

一种基于用户画像的物品推荐方法和装置

相关申请的交叉引用

5 本申请要求享有 2020 年 11 月 12 日提交的申请号为 202011264500.8 的中国发明专利申请的优先权，其全部内容通过引用并入本文。

技术领域

10 本公开涉及计算机技术领域，尤其涉及一种基于用户画像的物品推荐方法和装置。

背景技术

15 用户画像是电商作营销活动，个性化推荐，基础数据服务等的关键。只有实时的获取精准用户画像标签，才能在最短的时间内，以最小的成本争抢到质量最优，最精准的用户群体，进而作各种营销推广等活动，促进获客、留客。其中，用户画像为注册成功的用户 ID 所包括的一系列数据信息（例如购物信息、个人信息等），是一个虚拟的数据集合体。

20 在实现本公开的过程中，发明人发现现有技术中至少存在如下问题：

25 针对用户画像建模来说，现有的技术方案通常是利用大数据平台来存储用户购物等数据，然后通过人工分析并建模的方式来分类用户群体数据。传统的用户画像需要有打标数据，但现实是大多画像标签的打标数据不好获得，或者获取成本非常大，或者准确度低甚至于被视为噪声数据，即传统的用户画像只能处理单一业务场景（具有标签数据）的用户数据。因此，基于传统方式处理后的用户画像，进行用户推荐的效率不高、精准度低，用户体验也不好。

30 发明内容

有鉴于此，本公开实施例提供一种基于用户画像的物品推荐方法和装置，能够解决现有精准度低的用户画像所造成的营销活动效率不高的问题。

5 为实现上述目的，根据本公开实施例的一个方面，提供了一种基于用户画像的物品推荐方法，包括接收用户行为数据和用户属性信息，通过特征工程转换成待处理的用户行为数据和待处理的用户属性信息；获取当前标签计算任务，判断所述标签计算任务是否属于预测类任务，若是则调用预设的预测模型，若否则调用预设的统计规则模型；根据
10 所述预测模型或所述统计规则模型，基于待处理的用户行为数据和待处理的用户属性信息得到用户画像，进而根据用户画像向该用户推送物品信息。

 可选地，接收用户行为数据，包括
15 根据预设的数据数量和数据格式，获取不同类目下的用户行为数据。

 可选地，接收用户行为数据和用户属性信息之后，包括：
 调用预处理模型，对用户行为数据和用户属性信息进行预处理。

20 可选地，根据所述预测模型，基于待处理的用户行为数据和待处理的用户属性信息得到用户画像，包括：

 基于待处理的用户行为数据和待处理的用户属性信息，通过 Actor
 网络梯度融合的 DDPG 算法模型计算得到用户画像。

25 可选地，所述 Actor 网络梯度融合的 DDPG 算法模型，包括：
 将通过预设的监督学习模型计算得到的交叉熵损失值加入到 DDPG 算法的 Actor 中，评估 Actor 的输出值。

30 可选地，根据所述统计规则模型，基于待处理的用户行为数据和

待处理的属性信息得到用户画像，包括：

获取预设的深度强化学习模型和业务规则模型，通过待处理的属性数据和待处理的属性信息分别得到相应的第一用户画像和第二用户画像；

- 5 根据目标运营效果指标，确定第一用户画像和第二用户画像的权重，以将第一用户画像和第二用户画像融合得到最终用户画像。

可选地，包括：

所述的深度强化学习模型采用 Actor-Critic 算法。

10

另外，本公开还提供了一种基于用户画像的物品推荐装置，包括获取模块，用于接收属性数据和属性信息，通过特征工程转换成待处理的属性数据和待处理的属性信息；

- 15 处理模块，用于获取当前标签计算任务，判断所述标签计算任务是否属于预测类任务，若是则调用预设的预测模型，若否则调用预设的统计规则模型；根据所述预测模型或所述统计规则模型，基于待处理的属性数据和待处理的属性信息得到用户画像，进而根据用户画像向该用户推送物品信息。

- 20 上述发明中的一个实施例具有如下优点或有益效果：本公开针对不同特性的属性数据和属性信息，采用不同的任务处理方式，即支持不同场景的用户画像处理，支持场景包括：带打标数据标签，非打标数据以及业务规则统计计算类标签任务。并且，本公开能够利用深度强化学习，结合线上运营的真实数据实时训练用户画像的各类处理模型。更进一步地，本公开将监督学习和强化学习进行融合，从而能够在用户的生命周期里，用户画像处理在每次通过活动运营事实上会在下次以及下下次对用户的属性产生影响，即每次处理的结果间存在关联。

- 30 上述的非惯用的可选方式所具有的进一步效果将在下文中结合具

体实施方式加以说明。

附图说明

附图用于更好地理解本公开，不构成对本公开的不当限定。其中：

5 图 1 是根据本公开第一实施例的基于用户画像的物品推荐方法的主要流程的示意图；

图 2 是根据本公开另一实施例的基于用户画像的物品推荐方法的主要流程的示意图；

10 图 3 是根据本公开实施例的基于用户画像的物品推荐装置的主要模块的示意图；

图 4 是本公开实施例可以应用于其中的示例性系统架构图；

图 5 是适于用来实现本公开实施例的终端设备或服务器的计算机系统的结构示意图。

15 具体实施方式

以下结合附图对本公开的示范性实施例做出说明，其中包括本公开实施例的各种细节以助于理解，应当将它们认为仅仅是示范性的。因此，本领域普通技术人员应当认识到，可以对这里描述的实施例做出各种改变和修改，而不会背离本公开的范围和精神。同样，为了清楚和简明，以下的描述中省略了对公知功能和结构的描述。

20

图 1 是根据本公开第一实施例的基于用户画像的物品推荐方法的主要流程的示意图，如图 1 所示，所述基于用户画像的物品推荐方法包括：

25 步骤 S101，接收用户行为数据和用户属性信息，通过特征工程转换成待处理的用户行为数据和待处理的属性信息。

在一些实施例中，接收用户行为数据，包括：

根据预设的数据数量和数据格式，获取不同类目下的用户行为数据。例如：针对用户性别数据分析，首先需要分析用户性别预测涉及到的特征数据，用户在不同类目下的行为数据的完整性（即格式）、

30

大小以及数量等，为后面分析挖掘做基础准备。

5 作为另一些实施例，接收用户行为数据和用户属性信息之后，可以调用预处理模型，对用户行为数据和用户属性信息进行预处理。例如：对用户行为数据和用户属性信息等

10 例如：对用户行为数据和用户属性信息等进行初步预处理，包括缺失值、噪声、离群点、数据类型等处理，结合分析结果进行数据清理（包括缺失值、噪声和离群点处理等）、数据编码、数据变形（包括标准化、正则化、缩放等）等预处理操作。

值得说明的是，在步骤 S101 中所述的特征工程是对用户行为数据以及用户属性信息在不同类目下的记录经过变换、交叉、映射以及提取等操作加工成模型所需的数据的过程。

10

步骤 S102，获取当前标签计算任务，判断所述标签计算任务是否属于预测类任务，若是则调用预设的预测模型，若否则调用预设的统计规则模型。

15

在一些实施例中，基于待处理的用户行为数据和待处理的用户属性信息，通过 Actor 网络梯度融合的 DDPG 算法模型计算得到用户画像。进一步地，将通过预设的监督学习模型计算得到的交叉熵损失值加入到 DDPG 算法的 Actor 中，评估 Actor 的输出值。其中，DDPG 全称 Deep Deterministic Policy Gradient，是网络融合进 DPG 的策略学习方法，且融合了 Actor-Critic 框架。

20

步骤 S103，根据所述预测模型或所述统计规则模型，基于待处理的用户行为数据和待处理的用户属性信息得到用户画像，进而根据用户画像向该用户推送物品信息。

25

在一些实施例中，如果调用预设的统计规则模型，基于待处理的用户行为数据和待处理的用户属性信息得到用户画像，包括：

获取预设的深度强化学习模型和业务规则模型，通过待处理的用户行为数据和待处理的用户属性信息分别得到相应的第一用户画像和第二用户画像。然后，根据目标运营效果指标，确定第一用户画像和第二用户画像的权重，以将第一用户画像和第二用户画像融合得到最

30

终用户画像。

进一步地实施例,所述的深度强化学习模型采用 Actor-Critic 算法。其中,深度强化学习基本的三要素:状态(state)、行为(action)和奖赏(reward)。在本公开中对应的分别是:

5 state 为特征工程加工后的待处理的用户行为数据和待处理的用户属性信息,例如:销量,销售额,用户行为,商品属性,用户和商品交叉特征等等。

 action 可以为在机器学习中的 label 集合,聚类任务中的聚类类别集合,业务规则中的计算结果等等。

10 reward 为将要达到的目标(例如:提高点击率、提高 GMV 等)转化为具体的奖赏函数 R,在学习过程中引导 Agent 完成目标,本公开中将运营活动中用户的 pv(访问量)、销量、销售额等作为 reward。

 值得说明的是,本公开能够将画像模型看作智能体(Agent)、把
15 用户看作环境(Environment),则画像模型做营销等问题可以被视为顺序决策问题。其中,每次做营销活动时,Agent(画像模型)做预测,将预测结果给用户。用户根据 Agent(画像模型)的预测结果,给出点击、浏览等反馈信号。Agent(画像模型)接收反馈信号,在新营销活动时做出新的预测策略,向用户推荐物品。在这种反复不断地试错过程
20 中将进行持续优化,Agent 将逐步学习到最优的策略。

 综上所述,本公开能够利用深度强化学习,结合线上运营的真实数据实时训练用户画像的标签计算算法模型。与传统用户画像标签模型训练方式不同的是,将监督学习和强化学习进行融合,并针对不同的
25 计算场景采用不同的融合方式优化传统标签的计算方式,支持场景包括:带打标数据标签,非打标数据(聚类标签任务),及业务规则统计计算类标签任务。也就是说,本公开能够在用户的生命周期里,画像模型在每次通过活动运营事实上会在下次以及下下次对用户的行为产生影响,即画像模型每次预测的结果间存在关联(举例:本次活动筛选男性用户推销剃须刀,如果模型计算结果是准确的,则该用户
30

可能会对营销活动给予支持等行为，该行为即为用户的反馈数据，应及时加入到模型中进行训练；下次推送活动时模型就更有可能将该用户预测为男性；相反，若该用户为女性，推送的剃须刀更有可能被忽视，同样这种行为应被模型及时学习利用来更新模型）。

5

图 2 是根据本公开另一实施例的基于用户画像的物品推荐方法的主要流程的示意图（图 2 中的 DRL 为深度强化学习，ML 为机器学习），所述基于用户画像的物品推荐方法可以包括：

接收到的用户行为数据和用户属性信息依次通过数据分析、数据
10 预处理以及特征工程的处理，得到待处理的用户行为数据和待处理的
用户属性信息。其中，数据分析为根据预设的数据数量和数据格式，
获取不同类目下的用户行为数据。数据预处理为调用预处理模型，对
用户行为数据和用户属性信息进行预处理。特征工程是对用户行为数
据以及用户属性信息在不同类目下的记录经过变换、交叉、映射以及
15 提取等操作加工成模型所需的数据的过程。

完成特征工程后，判断当前标签计算任务是否属于预测类任务，
若是则调用预测模型（图 2 中左边任务流程），若否则调用统计规则
模型（图 2 中右边任务流程）。例如：如果当前标签计算任务可以简
单采用统计方式如用 sql 计算方式求得的，则调用统计规则模型。如果
20 当前标签计算任务是预测类的任务如有大部分性别标签需要采用机器
学习算法才能计算得到的而不能通过规则计算求得的，则预测模型）。

而左边任务流程是基于 DDPG 算法基础上增加了机器学习模型作
Actor 网络的梯度融合，使强化学习模型（例如 DDPG 算法）能从机器
学习模型中学习到梯度。较佳地，图 2 中左边任务流程中机器学习模
25 型 Model 可以采用监督学习模型，例如：svm（支持向量机），xgb（是
GBDT 的一种工业实现）等。具体地：

先对 DDPG 算法中的 Actor 和 Critic 分别构建两个网络：Actor 网络
为：Actor_eval_net 和 Actor_target_net。Critic 网络为：Critic_eval_net
和 Critic_target_net。其中，Actor_eval_net 输入 state（在本公开中即为
30 特征向量：待处理的用户行为数据和待处理用户属性信息），输出行

为 action。Actor_target_net 输入 next_state (action 作用在环境后得到 next_state 和对应的奖赏 reward), 输出行为 next_action。Critic_eval_net 输入 action, 输出行为对应的 Q 值(Q 值是指 $Q(\text{state}, \text{action})$ 函数(Quality) 值, 用来表示智能体在某状态下采取某个动作并在之后采取最优动作 5 条件下的打折的未来奖励)。Critic_target_net 输入 next_action 和 next_state, 输出行为对应的 Q 值。

例如: 以用户性别标签计算为例, 在 Actor_eval_net 预测网络中输入的状态 state 为用户性别对应的特征, 即特征工程计算的特征, 包括: 用户在不同类目的行为数据例如浏览, 下单, 加购, 点击, 评论等销量, 销售额, 商品属性, 用户和商品交叉特征等等。输出的是 action 10 对应用户性别预测标签是男(0), 女(1), 未知(-1)。选择了男或女的 action 动作后, 在环境中(对应营销活动中即: 业务人员根据营销场景条件选择对应所需的性别, 并作投放, 消息推送等)得到人群用户的反馈, 例如给男士推送了剃须刀, 则该男士用户可能会产生点击, 15 下单, 加购, 评论等行为, 这些在不同类目下的行为将作为下一次输入的状态(next_state)输入到 Actor_target_net, 再输出下一次的 行为(next_action), Critic_eval_net 网络同样的输入是根据当前状态得到的 action 动作(选择男, 女), 输出行为对应的 Q 值, Critic_target_net 输入的是下一次的 action 和下一次的状态, 输出的是对应的 Q 值。

20 较佳地, 本公开利用监督学习模型(例如监督学习模型 Model 可以是 svm(支持向量机)、XGBoost(Extreme Gradient Boosting 的简称, 是 GBDT 的一种高效实现)等), 输入对应的特征(用户行为数据和用户属性信息)及打标数据, 训练模型并对未打标数据进行预测, 25 补齐未打标数据。此时所有待输入数据全部完成打标, 再构建一层网络, 计算 Actor 的输出在监督学习模型中的交叉熵损失:

$$L(a, X, Y) = \frac{1}{n} \sum_i^n y_i \log(\sigma(a^T x_i)) + (1 - y_i) \log(1 - \sigma(a^T x_i))$$

其中, a: 网络输出的 action, x: 用户行为数据和用户属性信息, y: 打标数据, σ : 待训练网络参数。

这样 Actor 最终的梯度损失函数变成:

$$\nabla_{\theta^{\mu}} J = -\frac{1}{N} \sum_i^N \nabla_a Q(s, a | \theta^Q) |_{s=s_i, a=\mu(s_i)} \nabla_{\theta^{\mu}} \mu(s | \theta^{\mu}) |_{s_i} + \lambda \nabla_a L(a, X, Y) |_{a=\mu(s_i)} \nabla_{\theta^{\mu}} \mu(s | \theta^{\mu}) |_{s_i}$$

其中，s:代表某个时刻代理的状态；a: 代表在某个时刻下执行的动作；Q(s, a): 表示智能体在某状态下采取某个动作并在之后采取最优动作条件下的打折的未来奖励； λ : 待调的超参，是监督学习模型的权重； θ : Actor 的网络参数， μ : 在某时刻的状态下，对应 actor 的动作值，N: 模型训练迭代次数。

因此，本公开将监督学习模型计算的交叉熵损失 loss 加入到 Actor 中，评估 Actor 的输出值，增强强化学习模型的稳定性并充分利用打标数据，提升整个预测模型性能。

而右边任务流程是利用用户反馈数据进行深度强化学习模型计算（其中训练流程同左边任务流程（少了 Model 部分）），以及业务规则模型计算（例如：计算用户是否已婚，若没有打标数据作为训练集，则业务规则模型计算，如历史购买物品名称包括“怀孕”，“孩子”等已婚特征词。），最后再根据线上的运营效果指标（如：转化率，停留时长等）确定两种计算结果的权重，再进行融合。

图 3 是根据本公开实施例的基于用户画像的物品推荐装置的主要模块的示意图，如图 3 所示，所述基于用户画像的物品推荐装置 300 包括获取模块 301 和处理模块 302。其中，获取模块 301 接收用户行为数据和用户属性信息，通过特征工程转换成待处理的用户行为数据和待处理的用户属性信息。处理模块 302 获取当前标签计算任务，判断所述标签计算任务是否属于预测类任务，若是则调用预设的预测模型，若否则调用预设的统计规则模型；根据所述预测模型或所述统计规则模型，基于待处理的用户行为数据和待处理的用户属性信息得到用户画像，进而根据用户画像向该用户推送物品信息。

在一些实施例中，获取模块 301 接收用户行为数据，包括根据预设的数据数量和数据格式，获取不同类目下的用户行为数

据。

在一些实施例中，获取模块 301 接收用户行为数据和用户属性信息之后，包括：

调用预处理模型，对用户行为数据和用户属性信息进行预处理。

5 在一些实施例中，处理模块 302 根据所述预测模型，基于待处理的
的用户行为数据和待处理的
的用户属性信息得到用户画像，包括：

基于待处理的
的用户行为数据和待处理的
的用户属性信息，通过 Actor
网络梯度融合的 DDPG 算法模型计算得到用户画像。

10 在一些实施例中，所述 Actor 网络梯度融合的 DDPG 算法模型，
包括：

将通过预设的
的监督学习模型计算得到的交叉熵损失值加入到
DDPG 算法的 Actor 中，评估 Actor 的输出值。

在一些实施例中，处理模块 302 根据所述统计规则模型，基于待
处理的
的用户行为数据和待处理的
的用户属性信息得到用户画像，包括：

15 获取预设的
的深度强化学习模型和业务规则模型，通过待处理的
的用户行为数据和待处理的
的用户属性信息分别得到相应的第一用户画像和
第二用户画像；

根据目标运营效果指标，确定第一用户画像和第二用户画像的权
重，以将第一用户画像和第二用户画像融合得到最终用户画像。

20 在一些实施例中，所述的深度强化学习模型采用 Actor-Critic 算法。

需要说明的是，在本公开所述基于用户画像的物品推荐方法和所
述基于用户画像的物品推荐装置在具体实施内容上具有相应关系，故
重复内容不再说明。

25 图 4 示出了可以应用本公开实施例的基于用户画像的物品推荐方
法或基于用户画像的物品推荐装置的示例性系统架构 400。

如图 4 所示，系统架构 400 可以包括终端设备 401、402、403，网
络 404 和服务
器 405。网络 404 用以在终端设备 401、402、403 和服务
器 405 之间提供通信链路的介质。网络 404 可以包括各种连接类型，
30 例如有线、无线通信链路或者光纤电缆等等。

用户可以使用终端设备 401、402、403 通过网络 404 与服务器 405 交互，以接收或发送消息等。终端设备 401、402、403 上可以安装有各种通讯客户端应用，例如购物类应用、网页浏览器应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等（仅为示例）。

5 终端设备 401、402、403 可以是具有基于用户画像的物品推荐屏并且支持网页浏览的各种电子设备，包括但不限于智能手机、平板电脑、膝上型便携计算机和台式计算机等等。

服务器 405 可以是提供各种服务的服务器，例如对用户利用终端设备 401、402、403 所浏览的购物类网站提供支持的后台管理服务器（仅为示例）。后台管理服务器可以对接收到的产品信息查询请求等数据进行分析等处理，并将处理结果（例如目标推送信息、产品信息--

10 仅为示例）反馈给终端设备。

需要说明的是，本公开实施例所提供的基于用户画像的物品推荐方法一般由服务器 405 执行，相应地，计算装置一般设置于服务器 405

15 中。

应该理解，图 4 中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要，可以具有任意数目的终端设备、网络和服务器。

下面参考图 5，其示出了适于用来实现本公开实施例的终端设备的计算机系统 500 的结构示意图。图 5 示出的终端设备仅仅是一个示例，

20 不应对本公开实施例的功能和使用范围带来任何限制。

如图 5 所示，计算机系统 500 包括中央处理单元（CPU）501，其可以根据存储在只读存储器（ROM）502 中的程序或者从存储部分 508 加载到随机访问存储器（RAM）503 中的程序而执行各种适当的动作和处理。在 RAM503 中，还存储有计算机系统 500 操作所需的各种程序和数据。CPU501、ROM502 以及 RAM503 通过总线 504 彼此相连。输入/输出（I/O）接口 505 也连接至总线 504。

25

以下部件连接至 I/O 接口 505：包括键盘、鼠标等的输入部分 506；

30 包括诸如阴极射线管（CRT）、液晶基于用户画像的物品推荐器（LCD）

等以及扬声器等的输出部分 507；包括硬盘等的存储部分 508；以及包括诸如 LAN 卡、调制解调器等网络接口卡的通信部分 509。通信部分 509 经由诸如因特网的网络执行通信处理。驱动器 510 也需要连接至 I/O 接口 505。可拆卸介质 511，诸如磁盘、光盘、磁光盘、半导体存储器等等，根据需要安装在驱动器 510 上，以便于从其上读出的计算机程序根据需要被安装入存储部分 508。

特别地，根据本公开的实施例，上文参考流程图描述的过程可以被实现为计算机软件程序。例如，本公开的实施例包括一种计算机程序产品，其包括承载在计算机可读介质上的计算机程序，该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中，该计算机程序可以通过通信部分 509 从网络上被下载和安装，和/或从可拆卸介质 511 被安装。在该计算机程序被中央处理单元 (CPU) 501 执行时，执行本公开的系统限定上述功能。

需要说明的是，本公开所示的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是一但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于：具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器 (RAM)、只读存储器 (ROM)、可擦式可编程只读存储器 (EPROM 或闪存)、光纤、便携式紧凑磁盘只读存储器 (CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中，计算机可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本公开中，计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何

计算机可读介质，该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输，包括但不限于：无线、电线、光缆、RF 等等，或者上述的任意合适的组合。

5

附图中的流程图和框图，图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，上述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的，框图或流程图中的每个方框、以及框图或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

10

15

描述于本公开实施例中所涉及到的模块可以通过软件的方式实现，也可以通过硬件的方式来实现。所描述的模块也可以设置在处理器中，例如，可以描述为：一种处理器包括获取模块和处理模块。其中，这些模块的名称在某种情况下并不构成对该模块本身的限定。

20

作为另一方面，本公开还提供了一种计算机可读介质，该计算机可读介质可以是上述实施例中描述的设备中所包含的；也可以是单独存在，而未装配入该设备中。上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序被一个该设备执行时，使得该设备包括接收用户行为数据和用户属性信息，通过特征工程转换成待处理的用户行为数据和待处理的用户属性信息；获取当前标签计算任务，判断所述标签计算任务是否属于预测类任务，若是则调用预设的预测模型，若否则调用预设的统计规则模型；根据所述预测模型或所述统

25

30

计规则模型，基于待处理的用户行为数据和待处理的用户属性信息得到用户画像，进而根据用户画像向该用户推送物品信息。

5 根据本公开实施例的技术方案，能够解决现有精准度低的用户画像所造成的营销活动效率不高的问题。

10 上述具体实施方式，并不构成对本公开保护范围的限制。本领域技术人员应该明白的是，取决于设计要求和因素，可以发生各种各样的修改、组合、子组合和替代。任何在本公开的精神和原则之内所作的修改、等同替换和改进等，均应包含在本公开保护范围之内。

权 利 要 求 书

1. 一种基于用户画像的物品推荐方法，其包括：
接收用户行为数据和用户属性信息，通过特征工程转换成待处理的
5 的用户行为数据和待处理的用户属性信息；
获取当前标签计算任务，判断所述标签计算任务是否属于预测类
任务，若是则调用预设的预测模型，若否则调用预设的统计规则模型；
根据所述预测模型或所述统计规则模型，基于待处理的用户行为
数据和待处理的用户属性信息得到用户画像，进而根据用户画像向该
10 用户推送物品信息。
2. 根据权利要求 1 所述的方法，其中，接收用户行为数据，包括：
根据预设的数据数量和数据格式，获取不同类目下的用户行为数
据。
- 15 3. 根据权利要求 1 所述的方法，其中，接收用户行为数据和用户
属性信息之后，包括：
调用预处理模型，对用户行为数据和用户属性信息进行预处理。
- 20 4. 根据权利要求 1 所述的方法，其中，根据所述预测模型，基于待
处理的用户行为数据和待处理的用户属性信息得到用户画像，包括：
基于待处理的用户行为数据和待处理的用户属性信息，通过 Actor
网络梯度融合的 DDPG 算法模型计算得到用户画像。
- 25 5. 根据权利要求 4 所述的方法，其中，所述 Actor 网络梯度融合的
DDPG 算法模型，包括：
将通过预设的监督学习模型计算得到的交叉熵损失值加入到
DDPG 算法的 Actor 中，评估 Actor 的输出值。
- 30 6. 根据权利要求 1 所述的方法，其中，根据所述统计规则模型，

- 基于待处理的⁵用户行为数据和待处理的¹⁰用户属性信息得到用户画像，
包括：
获取预设的深度强化学习模型和业务规则模型，通过待处理的¹⁵用户行为数据和待处理的²⁰用户属性信息分别得到相应的第一用户画像和²⁵第二用户画像；
根据目标运营效果指标，确定第一用户画像和第二用户画像的权重，以将第一用户画像和第二用户画像融合得到最终用户画像。
7. 根据权利要求 6 所述的方法，还包括：
所述的深度强化学习模型采用 Actor-Critic 算法。
8. 一种基于用户画像的物品推荐装置，其包括：
获取模块，用于接收用户行为数据和用户属性信息，通过特征工程转换成待处理的¹⁵用户行为数据和待处理的²⁰用户属性信息；
处理模块，用于获取当前标签计算任务，判断所述标签计算任务是否属于预测类任务，若是则调用预设的预测模型，若否则调用预设的统计规则模型；根据所述预测模型或所述统计规则模型，基于待处理的²⁵用户行为数据和待处理的³⁰用户属性信息得到用户画像，进而根据用户画像向该用户推送物品信息。
9. 一种电子设备，其包括：
一个或多个处理器；
存储装置，用于存储一个或多个程序，
当所述一个或多个程序被所述一个或多个处理器执行，使得所述²⁵一个或多个处理器实现如权利要求 1-7 中任一所述的方法。
10. 一种计算机可读介质，其上存储有计算机程序，所述程序被处理器执行时实现如权利要求 1-7 中任一所述的方法。