



(12)发明专利申请

(10)申请公布号 CN 111292168 A

(43)申请公布日 2020.06.16

(21)申请号 202010081761.X

(22)申请日 2020.02.06

(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 潘颖吉

(74)专利代理机构 广州三环专利商标代理有限
公司 44202

代理人 熊永强 杜维

(51)Int.Cl.

G06Q 30/06(2012.01)

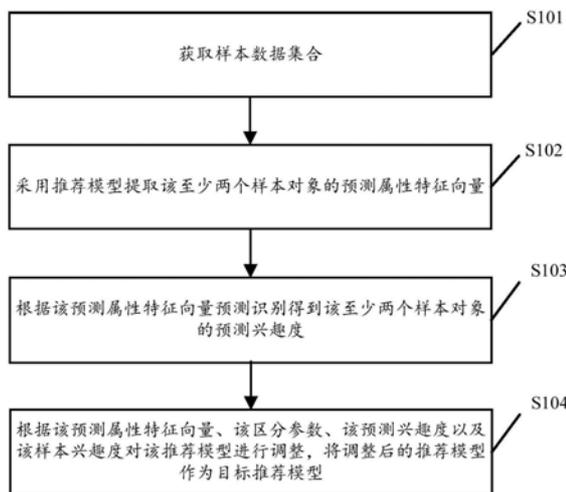
权利要求书4页 说明书25页 附图11页

(54)发明名称

数据处理方法、装置及设备

(57)摘要

本申请实施例公开了一种数据处理方法、装置及设备,属于云技术-人工智能服务领域。该方法包括:获取样本数据集合,样本数据集合包括至少两个样本对象的样本兴趣度以及用于区分至少两个样本对象的区分参数,采用推荐模型提取至少两个样本对象的预测属性特征向量,根据预测属性特征向量预测识别得到至少两个样本对象的预测兴趣度,根据预测属性特征向量、区分参数、预测兴趣度以及样本兴趣度对推荐模型进行调整,将调整后的推荐模型作为目标推荐模型,目标推荐模型所输出针对至少两个样本对象的更新预测属性特征向量之间的区分程度,与区分参数对应的区分程度相匹配。通过本申请实例可提高推荐模型对对象的区分度,可实现精准推荐。



1. 一种数据处理方法,其特征在于,包括:

获取样本数据集,所述样本数据集包括至少两个样本对象的样本兴趣度以及用于区分所述至少两个样本对象的区分参数;

采用推荐模型提取所述至少两个样本对象的预测属性特征向量;

根据所述预测属性特征向量预测识别得到所述至少两个样本对象的预测兴趣度;

根据所述预测属性特征向量、所述区分参数、所述预测兴趣度以及所述样本兴趣度对所述推荐模型进行调整,将调整后的推荐模型作为目标推荐模型,所述目标推荐模型所输出针对所述至少两个样本对象的更新预测属性特征向量之间的区分程度,与所述区分参数对应的区分程度相匹配。

2. 如权利要求1所述的方法,其特征在于,所述根据所述预测属性特征向量、所述区分参数、所述预测兴趣度以及所述样本兴趣度对所述推荐模型进行调整,将调整后的推荐模型作为目标推荐模型,包括:

采用所述预测兴趣度以及所述样本兴趣度生成所述推荐模型的原始损失值;

根据所述预测属性特征向量以及所述区分参数确定对象区分度;

采用所述对象区分度生成所述推荐模型的区分损失值;

采用所述区分损失值和所述原始损失值生成所述推荐模型的总损失值;

若所述推荐模型的总损失值不满足收敛条件,则根据所述推荐模型的总损失值对所述推荐模型进行调整,将调整后的推荐模型作为目标推荐模型。

3. 如权利要求2所述的方法,其特征在于,所述根据所述推荐模型的总损失值对所述推荐模型进行调整,将调整后的推荐模型作为目标推荐模型,包括:

据所述推荐模型的总损失值对所述推荐模型进行调整;

采用所述调整后的推荐模型提取所述至少两个样本对象的更新预测属性特征向量;

根据所述更新预测属性特征向量预测识别得到所述至少两个样本对象的更新预测兴趣度;

根据所述更新预测属性特征向量、所述区分参数、所述更新预测兴趣度以及所述样本兴趣度,生成所述调整后的推荐模型的总损失值;

获取所述更新预测属性特征向量之间的角度,根据所述更新预测属性特征向量之间的角度确定所述更新预测属性特征向量之间的区分程度;

若所述调整后的推荐模型的总损失值满足收敛条件,且所述更新预测属性特征向量之间的区分程度,与所述区分参数对应的区分程度之间的差值小于程度阈值,则将调整后的推荐模型作为目标推荐模型。

4. 如权利要求2所述的方法,其特征在于,所述区分参数为区分数值,所述根据所述预测属性特征向量以及所述区分参数确定对象区分度,包括:

获取所述预测属性特征向量的模长;

获取所述预测属性特征向量的模长与所述区分数值之间的差值;

根据所述差值确定所述对象区分度。

5. 如权利要求2所述的方法,其特征在于,所述区分参数为区分向量,所述根据所述预测属性特征向量以及所述区分参数确定对象区分度,包括:

对所述区分向量与所述预测属性特征向量进行乘积运算,得到向量角度;

根据所述向量角度确定所述对象区分度。

6. 如权利要求5所述的方法,其特征不在于,所述获取用于区分所述至少两个样本对象的区分参数,包括:

获取第一样本对象的语义信息,所述第一样本对象为所述至少两个样本对象的任一样本对象;

从数据库中筛选出与所述第一样本对象的语义信息匹配的目标待选区分向量,作为所述第一样本对象的区分向量,所述数据库中包括多个待选区分向量,一个所述待选区分向量与一种语义信息对应。

7. 如权利要求6所述的方法,其特征不在于,所述从数据库中筛选出与所述第一样本对象的语义信息匹配的目标待选区分向量,作为所述第一样本对象的区分向量,包括:

从数据库中筛选出与所述第一样本对象的语义信息匹配的目标待选区分向量;

对所述目标待选区分向量以及剩余待选区分向量进行降维处理;所述剩余待选区分向量为所述数据库中除所述目标待选区分向量以外的待选区分向量;

获取降维处理后的目标待选区分向量与降维处理后的剩余待选区分向量之间的距离;

若所述距离大于距离阈值,则将所述目标待选区分向量作为所述第一样本对象的区分向量。

8. 如权利要求2所述的方法,其特征不在于,所述至少两个样本对象包括第二样本对象和第三样本对象,所述采用所述预测兴趣度以及所述样本兴趣度生成所述推荐模型的原始损失值;包括:

采用所述第二样本对象的预测兴趣度以及所述第二样本对象的样本兴趣度生成第一原始损失值;

采用所述第三样本对象的预测兴趣度以及所述第三样本对象的样本兴趣度生成第二原始损失值;

对所述第一原始损失值与所述第二原始损失值进行求和运算,得到所述推荐模型的原始损失值;

所述采用所述对象区分度生成所述推荐模型的区分损失值,包括:

对所述第二样本对象的对象区分度和所述第三样本对象的对象区分度进行求和运算,得到所述推荐模型的区分损失值。

9. 如权利要求1所述的方法,其特征不在于,所述样本数据集包括访问所述至少两个样本对象的访问者的画像数据,所述根据所述预测属性特征向量预测识别得到所述至少两个样本对象的预测兴趣度;包括:

采用所述推荐模型对所述画像数据进行画像特征提取,得到预测画像特征向量;

采用所述推荐模型对所述预测画像特征向量和所述预测属性特征向量进行预测识别,得到所述至少两个样本对象的预测兴趣度。

10. 如权利要求9所述的方法,其特征不在于,所述采用所述推荐模型对所述画像数据进行画像特征提取,得到预测画像特征向量,包括:

获取所述至少两个样本对象中每个样本对象的对象类型;

从所述画像数据中筛选出与所述对象类型匹配的画像数据,作为有效画像数据;

采用所述推荐模型对所述有效画像数据进行画像特征提取,得到所述预测画像特征向

量。

11. 一种数据处理方法,其特征在于,包括:

获取待推荐的目标对象的属性信息,以及访问所述目标对象的目标访问者的画像数据;

采用目标推荐模型对所述目标对象的属性信息进行属性特征提取,得到所述目标对象的属性特征向量,所述目标推荐模型为采用如权利要求1-10任一项所述数据处理方法对目标推荐模型调整得到的;

采用所述目标推荐模型对所述目标访问者的画像数据进行画像特征提取,得到所述目标访问者的画像特征向量;

采用所述目标推荐模型对所述属性特征向量和所述画像特征向量进行兴趣度识别,得到所述目标访问者对所述目标对象的兴趣度;

根据所述目标访问者对所述目标对象的兴趣度向所述目标访问者推荐所述目标对象。

12. 如权利要求11所述的方法,其特征在于,所述根据所述目标访问者对所述目标对象的兴趣度向所述目标访问者推荐所述目标对象,包括:

获取与所述目标对象的对象类型相似度大于相似度阈值的待选对象;

根据所述目标访问者对所述目标对象的兴趣度向所述目标访问者推荐所述目标对象以及所述待选对象。

13. 一种数据处理装置,其特征在于,包括:

获取模块,用于获取样本数据集合,所述样本数据集合包括至少两个样本对象的样本兴趣度以及用于区分所述至少两个样本对象的区分参数;

提取模块,用于采用推荐模型提取所述至少两个样本对象的预测属性特征向量;

识别模块,用于根据所述预测属性特征向量预测识别得到所述至少两个样本对象的预测兴趣度;

调整模块,用于根据所述预测属性特征向量、所述区分参数、所述预测兴趣度以及所述样本兴趣度对所述推荐模型进行调整,将调整后的推荐模型作为目标推荐模型,所述目标推荐模型所输出针对所述至少两个样本对象的更新预测属性特征向量之间的区分程度,与所述区分参数对应的区分程度相匹配。

14. 一种数据处理装置,其特征在于,包括:

获取模块,用于获取待推荐的目标对象的属性信息,以及访问所述目标对象的目标访问者的画像数据;

提取模块,用于采用目标推荐模型对所述目标对象的属性信息进行属性特征提取,得到所述目标对象的属性特征向量,所述目标推荐模型为采用如权利要求1-10任一项所述数据处理方法对目标推荐模型调整得到的;采用所述目标推荐模型对所述目标访问者的画像数据进行画像特征提取,得到所述目标访问者的画像特征向量;

识别模块,用于采用所述目标推荐模型对所述属性特征向量和所述画像特征向量进行兴趣度识别,得到所述目标访问者对所述目标对象的兴趣度;

推荐模块,用于根据所述目标访问者对所述目标对象的兴趣度向所述目标访问者推荐所述目标对象。

15. 一种计算机设备,其特征在于,包括:处理器、存储器以及网络接口;

所述处理器与存储器、网络接口相连,其中,网络接口用于提供数据通信功能,所述存储器用于存储程序代码,所述处理器用于调用所述程序代码,以执行如权利要求1至12任一项所述的方法。

数据处理方法、装置及设备

技术领域

[0001] 本申请涉及云技术-人工智能服务领域,尤其涉及一种数据处理方法、装置及设备。

背景技术

[0002] 随着互联网和电子商务的发展,越来越多的商家采用线上销售方式来销售商品,即商家通过将商品的相关信息(如商品的类型、价格)发布至网站上,用户通过网站来购买想要的商品。但是,通常网站上包括大量商品,用户需要花费较长时间才能找到想要的商品,基于此,目前主要是根据用户的购买行为来为用户推荐商品,以辅助用户快速找到想要购买的商品。例如,用户A在网站上购买过啤酒和尿布,推荐终端确定啤酒和尿布具有相似性,会将啤酒和尿布推荐给用户A;如果用户B浏览过啤酒或尿布,推荐终端也会将啤酒和尿布推荐给用户B。由此可见,根据用户的购买行为来推荐商品,会出现为不同用户推荐相同商品的情况。但是,不同用户的兴趣点不一样,即不同用户想要购买的商品不一致,因此,基于用户购买行为的商品推荐方式的准确度比较低。

发明内容

[0003] 本申请实施例所要解决的技术问题在于,提供一种数据处理方法、装置及设备,通过提高推荐模型对对象的区分度,以提高推荐商品(即对象)的准确度。

[0004] 本申请实施例一方面提供一种数据处理方法,包括:

[0005] 获取样本数据集合,所述样本数据集合包括至少两个样本对象的样本兴趣度以及用于区分所述至少两个样本对象的区分参数;

[0006] 采用推荐模型提取所述至少两个样本对象的预测属性特征向量;

[0007] 根据所述预测属性特征向量预测识别得到所述至少两个样本对象的预测兴趣度;

[0008] 根据所述预测属性特征向量、所述区分参数、所述预测兴趣度以及所述样本兴趣度对所述推荐模型进行调整,将调整后的推荐模型作为目标推荐模型,所述目标推荐模型所输出针对所述至少两个样本对象的更新预测属性特征向量之间的区分程度,与所述区分参数对应的区分程度相匹配。

[0009] 本申请实施例一方面提供一种数据处理方法,包括:

[0010] 获取待推荐的目标对象的属性信息,以及访问所述目标对象的目标访问者的画像数据;

[0011] 采用目标推荐模型对所述目标对象的属性信息进行属性特征提取,得到所述目标对象的属性特征向量,所述目标推荐模型为采用上述数据处理方法对目标推荐模型调整得到的;

[0012] 采用所述目标推荐模型对所述目标访问者的画像数据进行画像特征提取,得到所述目标访问者的画像特征向量;

[0013] 采用所述目标推荐模型对所述属性特征向量和所述画像特征向量进行兴趣度识

别,得到所述目标访问者对所述目标对象的兴趣度;

[0014] 根据所述目标访问者对所述目标对象的兴趣度向所述目标访问者推荐所述目标对象。

[0015] 本申请实施例一方面提供一种数据处理装置,包括:

[0016] 获取模块,用于获取样本数据集合,所述样本数据集合包括至少两个样本对象的样本兴趣度以及用于区分所述至少两个样本对象的区分参数;

[0017] 提取模块,用于采用推荐模型提取所述至少两个样本对象的预测属性特征向量;

[0018] 识别模块,用于根据所述预测属性特征向量预测识别得到所述至少两个样本对象的预测兴趣度;

[0019] 调整模块,用于根据所述预测属性特征向量、所述区分参数、所述预测兴趣度以及所述样本兴趣度对所述推荐模型进行调整,将调整后的推荐模型作为目标推荐模型,所述目标推荐模型所输出针对所述至少两个样本对象的更新预测属性特征向量之间的区分程度,与所述区分参数对应的区分程度相匹配。

[0020] 可选的,调整模块,具体用于采用所述预测兴趣度以及所述样本兴趣度生成所述推荐模型的原始损失值;根据所述预测属性特征向量以及所述区分参数确定对象区分度;采用所述对象区分度生成所述推荐模型的区分损失值;采用所述区分损失值和所述原始损失值生成所述推荐模型的总损失值;若所述推荐模型的总损失值不满足收敛条件,则根据所述推荐模型的总损失值对所述推荐模型进行调整,将调整后的推荐模型作为目标推荐模型。

[0021] 可选的,调整模块,具体用于据所述推荐模型的总损失值对所述推荐模型进行调整;采用所述调整后的推荐模型提取所述至少两个样本对象的更新预测属性特征向量;根据所述更新预测属性特征向量预测识别得到所述至少两个样本对象的更新预测兴趣度;根据所述更新预测属性特征向量、所述区分参数、所述更新预测兴趣度以及所述样本兴趣度,生成所述调整后的推荐模型的总损失值;获取所述更新预测属性特征向量之间的角度,根据所述更新预测属性特征向量之间的角度确定所述更新预测属性特征向量之间的区分程度;若所述调整后的推荐模型的总损失值满足收敛条件,且所述更新预测属性特征向量之间的区分程度,与所述区分参数对应的区分程度之间的差值小于程度阈值,则将调整后的推荐模型作为目标推荐模型。

[0022] 可选的,所述区分参数为区分数值,调整模块,具体用于获取所述预测属性特征向量的模长;获取所述预测属性特征向量的模长与所述区分数值之间的差值;根据所述差值确定所述对象区分度。

[0023] 可选的,所述区分参数为区分向量,可选的,调整模块,具体用于对所述区分向量与所述预测属性特征向量进行乘积运算,得到向量角度;根据所述向量角度确定所述对象区分度。

[0024] 可选的,获取模块,具体用于获取第一样本对象的语义信息,所述第一样本对象为所述至少两个样本对象的任一样本对象;从数据库中筛选出与所述第一样本对象的语义信息匹配的目标待选区分向量,作为所述第一样本对象的区分向量,所述数据库中包括多个待选区分向量,一个所述待选区分向量与一种语义信息对应。

[0025] 可选的,获取模块,具体用于从数据库中筛选出与所述第一样本对象的语义信息

匹配的目标待选区分向量;对所述目标待选区分向量以及剩余待选区分向量进行降维处理;所述剩余待选区分向量为所述数据库中除所述目标待选区分向量以外的待选区分向量;获取降维处理后的目标待选区分向量与降维处理后的剩余待选区分向量之间的距离;若所述距离大于距离阈值,则将所述目标待选区分向量作为所述第一样本对象的区分向量。

[0026] 可选的,所述至少两个样本对象包括第二样本对象和第三样本对象,可选的,调整模块,具体用于采用所述第二样本对象的预测兴趣度以及所述第二样本对象的样本兴趣度生成第一原始损失值;采用所述第三样本对象的预测兴趣度以及所述第三样本对象的样本兴趣度生成第二原始损失值;对所述第一原始损失值与所述第二原始损失值进行求和运算,得到所述推荐模型的原始损失值;对所述第二样本对象的对象区分度和所述第三样本对象的对象区分度进行求和运算,得到所述推荐模型的区分损失值。

[0027] 可选的,所述样本数据集合包括访问所述至少两个样本对象的访问者的画像数据,识别模块,具体用于采用所述推荐模型对所述画像数据进行画像特征提取,得到预测画像特征向量;采用所述推荐模型对所述预测画像特征向量和所述预测属性特征向量进行预测识别,得到所述至少两个样本对象的预测兴趣度。

[0028] 可选的,识别模块,具体用于获取所述至少两个样本对象中每个样本对象的对象类型;从所述画像数据中筛选出与所述对象类型匹配的画像数据,作为有效画像数据;采用所述推荐模型对所述有效画像数据进行画像特征提取,得到所述预测画像特征向量。

[0029] 本申请实施例一方面提供一种数据处理装置,包括:

[0030] 获取模块,用于获取待推荐的目标对象的属性信息,以及访问所述目标对象的目标访问者的画像数据;

[0031] 提取模块,用于采用目标推荐模型对所述目标对象的属性信息进行属性特征提取,得到所述目标对象的属性特征向量,所述目标推荐模型为采用上述数据处理方法对目标推荐模型调整得到的;采用所述目标推荐模型对所述目标访问者的画像数据进行画像特征提取,得到所述目标访问者的画像特征向量;

[0032] 识别模块,用于采用所述目标推荐模型对所述属性特征向量和所述画像特征向量进行兴趣度识别,得到所述目标访问者对所述目标对象的兴趣度;

[0033] 推荐模块,用于根据所述目标访问者对所述目标对象的兴趣度向所述目标访问者推荐所述目标对象。

[0034] 可选的,推荐模块,用于获取与所述目标对象的对象类型相似度大于相似度阈值的待选对象;根据所述目标访问者对所述目标对象的兴趣度向所述目标访问者推荐所述目标对象以及所述待选对象。

[0035] 本申请一方面提供了一种计算机设备,包括:处理器、存储器、网络接口;

[0036] 上述处理器与存储器、网络接口相连,其中,网络接口用于提供数据通信功能,上述存储器用于存储计算机程序,上述处理器用于调用上述计算机程序,以执行如下步骤:

[0037] 获取样本数据集合,所述样本数据集合包括至少两个样本对象的样本兴趣度以及用于区分所述至少两个样本对象的区分参数;

[0038] 采用推荐模型提取所述至少两个样本对象的预测属性特征向量;

[0039] 根据所述预测属性特征向量预测识别得到所述至少两个样本对象的预测兴趣度;

[0040] 根据所述预测属性特征向量、所述区分参数、所述预测兴趣度以及所述样本兴趣度对所述推荐模型进行调整,将调整后的推荐模型作为目标推荐模型,所述目标推荐模型所输出针对所述至少两个样本对象的更新预测属性特征向量之间的区分程度,与所述区分参数对应的区分程度相匹配。

[0041] 可选的,上述处理器用于调用上述计算机程序,以执行如下步骤:

[0042] 获取待推荐的目标对象的属性信息,以及访问所述目标对象的目标访问者的画像数据;

[0043] 采用目标推荐模型对所述目标对象的属性信息进行属性特征提取,得到所述目标对象的属性特征向量,所述目标推荐模型为采用如权利要求1-10任一项所述数据处理方法对目标推荐模型调整得到的;

[0044] 采用所述目标推荐模型对所述目标访问者的画像数据进行画像特征提取,得到所述目标访问者的画像特征向量;

[0045] 采用所述目标推荐模型对所述属性特征向量和所述画像特征向量进行兴趣度识别,得到所述目标访问者对所述目标对象的兴趣度;

[0046] 根据所述目标访问者对所述目标对象的兴趣度向所述目标访问者推荐所述目标对象。

[0047] 本申请实施例一方面提供了一种计算机可读存储介质,上述计算机可读存储介质存储有计算机程序,上述计算机程序包括程序指令,上述程序指令当被处理器执行时,以执行如下步骤:

[0048] 获取样本数据集合,所述样本数据集合包括至少两个样本对象的样本兴趣度以及用于区分所述至少两个样本对象的区分参数;

[0049] 采用推荐模型提取所述至少两个样本对象的预测属性特征向量;

[0050] 根据所述预测属性特征向量预测识别得到所述至少两个样本对象的预测兴趣度;

[0051] 根据所述预测属性特征向量、所述区分参数、所述预测兴趣度以及所述样本兴趣度对所述推荐模型进行调整,将调整后的推荐模型作为目标推荐模型,所述目标推荐模型所输出针对所述至少两个样本对象的更新预测属性特征向量之间的区分程度,与所述区分参数对应的区分程度相匹配。

[0052] 可选的,上述程序指令当被处理器执行时,以执行如下步骤:

[0053] 获取待推荐的目标对象的属性信息,以及访问所述目标对象的目标访问者的画像数据;

[0054] 采用目标推荐模型对所述目标对象的属性信息进行属性特征提取,得到所述目标对象的属性特征向量,所述目标推荐模型为采用如权利要求1-10任一项所述数据处理方法对目标推荐模型调整得到的;

[0055] 采用所述目标推荐模型对所述目标访问者的画像数据进行画像特征提取,得到所述目标访问者的画像特征向量;

[0056] 采用所述目标推荐模型对所述属性特征向量和所述画像特征向量进行兴趣度识别,得到所述目标访问者对所述目标对象的兴趣度;

[0057] 根据所述目标访问者对所述目标对象的兴趣度向所述目标访问者推荐所述目标对象。

[0058] 本申请中, 计算机设备可以采用推荐模型提取该至少两个样本对象的预测属性特征向量, 根据该预测属性特征向量识别得到该至少两个样本对象的预测兴趣度, 根据该预测属性特征向量、该区分参数、该预测兴趣度以及样本兴趣度对该推荐模型进行调整, 将调整后的推荐模型作为目标推荐模型。即通过该预测兴趣度以及样本兴趣度对推荐模型进行调整, 可以提高推荐模型的兴趣度识别的准确性, 通过根据该预测属性特征向量、该区分参数对推荐模型进行调整, 可提高推荐模型对对象的区分度, 即提高调整后的推荐模型的对象区分能力。也就是说, 通过根据该预测属性特征向量、该区分参数、该预测兴趣度以及样本兴趣度对该推荐模型进行调整, 可提高推荐模型兴趣度识别的准确性, 可提高调整后的推荐模型对对象的区分度。通过提高推荐模型的兴趣度识别的准确度和对象区分能力, 可提高对象推荐准确度。计算设备获取到目标推荐模型后, 可采用目标推荐模型为用户推荐对象, 可缓解对象之间的集中度, 即提高对象之间的对象区分度, 实现多样化的推荐, 可满足用户个性化的需求。

附图说明

[0059] 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的前提下, 还可以根据这些附图获得其他的附图。

[0060] 图1a是本申请实施例提供一种数据处理系统的架构图;

[0061] 图1b是本申请实施例提供一种数据处理系统的交互示意图;

[0062] 图2是本申请实施例提供一种数据处理方法的应用场景图;

[0063] 图3是本申请实施例提供一种数据处理方法的流程示意图;

[0064] 图4是本申请实施例提供一种推荐模型的示意图;

[0065] 图5是本申请实施例提供一种关于教育程度的画像特征向量的示意图;

[0066] 图6是本申请实施例提供一种关于性别的画像特征向量的示意图;

[0067] 图7是本申请实施例提供一种推荐模型输出的预测属性特征向量的示意图;

[0068] 图8是本申请实施例提供一种调整后的推荐模型输出的更新预测属性特征向量的示意图;

[0069] 图9是本申请实施例提供一种数据处理方法的流程示意图;

[0070] 图10是本申请实施例提供一种数据处理方法的应用场景图;

[0071] 图11是本申请实施例提供一种数据处理装置的结构示意图;

[0072] 图12是本申请实施例提供一种数据处理装置的结构示意图;

[0073] 图13是本申请实施例提供一种计算机设备的结构示意图。

具体实施方式

[0074] 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, 都属于本申请保护的范围。

[0075] 云技术(Cloud technology)基于云计算商业模式应用的网络技术、信息技术、整合技术、管理平台技术、应用技术等的总称,可以组成资源池,按需所用,灵活便利。云计算技术将变成重要支撑。技术网络系统的后台服务需要大量的计算、存储资源,如视频网站、图片类网站和更多的门户网站。伴随着互联网行业的高度发展和应用,将来每个物品都有可能存在自己的识别标志,都需要传输到后台系统进行逻辑处理,不同程度级别的数据将会分开处理,各类行业数据皆需要强大的系统后盾支撑,只能通过云计算来实现。云技术被广泛应用于医疗云、云物联、云安全、云呼叫、云社交、人工智能云服务等领域。

[0076] 所谓人工智能云服务,一般也被称作是AIaaS(AI as a Service,中文为“AI即服务”)。这是目前主流的一种人工智能平台的服务方式,具体来说AIaaS平台会把几类常见的AI服务进行拆分,并在云端提供独立或者打包的服务。这种服务模式类似于开了一个AI主题商城:所有的开发者都可以通过API接口的方式来接入使用平台提供的一种或者是多种人工智能服务,部分资深的开发者还可以使用平台提供的AI框架和AI基础设施来部署和运维自己专属的云人工智能服务。

[0077] 所谓人工智能(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。

[0078] 人工智能技术涉及的技术领域广泛,既有硬件层面的技术也有软件层面的技术。人工智能硬件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。其中,机器学习(Machine Learning, ML)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,例如,对象推荐、智能分类、交通预测等领域。对象推荐是指根据对象的属性信息以及用户信息等来为用户推荐对象;此处对象可以是指物品、新闻、文章、应用程序等等,物品可以是指被发布于互联网上,且能够被用户购买的商品,包括书籍、衣物、器材等。通常采用推荐模型根据用户购买行为来预测用户想要购买或查看的对象,根据预测结果为用户推荐对象。推荐模型是指用于预测用户对对象的兴趣度,根据兴趣度为用户推荐对象的模型,兴趣度可用于表示用户对对象的喜好程度,也可以用于表示用户购买或查看对象的概率。即用户对某一对象的兴趣度越高,用户购买或查看该对象的概率越高;兴趣度越低,用户购买或查看该对象的概率越低。实践中发现,推荐模型的预测准确度会影响用户对推荐对象的购买和查看情况。例如,如果推荐模型对对象的区分度较低,这时推荐模型会为不同用户推荐相同的对象。但是,不同用户所感兴趣的商品不一致,因此,会导致用户难以找到想要查看或购买的对象,进而影响用户对推荐对象的购买和查看。基于此,本申请提供一种数据处理方法,通过提高推荐模型对对象的区分度,以提高推荐模型的预测准确度。首先介绍本申请实施例提供的数据处理系统,图1a示出了本申请一个示例性实施例提供的一种数据处理系统的结构示意图

图,数据处理系统包括服务器10及至少一个终端,图1a中以两个终端为例,分别为终端11和终端12。

[0079] 其中,终端11和终端12均是指面向用户的终端,具体可以是智能手机、平板电脑、便携式个人计算机、台式计算机、智能音箱、智能手表、手环及智能电视等智能设备。服务器10可以独立的物理服务器,也可以是多个物理服务器构成的服务器集群或者分布式系统,还可以是提供云服务、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。终端11和终端12以及服务器10可以通过有线或无线通信方式进行直接或间接地连接,本申请在此不做限制。

[0080] 请参见图1a,终端11和终端12均可以用于获取样本数据集,将样本数据集发送至服务器10,由服务器10对推荐模型进行调整,得到目标推荐模型。此处推荐模型是指对对象的区分度较低的推荐模型,即推荐模型的对象区分能力较弱;目标推荐模型是指对对象的区分度较高的推荐模型,即目标推荐模型区分对象的能力较强。以终端11为例,请一并参见图1b,是本申请实施例提供的一种数据处理系统的交互示意图,如图1b所示,终端11可以从网络中获取样本数据集,或者,从本地存储中获取样本数据集,将该样本数据集发送至服务器10。该样本数据集包括至少两个样本对象、至少两个样本对象的样本兴趣度、用于区分至少两个样本对象的区分参数。样本对象的样本兴趣度用于指示对应访问者购买或查看该样本对象的概率,样本对象的样本兴趣度可以是指人工标注得到的,或者,样本对象的样本兴趣度可以是指通过识别模型识别得到的,识别模型可以是指获取样本对象的样本兴趣度的准确度大于准确度阈值的模型。样本对象的区分参数可以是指用于区分样本对象参数,即样本对象的区分参数用于提高该样本对象与至少两个样本对象中除该样本对象以外的样本对象之间的区分度的参数,具体可以是一个区分数值或一个区分向量。服务器10接收到样本数据集后,采用推荐模型提取该至少两样本对象的预测属性特征向量,根据该预测属性特征向量预测识别得到该至少两个样本对象的预测兴趣度。服务器10可根据该预测属性特征向量、区分参数、预测兴趣度以及样本兴趣度对推荐模型进行调整,将调整后的推荐模型作为目标推荐模型。目标推荐模型所输出针对至少两个样本对象的更新预测属性特征向量之间的区分程度,与区分参数对应的区分程度相匹配,即该目标推荐模型所输出样本对象的更新预测属性特征向量更加趋近与对应样本对象的区分参数,也即目标推荐模型对对象的区分度比较高,对象区分能力较强。

[0081] 请参见图1b,服务器10获取到目标推荐模型后,可以将该目标推荐模型发送至终端11,由终端11采用目标推荐模型来推荐对象。或者,服务器10可以采用目标推荐模型来推荐对象,将推荐结果发送至终端11。下面以服务器10采用目标推荐模型来推荐对象,将推荐结果发送至终端11为例进行说明。请一并参见图2,是本申请实施例提供的一种数据处理的应用场景示意图。如图2所示,当服务器10检测到目标访问者访问目标网页时,可以从终端11获取对象数据集合。目标网页可以是用于发布待推荐对象信息以及购买待推荐对象的网页,具体可以是浏览器上的购物网页、购物应用程序的页面、公众号的页面以及社交空间的页面等等,此处目标访问者可以是指操作终端11的用户。该对象数据集合可包括多个待推荐对象的属性信息、目标访问者的画像数据,画像数据包括用户信息和目标访问者历史访问该多个待推荐对象的访问信息。访问信息包括访问次数、应用程序的标识、以及访问时

间,此处应用程序可以是指目标网页对应的应用程序,可以是购物应用程序、浏览器应用程序以及社交应用程序等等,应用程序的标识可以是指版本号、名称等等。访问可以是指浏览、购买、收藏、加入购物车、关注等操作;用户信息可包括目标访问者的标识、性别、教育程度、年龄以及位置等信息。服务器10可以采用目标推荐模型对多个待推荐对象的属性信息进行属性特征提取,得到多个待推荐对象的属性特征向量;采用目标推荐模型对目标访问者的画像数据进行画像特征提取,得到画像特征向量。进一步,服务器10可采用目标推荐模型对多个待推荐对象的属性特征向量以及画像特征向量进行识别,得到目标访问者对每个待推荐对象的兴趣度,可根据目标访问者对每个待推荐对象的兴趣度为目标访问者推荐该多个待推荐对象。具体的,服务器10可根据目标访问者对每个待推荐对象的兴趣度对各个待推荐对象进行排序,则可将排序后的多个待推荐对象推送给终端11,如将排序靠前的N个待推荐对象推送给终端11,N为大于等于2的正整数。如图2中,多个待推荐对象包括待推荐对象1、待推荐对象2、待推荐对象3……,获取到各个待推荐对象的兴趣度后,可以按照兴趣度由高到低的顺序排列该多个待推荐对象,排序后的多个待推荐对象可以为:待推荐对象1、待推荐对象5、待推荐对象2、待推荐对象3……,即目标访问者对待推荐对象1的兴趣度最高,对待推荐对象2的兴趣度次之。服务器可以从该多个待推荐对象中选择出4个待推荐对象,即选择排序最靠前的4个待推荐对象,作为推荐对象,将推荐对象发送至终端11。终端11接收到推荐对象后,可以按照各个推荐对象的兴趣度由高到低的顺序展示各个推荐对象。这样目标访问者能够快速浏览到感兴趣的对象,即能够快速浏览到想要购买或查看的对象,提高目标访问者查找对象的效率,以及提高目标访问者购买或查看对象的概率。同时,目标推荐模型具有对象区分能力,能够为不同访问者推荐不一样的对象,实现多样化推荐,满足访问者的个性化需求。

[0082] 需要说明的是,上述对推荐模型的调整步骤也可以由终端11或终端12来执行,上述采用目标推荐模型为目标访问者推荐对象的步骤也可以由终端11或终端12来执行,本申请实施例对此不做限定。

[0083] 基于上述的描述,请参见图3,是本申请实施例提供的一种数据处理方法的流程示意图。该方法可由计算机设备来执行,该计算机设备可以是指图1a中的服务器10、终端11或终端12,如图3所示,上述方法可以包括:

[0084] S101、获取样本数据集合,该样本数据集合包括至少两个样本对象的样本兴趣度以及用于区分该至少两个样本对象的区分参数。

[0085] 为了提高推荐模型对对象的区分度,计算机设备可以从本地存储或第三方获取样本数据集合,该样本数据集合包括至少两个样本对象的样本兴趣度以及用于区分该至少两个样本对象的区分参数。其中,区分参数是用于区分样本对象的参数,该区分参数可以根据样本对象的对象类型随机生成的,或者,是根据样本对象的语义信息生成的;该区分参数可以是一个区分数值或区分向量。样本数据集合还可以包括每个样本对象的属性信息、访问每个样本对象的访问者的画像数据。此处对样本数据集合的解释说明可以参见图2中对样本数据集合的解释说明,重复之处不再赘述。

[0086] S102、采用推荐模型提取该至少两个样本对象的预测属性特征向量。

[0087] S103、根据该预测属性特征向量预测识别得到该至少两个样本对象的预测兴趣度。

[0088] 步骤S102和步骤S103中,计算机设备可以采用该推荐模型对至少两个样本对象的属性信息进行属性特征提取,得到至少两个样本对象的预测属性特征向量。该预测属性特征向量用于描述样本对象的特征信息,即该预测属性特征向量可用于标识样本对象,也即该预测属性特征向量可用于区分样本对象。计算机设备获取到预测属性特征向量后,可以采用推荐模型根据该预测属性特征向量预测识别得到该至少两个样本对象的预测兴趣度。该至少两个样本对象的预测兴趣度可以用于确定推荐模型的兴趣度识别的准确度,如果样本对象的预测兴趣度与样本兴趣度相似或相同,表明该推荐模型的兴趣度识别的准确度比较高;如果样本对象的预测兴趣度与样本兴趣度不相似,表明该推荐模型的兴趣度识别的准确度比较低。

[0089] 其中,推荐模型具体可以是指基于卷积神经网络(Convolutional Neural Networks, CNN)所构建的推荐模型,例如:TextCNN(文本卷积神经网络)、VGGNet网络(Visual Geometry Group Network,一种卷积神经网络)、ResNet网络(残差网络)以及AlexNet网络(一种卷积神经网络),等等;也可以是基于FCN(Fully Convolutional Networks,全神经网络)所构建的推荐模型,对此不作限定。

[0090] 可选的,该样本数据集合包括访问该至少两个样本对象的访问者的画像数据,步骤S103可包括如下步骤s11~s12。

[0091] s11、采用该推荐模型对该画像数据进行画像特征提取,得到预测画像特征向量。

[0092] s12、采用该推荐模型对该预测画像特征向量和该预测属性特征向量进行预测识别,得到该至少两个样本对象的预测兴趣度。

[0093] 在步骤s11~s12中,为了使预测兴趣度拟合访问者对样本对象的行为习惯、喜欢,计算机设备可以采用该推荐模型对该画像数据进行画像特征提取,得到预测画像特征向量,该画像特征向量用于描述访问者访问样本对象的行为习惯、喜好。进一步,采用该推荐模型对该预测画像特征向量和该预测属性特征向量进行预测识别,得到该至少两个样本对象的预测兴趣度。

[0094] 例如,该推荐模型的层级结构如图4所示,在图4中,该推荐模型包括对象侧、用户侧以、输出层以及拼接层,对象侧可用于获取样本对象的预测属性特征向量,用户侧可以用于获取访问样本对象的访问者的画像特征向量。对象侧可包括属性识别层和属性隐藏层,属性识别层对样本对象的属性信息中的每一项进行识别,得到每一项对应的属性向量。如,采用属性识别层对样本对象的对象类型进行识别,得到类型向量,采用属性识别层对样本对象的名称进行识别,得到名称向量,采用属性识别层对样本对象的图像进行识别,得到图像向量。进一步,采用属性隐藏层对样本对象的属性信息中各项对应的属性向量进行拼接处理,得到样本对象的预测属性特征向量。同理,用户侧可包括画像识别层和画像隐藏层,画像识别层用于对访问样本对象的访问者的画像数据中的每一项进行识别,得到每一项对应的画像向量。如,可采用画像识别层对访问信息进行识别,得到访问对象向量,采用画像识别层对访问者的位置进行识别,得到位置向量;采用画像识别层对访问者的性别进行识别,得到性别向量,采用画像识别层对应用程序进行识别,得到应用程序向量。进一步,采用画像隐藏层对画像数据中各项数据对应的画像向量进行拼接处理,得到预测画像特征向量。该推荐模型的拼接层用于对预测画像特征向量和预测属性特征向量进行拼接处理,得到目标特征向量,对该目标特征向量进行识别,得到样本对象的预测兴趣度,由输出层输出

样本对象的预测兴趣度。

[0095] 可选的,步骤s12可包括如下步骤s21~s23。

[0096] s21、获取该至少两个样本对象中每个样本对象的对象类型。

[0097] s22、从该画像数据中筛选出与该对象类型匹配的画像数据,作为有效画像数据。

[0098] s23、采用该推荐模型对该有效画像数据进行画像特征提取,得到该预测画像特征向量。

[0099] 在步骤s21~s23中,对象类型可以用于指示样本对象的类别,按照对象属性划分,对象类型可包括书籍、衣物、食品以及生活用品等;按照对象的适用人群划分,对象类型可包括女士用品、男士用品等。针对不同对象类型的样本对象,画像数据中的各项数据对样本对象的预测兴趣度的影响程度不一致。例如,该样本对象的对象类型为衣物,则画像数据中的性别对该样本对象的预测兴趣度的影响程度较高,画像数据中的教育程度对该样本对象的预测兴趣度的影响程度较低;此时可称为教育程度为无效画像数据,性别为有效画像数据。同时,如果对于不同对象类型的样本对象,均采用相同的画像数据进行兴趣度识别,则会导致样本对象的对象区分度较低。例如,假设样本数据集中包括三种对象类型的样本对象,分别为对象类型1的样本对象、对象类型2的样本对象以及对象类型3的样本对象;画像数据中可包括访问样本对象的访问者的教育程度以及性别。计算机设备可以对访问者的教育程度进行画像特征提取,得到教育程度对应的画像特征向量,对教育程度对应的画像特征向量进行降维处理,并将降维处理后的教育程度对应的画像特征向量进行可视化。可视化后的画像特征向量如图5所示,图5中三角形表示对象类型1的样本对象的对应的画像特征向量(即访问对象类型1的样本对象的访问者的教育程度对应的画像特征向量),五角星表示对象类型2的样本对象对应的画像特征向量,圆形表示对象类型3的样本对象对应的画像特征向量。由图5中可知,不同对象类型的样本对象对应的画像特征向量之间的距离比较近,即不同对象类型的样本对象对应的画像特征向量比较集中。因此,不同对象类型的样本对象对应的画像特征向量之间的区分程度较低,进而导致样本对象的区分度较低。也就是说,画像数据中的教育程度不能较好的区分对象类型1、对象类型2以及对象类型3的样本对象,即教育程度对对象类型1、对象类型2以及对象类型3的样本对象的预测兴趣度的影响程度较低,也即教育程度对于对象类型1、对象类型2以及对象类型3的样本对象是无效画像数据。

[0100] 同理,计算机设备可以对访问者的性别进行画像特征提取,得到性别对应的画像特征向量,对性别对应的画像特征向量进行降维处理,并将降维处理后的性别对应的画像特征向量进行可视化。可视化后的各个样本对象对应的画像特征向量如图6所示,图6中三角形表示对象类型1的样本对象的对应的画像特征向量(即访问对象类型1的样本对象的访问者的性别对应的画像特征向量),五角星表示对象类型2的样本对象对应的画像特征向量,圆形表示对象类型3的样本对象对应的画像特征向量。由图6中可知,不同对象类型的样本对象对应的画像特征向量之间的距离比较远,即不同对象类型的样本对象对应的画像特征向量比较分散。因此,不同对象类型的样本对象对应的画像特征向量之间的区分程度较高,使样本对象的区分度较高。也就是说,画像数据中的性别能够较好的区分对象类型1、对象类型2以及对象类型3的样本对象,即性别对对象类型1、对象类型2以及对象类型3的样本对象的预测兴趣度的影响程度较高,也即性别对于对象类型1、对象类型2以及对象类型3的

样本对象是有效画像数据。

[0101] 对比图5和图6可知,有效画像数据能够提高样本对象的区分度,无效画像数据会降低样本对象的区分度。因此,计算设备可以获取该至少两个样本对象中每个样本对象的对象类型,从画像数据中筛选出与该对象类型匹配的画像数据,作为有效画像数据。匹配的画像数据是指对该样本对象的预测兴趣度的影响程度较高的画像数据,即能够区分样本对象的画像数据。进一步,可采用该推荐模型对该有效画像数据进行画像特征提取,得到该预测画像特征向量。即通过对画像数据进行筛选,可提高对象的区分度,即针对不同的用户能够推荐不一样的对象,满足用户的个性化需求,并提高对象推荐的多样性。

[0102] S104、根据该预测属性特征向量、该区分参数、该预测兴趣度以及该样本兴趣度对该推荐模型进行调整,将调整后的推荐模型作为目标推荐模型,该目标推荐模型所输出针对该至少两个样本对象的更新预测属性特征向量之间的区分程度,与该区分参数对应的区分程度相匹配。

[0103] 通过该预测兴趣度以及样本兴趣度对推荐模型进行调整可以确保推荐模型的兴趣度识别的准确性;根据该预测属性特征向量、该区分参数对推荐模型进行调整,可提高推荐模型对对象的区分度,即提高调整后的推荐模型的对象区分能力。因此,根据该预测属性特征向量、该区分参数、该预测兴趣度以及该样本兴趣度对该推荐模型进行调整,将调整后的推荐模型作为目标推荐模型,可确保推荐模型的兴趣度识别的准确度,以及提高推荐模型的对象区分能力。

[0104] 在对推荐模型调整后,可以验证调整后的推荐模型的性能。可选的,计算机设备可以根据调整后的推荐模型输出的更新预测属性特征向量来验证调整后的推荐模型的性能。具体的,可采用调整后的推荐模型提取至少两个样本对象的更新预测属性特征向量,如果至少两个样本对象的更新预测属性特征向量之间的区分程度,与该区分参数之间的区分程度不匹配,表明调整后的推荐模型输出的更新预测属性特征向量不能够区分样本对象,即表明调整后的推荐模型的对象区分能力较弱,则可以对调整后的推荐模型进行调整。如果至少两个样本对象的更新预测属性特征向量之间的区分程度,与该区分参数之间的区分程度匹配,表明调整后的推荐模型输出的更新预测属性特征向量能够区分样本对象,即表明调整后的推荐模型的对象区分能力较强,则可以暂停对推荐模型的调整,将调整后的推荐模型作为目标推荐模型。可选的,计算机设备可以根据调整后的推荐模型的总损失值来验证推荐模型的性能,调整后的推荐模型的总损失值用于指示调整后的推荐模型的兴趣度识别的准确度以及对象区分准确度。具体的,如果调整后的推荐模型的总损失值处于收敛状态,此处收敛状态是指该调整后的推荐模型的兴趣度识别准确度以及对象区分度均达到最佳,即如果调整后的推荐模型的总损失值处于收敛状态,仍对调整后的推荐模型进行调整,会使调整后的推荐模型的兴趣度识别的准确度以及对象区分准确度均变低。因此,可以暂停对推荐模型的调整,将调整后的推荐模型作为目标推荐模型。如果调整后的推荐模型的总损失值未处于收敛状态,表明该调整后的推荐模型的兴趣度识别准确度以及对象区分度未达到最佳,则可以对调整后的推荐模型进行调整。可选的,计算机设备可以根据调整后的推荐模型输出的更新预测属性特征向量,以及调整后的推荐模型的总损失值来验证调整后的推荐模型的性能。

[0105] 其中,该至少两个样本对象的更新预测属性特征向量之间的区分程度,与该区分

参数对应的区分程度相匹配可以是指：该至少两个样本对象的更新预测属性特征向量之间的区分程度，与该区分参数对应的区分程度之间的差值小于程度阈值，或者，该至少两个样本对象的更新预测属性特征向量之间的区分程度，与该区分参数对应的区分程度相匹配之间的比值小于预设数值。

[0106] 可选的，该至少两个样本对象的更新预测属性特征向量之间的区分程度可以是根据至少两个样本对象的更新预测属性特征向量之间的角度确定的；即若至少两个样本对象的更新预测属性特征向量之间的角度越大，则至少两个样本对象的更新预测属性特征向量之间区分程度越大；若至少两个样本对象的更新预测属性特征向量之间的角度越小，则至少两个样本对象的更新预测属性特征向量之间区分程度越小。可选的，该至少两个样本对象的更新预测属性特征向量之间的区分程度是根据至少两个样本对象的更新预测属性特征向量之间的距离确定的；即若至少两个样本对象的更新预测属性特征向量之间的距离越大，则至少两个样本对象的更新预测属性特征向量之间区分程度越大；若至少两个样本对象的更新预测属性特征向量之间的距离越小，则至少两个样本对象的更新预测属性特征向量之间区分程度越小。至少两个样本对象的更新预测属性特征向量之间的距离可以包括但不限于：欧式距离、切比雪夫距离以及曼哈顿距离等。

[0107] 可选的，当该区分参数区分数值时，该区分参数之间的区分程序可以是指区分数值之间的差值；或者，当该区分参数为区分向量时，该区分参数之间的区分程序可以是根据区分向量之间的角度或距离确定的。

[0108] 可选的，计算设备在调整推荐模型的过程中，可以将推荐模型输出的预测属性特征向量，以及推荐模型输出的更新预测属性特征向量进行可视化，以使用户确定调整后的推荐模型的对象区分性能是否满足需求。具体的，假设样本数据集中包括三种对象类型的样本对象，分别为对象类型1的样本对象、对象类型2的样本对象以及对象类型3的样本对象。如图7所示，计算设备可以将推荐模型输出的预测属性特征向量进行降维处理，对降维处理后的预测属性特征向量进行可视化，并在计算机设备的界面15中输出各个降维处理后的预测属性特征向量。在界面15中的五边形表示对象类型1的样本对象的降维后的预测属性特征向量，八边形表示对象类型2的样本对象的降维处理后的预测属性特征向量，八角星表示对象类型3的样本对象的降维处理后的预测属性特征向量。根据界面15可知，不同对象类型的样本对象的降维处理后的预测属性特征向量之间的距离比较接近，即不同对象类型的样本对象的降维处理后的预测属性特征向量比较集中。因此，不同对象类型的样本对象对应的预测属性特征向量之间的区分程度较低，则确定推荐模型的对象区分能力比较弱。

[0109] 同理，如图8所示，计算设备可以将调整后的推荐模型输出的更新预测属性特征向量进行降维处理，对降维处理后的更新预测属性特征向量进行可视化，并在计算机设备的界面16中输出的各个降维处理后的更新预测属性特征向量。在界面16中的五边形表示对象类型1的样本对象的降维后的更新预测属性特征向量，八边形表示对象类型2的样本对象的降维处理后的更新预测属性特征向量，八角星表示对象类型3的样本对象的降维处理后的更新预测属性特征向量。根据界面16可知，相同对象类型的样本对象的降维处理后的更新预测属性特征向量之间的距离比较近，不同对象类型的样本对象的降维处理后的更新预测属性特征向量之间的距离比较远，即不同对象类型的样本对象的降维处理后的预测属性特征向量比较分散。因此，不同对象类型的样本对象对应的更新预测属性特征向量之间的区

分程度较高,则确定调整后的推荐模型的对象区分能力比较强。

[0110] 本申请中,计算机设备可以采用推荐模型提取该至少两个样本对象的预测属性特征向量,根据该预测属性特征向量识别得到该至少两个样本对象的预测兴趣度,根据该预测属性特征向量、该区分参数、该预测兴趣度以及样本兴趣度对该推荐模型进行调整,将调整后的推荐模型作为目标推荐模型。即通过该预测兴趣度以及样本兴趣度对推荐模型进行调整,可以提高推荐模型的兴趣度识别的准确性,通过根据该预测属性特征向量、该区分参数对推荐模型进行调整,可提高推荐模型对对象的区分度,即提高调整后的推荐模型的对象区分能力。也就是说,通过根据该预测属性特征向量、该区分参数、该预测兴趣度以及样本兴趣度对该推荐模型进行调整,可提高推荐模型兴趣度识别的准确性,可提高调整后的推荐模型对对象的区分度。通过提高推荐模型的兴趣度识别的准确度和对象区分能力,可提高对象推荐准确度。计算设备获取到目标推荐模型后,可采用目标推荐模型为用户推荐对象,可缓解对象之间的集中度,即提高对象之间的对象区分度,实现多样化的推荐,可满足用户个性化的需求。

[0111] 可选的,区分参数可以为区分向量,步骤S101可包括如下步骤s31~s32。

[0112] s31、获取第一样本对象的语义信息,该第一样本对象为该至少两个样本对象的任一样本对象。

[0113] s32、从数据库中筛选出与该第一样本对象的语义信息匹配的目标待选区分向量,作为该第一样本对象的区分向量,该数据库中包括多个待选区分向量,一个该待选区分向量与一种语义信息对应。

[0114] 在步骤s31~s32中,样本对象的语义信息是指用于标识样本对象的对象类型的信息,不同对象类型的样本对象对应的语义信息不同,因此,可根据语义信息获取区分向量。具体的,计算机设备可以获取第一样本对象的语义信息,从该数据库中筛选出与该第一样本对象的语义信息匹配的目标待选区分向量,作为第一样本对象的区分向量。该数据库中可包括多个待选区分向量,一个待选区分向量与一种语义信息对应,即每个待选区分向量可以是根据样本对象的语义信息生成的。

[0115] 可选的,步骤s32可包括如下步骤s41~s44。

[0116] s41、从数据库中筛选出与该第一样本对象的语义信息匹配的目标待选区分向量。

[0117] s42、对该目标待选区分向量以及剩余待选区分向量进行降维处理;该剩余待选区分向量为该数据库中除该目标待选区分向量以外的待选区分向量。

[0118] s43、获取降维处理后的目标待选区分向量与降维处理后的剩余待选区分向量之间的距离。

[0119] s44、若该距离大于距离阈值,则将该目标待选区分向量作为该第一样本对象的区分向量。

[0120] 在步骤s41~s44中,计算机设备可以对目标待选区分向量的有效性进行验证,以提高区分向量的准确度和有效性。具体的,计算设备可以从数据库中筛选出与该第一样本对象的语义信息匹配的目标待选区分向量,该目标待选区分向量可以是指多维的向量。为了降低计算量,以及为了便于可视化区分向量,可对该目标待选区分向量以及剩余待选区分向量进行降维处理,计算降维处理后的目标待选区分向量与降维处理后的其他待选区分向量之间的距离;此处距离可以包括但不限于:欧式距离、切比雪夫距离以及曼哈顿距离

等。如果距离大于距离阈值,表明目标待选区分向量与降维处理后的剩余待选区分向量之间的区分程度较高,即目标待选区分向量能够用于区分第一样本对象,因此可将该目标待选区分向量作为第一样本对象的区分向量。如果距离小于或等于距离阈值,表明目标待选区分向量与降维处理后的剩余待选区分向量之间的区分程度较低,即目标待选区分向量不能够用于区分第一样本对象,因此可对该目标待选区分向量进行调整,或,采用其他方式获取第一样本对象的区分向量。

[0121] 可选的,步骤S104可包括如下步骤s51~s55。

[0122] s51、采用该预测兴趣度以及该样本兴趣度生成该推荐模型的原始损失值。

[0123] s52、根据该预测属性特征向量以及该区分参数确定对象区分度。

[0124] s53、采用该对象区分度生成该推荐模型的区分损失值。

[0125] s54、采用该区分损失值和该原始损失值生成该推荐模型的总损失值。

[0126] s55、若该推荐模型的总损失值不满足收敛条件,则根据该推荐模型的总损失值对该推荐模型进行调整,将调整后的推荐模型作为目标推荐模型。

[0127] 在步骤s51~s55中,计算机设备可以根据推荐模型的总损失值对推荐模型进行调整,可确保推荐模型的兴趣度识别的准确度,并提高推荐模型的对象区分能力。具体的,计算机设备可以采用该预测兴趣度以及该样本兴趣度生成该推荐模型的原始损失值,即该原始损失值用于指示推荐模型的兴趣度识别的准确度,若该原始损失值越大,则推荐模型的兴趣度识别的准确度越低;若该原始损失值越小,则该推荐模型的兴趣度识别的准确度越高。计算机设备获取到推荐模型的原始损失值后,可以根据该预测属性特征向量以及该区分参数确定对象区分度,采用该对象区分度生成该推荐模型的区分损失值。此处区分损失值用于衡量推荐模型的对象区分能力,若区分损失值越大,则推荐模型的对象区分能力越弱,若区分损失值越小,则推荐模型的对象区分能力越强。

[0128] 计算机设备获取到区分损失值和原始损失值后,可采用该区分损失值和该原始损失值生成该推荐模型的总损失值;若该推荐模型的总损失值满足收敛条件,则表明推荐模型的总损失值比较小,即推荐模型的对象区分能力比较强,且推荐模型的兴趣度识别的准确度较高,可以将该推荐模型作为目标推荐模型。若该推荐模型的总损失值不满足收敛条件,则表明推荐模型的总损失值比较大,即推荐模型的对象区分能力比较弱,且推荐模型的兴趣度识别的准确度较低,可根据该推荐模型的总损失值对该推荐模型进行调整,将调整后的推荐模型作为目标推荐模型。

[0129] 例如,当该区分参数为区分数值时,该推荐模型的总损失值可以采用如下公式(1)表示:

$$[0130] \quad Loss_z = Loss_1 + \alpha \sum_{i=1}^M |F_i - C_i| \quad (1)$$

[0131] 其中,公式(1)中的 $Loss_z$ 表示推荐模型的总损失值, $Loss_1$ 表示推荐模型的原始损失值, $\alpha \sum_{i=1}^M |F_i - C_i|$ 表示推荐模型的区分损失度。 M 表示样数据集集合中的样本对象的数量, F_i 表示第*i*个样本对象的预测属性特征向量的模长, C_i 表示第*i*个样本对象的区分数值, α 表示权重参数,该权重参数可以为经验值,如可以为0.2。

[0132] 再例如,当该区分参数为区分向量时,该推荐模型的总损失值可以采用如下公式

(2) 表示:

$$[0133] \quad Loss_z = Loss_1 - \beta \sum_{i=1}^M P_i * Q_i \quad (2)$$

[0134] 其中,公式(2)中的 $\beta \sum_{i=1}^M P_i * Q_i$ 表示推荐模型的区分损失值, P_i 表示第 i 个样本对象的区分向量, Q_i 表示第 i 个样本对象的预测属性特征向量, β 表示权重参数, 该权重参数可以为经验值, 如 0.3。

[0135] 在此实施例中, 该区分参数为区分数值, 步骤 s52 可包括如下步骤 s61~s63。

[0136] s61、获取该预测属性特征向量的模长。

[0137] s62、获取该预测属性特征向量的模长与该区分数值之间的差值。

[0138] s63、根据该差值确定该对象区分度。

[0139] 在步骤 s61~s63 中, 计算设备可以获取该预测属性特征向量的模长, 计算该预测属性特征向量的模长与该区分数值之间的差值, 根据该差值确定该对象区分度。该差值与对象区分度具有负相关关系, 即该差值越小, 表明预测属性特征向量的模长更加接近区分数值, 即对象区分度较大; 反之, 即该差值越大, 表明预测属性特征向量的模长与区分数值的差距较大, 即对象区分度较小。

[0140] 在此实施例中, 该区分参数为区分向量, 步骤 s52 可包括如下步骤 s71~s72。

[0141] s71、对该区分向量与该预测属性特征向量进行乘积运算, 得到向量角度。

[0142] s72、根据该向量角度确定该对象区分度。

[0143] 在步骤 s71~s72 中, 计算设备可以对该区分向量与该预测属性特征向量进行乘积运算, 得到向量角度, 根据该向量角度确定该对象区分度。该向量角度与对象区分度具有负相关关系, 即该向量角度越大, 表明预测属性特征向量与区分向量之间的差距较大, 即对象区分度较小; 反之, 即该向量角度越小, 表明预测属性特征向量更加接近区分向量, 即对象区分度较大。

[0144] 在此实施例中, 步骤 s55 可包括如下步骤 s81~s86。

[0145] s81、据该推荐模型的总损失值对该推荐模型进行调整。

[0146] s82、采用该调整后的推荐模型提取该至少两个样本对象的更新预测属性特征向量。

[0147] s83、根据该更新预测属性特征向量预测识别得到该至少两个样本对象的更新预测兴趣度。

[0148] s84、根据该更新预测属性特征向量、该区分参数、该更新预测兴趣度以及该样本兴趣度, 生成该调整后的推荐模型的总损失值。

[0149] s85、获取该更新预测属性特征向量之间的角度, 根据该更新预测属性特征向量之间的角度确定该更新预测属性特征向量之间的区分程度。

[0150] s86、若该调整后的推荐模型的总损失值满足收敛条件, 且该更新预测属性特征向量之间的区分程度, 与该区分参数对应的区分程度之间的差值小于程度阈值, 则将调整后的推荐模型作为目标推荐模型。

[0151] 在步骤 s81~s86 中, 计算机设备可以降低推荐模型的总损失值为目标, 对该推荐模型进行调整, 采用该调整后的推荐模型提取该至少两个样本对象的更新预测属性特征向

量,根据该更新预测属性特征向量预测识别得到该至少两个样本对象的更新预测兴趣度。进一步,可根据该更新预测属性特征向量、该区分参数确定调整后的推荐模型的区分损失值,可根据该更新预测兴趣度以及该样本兴趣度确定调整后的推荐模型的原始损失值;根据调整后的推荐模型的区分损失值以及调整后的推荐模型的原始损失值生成该调整后的推荐模型的总损失值。即该调整后的推荐模型的总损失值用于指示调整后的推荐模型的兴趣度识别的准确度,和调整后的推荐模型的对象区分能力。若该调整后的推荐模型的总损失值不满足收敛条件,或该更新预测属性特征向量之间的区分程度,与该区分参数对应的区分程度之间的差值大于程度阈值,则表明调整后的推荐模型的兴趣度识别的准确度较低,或,调整后的推荐模型的对象区分能力较弱,因此,可采用调整后的推荐模型的总损失值对调整后推荐模型进行调整。若该调整后的推荐模型的总损失值满足收敛条件,且该更新预测属性特征向量之间的区分程度,与该区分参数对应的区分程度之间的差值小于程度阈值,则表明调整后的推荐模型的兴趣度识别的准确度较高,以及调整后的推荐模型的对象区分能力较强,因此,可将调整后的推荐模型作为目标推荐模型。

[0152] 在此实施例中,该至少两个样本对象包括第二样本对象和第三样本对象,步骤s51可包括如下步骤s91~s93。

[0153] s91、采用该第二样本对象的预测兴趣度以及该第二样本对象的样本兴趣度生成第一原始损失值。

[0154] s92、采用该第三样本对象的预测兴趣度以及该第三样本对象的样本兴趣度生成第二原始损失值。

[0155] s93、对该第一原始损失值与该第二原始损失值进行求和运算,得到该推荐模型的原始损失值。

[0156] 在步骤s91~s93中,计算设备可以每个样本对象的原始损失值计算推荐模型的原始损失值。具体的,可以采用该第二样本对象的预测兴趣度以及该第二样本对象的样本兴趣度生成第一原始损失值,并采用该第三样本对象的预测兴趣度以及该第三样本对象的样本兴趣度生成第二原始损失值。对该第一原始损失值与第二原始损失值进行求和运算,得到该推荐模型的原始损失值。可选的,步骤s53可包括:对该第二样本对象的对象区分度和该第三样本对象的对象区分度进行求和运算,得到该推荐模型的区分损失值。

[0157] 计算机设备可以对至少两个样本对象中各个样本对象的对象区分度进行累加,得到推荐模型的区分损失值,如该至少两个样本对象中包括第二样本对象和第三样本对象,可对该第二样本对象的对象区分度和该第三样本对象的对象区分度进行求和运算,得到该推荐模型的区分损失值。

[0158] 基于上述的描述,请参见图9,是本申请实施例提供的一种数据处理方法的流程示意图。该方法可由计算机设备来执行,该计算机设备可以是指图1a中的服务器10、终端11或终端12,如图9所示,上述方法可以包括:

[0159] S201、获取待推荐的目标对象的属性信息,以及访问该目标对象的目标访问者的画像数据。

[0160] 目标对象可以为目标网页中多个待推荐对象中的任一待推荐对象,目标网页具体可以是浏览器上的购物网页、购物应用程序的页面、公众号的页面以及社交空间页面等等。如果用户想要从目标网页中购买或查看对象时,该用户可以是指目标访问者,该目标访问

者可以对目标网页执行触控操作。当计算机设备检测针对目标网页的触控操作时,可以从本地存储或目标网页的后台服务器中获取目标对象的属性信息,以及目标访问者的画像数据。

[0161] S202、采用目标推荐模型对该目标对象的属性信息进行属性特征提取,得到该目标对象的属性特征向量,该目标推荐模型为采用如图3所示的该数据处理方法对目标推荐模型调整得到的。

[0162] S203、采用该目标推荐模型对该目标访问者的画像数据进行画像特征提取,得到该目标访问者的画像特征向量。

[0163] S204、采用该目标推荐模型对该属性特征向量和该画像特征向量进行兴趣度识别,得到该目标访问者对该目标对象的兴趣度。

[0164] 步骤S202~S204中,由于该目标推荐模型具有对象区分能力,因此该目标推荐模型能够目标访问者推荐多样化的对象。因此,计算机设备可以采用目标推荐模型识别得到目标访问者对该目标对象的兴趣度,具体的,可以采用目标推荐模型对该目标对象的属性信息进行属性特征提取,得到该目标对象的属性特征向量,该属性特性向量可以用于区分目标对象,以及用于描述目标对象的属性特征。进一步,可采用该目标推荐模型对该目标访问者的画像数据进行画像特征提取,得到该目标访问者的画像特征向量,该画像特征向量用于描述目标访问者的行为习惯、喜好等,采用该目标推荐模型对该属性特征向量和该画像特征向量进行拼接处理,得到拼接处理后的特征向量,对该拼接处理后的特征向量进行兴趣度识别,得到目标访问者对该目标对象的兴趣度。由于目标对象的属性特征向量可以用于区分目标对象,即不同对象对应的属性特征向量之间具有较大区分程度,因此,不同对象对应的兴趣度不一致,可实现多样化推荐。

[0165] S205、根据该目标访问者对该目标对象的兴趣度向该目标访问者推荐该目标对象。

[0166] 计算机设备可以根据该目标访问者对该目标对象的兴趣度向该目标访问者推荐目标对象,即判断该目标对象的兴趣度是否大于兴趣度阈值,若该目标对象的兴趣度小于或等于兴趣度阈值,表明目标访问者购买或查看该目标对象的概率比较低,因此,可以不将该目标对象推荐给该目标访问者,或者将该目标对象排所推荐对象中比较靠后的位置;若是表明目标访问者想要购买或查看该目标对象的概率比较高,因此,可以将该目标对象推荐给该目标访问者,将该目标对象排所推荐对象中比较靠前的位置。

[0167] 可选的,计算机设备采用目标推荐模型为目标访问者推荐对象后,可以再次获取目标访问者关于推荐对象的访问数据,根据访问数据来优化目标推荐模型,以提高目标推荐模型的推荐效果。

[0168] 本申请实施例中,计算机设备可以采用目标推荐模型对目标对象的属性信息进行属性特征提取,得到目标对象的属性特征向量,采用目标推荐模型对目标访问者的画像数据进行画像特征提取,得到画像特征向量。由于该目标推荐模型具有对象区分能力,因此,该目标对象的属性特征向量可用于区分目标对象。计算设备可以采用该目标推荐模型对属性特征向量和画像特征向量进行识别,得到目标访问者对该目标对象的兴趣度,根据该目标对象的兴趣度向目标访问者推荐目标对象。通过采用目标推荐模型向目标访问者推荐目标对象,可提高对象的区分度,可为目标访问者提供多样化的推荐,满足目标访问者个性化

的需求。

[0169] 在一个实施例中,步骤S204包括:获取与该目标对象的对象类型相似度大于相似度阈值的待选对象,根据该目标访问者对该目标对象的兴趣度向该目标访问者推荐该目标对象以及该待选对象。

[0170] 目标访问者在查看对象的过程中,通常习惯浏览相似对象类型的对象,例如,目标访问者在查看手机时,同时会浏览手机壳,因此,为了更加精准地为目标访问者推荐对象,计算机设备可以获取与该目标对象的对象类型相似度大于相似度阈值的待选对象,根据该目标访问者对该目标对象的兴趣度向该目标访问者推荐该目标对象以及该待选对象。即在向目标访问者推荐目标对象的同时,还推荐与目标对象的对象类型相似的待选对象,可使目标访问者快速地查找到想要购买或查看的对象。

[0171] 本申请中的数据处理方法可以应用于推荐新闻、推荐商品、推荐应用程序等等场景,以该方法应用于推荐商品的场景为例进行说明。该场景中,待推荐对象可以是指商品,该方法可由服务器10执行,请参见图10,该方法包括:

[0172] 如图10所示,服务器10可以输出推荐界面17,推荐界面17可用于选择推荐方式,推荐界面17包括用于配置商品推荐相关的参数,商品推荐相关的参数包括推荐方式。推荐形式可以用于配置推荐的商品数量,包括单商品以及多商品,单商品是指一次推荐一个商品,多商品是指一次推荐多个商品。推荐的商品用于配置所要推荐的商品的类型,商品的类型包括零食饮料、衣物等。推荐方式包括普通推荐方式以及模型推荐方式;普通推荐方式是指为每个用户推荐的商品均相同的推荐方式,模型推荐方式可以是指采用目标推荐模型为用户推荐商品的方式,模型推荐方式为不同用户推荐的商品不相同。商家或目标网页的管理者可以从推荐界面17选择推荐方式,若服务器10检测到针对模型推荐方式的选择操作,则可以采用模型推荐方式为目标访问者推荐商品。通过在推荐界面17中增加模型推荐方式选项,商家或管理者可便捷地从推荐界面17中选择推荐方式,不需要对访问者所访问的目标网页进行改进,降低了推荐方式选择的难度。另外,商品推荐相关的参数还可包括推荐形式以及推荐的商品(即商品系列)等等。推荐形式可以用于配置推荐的商品数量,包括单商品以及多商品,单商品是指一次推荐一个商品,多商品是指一次推荐多个商品。推荐的商品用于配置所要推荐的商品的类型,商品的类型包括零食饮料、衣物等,可选择一种或多种类型的商品进行推荐。其中,推荐界面17可以是指面向商家或目标网页的管理者的界面,具体可以是指广告投放界面或目标网页对应的后台管理界面,目标网页可以是指面向访问者的界面,目标网页对应的前端呈现页面。当然,推荐界面也可以是指面向访问者的界面,如推荐界面可以目标网页的子页面。

[0173] 如图10所示,服务器10获取到推荐方式后,可以获取商品数据集合,商品数据集合包括待推荐的多个商品的属性信息,以及目标访问者访问各个商品的画像数据,计算设备18可以采用目标推荐模型对每个商品的属性信息进行属性特征提取,得到每个商品的属性特征向量,采用目标推荐模型对目标访问者的画像数据进行画像特征提取,得到画像特征向量。可采用目标推荐模型对属性特征向量以及画像特征向量进行识别,得到目标访问者对每个商品的兴趣度。

[0174] 如图10所示,服务器10获取到目标访问者对各个商品的兴趣度后,可按照兴趣度由高到低的顺序对多个商品进行排序,排序后的多个商品可以为:商品1、商品5、商品2、商

品3……,即目标访问者对商品1的兴趣度最高,对商品2的兴趣度次之。服务器10可以从该多个商品中选择出4个商品,即选择排序最靠前的4个商品,作为推荐商品,将推荐商品发送至终端11。终端11接收到推荐商品后,可按照4个商品对应的兴趣度由高到低地顺序将4个商品展示在目标网页18上。这样目标访问者能够从目标网页18中快速地浏览到感兴趣的物品,即能够快速浏览到想要购买或查看的商品,提高目标访问者查找商品的效率,以及提高目标访问者购买或查看商品的概率。同时,目标推荐模型具有商品区分能力,能够为不同访问者推荐不一样的商品,实现多样化推荐,满足访问者的个性化需求。

[0175] 请参见图11,是本申请实施例提供的一种数据处理装置的结构示意图。该数据处理装置可以是运行于计算机设备中的一个计算机程序(包括程序代码),例如该数据处理装置为一个应用软件;该装置可以用于执行本申请实施例提供的方法中的相应步骤。如图11所示,该数据处理装置1可以包括:

[0176] 获取模块21,用于获取样本数据集合,所述样本数据集合包括至少两个样本对象的样本兴趣度以及用于区分所述至少两个样本对象的区分参数;

[0177] 提取模块22,用于采用推荐模型提取所述至少两个样本对象的预测属性特征向量;

[0178] 识别模块23,用于根据所述预测属性特征向量预测识别得到所述至少两个样本对象的预测兴趣度;

[0179] 调整模块24,用于根据所述预测属性特征向量、所述区分参数、所述预测兴趣度以及所述样本兴趣度对所述推荐模型进行调整,将调整后的推荐模型作为目标推荐模型,所述目标推荐模型所输出针对所述至少两个样本对象的更新预测属性特征向量之间的区分程度,与所述区分参数对应的区分程度相匹配。

[0180] 可选的,调整模块24,具体用于采用所述预测兴趣度以及所述样本兴趣度生成所述推荐模型的原始损失值;根据所述预测属性特征向量以及所述区分参数确定对象区分度;采用所述对象区分度生成所述推荐模型的区分损失值;采用所述区分损失值和所述原始损失值生成所述推荐模型的总损失值;若所述推荐模型的总损失值不满足收敛条件,则根据所述推荐模型的总损失值对所述推荐模型进行调整,将调整后的推荐模型作为目标推荐模型。

[0181] 可选的,调整模块24,具体用于据所述推荐模型的总损失值对所述推荐模型进行调整;采用所述调整后的推荐模型提取所述至少两个样本对象的更新预测属性特征向量;根据所述更新预测属性特征向量预测识别得到所述至少两个样本对象的更新预测兴趣度;根据所述更新预测属性特征向量、所述区分参数、所述更新预测兴趣度以及所述样本兴趣度,生成所述调整后的推荐模型的总损失值;获取所述更新预测属性特征向量之间的角度,根据所述更新预测属性特征向量之间的角度确定所述更新预测属性特征向量之间的区分程度;若所述调整后的推荐模型的总损失值满足收敛条件,且所述更新预测属性特征向量之间的区分程度,与所述区分参数对应的区分程度之间的差值小于程度阈值,则将调整后的推荐模型作为目标推荐模型。

[0182] 可选的,所述区分参数为区分数值,调整模块24,具体用于获取所述预测属性特征向量的模长;获取所述预测属性特征向量的模长与所述区分数值之间的差值;根据所述差值确定所述对象区分度。

[0183] 可选的,所述区分参数为区分向量,可选的,调整模块24,具体用于对所述区分向量与所述预测属性特征向量进行乘积运算,得到向量角度;根据所述向量角度确定所述对象区分度。

[0184] 可选的,获取模块21,具体用于获取第一样本对象的语义信息,所述第一样本对象为所述至少两个样本对象的任一样本对象;从数据库中筛选出与所述第一样本对象的语义信息匹配的目标待选区分向量,作为所述第一样本对象的区分向量,所述数据库中包括多个待选区分向量,一个所述待选区分向量与一种语义信息对应。

[0185] 可选的,获取模块21,具体用于从数据库中筛选出与所述第一样本对象的语义信息匹配的目标待选区分向量;对所述目标待选区分向量以及剩余待选区分向量进行降维处理;所述剩余待选区分向量为所述数据库中除所述目标待选区分向量以外的待选区分向量;获取降维处理后的目标待选区分向量与降维处理后的剩余待选区分向量之间的距离;若所述距离大于距离阈值,则将所述目标待选区分向量作为所述第一样本对象的区分向量。

[0186] 可选的,所述至少两个样本对象包括第二样本对象和第三样本对象,可选的,调整模块24,具体用于采用所述第二样本对象的预测兴趣度以及所述第二样本对象的样本兴趣度生成第一原始损失值;采用所述第三样本对象的预测兴趣度以及所述第三样本对象的样本兴趣度生成第二原始损失值;对所述第一原始损失值与所述第二原始损失值进行求和运算,得到所述推荐模型的原始损失值;对所述第二样本对象的对象区分度和所述第三样本对象的对象区分度进行求和运算,得到所述推荐模型的对象区分损失值。

[0187] 可选的,所述样本数据集包括访问所述至少两个样本对象的访问者的画像数据,识别模块23,具体用于采用所述推荐模型对所述画像数据进行画像特征提取,得到预测画像特征向量;采用所述推荐模型对所述预测画像特征向量和所述预测属性特征向量进行预测识别,得到所述至少两个样本对象的预测兴趣度。

[0188] 可选的,识别模块23,具体用于获取所述至少两个样本对象中每个样本对象的对象类型;从所述画像数据中筛选出与所述对象类型匹配的画像数据,作为有效画像数据;采用所述推荐模型对所述有效画像数据进行画像特征提取,得到所述预测画像特征向量。

[0189] 本申请中,计算机设备可以采用推荐模型提取该至少两个样本对象的预测属性特征向量,根据该预测属性特征向量识别得到该至少两个样本对象的预测兴趣度,根据该预测属性特征向量、该区分参数、该预测兴趣度以及样本兴趣度对该推荐模型进行调整,将调整后的推荐模型作为目标推荐模型。即通过该预测兴趣度以及样本兴趣度对推荐模型进行调整,可以提高推荐模型的兴趣度识别的准确性,通过根据该预测属性特征向量、该区分参数对推荐模型进行调整,可提高推荐模型对对象的区分度,即提高调整后的推荐模型的对象区分能力。也就是说,通过根据该预测属性特征向量、该区分参数、该预测兴趣度以及样本兴趣度对该推荐模型进行调整,可提高推荐模型兴趣度识别的准确性,可提高调整后的推荐模型对对象的区分度。通过提高推荐模型的兴趣度识别的准确度和对象区分能力,可提高对象推荐准确度。计算设备获取到目标推荐模型后,可采用目标推荐模型为用户推荐对象,可缓解对象之间的集中度,即提高对象之间的对象区分度,实现多样化的推荐,可满足用户个性化的需求。

[0190] 请参见图12,是本申请实施例提供的一种数据处理装置的结构示意图。该数据处

理装置可以是运行于计算机设备中的一个计算机程序(包括程序代码),例如该数据处理装置为一个应用软件;该装置可以用于执行本申请实施例提供的方法中的相应步骤。如图12所示,该数据处理装置2可以包括:

[0191] 获取模块31,用于获取待推荐的目标对象的属性信息,以及访问所述目标对象的目标访问者的画像数据;

[0192] 提取模块32,用于采用目标推荐模型对所述目标对象的属性信息进行属性特征提取,得到所述目标对象的属性特征向量,所述目标推荐模型为采用上述数据处理方法对目标推荐模型调整得到的;采用所述目标推荐模型对所述目标访问者的画像数据进行画像特征提取,得到所述目标访问者的画像特征向量;

[0193] 识别模块33,用于采用所述目标推荐模型对所述属性特征向量和所述画像特征向量进行兴趣度识别,得到所述目标访问者对所述目标对象的兴趣度;

[0194] 推荐模块34,用于根据所述目标访问者对所述目标对象的兴趣度向所述目标访问者推荐所述目标对象。

[0195] 可选的,推荐模块34,用于获取与所述目标对象的对象类型相似度大于相似度阈值的待选对象;根据所述目标访问者对所述目标对象的兴趣度向所述目标访问者推荐所述目标对象以及所述待选对象。

[0196] 本申请实施例中,计算机设备可以采用目标推荐模型对目标对象的属性信息进行属性特征提取,得到目标对象的属性特征向量,采用目标推荐模型对目标访问者的画像数据进行画像特征提取,得到画像特征向量。由于该目标推荐模型具有对象区分能力,因此,该目标对象的属性特征向量可用于区分目标对象。计算设备可以采用该目标推荐模型对属性特征向量和画像特征向量进行识别,得到目标访问者对该目标对象的兴趣度,根据该目标对象的兴趣度向目标访问者推荐目标对象。通过采用目标推荐模型向目标访问者推荐目标对象,可提高对象的区分度,可为目标访问者提供多样化的推荐,满足目标访问者个性化的需求。

[0197] 请参见图13,是本申请实施例提供的一种计算机设备的结构示意图。如图13所示,该计算机设备可以为图1a中的终端或服务器,上述计算机设备2000可以包括:处理器2001,网络接口2004和存储器2005,此外,上述计算机设备2000还可以包括:用户接口2003,和至少一个通信总线2002。其中,通信总线2002用于实现这些组件之间的连接通信。其中,用户接口2003可以包括显示屏(Display)、键盘(Keyboard),可选用户接口2003还可以包括标准的有线接口、无线接口。网络接口2004可选的可以包括标准的有线接口、无线接口(如WI-FI接口)。存储器2005可以是高速RAM存储器,也可以是非易失性的存储器(non-volatile memory),例如至少一个磁盘存储器。存储器2005可选的还可以是至少一个位于远离前述处理器2001的存储装置。如图12所示,作为一种计算机可读存储介质的存储器2005中可以包括操作系统、网络通信模块、用户接口模块以及设备控制应用程序。

[0198] 在图12所示的计算机设备2000中,网络接口2004可提供网络通讯功能;而用户接口2003主要用于为用户提供输入的接口;而处理器2001可以用于调用存储器2005中存储的设备控制应用程序,以实现:

[0199] 获取样本数据集合,所述样本数据集合包括至少两个样本对象的样本兴趣度以及用于区分所述至少两个样本对象的区分参数;

- [0200] 采用推荐模型提取所述至少两个样本对象的预测属性特征向量；
- [0201] 根据所述预测属性特征向量预测识别得到所述至少两个样本对象的预测兴趣度；
- [0202] 根据所述预测属性特征向量、所述区分参数、所述预测兴趣度以及所述样本兴趣度对所述推荐模型进行调整，将调整后的推荐模型作为目标推荐模型，所述目标推荐模型所输出针对所述至少两个样本对象的更新预测属性特征向量之间的区分程度，与所述区分参数对应的区分程度相匹配。
- [0203] 可选的，处理器2001可以用于调用存储器2005中存储的设备控制应用程序，以实现：
- [0204] 采用所述预测兴趣度以及所述样本兴趣度生成所述推荐模型的原始损失值；
- [0205] 根据所述预测属性特征向量以及所述区分参数确定对象区分度；
- [0206] 采用所述对象区分度生成所述推荐模型的区分损失值；
- [0207] 采用所述区分损失值和所述原始损失值生成所述推荐模型的总损失值；
- [0208] 若所述推荐模型的总损失值不满足收敛条件，则根据所述推荐模型的总损失值对所述推荐模型进行调整，将调整后的推荐模型作为目标推荐模型。
- [0209] 可选的，处理器2001可以用于调用存储器2005中存储的设备控制应用程序，以实现：
- [0210] 据所述推荐模型的总损失值对所述推荐模型进行调整；
- [0211] 采用所述调整后的推荐模型提取所述至少两个样本对象的更新预测属性特征向量；
- [0212] 根据所述更新预测属性特征向量预测识别得到所述至少两个样本对象的更新预测兴趣度；
- [0213] 根据所述更新预测属性特征向量、所述区分参数、所述更新预测兴趣度以及所述样本兴趣度，生成所述调整后的推荐模型的总损失值；
- [0214] 获取所述更新预测属性特征向量之间的角度，根据所述更新预测属性特征向量之间的角度确定所述更新预测属性特征向量之间的区分程度；
- [0215] 若所述调整后的推荐模型的总损失值满足收敛条件，且所述更新预测属性特征向量之间的区分程度，与所述区分参数对应的区分程度之间的差值小于程度阈值，则将调整后的推荐模型作为目标推荐模型。
- [0216] 可选的，处理器2001可以用于调用存储器2005中存储的设备控制应用程序，以实现：
- [0217] 获取所述预测属性特征向量的模长；
- [0218] 获取所述预测属性特征向量的模长与所述区分数值之间的差值；
- [0219] 根据所述差值确定所述对象区分度。
- [0220] 可选的，处理器2001可以用于调用存储器2005中存储的设备控制应用程序，以实现：
- [0221] 对所述区分向量与所述预测属性特征向量进行乘积运算，得到向量角度；
- [0222] 根据所述向量角度确定所述对象区分度。
- [0223] 可选的，处理器2001可以用于调用存储器2005中存储的设备控制应用程序，以实现：

- [0224] 获取第一样本对象的语义信息,所述第一样本对象为所述至少两个样本对象的任一样本对象;
- [0225] 从数据库中筛选出与所述第一样本对象的语义信息匹配的目标待选区分向量,作为所述第一样本对象的区分向量,所述数据库中包括多个待选区分向量,一个所述待选区分向量与一种语义信息对应。
- [0226] 可选的,处理器2001可以用于调用存储器2005中存储的设备控制应用程序,以实现:
- [0227] 从数据库中筛选出与所述第一样本对象的语义信息匹配的目标待选区分向量;
- [0228] 对所述目标待选区分向量以及剩余待选区分向量进行降维处理;所述剩余待选区分向量为所述数据库中除所述目标待选区分向量以外的待选区分向量;
- [0229] 获取降维处理后的目标待选区分向量与降维处理后的剩余待选区分向量之间的距离;
- [0230] 若所述距离大于距离阈值,则将所述目标待选区分向量作为所述第一样本对象的区分向量。
- [0231] 可选的,处理器2001可以用于调用存储器2005中存储的设备控制应用程序,以实现:
- [0232] 采用所述第二样本对象的预测兴趣度以及所述第二样本对象的样本兴趣度生成第一原始损失值;
- [0233] 采用所述第三样本对象的预测兴趣度以及所述第三样本对象的样本兴趣度生成第二原始损失值;
- [0234] 对所述第一原始损失值与所述第二原始损失值进行求和运算,得到所述推荐模型的原始损失值;
- [0235] 所述采用所述对象区分度生成所述推荐模型的区分损失值,包括:
- [0236] 对所述第二样本对象的对象区分度和所述第三样本对象的对象区分度进行求和运算,得到所述推荐模型的区分损失值。
- [0237] 可选的,处理器2001可以用于调用存储器2005中存储的设备控制应用程序,以实现:
- [0238] 采用所述推荐模型对所述画像数据进行画像特征提取,得到预测画像特征向量;
- [0239] 采用所述推荐模型对所述预测画像特征向量和所述预测属性特征向量进行预测识别,得到所述至少两个样本对象的预测兴趣度。
- [0240] 可选的,处理器2001可以用于调用存储器2005中存储的设备控制应用程序,以实现:
- [0241] 获取所述至少两个样本对象中每个样本对象的对象类型;
- [0242] 从所述画像数据中筛选出与所述对象类型匹配的画像数据,作为有效画像数据;
- [0243] 采用所述推荐模型对所述有效画像数据进行画像特征提取,得到所述预测画像特征向量。
- [0244] 可选的,处理器2001可以用于调用存储器2005中存储的设备控制应用程序,以实现:
- [0245] 获取待推荐的目标对象的属性信息,以及访问所述目标对象的目标访问者的画像

数据；

[0246] 采用目标推荐模型对所述目标对象的属性信息进行属性特征提取，得到所述目标对象的属性特征向量，所述目标推荐模型为采用如权利要求1-10任一项所述数据处理方法对目标推荐模型调整得到的；

[0247] 采用所述目标推荐模型对所述目标访问者的画像数据进行画像特征提取，得到所述目标访问者的画像特征向量；

[0248] 采用所述目标推荐模型对所述属性特征向量和所述画像特征向量进行兴趣度识别，得到所述目标访问者对所述目标对象的兴趣度；

[0249] 根据所述目标访问者对所述目标对象的兴趣度向所述目标访问者推荐所述目标对象。

[0250] 可选的，处理器2001可以用于调用存储器2005中存储的设备控制应用程序，以实现：

[0251] 获取与所述目标对象的对象类型相似度大于相似度阈值的待选对象；

[0252] 根据所述目标访问者对所述目标对象的兴趣度向所述目标访问者推荐所述目标对象以及所述待选对象。

[0253] 本申请中，计算机设备可以采用推荐模型提取该至少两个样本对象的预测属性特征向量，根据该预测属性特征向量识别得到该至少两个样本对象的预测兴趣度，根据该预测属性特征向量、该区分参数、该预测兴趣度以及样本兴趣度对该推荐模型进行调整，将调整后的推荐模型作为目标推荐模型。即通过该预测兴趣度以及样本兴趣度对推荐模型进行调整，可以提高推荐模型的兴趣度识别的准确性，通过根据该预测属性特征向量、该区分参数对推荐模型进行调整，可提高推荐模型对对象的区分度，即提高调整后的推荐模型的对象区分能力。也就是说，通过根据该预测属性特征向量、该区分参数、该预测兴趣度以及样本兴趣度对该推荐模型进行调整，可提高推荐模型兴趣度识别的准确性，可提高调整后的推荐模型对对象的区分度。通过提高推荐模型的兴趣度识别的准确度和对象区分能力，可提高对象推荐准确度。计算设备获取到目标推荐模型后，可采用目标推荐模型为用户推荐对象，可缓解对象之间的集中度，即提高对象之间的对象区分度，实现多样化的推荐，可满足用户个性化的需求。

[0254] 应当理解，本申请实施例中所描述的计算机设备2000可执行前文图3和图9所对应实施例中对上述数据处理方法的描述，也可执行前文图11或12所对应实施例中对上述基于数据处理装置的描述，在此不再赘述。另外，对采用相同方法的有益效果描述，也不再赘述。

[0255] 本申请中，计算机设备可以采用推荐模型提取该至少两个样本对象的预测属性特征向量，根据该预测属性特征向量识别得到该至少两个样本对象的预测兴趣度，根据该预测属性特征向量、该区分参数、该预测兴趣度以及样本兴趣度对该推荐模型进行调整，将调整后的推荐模型作为目标推荐模型。即通过该预测兴趣度以及样本兴趣度对推荐模型进行调整，可以确保推荐模型的兴趣度识别的准确性，通过根据该预测属性特征向量、该区分参数对推荐模型进行调整，可提高推荐模型对对象的区分度，即提高调整后的推荐模型的对象区分能力。也就是说，通过根据该预测属性特征向量、该区分参数、该预测兴趣度以及样本兴趣度对该推荐模型进行调整，可确保推荐模型兴趣度识别的准确性，可提高调整后的

推荐模型对对象的区分度,可提高推荐模型的预测准确度。计算设备获取到目标推荐模型后,可采用目标推荐模型为用户推荐对象,可缓解对象之间的集中度,即提高对象之间的对象区分度,实现多样化的推荐,可满足用户个性化的需求。

[0256] 此外,这里需要指出的是:本申请实施例还提供了一种计算机可读存储介质,且上述计算机可读存储介质中存储有前文提及的数据处理装置1以及基于数据处理装置2所执行的计算机程序,且上述计算机程序包括程序指令,当上述处理器执行上述程序指令时,能够执行前文图3和图9对应实施例中对上述数据处理方法的描述,因此,这里将不再进行赘述。另外,对采用相同方法的有益效果描述,也不再赘述。对于本申请所涉及的计算机可读存储介质实施例中未披露的技术细节,请参照本申请方法实施例的描述。

[0257] 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,上述的程序可存储于一计算机可读取存储介质中,该程序在执行时,可包括如上述各方法的实施例的流程。其中,上述的存储介质可为磁盘、光盘、只读存储器(Read-Only Memory,ROM)或随机存储器(Random Access Memory,RAM)等。

[0258] 以上所揭露的仅为本申请较佳实施例而已,当然不能以此来限定本申请之权利范围,因此依本申请权利要求所作的等同变化,仍属本申请所涵盖的范围。

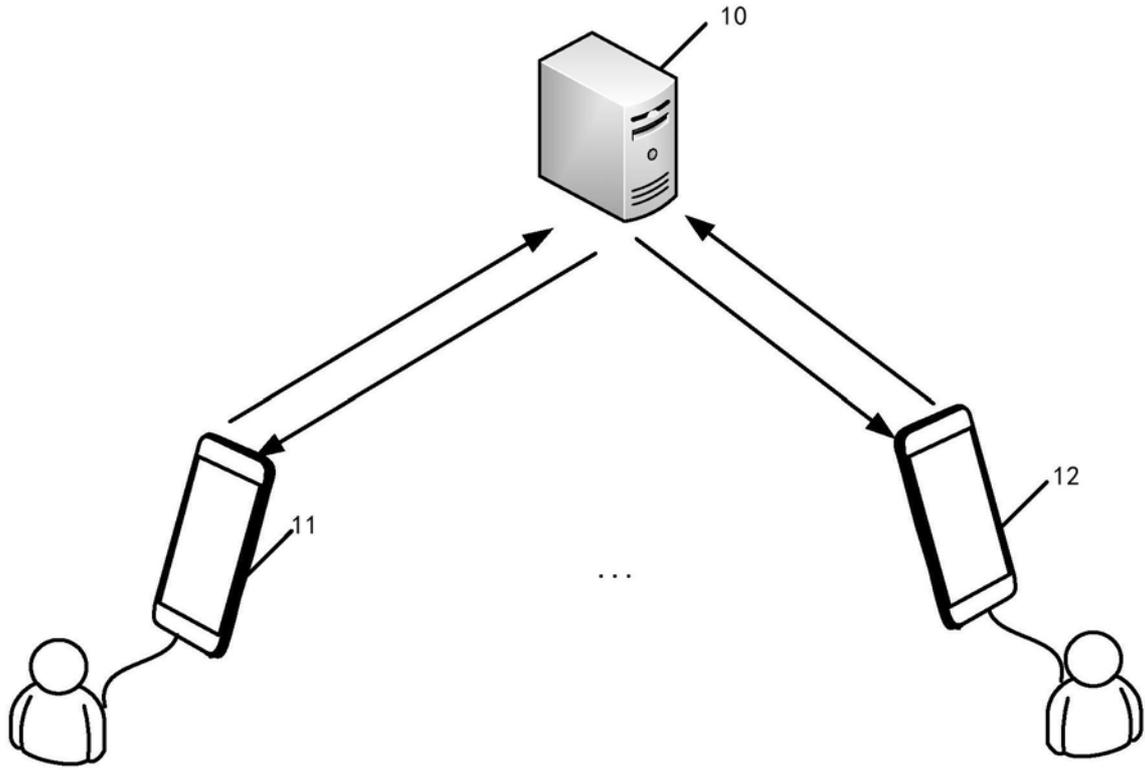


图1a

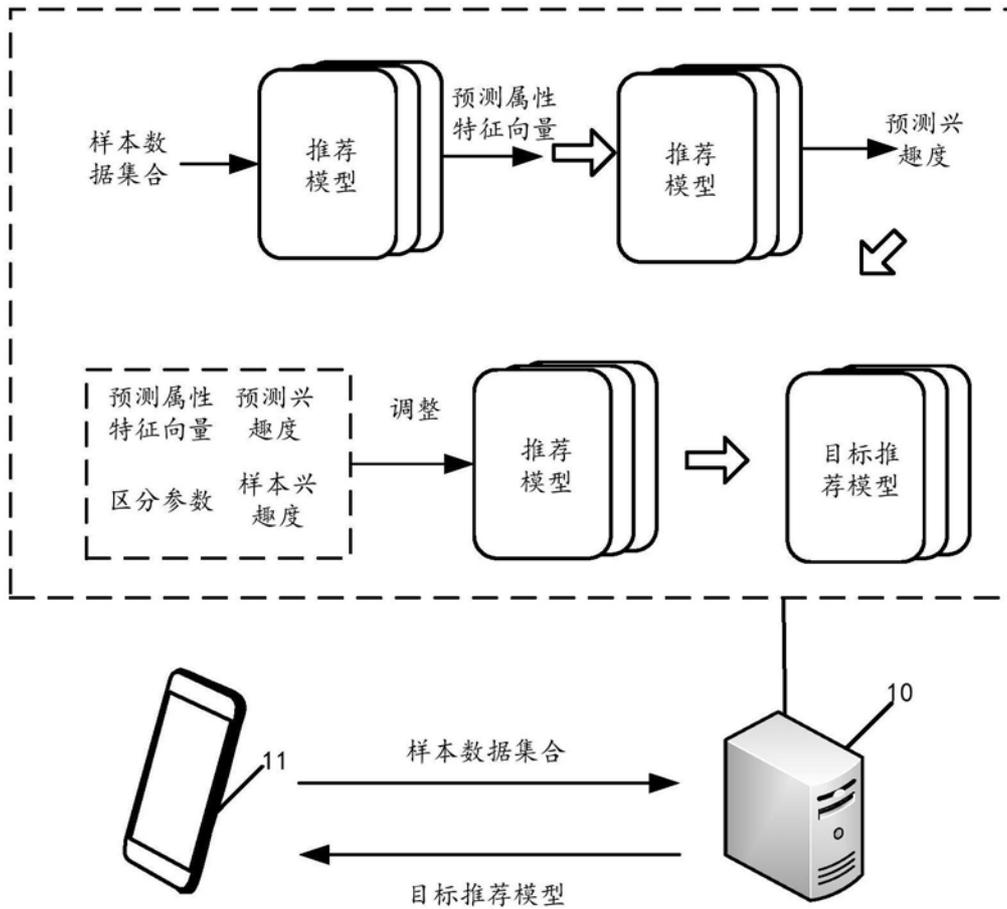


图1b

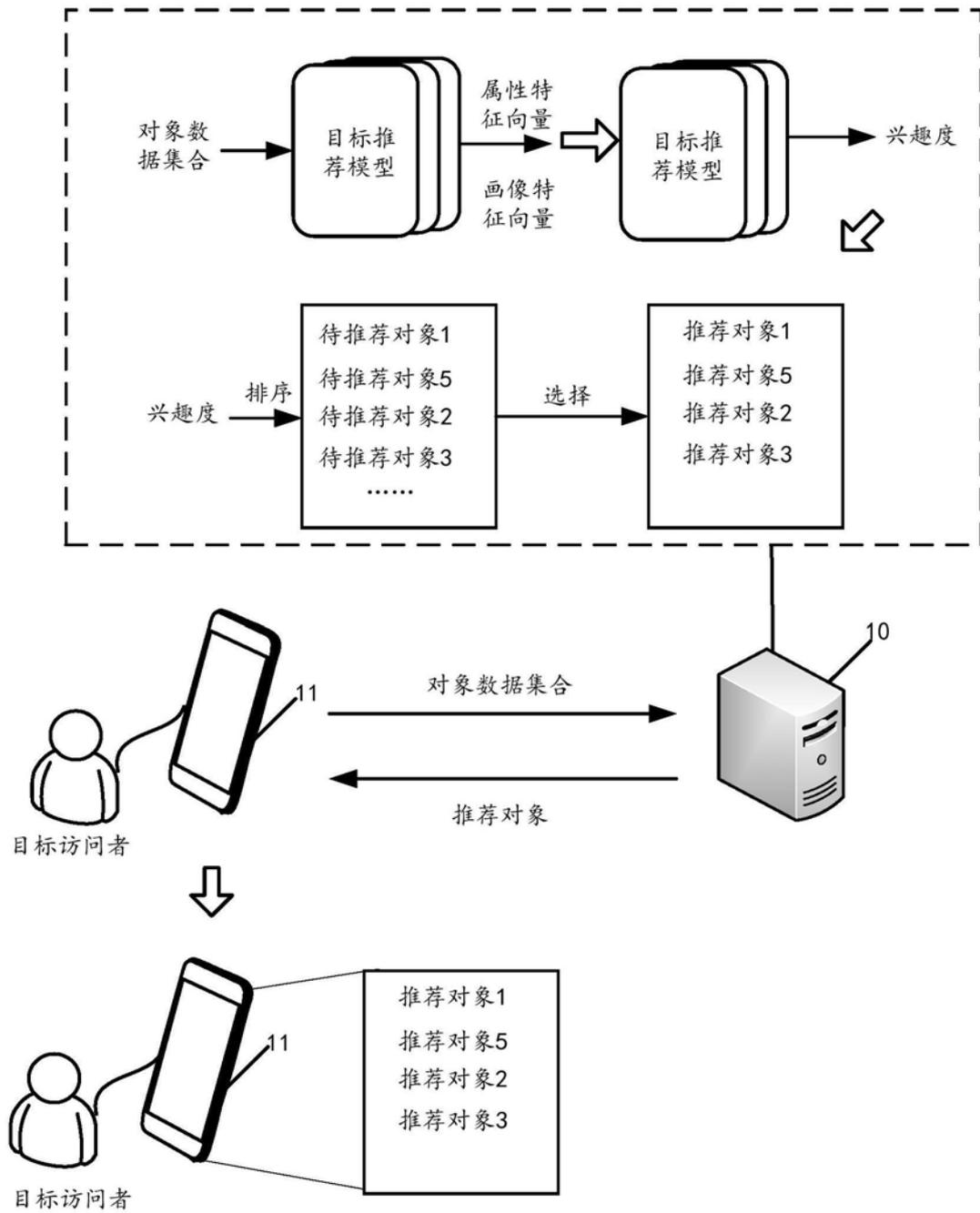


图2

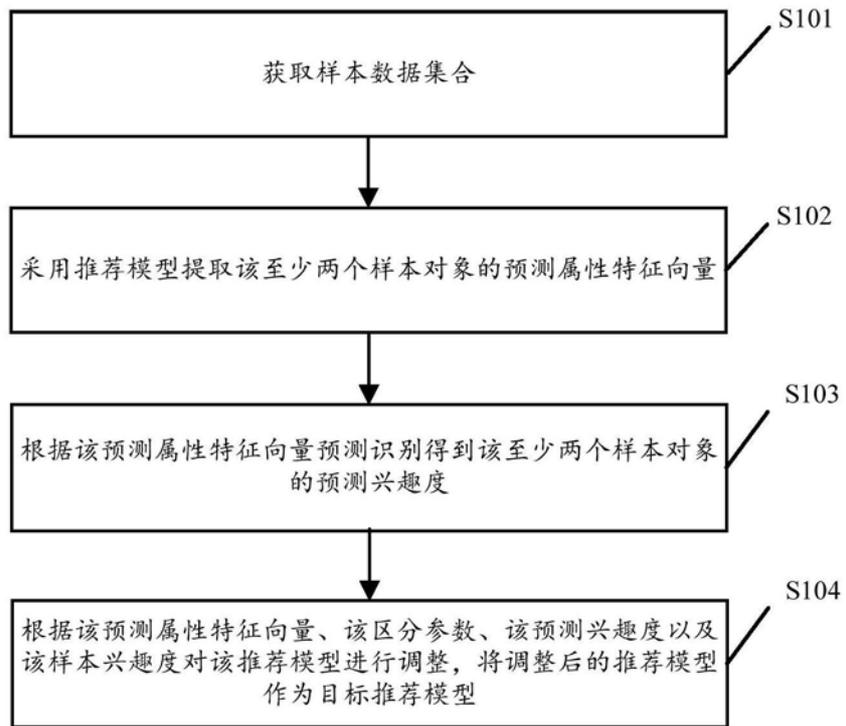


图3

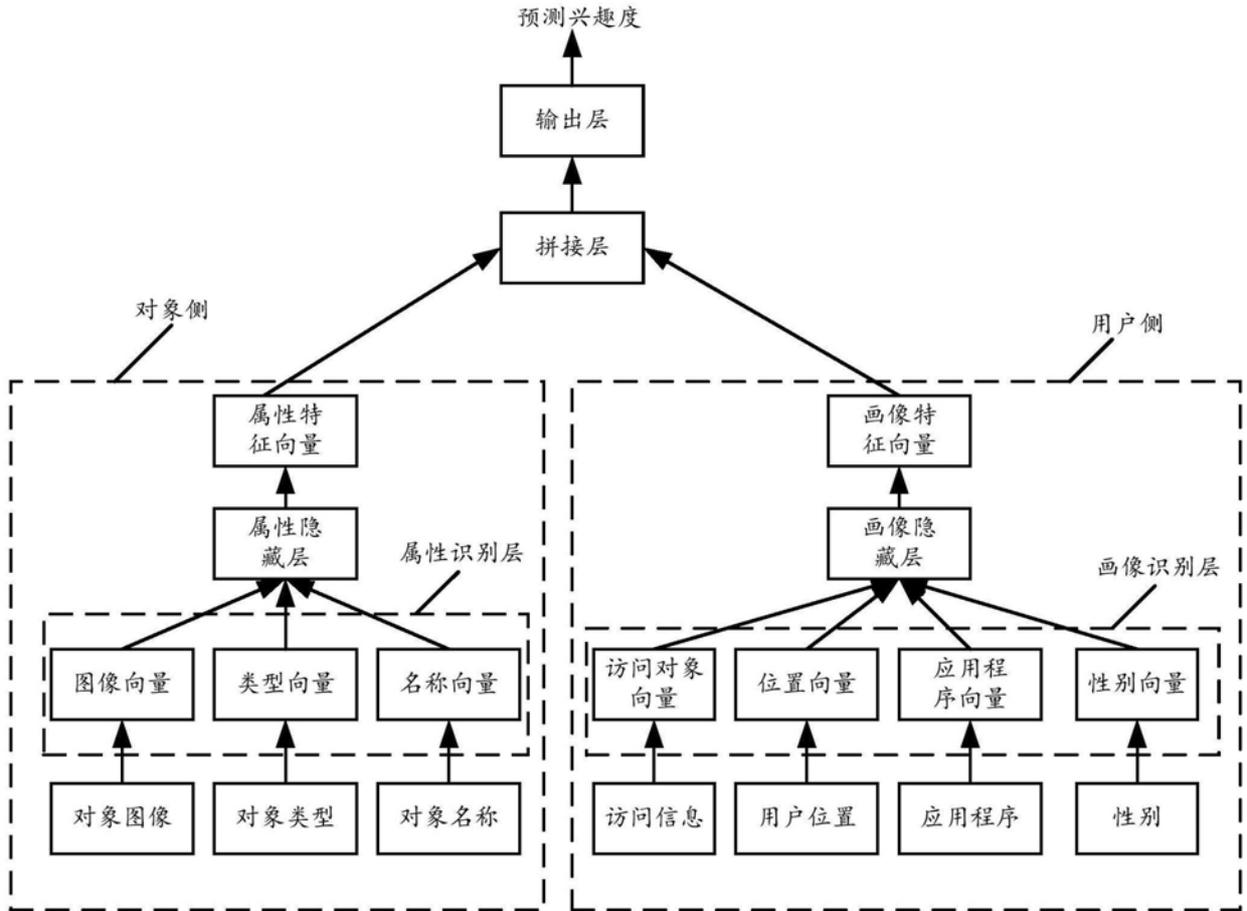


图4

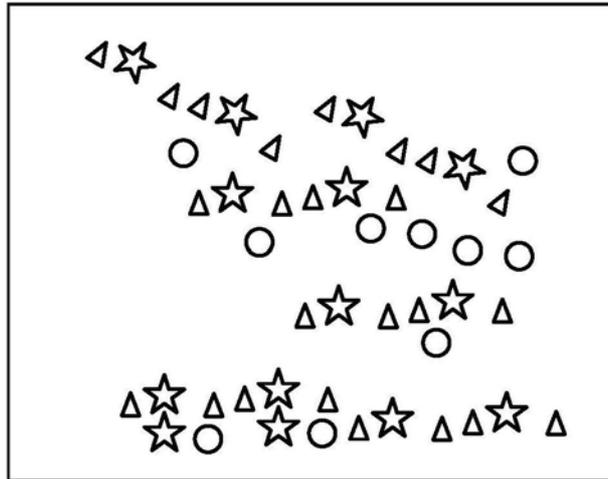


图5

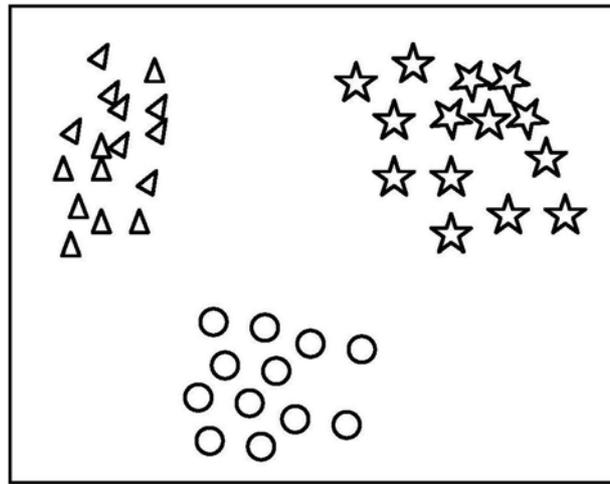


图6

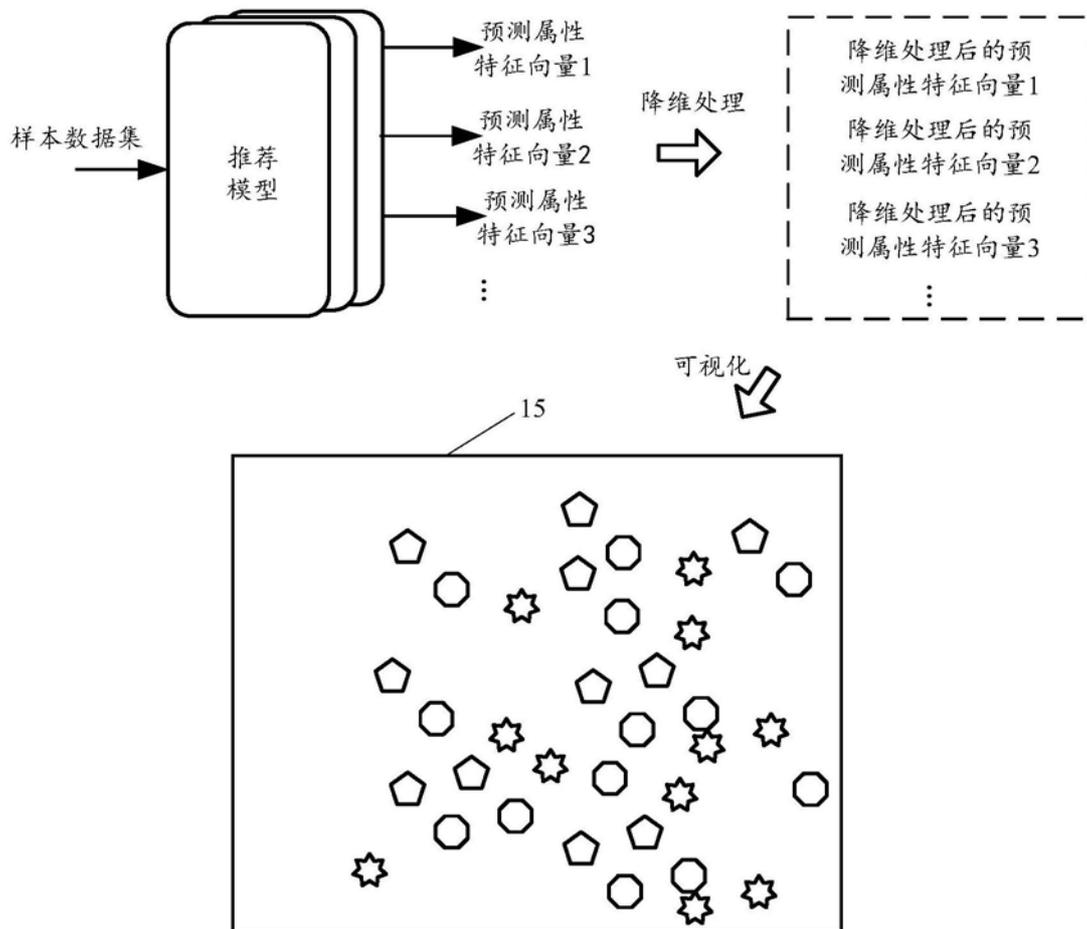


图7

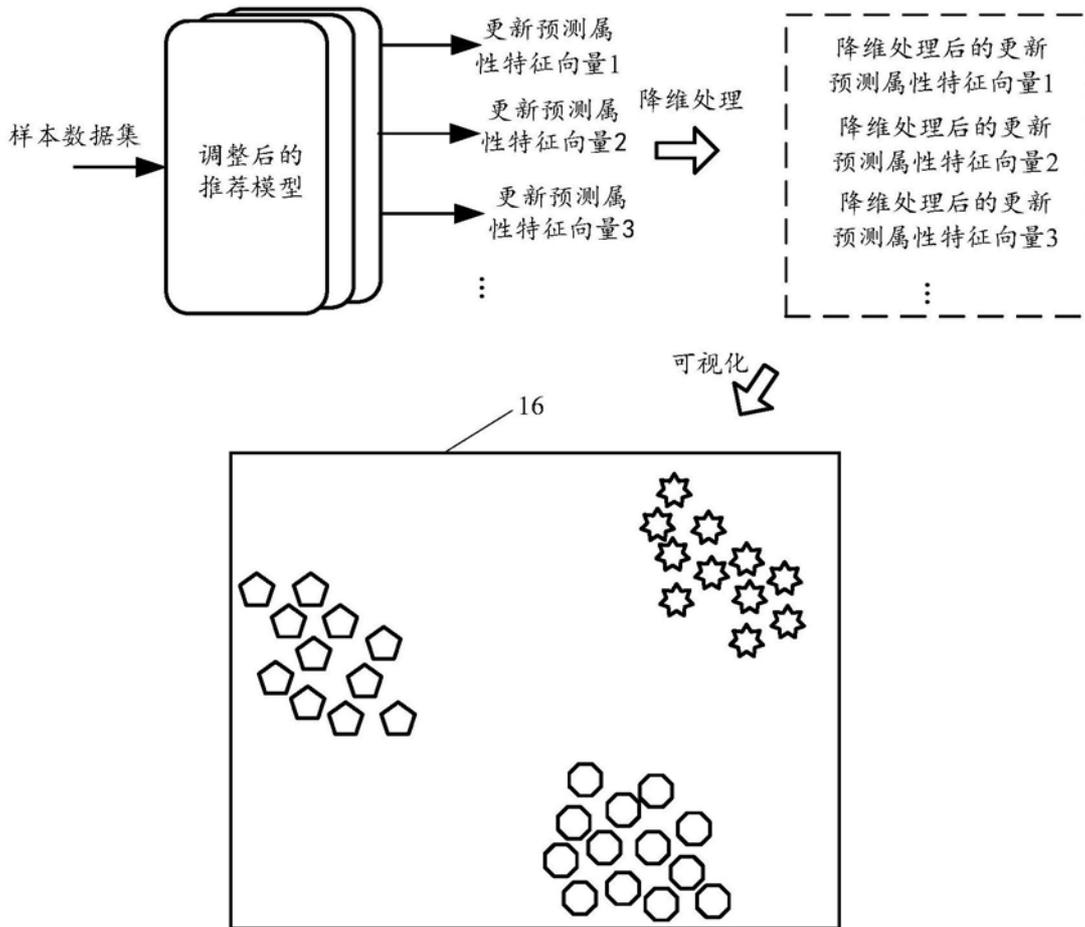


图8

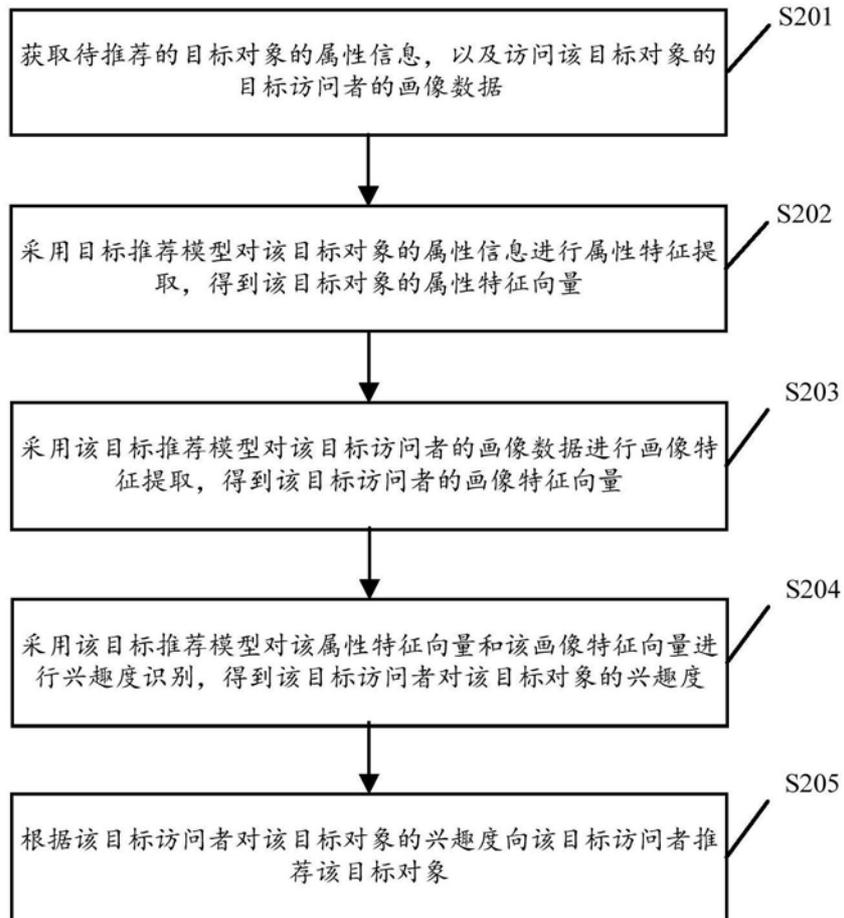


图9

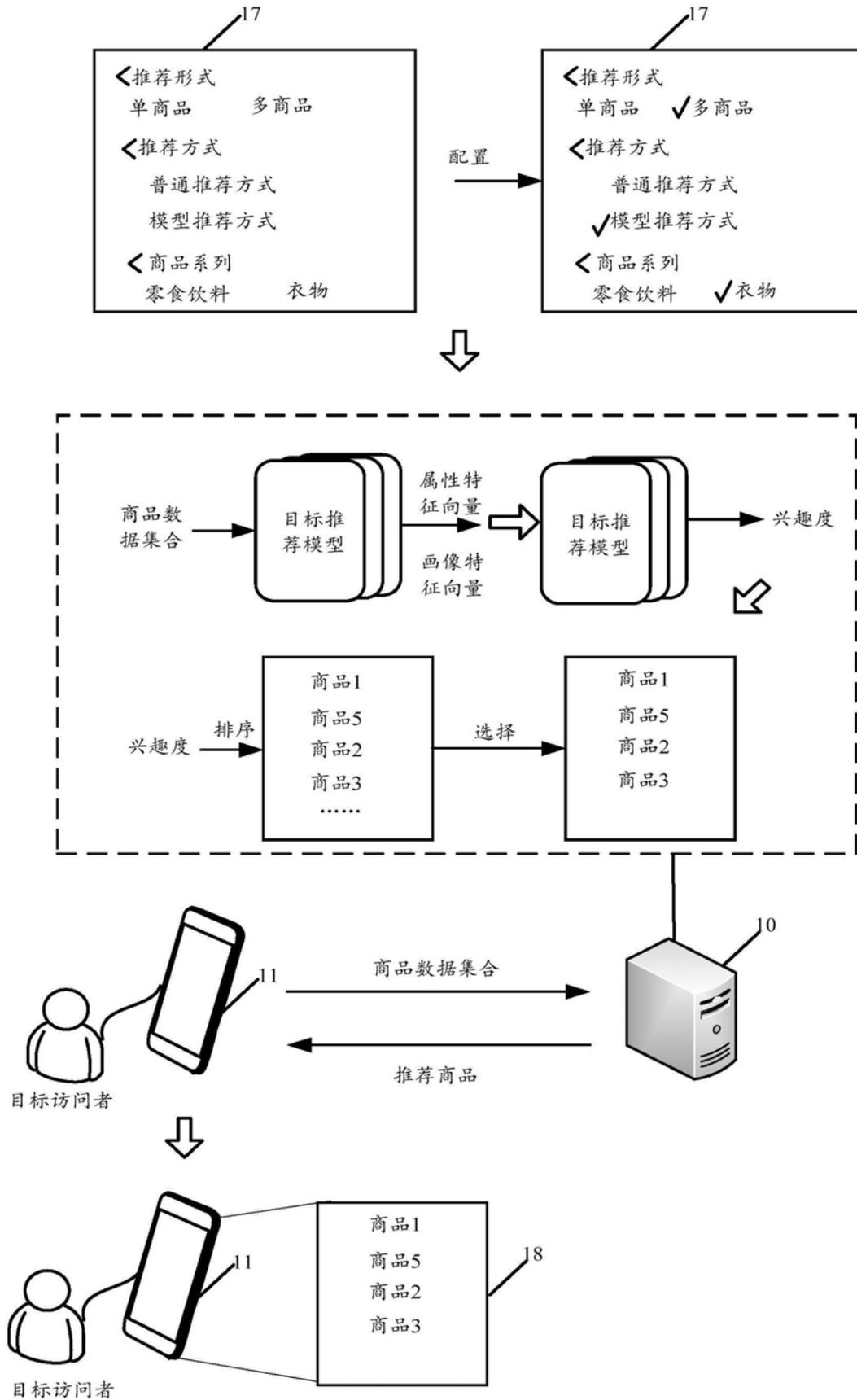


图10

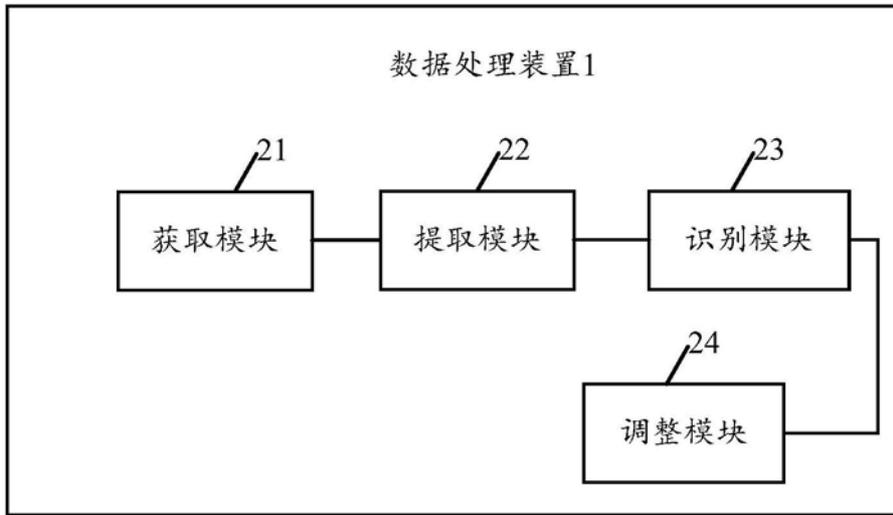


图11

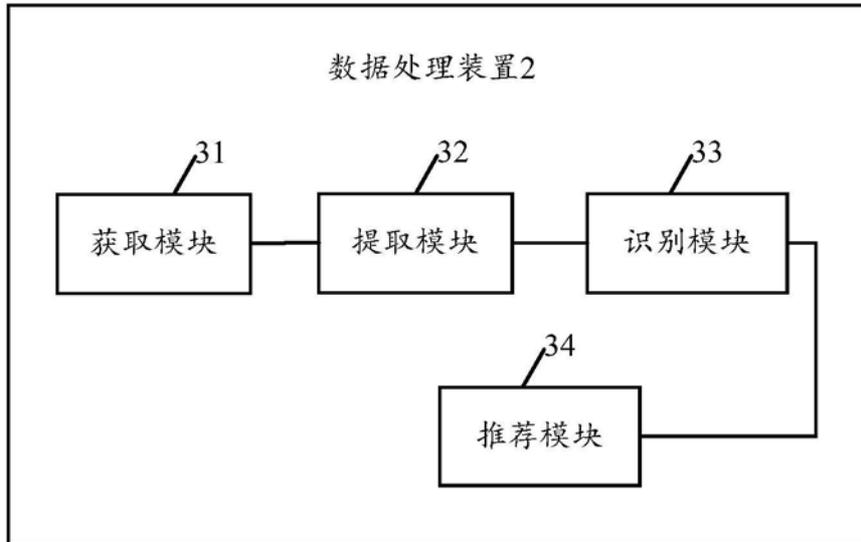


图12

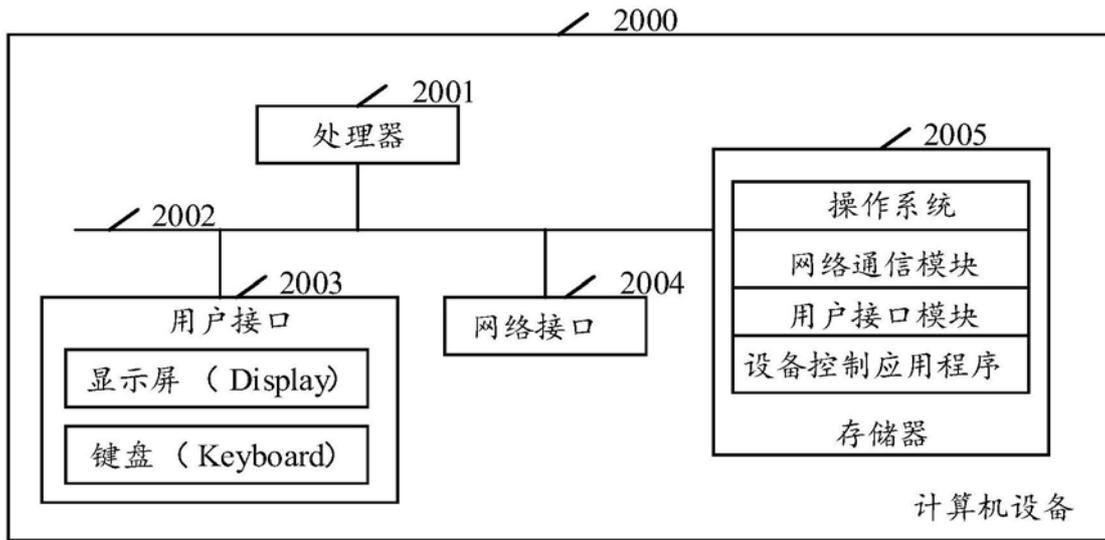


图13