结果更新第二流量分配并训练第三xgboost模型 (S16); 当第三xgboost模型的第三输出结果大于等于预设结果阈值时, 确定第三输出结果为AB测试的目标结果 (S17)。本申请能够提高AB测试的测试效率, 且保证AB测试结果的长期效应与短期效应的一致性。

基于人工智能的 AB 测试方法、装置、计算机设备及介质

本申请要求于 2020 年 12 月 30 日提交中国专利局、申请号为 202011614892.6，发明名称为“基于人工智能的 AB 测试方法、装置、计算机设备及介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及人工智能技术领域，具体涉及一种基于人工智能的 AB 测试方法、装置、计算机设备及介质。

背景技术

AB 测试已成为企业业务评估与业务增长效能的必备工具，通过保持确定维度变量，改变若干因素来得出不同方案，对实验用户分群分别投放不通方案，并获取反馈信息和相关数据。

发明人在实现本发明的过程中发现，传统的 AB 测试，需要进行 AB 测试方案设计、AB 测试多方案场景代码编写等一系列繁琐的过程，导致 AB 测试效率较低；此外，通过多长期多次的 AB 测试结果的分析，发现 UI/UE 的变更，经常会对用户的行为产生较大的影响，但当用户习惯新的业务流程后，带来的业务效果未必与功能初期上线的结果一致，即长期 AB 测试结果与短期 AB 测试结果不统一，AB 测试准确率较差。

发明内容

鉴于以上内容，有必要提出一种基于人工智能的 AB 测试方法、装置、计算机设备及介质，能够提高 AB 测试的测试效率，且保证 AB 测试结果的长期效应与短期效应的一致性。

本申请的第一方面提供一种基于人工智能的 AB 测试方法，所述方法包括：

采集用户多维度特征数据并对所述多维度特征数据进行数据清洗得到多维度标准数据；根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值；

基于所述初始流量分配对应的多维度标准数据及所述功效指标值训练第一 xgboost 模型，并根据所述第一 xgboost 模型输出的预测结果计算得到第一输出结果；

根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配，并基于所述第一流量分配对应的多维度标准数据训练第二 xgboost 模型，并根据所述第二 xgboost 模型输出的预测结果计算得到第二输出结果；

判断所述第二输出结果是否优于所述第一输出结果；

当确定所述第二输出结果优于所述第一输出结果时，迭代根据所述第二输出结果对所述第一流量分配进行更新得到第二流量分配，并基于所述第二流量分配对应的多维度标准数据训练第三 xgboost 模型，直至根据所述第三 xgboost 模型输出的预测结果计算得到的第三输出结果大于或者等于预设结果阈值时，停止所述第三 xgboost 模型的训练；

确定所述第三输出结果为 AB 测试的目标结果。

本申请的第二方面提供一种基于人工智能的 AB 测试装置，所述装置包括：

采集模块，用于采集用户多维度特征数据并对所述多维度特征数据进行数据清洗得到多维度标准数据；

测试模块，用于根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值；

训练模块，用于基于所述初始流量分配对应的多维度标准数据及所述功效指标值训练第一 xgboost 模型，并根据所述第一 xgboost 模型输出的预测结果计算得到第一输出结果；

所述训练模块，还用于根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配，并基于所述第一流量分配对应的多维度标准数据训练第二 xgboost 模型，并根据所述第二 xgboost 模型输出的预测结果计算得到第二输出结果；

判断模块，用于判断所述第二输出结果是否优于所述第一输出结果；

所述训练模块，还用于当确定所述第二输出结果优于所述第一输出结果时，迭代根据所述第二输出结果对所述第一流量分配进行更新得到第二流量分配，并基于所述第二流量分配对应的多维度标准数据训练第三 xgboost 模型，直至根据所述第三 xgboost 模型输出的预测结果计算得到的第三输出结果大于或者等于预设结果阈值时，停止所述第三 xgboost 模型的训练；

确定模块，用于确定所述第三输出结果为 AB 测试的目标结果。

本申请的第三方面提供一种计算机设备，所述计算机设备包括处理器，所述处理器用于执行存储器中存储的计算机可读指令时实现以下步骤：

采集用户多维度特征数据并对所述多维度特征数据进行数据清洗得到多维度标准数据；

根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值；

基于所述初始流量分配对应的多维度标准数据及所述功效指标值训练第一 xgboost 模型，并根据所述第一 xgboost 模型输出的预测结果计算得到第一输出结果；

根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配，并基于所述第一流量分配对应的多维度标准数据训练第二 xgboost 模型，并根据所述第二 xgboost 模型输出的预测结果计算得到第二输出结果；

判断所述第二输出结果是否优于所述第一输出结果；

当确定所述第二输出结果优于所述第一输出结果时，迭代根据所述第二输出结果对所述第一流量分配进行更新得到第二流量分配，并基于所述第二流量分配对应的多维度标准数据训练第三 xgboost 模型，直至根据所述第三 xgboost 模型输出的预测结果计算得到的第三输出结果大于或者等于预设结果阈值时，停止所述第三 xgboost 模型的训练；

确定所述第三输出结果为 AB 测试的目标结果。

本申请的第四方面提供一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机可读指令，所述计算机可读指令被处理器执行时实现以下步骤：

采集用户多维度特征数据并对所述多维度特征数据进行数据清洗得到多维度标准数据；

根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值；

基于所述初始流量分配对应的多维度标准数据及所述功效指标值训练第一 xgboost 模型，并根据所述第一 xgboost 模型输出的预测结果计算得到第一输出结果；

根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配，并基于所述第一流量分配对应的多维度标准数据训练第二 xgboost 模型，并根据所述第二 xgboost 模型输出的预测结果计算得到第二输出结果；

判断所述第二输出结果是否优于所述第一输出结果；

当确定所述第二输出结果优于所述第一输出结果时，迭代根据所述第二输出结果对所述第一流量分配进行更新得到第二流量分配，并基于所述第二流量分配对应的多维度标准数据训练第三 xgboost 模型，直至根据所述第三 xgboost 模型输出的预测结果计算得到的第三输出结果大于或者等于预设结果阈值时，停止所述第三 xgboost 模型的训练；

确定所述第三输出结果为 AB 测试的目标结果。

综上所述，本申请所述的基于人工智能的 AB 测试方法、装置、计算机设备及介质，通过初始流量分配执行 AB 测试并获取 AB 测试的功效指标值；基于所述初始流量分配对应的多维度标准数据及所述功效指标值训练 xgboost 模型，并获取所述 xgboost 模型的输出结果；根据所述输出结果对所述初始流量分配进行更新得到第一流量分配，并迭代训练 xgboost 模型直至 xgboost 模型的输出结果大于或者等于预设结果阈值时，将 xgboost 模型的输出结果确定为 AB 测试的目标结果。通过不断迭代训练 xgboost 模型从而自动模拟 AB 测试，提高了 AB 测试的测试效率；此外，只有在 xgboost 模型的输出结果大于或者等于预设结果阈值时，将 xgboost 模型的输出结果确定为 AB 测试的目标结果，能够保证 AB 测试结果的长期效应与短期效应的一致性。

附图说明

图 1 是本申请实施例一提供的基于人工智能的 AB 测试方法的流程图。

图 2 是本申请实施例二提供的基于人工智能的 AB 测试装置的结构图。

图 3 是本申请实施例三提供的计算机设备的结构示意图。

具体实施方式

为了能够更清楚地理解本申请的上述目的、特征和优点，下面结合附图和具体实施例对本申请进行详细描述。需要说明的是，在不冲突的情况下，本申请的实施例及实施例中的特征可以相互组合。

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申请的技术领域的技术人员通常理解的含义相同。本文中在本申请的说明书中所使用的术语只是为了描述具体的实施例的目的，不是旨在于限制本申请。

本申请实施例提供的基于人工智能的 AB 测试方法由计算机设备执行，相应地，基于人工智能的 AB 测试装置运行于计算机设备中。

图 1 是本申请实施例一提供的基于人工智能的 AB 测试方法的流程图。所述基于人工智能的 AB 测试方法具体包括以下步骤，根据不同的需求，该流程图中步骤的顺序可以改变，某些可以省略。

S11，采集用户多维度特征数据并对所述多维度特征数据进行数据清洗得到多维度标准数据。

可以从用户客户端中采集用户的多维度特征数据。

由于不同用户的同一维度特征数据可能相差较大，因此需要对采集到的多维度特征数据进行数据清洗得到多维度标准数据，便于后续基于所述多维度标准数据进行机器学习与训练，从而主动的进行 AB 测试。

在一个可选的实施例中，所述采集用户多维度特征数据包括：

定义多维度关键绩效指标并发送所述多维度关键绩效指标至所述多个用户客户端；

获取所述多个用户客户端根据所述多维度关键绩效指标上报的埋点数据；

从所述埋点数据中采集所述用户多维度特征数据。

该可选的实施例中，为了精准且高效的获取用户多维度特征数据，事先定义多个维度的关键绩效指标（Key Performance Indicator, KPI），使得用户客户端能够根据多个维度的 KPI 自主上报埋点数据，从而使得计算机设备从所述埋点数据中采集用户多维度特征数据。其中，所述多个维度的关键绩效指标可以包括，但不限于：用户客户端的系统版本、用户客户端的 GPS 信息、应用的版本、行为数据等。所述行为数据是指用户在用户客户端上的 APP 的操作行为，例如，点击行为、关闭行为、分享行为等。

计算机设备发送携带有多个维度 KPI 及每个维度 KPI 对应的采集序号的采集信号至用户客户端。用户客户端响应于接收到的采集信号，将每个维度 KPI 对应的特征数据通过埋点的形式发送给计算机设备。计算机设备响应于接收到的埋点数据，对埋点数据进行解析得到解析结果，并按照多个维度 KPI 对应的采集序号从所述解析结果中依次获取特征数据，从而得到用户多维度特征数据。

在上述实施例中，通过发送多维度 KPI 至多个用户客户端，使得多个用户客户端上报的埋点数据中包括且仅包括了与多维度 KPI 对应的特征数据，因而用户多维度特征数据的采集效率高。

在一个可选的实施例中，所述对所述多维度特征数据进行数据清洗得到多维度标准数据包括：

对所述多维度特征数据进行去重处理；

采用随机森林模型检测去重处理后的多维度特征数据中的异常数据；

删除所述去重处理后的多维度特征数据中的异常数据得到多维度标准数据。

该可选的实施例中，用户客户端在上报埋点数据时会受到网络质量的影响，当网络质量较差时，可能会重复多次上报同一个埋点数据，因此，需要对同一个用户的同维度特征数据进行去重处理，去重处理有助于提高多维度特征数据的数据质量，有助于提高训练 xgboost 模型的性能。

在风控场景中，有些用户客户端上报的埋点数据会存在造假或者失真，采用随机森林模型检测去重处理后的多维度特征数据中的异常数据。随机森林是指利用随机森林算法生成树，异常点会因为数值异常很容易被检测出来。基于去掉异常数据后的多维度特征数据训练 xgboost 模型，能够提高 xgboost 模型的稳定性。

S12，根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值。

AB 是 Apache 超文本传输协议（HTTP）的性能测试工具。其设计意图是描绘当前所安装的 Apache 的执行性能，主要是显示安装的 Apache 每秒可以处理多少个请求。AB 还可以针对某个特定的 URL 仿真出连续的联机请求，同时还可以仿真出同时间点数个相同的联机请求。

在一个可选的实施例中，所述根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值包括：

判断多维度标准数据的数据量；

根据预设数据量范围与流量分配比例之间的映射表确定与所述数据量对应的目标流量分配比例，作为所述初始流量分配；

按照所述初始流量分配切分所述多维度标准数据为第一多维度标准数据和第二多维度标准数据；

将所述第一多维度标准数据分配给 A 版本及将所述第二多维度标准数据分配给 B 版本并执行模拟测试过程；

获取模拟测试结束后的测试数据并根据所述测试数据计算功效指标。

示例性的，假设多维度标准数据的数据量为 1000，则初始流量分配为 8:2，即将所述多维度标准数据中的 800 条（80%）数据分配给 A 版本，将所述多维度标准数据中剩余的 200 条（20%）的数据分配给 B 版本；假设多维度标准数据的数据量为 10000，则初始流量分配为 9:1，即将所述多维度标准数据中的 9000 条（90%）的数据分配给 A 版本，将所述多维度标准数据中剩余的 1000 条（10%）的数据分配给 B 版本。其中，A 版本为旧版本，B 版本为新版本。

其中，功效指标可以包括，但不限于：转化率，成交率，点击率等。依据实际情况，可以对功效指令进行调整与修改。

该可选的实施例中，由于应用更新版本后，并不能确定新版本应用是否会带来好的或者不好的影响，多维度标准数据的数据量越大，则将较多的数据分配给 A 版本，将较少的数据分配给 B 版本；多维度标准数据的数据量越大，则将较多的数据分配给 A 版本，将较少的数据分配给 B 版本。本实施例中通过根据多维度标准数据的数据量确定初始流量分配并执行 AB 测试，能够使得 AB 测试更贴近实际业务场景，避免一刀切的现象，AB 测试效率更高。

S13，基于所述初始流量分配对应的多维度标准数据及所述功效指标值训练第一 xgboost 模型，并根据所述第一 xgboost 模型输出的预测结果计算得到第一输出结果。

将分配给 A 版本的多维度标准数据打上第一标签，将分配给 B 版本的多维度标准数据打上第二标签。将所述第一标签及对应的第一多维度标准数据、所述第二标签及对应的第二多维度标准数据和功效指标值作为数据集，采用有监督的学习方式训练第一 xgboost 模型。具体实施时，首先定义训练结束条件，所述训练结束条件可以是迭代训练的次数超过预设次数阈值；接着输入所述数据集至所述第一 xgboost 模型中进行迭代训练；当迭代训练的次数超过预设次数阈值时，停止对所述第一 xgboost 模型的训练。所述第一 xgboost 模型在训练结束后会输出预测结果，根据所述第一 xgboost 模型输出的预测结果计算得到所述第一输出结果。其中，所述预测结果包括：分配给 A 版本的每一个用户对应的多维度标准数据的第一预测标签

及第一预测概率，分配给 B 版本的每一个用户对应的多维度标准数据的第二预测标签及第二预测概率。

所述第一预测概率用以表示第一 xgboost 模型预测分配给 A 版本的每一个用户对应的多维度标准数据的第一预测标签的最大可能性，所述第二预测概率用以表示第一 xgboost 模型预测分配给 B 版本的每一个用户对应的多维度标准数据的第二预测标签的最大可能性。所述第一预测标签和所述第二预测标签用以预测每一个用户对应的多维度标准数据属于所述第一标签还是所述第二标签。

所述根据所述第一 xgboost 模型输出的预测结果计算得到所述第一输出结果可以包括：获取所述预测结果中的第一预测标签及第一预测概率，第二预测标签及第二预测概率；获取所述第一预测标签及所述第二预测标签中第一标签的第一个数；获取所述第一预测标签及所述第二预测标签中第二标签的第二个数；计算所述第一个数与所述第二个数之和得到个数和；计算所述第一个数与所述个数和得到第一比例值；计算所述第二个数和所述个数和得到第二比例值。

所述第一输出结果包括：第一标签，第一比例值；第二标签，第二比例值。所述第一比例值与所述第二比例值之和为 1，第一比例值代表第一标签所占的比例，第二比例值代表第二标签所占的比例。所述第一标签与 A 版本对应，所述第二标签与 B 版本对应。

S14，根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配，并基于所述第一流量分配对应的多维度标准数据训练第二 xgboost 模型，并根据所述第二 xgboost 模型输出的预测结果计算得到第二输出结果。

在第一轮训练第一 xgboost 模型之后，需要对初始流量分配进行更新，以便进行第二轮 xgboost 模型的训练。

所述第二输出结果的计算过程同所述第一输出结果的计算过程。所述第二输出结果包括：第一标签，第一比例值；第二标签，第二比例值。所述第一比例值与所述第二比例值之和为 1，第一比例值代表第一标签所占的比例，第二比例值代表第二标签所占的比例。所述第一标签与 A 版本对应，所述第二标签与 B 版本对应。

在一个可选的实施例中，所述根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配包括：

判断所述第一比例值是否小于所述第二比例值；

当确定所述第一比例值小于所述第二比例值时，在所述初始流量分配的基础上增加第一值得到第一流量分配；

当确定所述第一比例值大于所述第二比例值时，在所述初始流量分配的基础上减少第二值得到第一流量分配。

该可选的实施例中，第一比例值小于所述第二比例值，表明 B 版本表现出了优于 A 版本的效应，则增加初始流量分配，使得分配给 A 版本的第一多维度标准数据的数据量减少，分配给 B 版本的第二多维度标准数据的数据量增加。第一比例值大于所述第二比例值，表明 A 版本表现出了优于 B 版本的效应，则减少初始流量分配，使得分配给 A 版本的第一多维度标准数据的数据量增加，分配给 B 版本的第二多维度标准数据的数据量减少。

S15，判断所述第二输出结果是否优于所述第一输出结果。

其中，所述预设结果阈值为预先设置的临界值，用以表示 xgboost 模型的结束条件。

如果第一输出结果中的第一比例值小于第二输出结果中的第一比例值，则表明第二输出结果优于第一输出结果。如果第一输出结果中的第一比例值大于第二输出结果中的第一比例值，则表明第二输出结果差于第一输出结果。

S16，当确定所述第二输出结果优于所述第一输出结果时，根据所述第二输出结果对所述第一流量分配进行更新得到第二流量分配，并基于所述第二流量分配对应的多维度标准数据训练第三 xgboost 模型。

当确定所述第二输出结果优于所述第一输出结果时，在所述第一流量分配的基础上增加

第一值得到第二流量分配；当确定所述第一输出结果优于所述第二输出结果时，在所述第一流量分配的基础上减少第二值得到第二流量分配。

在根据所述第二输出结果对所述第一流量分配进行更新得到第二流量分配之后，基于所述第二流量分配对应的多维度标准数据训练第三 xgboost 模型，并根据所述第三 xgboost 模型输出的预测结果计算得到第三输出结果。判断所述第二输出结果是否优于所述第二输出结果及判断第三输出结果是否大于预设结果阈值。当判断所述第二输出结果优于所述第二输出结果且第三输出结果小于预设结果阈值时，根据所述第三输出结果对所述第二流量分配进行更新得到第三流量分配，并基于所述第三流量分配对应的多维度标准数据训练第四 xgboost 模型，以此类推。当判断所述第二输出结果优于所述第二输出结果且判断第三输出结果大于或者等于预设结果阈值时，停止所述第三 xgboost 模型的训练。

对于统计功效更优的版本，进行流量动态变更，并使用训练结果模型验证数据结果一致性，若实验结果出现不一致结果集，将重新进行自学习，直至模型结果收敛，并且实验结果中将进行多次流量自动调整。

S17，当根据所述第三 xgboost 模型输出的预测结果计算得到的第三输出结果大于或者等于预设结果阈值时，确定所述第三输出结果为 AB 测试的目标结果。

停止对所述第三 xgboost 模型的训练，表明 AB 测试已经结束，可以将第三 xgboost 模型的第三输出结果确定为目标结果。其中，所述第三输出结果的计算过程同所述第一输出结果的计算过程。

在一个可选的实施例中，在所述确定所述第三输出结果为 AB 测试的目标结果之后，所述方法还包括：

获取所述目标结果中的 A 结果及 B 结果；

比较所述 A 结果及所述 B 结果；

当确定所述 A 结果小于所述 B 结果，发送通知信息至多个用户客户端，使得所述多个用户客户端更新应用为 B 版本。

所述第三输出结果包括：第一标签，第一比例值；第二标签，第二比例值。所述第一比例值与所述第二比例值之和为 1，第一比例值代表第一标签所占的比例，第二比例值代表第二标签所占的比例。所述第一标签与 A 版本对应，所述第二标签与 B 版本对应。当第一比例值小于第二比例值时，表示 A 结果小于 B 结果。当第一比例值大于第二比例值时，表示 A 结果大于 B 结果。

在多次训练 xgboost 模型之后，即可模拟 AB 测试的长期效应。当长期效应表示 A 结果小于 B 结果，则表明 B 版本应用的效应优于 A 版本应用的效应；当长期效应表示 A 结果大于 B 结果，则表明 A 版本应用的效应优于 B 版本应用的效应。

当长期效应表明某个版本优于另一个版本，将可以发送通知信息至用户客户端，通知用户客户端更新版本至更优版本。

需要说明的是，本申请提供的方法适用于同一个应用的两个不同的版本的测试，测试哪一个版本对用户更有效；本申请提供的方法还可适用于两个不同的方案的测试，测试哪一个方案对用户更有效；本申请提供的方法还可适用于同一个应用的多个不同的版本的测试或者用于多个不同的方案的测试。

综上，本申请可应用于智慧政务等智慧场景中，从而推动智慧城市的建设。本申请通过初始流量分配执行 AB 测试并获取 AB 测试的功效指标值；基于所述初始流量分配对应的多维度标准数据及所述功效指标值训练 xgboost 模型，并获取所述 xgboost 模型的输出结果；根据所述输出结果对所述初始流量分配进行更新得到第一流量分配，并迭代训练 xgboost 模型直至 xgboost 模型的输出结果大于或者等于预设结果阈值时，将 xgboost 模型的输出结果确定为 AB 测试的目标结果。通过不断迭代训练 xgboost 模型从而自动模拟 AB 测试，提高了 AB 测试的测试效率；此外，只有在 xgboost 模型的输出结果大于或者等于预设结果阈值时，将 xgboost 模型的输出结果确定为 AB 测试的目标结果，能够保证 AB 测试结果的长期效应与短

期效应的一致性。

需要强调的是，为进一步保证上述目标结果的私密性和安全性，上述目标结果可存储于区块链的节点中。

图 2 是本申请实施例二提供的基于人工智能的 AB 测试装置的结构图。

在一些实施例中，所述基于人工智能的 AB 测试装置 20 可以包括多个由计算机可读指令段所组成的功能模块。所述基于人工智能的 AB 测试装置 20 中的各个程序段的计算机可读指令可以存储于计算机设备的存储器中，并由至少一个处理器所执行，以执行（详见图 1 描述）基于人工智能的 AB 测试的功能。

本实施例中，所述基于人工智能的 AB 测试装置 20 根据其所执行的功能，可以被划分为多个功能模块。所述功能模块可以包括：采集模块 201、测试模块 202、训练模块 203、判断模块 204、确定模块 205 及通知模块 206。本申请所称的模块是指一种能够被至少一个处理器所执行并且能够完成固定功能的一系列计算机可读指令段，其存储在存储器中。在本实施例中，关于各模块的功能将在后续的实施例中详述。

所述采集模块 201，用于采集用户多维度特征数据并对所述多维度特征数据进行数据清洗得到多维度标准数据。

可以从用户客户端中采集用户的多维度特征数据。

由于不同用户的同一维度特征数据可能相差较大，因此需要对采集到的多维度特征数据进行数据清洗得到多维度标准数据，便于后续基于所述多维度标准数据进行机器学习与训练，从而主动的进行 AB 测试。

在一个可选的实施例中，所述采集模块 201 采集用户多维度特征数据包括：

定义多维度关键绩效指标并发送所述多维度关键绩效指标至所述多个用户客户端；

获取所述多个用户客户端根据所述多维度关键绩效指标上报的埋点数据；

从所述埋点数据中采集所述用户多维度特征数据。

该可选的实施例中，为了精准且高效的获取用户多维度特征数据，事先定义多个维度的关键绩效指标（Key Performance Indicator, KPI），使得用户客户端能够根据多个维度的 KPI 自主上报埋点数据，从而使得计算机设备从所述埋点数据中采集用户多维度特征数据。其中，所述多个维度的关键绩效指标可以包括，但不限于：用户客户端的系统版本、用户客户端的 GPS 信息、应用的版本、行为数据等。所述行为数据是指用户在用户客户端上的 APP 的操作行为，例如，点击行为、关闭行为、分享行为等。

计算机设备发送携带有多个维度 KPI 及每个维度 KPI 对应的采集序号的采集信号至用户客户端。用户客户端响应于接收到的采集信号，将每个维度 KPI 对应的特征数据通过埋点的形式发送给计算机设备。计算机设备响应于接收到的埋点数据，对埋点数据进行解析得到解析结果，并按照多个维度 KPI 对应的采集序号从所述解析结果中依次获取特征数据，从而得到用户多维度特征数据。

在上述实施例中，通过发送多维度 KPI 至多个用户客户端，使得多个用户客户端上报的埋点数据中包括且仅包括了与多维度 KPI 对应的特征数据，因而用户多维度特征数据的采集效率高。

在一个可选的实施例中，所述采集模块 201 对所述多维度特征数据进行数据清洗得到多维度标准数据包括：

对所述多维度特征数据进行去重处理；

采用随机森林模型检测去重处理后的多维度特征数据中的异常数据；

删除所述去重处理后的多维度特征数据中的异常数据得到多维度标准数据。

该可选的实施例中，用户客户端在上报埋点数据时会受到网络质量的影响，当网络质量较差时，可能会重复多次上报同一个埋点数据，因此，需要对同一个用户的同维度特征数据进行去重处理，去重处理有助于提高多维度特征数据的数据质量，有助于提高训练 xgboost 模型的性能。

在风控场景中，有些用户客户端上报的埋点数据会存在造假或者失真，采用随机森林模型检测去重处理后的多维度特征数据中的异常数据。随机森林是指利用随机森林算法生成树，异常点会因为数值异常很容易被检测出来。基于去掉异常数据后的多维度特征数据训练 xgboost 模型，能够提高 xgboost 模型的稳定性。

所述测试模块 202，用于根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值。

AB 是 Apache 超文本传输协议（HTTP）的性能测试工具。其设计意图是描绘当前所安装的 Apache 的执行性能，主要是显示安装的 Apache 每秒可以处理多少个请求。AB 还可以针对某个特定的 URL 仿真出连续的联机请求，同时还可以仿真出同时间点数个相同的联机请求。

在一个可选的实施例中，所述测试模块 202 根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值包括：

判断多维度标准数据的数据量；

根据预设数据量范围与流量分配比例之间的映射表确定与所述数据量对应的目标流量分配比例，作为所述初始流量分配；

按照所述初始流量分配切分所述多维度标准数据为第一多维度标准数据和第二多维度标准数据；

将所述第一多维度标准数据分配给 A 版本及将所述第二多维度标准数据分配给 B 版本并执行模拟测试过程；

获取模拟测试结束后的测试数据并根据所述测试数据计算功效指标。

示例性的，假设多维度标准数据的数据量为 1000，则初始流量分配为 8:2，即将所述多维度标准数据中的 800 条（80%）数据分配给 A 版本，将所述多维度标准数据中剩余的 200 条（20%）的数据分配给 B 版本；假设多维度标准数据的数据量为 10000，则初始流量分配为 9:1，即将所述多维度标准数据中的 9000 条（90%）的数据分配给 A 版本，将所述多维度标准数据中剩余的 1000 条（10%）的数据分配给 B 版本。其中，A 版本为旧版本，B 版本为新版本。

其中，功效指标可以包括，但不限于：转化率，成交率，点击率等。依据实际情况，可以对功效指令进行调整与修改。

该可选的实施例中，由于应用更新版本后，并不能确定新版本应用是否会带来好的或者不好的影响，多维度标准数据的数据量越大，则将较多的数据分配给 A 版本，将较少的数据分配给 B 版本；多维度标准数据的数据量越大，则将较多的数据分配给 A 版本，将较少的数据分配给 B 版本。本实施例中通过根据多维度标准数据的数据量确定初始流量分配并执行 AB 测试，能够使得 AB 测试更贴近实际业务场景，避免一刀切的现象，AB 测试效率更高。

所述训练模块 203，用于基于所述初始流量分配对应的多维度标准数据及所述功效指标值训练第一 xgboost 模型，并根据所述第一 xgboost 模型输出的预测结果计算得到第一输出结果。

将分配给 A 版本的多维度标准数据打上第一标签，将分配给 B 版本的多维度标准数据打上第二标签。将所述第一标签及对应的第一多维度标准数据、所述第二标签及对应的第二多维度标准数据和功效指标值作为数据集，采用有监督的学习方式训练第一 xgboost 模型。具体实施时，首先定义训练结束条件，所述训练结束条件可以是迭代训练的次数超过预设次数阈值；接着输入所述数据集至所述第一 xgboost 模型中进行迭代训练；当迭代训练的次数超过预设次数阈值时，停止对所述第一 xgboost 模型的训练。所述第一 xgboost 模型在训练结束后会输出预测结果，根据所述第一 xgboost 模型输出的预测结果计算得到所述第一输出结果。其中，所述预测结果包括：分配给 A 版本的每一个用户对应的多维度标准数据的第一预测标签及第一预测概率，分配给 B 版本的每一个用户对应的多维度标准数据的第二预测标签及第二预测概率。

所述第一预测概率用以表示第一 xgboost 模型预测分配给 A 版本的每一个用户对应的多

维度标准数据的第一预测标签的最大可能性，所述第二预测概率用以表示第一 xgboost 模型预测分配给 B 版本的每一个用户对应的多维度标准数据的第二预测标签的最大可能性。所述第一预测标签和所述第二预测标签用以预测每一个用户对应的多维度标准数据属于所述第一标签还是所述第二标签。

所述根据所述第一 xgboost 模型输出的预测结果计算得到所述第一输出结果可以包括：获取所述预测结果中的第一预测标签及第一预测概率，第二预测标签及第二预测概率；获取所述第一预测标签及所述第二预测标签中第一标签的第一个数；获取所述第一预测标签及所述第二预测标签中第二标签的第二个数；计算所述第一个数与所述第二个数之和得到个数和；计算所述第一个数与所述个数和得到第一比例值；计算所述第二个数和所述个数和得到第二比例值。所述第一输出结果包括：第一标签，第一比例值；第二标签，第二比例值。所述第一比例值与所述第二比例值之和为 1，第一比例值代表第一标签所占的比例，第二比例值代表第二标签所占的比例。所述第一标签与 A 版本对应，所述第二标签与 B 版本对应。

所述训练模块 203，还用于根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配，并基于所述第一流量分配对应的多维度标准数据训练第二 xgboost 模型，并根据所述第二 xgboost 模型输出的预测结果计算得到第二输出结果。

在第一轮训练第一 xgboost 模型之后，需要对初始流量分配进行更新，以便进行第二轮 xgboost 模型的训练。

所述第二输出结果的计算过程同所述第一输出结果的计算过程。所述第二输出结果包括：第一标签，第一比例值；第二标签，第二比例值。所述第一比例值与所述第二比例值之和为 1，第一比例值代表第一标签所占的比例，第二比例值代表第二标签所占的比例。所述第一标签与 A 版本对应，所述第二标签与 B 版本对应。

在一个可选的实施例中，所述根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配包括：

判断所述第一比例值是否小于所述第二比例值；

当确定所述第一比例值小于所述第二比例值时，在所述初始流量分配的基础上增加第一值得到第一流量分配；

当确定所述第一比例值大于所述第二比例值时，在所述初始流量分配的基础上减少第二值得到第一流量分配。

该可选的实施例中，第一比例值小于所述第二比例值，表明 B 版本表现出了优于 A 版本的效应，则增加初始流量分配，使得分配给 A 版本的第一多维度标准数据的数据量减少，分配给 B 版本的第二多维度标准数据的数据量增加。第一比例值大于所述第二比例值，表明 A 版本表现出了优于 B 版本的效应，则减少初始流量分配，使得分配给 A 版本的第一多维度标准数据的数据量增加，分配给 B 版本的第二多维度标准数据的数据量减少。

所述判断模块 204，还用于判断所述第二输出结果是否优于所述第一输出结果。

其中，所述预设结果阈值为预先设置的临界值，用以表示 xgboost 模型的结束条件。

如果第一输出结果中的第一比例值小于第二输出结果中的第一比例值，则表明第二输出结果优于第一输出结果。如果第一输出结果中的第一比例值大于第二输出结果中的第一比例值，则表明第二输出结果差于第一输出结果。

所述训练模块 203，还用于当确定所述第二输出结果优于所述第一输出结果时，根据所述第二输出结果对所述第一流量分配进行更新得到第二流量分配，并基于所述第二流量分配对应的多维度标准数据训练第三 xgboost 模型。

当确定所述第二输出结果优于所述第一输出结果时，在所述第一流量分配的基础上增加第一值得到第二流量分配；当确定所述第一输出结果优于所述第二输出结果时，在所述第一流量分配的基础上减少第二值得到第二流量分配。

在根据所述第二输出结果对所述第一流量分配进行更新得到第二流量分配之后，基于所述第二流量分配对应的多维度标准数据训练第三 xgboost 模型，并根据所述第三 xgboost 模型

输出的预测结果计算得到第三输出结果。判断所述第二输出结果是否优于所述第二输出结果及判断第三输出结果是否大于预设结果阈值。当判断所述第二输出结果优于所述第二输出结果且第三输出结果小于预设结果阈值时，根据所述第三输出结果对所述第二流量分配进行更新得到第三流量分配，并基于所述第三流量分配对应的多维度标准数据训练第四 xgboost 模型，以此类推。当判断所述第二输出结果优于所述第二输出结果且判断第三输出结果大于或者等于预设结果阈值时，停止所述第三 xgboost 模型的训练。

对于统计功效更优的版本，进行流量动态变更，并使用训练结果模型验证数据结果一致性，若实验结果出现不一致结果集，将重新进行自学习，直至模型结果收敛，并且实验结果中将进行多次流量自动调整。

所述确定模块 205，用于当根据所述第三 xgboost 模型输出的预测结果计算得到的第三输出结果大于或者等于预设结果阈值时，确定所述第三输出结果为 AB 测试的目标结果。

停止对所述第三 xgboost 模型的训练，表明 AB 测试已经结束，可以将第三 xgboost 模型的第三输出结果确定为目标结果。其中，所述第三输出结果的计算过程同所述第一输出结果的计算过程。

所述通知模块 206，还用于在所述确定所述第三输出结果为 AB 测试的目标结果之后，获取所述目标结果中的 A 结果及 B 结果；比较所述 A 结果及所述 B 结果；当确定所述 A 结果小于所述 B 结果，发送通知信息至多个用户客户端，使得所述多个用户客户端更新应用为 B 版本。

所述第三输出结果包括：第一标签，第一比例值；第二标签，第二比例值。所述第一比例值与所述第二比例值之和为 1，第一比例值代表第一标签所占的比例，第二比例值代表第二标签所占的比例。所述第一标签与 A 版本对应，所述第二标签与 B 版本对应。当第一比例值小于第二比例值时，表示 A 结果小于 B 结果。当第一比例值大于第二比例值时，表示 A 结果大于 B 结果。

在多次训练 xgboost 模型之后，即可模拟 AB 测试的长期效应。当长期效应表示 A 结果小于 B 结果，则表明 B 版本应用的效应优于 A 版本应用的效应；当长期效应表示 A 结果大于 B 结果，则表明 A 版本应用的效应优于 B 版本应用的效应。

当长期效应表明某个版本优于另一个版本，将可以发送通知信息至用户客户端，通知用户客户端更新版本至更优版本。

需要说明的是，本申请提供的装置适用于同一个应用的两个不同的版本的测试，测试哪一个版本对用户更有效；本申请提供的装置还可适用于两个不同的方案的测试，测试哪一个方案对用户更有效；本申请提供的装置还可适用于同一个应用的多个不同的版本的测试或者用于多个不同的方案的测试。

综上，本申请可应用于智慧政务等智慧场景中，从而推动智慧城市的建设。本申请通过初始流量分配执行 AB 测试并获取 AB 测试的功效指标值；基于所述初始流量分配对应的多维度标准数据及所述功效指标值训练 xgboost 模型，并获取所述 xgboost 模型的输出结果；根据所述输出结果对所述初始流量分配进行更新得到第一流量分配，并迭代训练 xgboost 模型直至 xgboost 模型的输出结果大于或者等于预设结果阈值时，将 xgboost 模型的输出结果确定为 AB 测试的目标结果。通过不断迭代训练 xgboost 模型从而自动模拟 AB 测试，提高了 AB 测试的测试效率；此外，只有在 xgboost 模型的输出结果大于或者等于预设结果阈值时，将 xgboost 模型的输出结果确定为 AB 测试的目标结果，能够保证 AB 测试结果的长期效应与短期效应的一致性。

需要强调的是，为进一步保证上述目标结果的私密性和安全性，上述目标结果可存储于区块链的节点中。

参阅图 3 所示，为本申请实施例三提供的计算机设备的结构示意图。在本申请较佳实施例中，所述计算机设备 3 包括存储器 31、至少一个处理器 32、至少一条通信总线 33 及收发器 34。

本领域技术人员应该了解，图3示出的计算机设备的结构并不构成本申请实施例的限定，既可以是总线型结构，也可以是星形结构，所述计算机设备3还可以包括比图示更多或更少的其他硬件或者软件，或者不同的部件布置。

在一些实施例中，所述计算机设备3是一种能够按照事先设定或存储的指令，自动进行数值计算和/或信息处理的计算机设备，其硬件包括但不限于微处理器、专用集成电路、可编程门阵列、数字处理器及嵌入式设备等。所述计算机设备3还可包括客户设备，所述客户设备包括但不限于任何一种可与客户通过键盘、鼠标、遥控器、触摸板或声控设备等方式进行人机交互的电子产品，例如，个人计算机、平板电脑、智能手机、数码相机等。

需要说明的是，所述计算机设备3仅为举例，其他现有的或今后可能出现的电子产品如可适应于本申请，也应包含在本申请的保护范围以内，并以引用方式包含于此。

在一些实施例中，所述存储器31中存储有计算机可读指令，所述计算机可读指令被所述至少一个处理器32执行时实现如所述的基于人工智能的AB测试方法中的全部或者部分步骤。所述存储器31包括只读存储器(Read-Only Memory, ROM)、可编程只读存储器(Programmable Read-Only Memory, PROM)、可擦除可编程只读存储器(Erasable Programmable Read-Only Memory, EPROM)、一次可编程只读存储器(One-time Programmable Read-Only Memory, OTPROM)、电子擦除式可复写只读存储器(Electrically-Erasable Programmable Read-Only Memory, EEPROM)、只读光盘(Compact Disc Read-Only Memory, CD-ROM)或其他光盘存储器、磁盘存储器、磁带存储器、或者能够用于携带或存储数据的计算机可读的任何其他介质。所述计算机可读存储介质可以是非易失性，也可以是易失性的。

进一步地，所述计算机可读存储介质可主要包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序等；存储数据区可存储根据区块链节点的使用所创建的数据等。

本申请所指区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链(Blockchain)，本质上是一个去中心化的数据库，是一串使用密码学方法相关联产生的数据块，每一个数据块中包含了一批次网络交易的信息，用于验证其信息的有效性(防伪)和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层等。

在一些实施例中，所述至少一个处理器32是所述计算机设备3的控制核心(Control Unit)，利用各种接口和线路连接整个计算机设备3的各个部件，通过运行或执行存储在所述存储器31内的程序或者模块，以及调用存储在所述存储器31内的数据，以执行计算机设备3的各种功能和处理数据。例如，所述至少一个处理器32执行所述存储器中存储的计算机可读指令时实现本申请实施例中所述的基于人工智能的AB测试方法的全部或者部分步骤；或者实现基于人工智能的AB测试装置的全部或者部分功能。所述至少一个处理器32可以由集成电路组成，例如可以由单个封装的集成电路所组成，也可以是由多个相同功能或不同功能封装的集成电路所组成，包括一个或者多个中央处理器(Central Processing unit, CPU)、微处理器、数字处理芯片、图形处理器及各种控制芯片的组合等。

在一些实施例中，所述至少一条通信总线33被设置为实现所述存储器31以及所述至少一个处理器32等之间的连接通信。

尽管未示出，所述计算机设备3还可以包括给各个部件供电的电源(比如电池)，优选的，电源可以通过电源管理装置与所述至少一个处理器32逻辑相连，从而通过电源管理装置实现管理充电、放电、以及功耗管理等功能。电源还可以包括一个或一个以上的直流或交流电源、再充电装置、电源故障检测电路、电源转换器或者逆变器、电源状态指示器等任意组件。所述计算机设备3还可以包括多种传感器、蓝牙模块、Wi-Fi模块等，在此不再赘述。

上述以软件功能模块的形式实现的集成的单元，可以存储在一个计算机可读取存储介质中。上述软件功能模块存储在一个存储介质中，包括若干指令用以使得一台计算机设备(可以是个人计算机、计算机设备、或者网络设备等)或处理器(processor)执行本申请各个实施

例所述方法的部分。

在本申请所提供的几个实施例中，应该理解到，所揭露的装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述模块的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式。

所述作为分离部件说明的模块可以是或者也可以不是物理上分开的，作为模块显示的部件可以是或者也可以不是物理单元，既可以位于一个地方，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者全部模块来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能模块可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用硬件加软件功能模块的形式实现。

对于本领域技术人员而言，显然本申请不限于上述示范性实施例的细节，而且在不背离本申请的精神或基本特征的情况下，能够以其他的具体形式实现本申请。因此，无论从哪一点来看，均应将实施例看作是示范性的，而且是非限制性的，本申请的范围由所附权利要求而不是上述说明限定，因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本申请内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。此外，显然“包括”一词不排除其他单元或，单数不排除复数。说明书中陈述的多个单元或装置也可以由一个单元或装置通过软件或者硬件来实现。第一，第二等词语用来表示名称，而并不表示任何特定的顺序。

最后应说明的是，以上实施例仅用以说明本申请的技术方案而非限制，尽管参照较佳实施例对本申请进行了详细说明，本领域的普通技术人员应当理解，可以对本申请的技术方案进行修改或等同替换，而不脱离本申请技术方案的精神和范围。

权 利 要 求 书

1. 一种基于人工智能的 AB 测试方法，其中，所述方法包括：

采集用户多维度特征数据并对所述多维度特征数据进行数据清洗得到多维度标准数据；

根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值；

基于所述初始流量分配对应的多维度标准数据及所述功效指标值训练第一 xgboost 模型，并根据所述第一 xgboost 模型输出的预测结果计算得到第一输出结果；

根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配，并基于所述第一流量分配对应的多维度标准数据训练第二 xgboost 模型，并根据所述第二 xgboost 模型输出的预测结果计算得到第二输出结果；

判断所述第二输出结果是否优于所述第一输出结果；

当确定所述第二输出结果优于所述第一输出结果时，迭代根据所述第二输出结果对所述第一流量分配进行更新得到第二流量分配，并基于所述第二流量分配对应的多维度标准数据训练第三 xgboost 模型，直至根据所述第三 xgboost 模型输出的预测结果计算得到的第三输出结果大于或者等于预设结果阈值时，停止所述第三 xgboost 模型的训练；

确定所述第三输出结果为 AB 测试的目标结果。

2. 如权利要求 1 所述的基于人工智能的 AB 测试方法，其中，所述根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值包括：

判断多维度标准数据的数据量；

根据预设数据量范围与流量分配比例之间的映射表确定与所述数据量对应的目标流量分配比例，作为所述初始流量分配；

按照所述初始流量分配切分所述多维度标准数据为第一多维度标准数据和第二多维度标准数据；

将所述第一多维度标准数据分配给 A 版本及将所述第二多维度标准数据分配给 B 版本并执行模拟测试过程；

获取模拟测试结束后的测试数据并根据所述测试数据计算功效指标值。

3. 如权利要求 1 所述的基于人工智能的 AB 测试方法，其中，所述根据所述第一 xgboost 模型输出的预测结果计算得到第一输出结果包括：

获取所述第一 xgboost 模型输出的预测结果；

获取所述预测结果中的第一预测标签及第一预测概率，第二预测标签及第二预测概率；

获取所述第一预测标签及所述第二预测标签中第一标签的第一个数；

获取所述第一预测标签及所述第二预测标签中第二标签的第二个数；

计算所述第一个数与所述第二个数之和得到个数和；

计算所述第一个数与所述个数和得到第一比例值；

计算所述第二个数和所述个数和得到第二比例值。

4. 如权利要求 3 所述的基于人工智能的 AB 测试方法，其中，所述根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配包括：

判断所述第一比例值是否小于所述第二比例值；

当确定所述第一比例值小于所述第二比例值时，在所述初始流量分配的基础上增加第一值得到第一流量分配；

当确定所述第一比例值大于所述第二比例值时，在所述初始流量分配的基础上减少第二值得到第一流量分配。

5. 如权利要求 1 所述的基于人工智能的 AB 测试方法，其中，在所述确定所述第三输出结果为 AB 测试的目标结果之后，所述方法还包括：

获取所述目标结果中的 A 结果及 B 结果；

比较所述 A 结果及所述 B 结果；

当确定所述 A 结果小于所述 B 结果，发送通知信息至多个用户客户端，使得所述多个用

户客户端更新应用为 B 版本。

6. 如权利要求 5 所述的基于人工智能的 AB 测试方法，其中，所述采集用户多维度特征数据包括：

 定义多维度关键绩效指标并发送所述多维度关键绩效指标至所述多个用户客户端；

 获取所述多个用户客户端上报的埋点数据；

 从所述埋点数据中采集所述用户多维度特征数据。

7. 如权利要求 6 所述的基于人工智能的 AB 测试方法，其中，所述对所述多维度特征数据进行数据清洗得到多维度标准数据包括：

 对所述多维度特征数据进行去重处理；

 采用随机森林模型检测去重处理后的多维度特征数据中的异常数据；

 删除所述去重处理后的多维度特征数据中的异常数据得到多维度标准数据。

8. 一种基于人工智能的 AB 测试装置，其中，所述装置包括：

 采集模块，用于采集用户多维度特征数据并对所述多维度特征数据进行数据清洗得到多维度标准数据；

 测试模块，用于根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值；

 训练模块，用于基于所述初始流量分配对应的多维度标准数据及所述功效指标值训练第一 xgboost 模型，并根据所述第一 xgboost 模型输出的预测结果计算得到第一输出结果；

 所述训练模块，还用于根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配，并基于所述第一流量分配对应的多维度标准数据训练第二 xgboost 模型，并根据所述第二 xgboost 模型输出的预测结果计算得到第二输出结果；

 判断模块，用于判断所述第二输出结果是否优于所述第一输出结果；

 所述训练模块，还用于当确定所述第二输出结果优于所述第一输出结果时，迭代根据所述第二输出结果对所述第一流量分配进行更新得到第二流量分配，并基于所述第二流量分配对应的多维度标准数据训练第三 xgboost 模型，直至根据所述第三 xgboost 模型输出的预测结果计算得到的第三输出结果大于或者等于预设结果阈值时，停止所述第三 xgboost 模型的训练；

 确定模块，用于确定所述第三输出结果为 AB 测试的目标结果。

9. 一种计算机设备，其中，所述计算机设备包括：

 处理器，所述处理器用于执行存储器中存储的计算机可读指令时实现以下步骤：

 采集用户多维度特征数据并对所述多维度特征数据进行数据清洗得到多维度标准数据；

 根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值；

 基于所述初始流量分配对应的多维度标准数据及所述功效指标值训练第一 xgboost 模型，并根据所述第一 xgboost 模型输出的预测结果计算得到第一输出结果；

 根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配，并基于所述第一流量分配对应的多维度标准数据训练第二 xgboost 模型，并根据所述第二 xgboost 模型输出的预测结果计算得到第二输出结果；

 判断所述第二输出结果是否优于所述第一输出结果；

 当确定所述第二输出结果优于所述第一输出结果时，迭代根据所述第二输出结果对所述第一流量分配进行更新得到第二流量分配，并基于所述第二流量分配对应的多维度标准数据训练第三 xgboost 模型，直至根据所述第三 xgboost 模型输出的预测结果计算得到的第三输出结果大于或者等于预设结果阈值时，停止所述第三 xgboost 模型的训练；

 确定所述第三输出结果为 AB 测试的目标结果。

10. 如权利要求 9 所述的计算机设备，其中，所述处理器执行所述计算机可读指令以实现根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值时，具体包括：

 判断多维度标准数据的数据量；

 根据预设数据量范围与流量分配比例之间的映射表确定与所述数据量对应的目标流量分

配比例，作为所述初始流量分配；

按照所述初始流量分配切分所述多维度标准数据为第一多维度标准数据和第二多维度标准数据；

将所述第一多维度标准数据分配给 A 版本及将所述第二多维度标准数据分配给 B 版本并执行模拟测试过程；

获取模拟测试结束后的测试数据并根据所述测试数据计算功效指标值。

11. 如权利要求 9 所述的计算机设备，其中，所述处理器执行所述计算机可读指令以实现根据所述第一 xgboost 模型输出的预测结果计算得到第一输出结果时，具体包括：

获取所述第一 xgboost 模型输出的预测结果；

获取所述预测结果中的第一预测标签及第一预测概率，第二预测标签及第二预测概率；

获取所述第一预测标签及所述第二预测标签中第一标签的第一个数；

获取所述第一预测标签及所述第二预测标签中第二标签的第二个数；

计算所述第一个数与所述第二个数之和得到个数和；

计算所述第一个数与所述个数和得到第一比例值；

计算所述第二个数和所述个数和得到第二比例值。

12. 如权利要求 11 所述的计算机设备，其中，所述处理器执行所述计算机可读指令以实现根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配时，具体包括：

判断所述第一比例值是否小于所述第二比例值；

当确定所述第一比例值小于所述第二比例值时，在所述初始流量分配的基础上增加第一值得到第一流量分配；

当确定所述第一比例值大于所述第二比例值时，在所述初始流量分配的基础上减少第二值得到第一流量分配。

13. 如权利要求 9 所述的计算机设备，其中，在所述确定所述第三输出结果为 AB 测试的目标结果之后，所述处理器执行所述计算机可读指令还用以实现以下步骤：

获取所述目标结果中的 A 结果及 B 结果；

比较所述 A 结果及所述 B 结果；

当确定所述 A 结果小于所述 B 结果，发送通知信息至多个用户客户端，使得所述多个用户客户端更新应用为 B 版本。

14. 如权利要求 13 所述的计算机设备，其中，所述处理器执行所述计算机可读指令以实现采集用户多维度特征数据时，具体包括：

定义多维度关键绩效指标并发送所述多维度关键绩效指标至所述多个用户客户端；

获取所述多个用户客户端上报的埋点数据；

从所述埋点数据中采集所述用户多维度特征数据。

15. 如权利要求 14 所述的计算机设备，其中，所述处理器执行所述计算机可读指令以实现对所述多维度特征数据进行数据清洗得到多维度标准数据时，具体包括：

对所述多维度特征数据进行去重处理；

采用随机森林模型检测去重处理后的多维度特征数据中的异常数据；

删除所述去重处理后的多维度特征数据中的异常数据得到多维度标准数据。

16. 一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机可读指令，其中，所述计算机可读指令被处理器执行时实现以下步骤：

采集用户多维度特征数据并对所述多维度特征数据进行数据清洗得到多维度标准数据；

根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值；

基于所述初始流量分配对应的多维度标准数据及所述功效指标值训练第一 xgboost 模型，并根据所述第一 xgboost 模型输出的预测结果计算得到第一输出结果；

根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配，并基于所述第一流量分配对应的多维度标准数据训练第二 xgboost 模型，并根据所述第二 xgboost 模型输出

的预测结果计算得到第二输出结果；

判断所述第二输出结果是否优于所述第一输出结果；

当确定所述第二输出结果优于所述第一输出结果时，迭代根据所述第二输出结果对所述第一流量分配进行更新得到第二流量分配，并基于所述第二流量分配对应的多维度标准数据训练第三 xgboost 模型，直至根据所述第三 xgboost 模型输出的预测结果计算得到的第三输出结果大于或者等于预设结果阈值时，停止所述第三 xgboost 模型的训练；

确定所述第三输出结果为 AB 测试的目标结果。

17. 如权利要求 16 所述的计算机可读存储介质，其中，所述计算机可读指令被所述处理器执行以实现根据初始流量分配执行 AB 测试并获取 AB 测试的功效指标值时，具体包括：

判断多维度标准数据的数据量；

根据预设数据量范围与流量分配比例之间的映射表确定与所述数据量对应的目标流量分配比例，作为所述初始流量分配；

按照所述初始流量分配切分所述多维度标准数据为第一多维度标准数据和第二多维度标准数据；

将所述第一多维度标准数据分配给 A 版本及将所述第二多维度标准数据分配给 B 版本并执行模拟测试过程；

获取模拟测试结束后的测试数据并根据所述测试数据计算功效指标值。

18. 如权利要求 16 所述的计算机可读存储介质，其中，所述计算机可读指令被所述处理器执行以实现根据所述第一 xgboost 模型输出的预测结果计算得到第一输出结果时，具体包括：

获取所述第一 xgboost 模型输出的预测结果；

获取所述预测结果中的第一预测标签及第一预测概率，第二预测标签及第二预测概率；

获取所述第一预测标签及所述第二预测标签中第一标签的第一个数；

获取所述第一预测标签及所述第二预测标签中第二标签的第二个数；

计算所述第一个数与所述第二个数之和得到个数和；

计算所述第一个数与所述个数和得到第一比例值；

计算所述第二个数和所述个数和得到第二比例值。

19. 如权利要求 18 所述的计算机可读存储介质，其中，所述计算机可读指令被所述处理器执行以实现根据所述第一输出结果对所述初始流量分配进行更新得到第一流量分配时，具体包括：

判断所述第一比例值是否小于所述第二比例值；

当确定所述第一比例值小于所述第二比例值时，在所述初始流量分配的基础上增加第一值得到第一流量分配；

当确定所述第一比例值大于所述第二比例值时，在所述初始流量分配的基础上减少第二值得到第一流量分配。

20. 如权利要求 16 所述的计算机可读存储介质，其中，在所述确定所述第三输出结果为 AB 测试的目标结果之后，所述计算机可读指令被所述处理器执行还用以实现以下步骤：

获取所述目标结果中的 A 结果及 B 结果；

比较所述 A 结果及所述 B 结果；

当确定所述 A 结果小于所述 B 结果，发送通知信息至多个用户客户端，使得所述多个用户客户端更新应用为 B 版本。

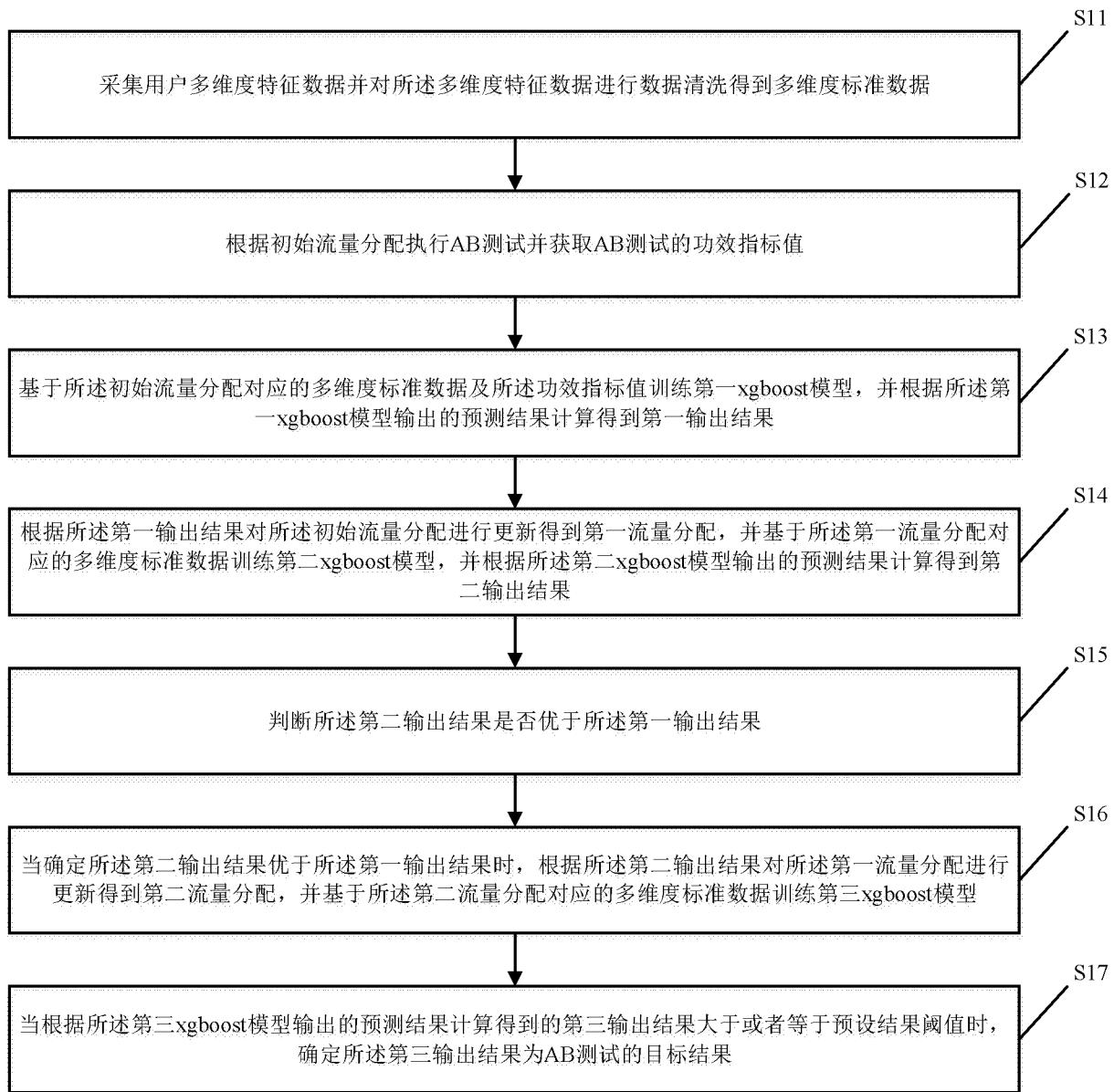


图 1

2/2

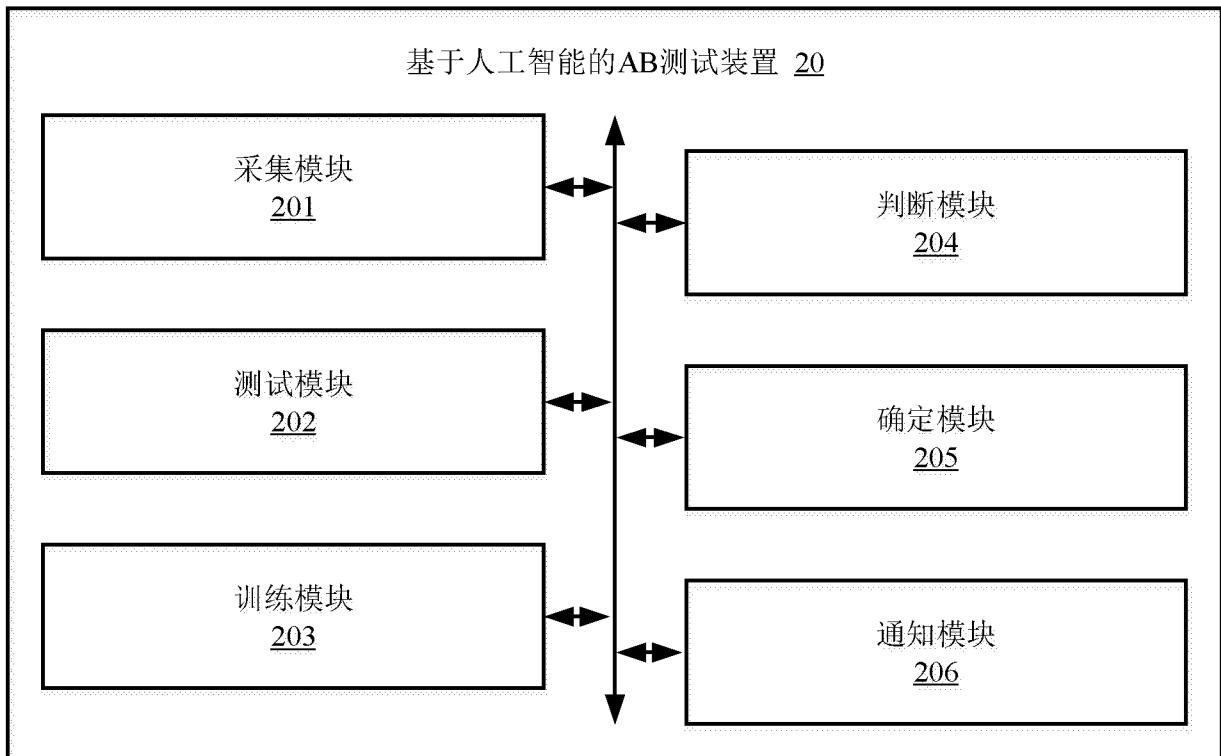


图 2

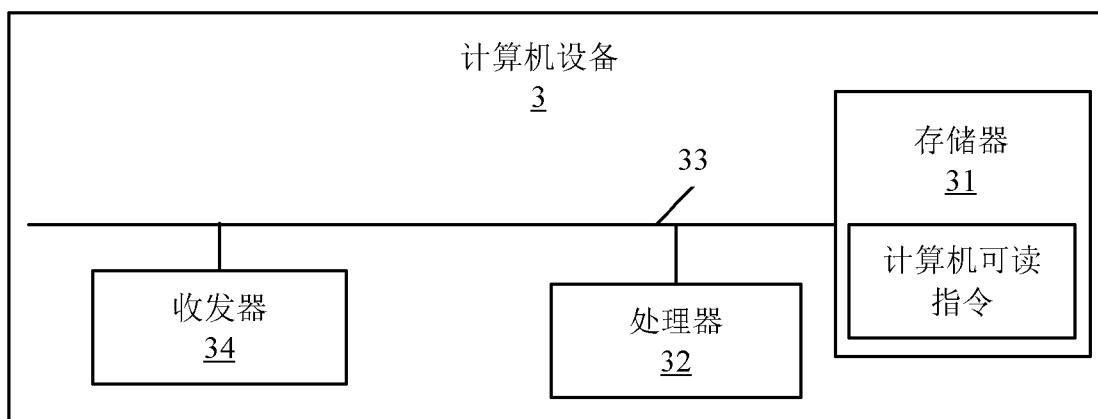


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/090460

A. CLASSIFICATION OF SUBJECT MATTER

G06Q 10/06(2012.01)i; G06F 11/36(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06Q; G06F; H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, CNKI, VEN, USTXT, WOTXT, EPTXT, IEEE: AB测试, 机器学习, 模型, 迭代, 训练, 流量, 分配, AB test+, A/B test+, controlled experiment?, split test+, bucket+ test+, xgboost, machine learning, model, iterat+, flow, distribut+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|--|-----------------------|
| PX | CN 112700131 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 23 April 2021 (2021-04-23) description, paragraphs 59-177, and figures 1-3 | 1-20 |
| A | CN 111008263 A (ALIPAY (HANGZHOU) INFORMATION TECHNOLOGY CO., LTD.) 14 April 2020 (2020-04-14) description, paragraphs 45-75, and figures 1-5 | 1-20 |
| A | CN 111950622 A (PING AN LIFE INSURANCE COMPANY OF CHINA, LTD.) 17 November 2020 (2020-11-17) description paragraphs 59-121, paragraphs 124-176, figures 1-3 | 1-20 |
| A | CN 110168595 A (SHANGHAI UPLTV TECHNOLOGY CO., LTD.) 23 August 2019 (2019-08-23) description, paragraphs 43-75, and figures 1-4 | 1-20 |
| A | WO 2019019958 A1 (BEIJING DIDI INFINITY TECHNOLOGY AND DEVELOPMENT CO., LTD.) 31 January 2019 (2019-01-31) entire document | 1-20 |

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

| | |
|--|---|
| Date of the actual completion of the international search 13 July 2021 | Date of mailing of the international search report 25 August 2021 |
|--|---|

| | |
|--|--------------------|
| Name and mailing address of the ISA/CN | Authorized officer |
|--|--------------------|

China National Intellectual Property Administration (ISA/CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088, China

| | |
|--------------------------------------|---------------|
| Facsimile No. (86-10)62019451 | Telephone No. |
|--------------------------------------|---------------|

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/CN2021/090460

| Patent document cited in search report | | | Publication date (day/month/year) | Patent family member(s) | | | Publication date (day/month/year) |
|--|------------|----|-----------------------------------|-------------------------|---------------|----|-----------------------------------|
| CN | 112700131 | A | 23 April 2021 | None | | | |
| CN | 111008263 | A | 14 April 2020 | None | | | |
| CN | 111950622 | A | 17 November 2020 | None | | | |
| CN | 110168595 | A | 23 August 2019 | WO | 2019200600 | A1 | 24 October 2019 |
| WO | 2019019958 | A1 | 31 January 2019 | US | 2019130333 | A1 | 02 May 2019 |
| | | | | EP | 3459208 | A4 | 24 July 2019 |
| | | | | CN | 109565452 | A | 02 April 2019 |
| | | | | CA | 3028291 | A1 | 19 March 2019 |
| | | | | JP | 2019534487 | A | 28 November 2019 |
| | | | | AU | 2018282441 | A1 | 14 February 2019 |
| | | | | EP | 3459208 | A1 | 27 March 2019 |
| | | | | CN | 109302301 | A | 01 February 2019 |
| | | | | IN | 201817048691A | A | 24 January 2020 |
| | | | | CN | 109308552 | A | 05 February 2019 |
| | | | | SG | 11201811512 | A1 | 27 February 2019 |

国际检索报告

国际申请号

PCT/CN2021/090460

A. 主题的分类

G06Q 10/06 (2012. 01) i; G06F 11/36 (2006. 01) i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06Q; G06F; H04L

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS, CNTXT, CNKI, VEN, USTXT, WOTXT, EPTXT, IEEE: AB测试, 机器学习, 模型, 迭代, 叠代, 训练, 流量, 分配, AB test+, A/B test+, controlled experiment?, split test+, bucket+ test+, xgboost, machine learning, model, iterat+, flow, distribut+

C. 相关文件

| 类 型* | 引用文件, 必要时, 指明相关段落 | 相关的权利要求 |
|------|--|---------|
| PX | CN 112700131 A (平安科技深圳有限公司) 2021年 4月 23日 (2021 - 04 - 23) 说明书第59-177段, 附图1-3 | 1-20 |
| A | CN 111008263 A (支付宝杭州信息技术有限公司) 2020年 4月 14日 (2020 - 04 - 14) 说明书第45-75段, 附图1-5 | 1-20 |
| A | CN 111950622 A (中国平安人寿保险股份有限公司) 2020年 11月 17日 (2020 - 11 - 17) 说明书第59-121段, 第124-176段, 附图1-3 | 1-20 |
| A | CN 110168595 A (上海荟萃网络科技有限公司) 2019年 8月 23日 (2019 - 08 - 23) 说明书第43-75段, 附图1-4 | 1-20 |
| A | WO 2019019958 A1 (BEIJING DIDI INFINITY TECHNOLOGY & DEV CO LTD) 2019年 1月 31日 (2019 - 01 - 31) 全文 | 1-20 |

 其余文件在C栏的续页中列出。 见同族专利附件。

- * 引用文件的具体类型:
- "A" 认为不特别相关的表示了现有技术一般状态的文件
- "E" 在国际申请日的当天或之后公布的在先申请或专利
- "L" 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)
- "O" 涉及口头公开、使用、展览或其他方式公开的文件
- "P" 公布日先于国际申请日但迟于所要求的优先权日的文件

- "T" 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件
- "X" 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性
- "Y" 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性
- "&" 同族专利的文件

国际检索实际完成的日期

2021年 7月 13日

国际检索报告邮寄日期

2021年 8月 25日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)
 中国 北京市海淀区蓟门桥西土城路6号 100088
 传真号 (86-10)62019451

受权官员

钱紫娟

电话号码 (86-27)59371331

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/090460

| 检索报告引用的专利文件 | | 公布日 (年/月/日) | | 同族专利 | | 公布日 (年/月/日) | |
|-------------|------------|----------------|---------------|------|---------------|----------------|---------------|
| CN | 112700131 | A | 2021年 4月 23日 | 无 | | | |
| CN | 111008263 | A | 2020年 4月 14日 | 无 | | | |
| CN | 111950622 | A | 2020年 11月 17日 | 无 | | | |
| CN | 110168595 | A | 2019年 8月 23日 | WO | 2019200600 | A1 | 2019年 10月 24日 |
| W0 | 2019019958 | A1 | 2019年 1月 31日 | US | 2019130333 | A1 | 2019年 5月 2日 |
| | | | | EP | 3459208 | A4 | 2019年 7月 24日 |
| | | | | CN | 109565452 | A | 2019年 4月 2日 |
| | | | | CA | 3028291 | A1 | 2019年 3月 19日 |
| | | | | JP | 2019534487 | A | 2019年 11月 28日 |
| | | | | AU | 2018282441 | A1 | 2019年 2月 14日 |
| | | | | EP | 3459208 | A1 | 2019年 3月 27日 |
| | | | | CN | 109302301 | A | 2019年 2月 1日 |
| | | | | IN | 201817048691A | A | 2020年 1月 24日 |
| | | | | CN | 109308552 | A | 2019年 2月 5日 |
| | | | | SG | 11201811512 | A1 | 2019年 2月 27日 |