



(12) 发明专利申请

(10) 申请公布号 CN 111767466 A

(43) 申请公布日 2020.10.13

(21) 申请号 202010905089.1

(22) 申请日 2020.09.01

(71) 申请人 腾讯科技(深圳)有限公司

地址 518000 广东省深圳市南山区高新区  
科技中一路腾讯大厦35层

(72) 发明人 严超

(74) 专利代理机构 北京派特恩知识产权代理有  
限公司 11270

代理人 赵翠萍 张颖玲

(51) Int. Cl.

G06F 16/9535 (2019.01)

G06N 20/00 (2019.01)

G06Q 30/02 (2012.01)

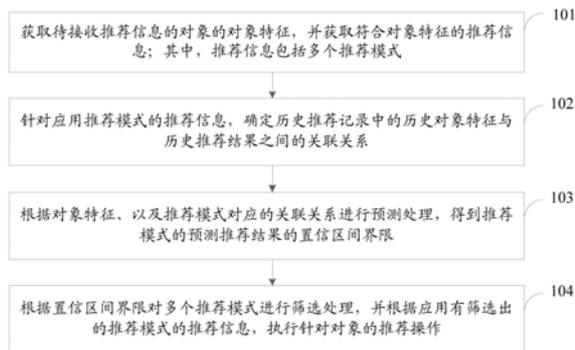
权利要求书4页 说明书24页 附图8页

(54) 发明名称

基于人工智能的推荐信息推荐方法、装置及  
电子设备

(57) 摘要

本申请提供了一种基于人工智能的推荐信息推荐方法、装置、电子设备及计算机可读存储介质;涉及大数据技术领域的精准推荐;方法包括:获取待接收推荐信息的对象的对象特征,并获取符合对象特征的推荐信息;其中,推荐信息包括多个推荐模式;针对应用推荐模式的推荐信息,确定历史推荐记录中的历史对象特征与历史推荐结果之间的关联关系;根据对象特征、以及推荐模式对应的关联关系进行预测处理,得到推荐模式的预测推荐结果的置信区间界限;根据置信区间界限对多个推荐模式进行筛选处理,并根据应用有筛选出的推荐模式的推荐信息,执行针对对象的推荐操作。通过本申请,能够提升推荐的针对性和精准性。



1. 一种基于人工智能的推荐信息推荐方法,其特征在于,包括:

获取待接收推荐信息的对象的对象特征,并获取符合所述对象特征的推荐信息;其中,所述推荐信息包括多个推荐模式;

针对应用所述推荐模式的所述推荐信息,确定历史推荐记录中的历史对象特征与历史推荐结果之间的关联关系;

根据所述对象特征、以及所述推荐模式对应的关联关系进行预测处理,得到所述推荐模式的预测推荐结果的置信区间界限;

根据所述置信区间界限对多个所述推荐模式进行筛选处理,并

根据应用有筛选出的推荐模式的所述推荐信息,执行针对所述对象的推荐操作。

2. 根据权利要求1所述的推荐信息推荐方法,其特征在于,所述根据所述置信区间界限对多个所述推荐模式进行筛选处理,包括:

在多个所述推荐模式的置信区间界限的降序排序中,将排序在前的第一设定数量的推荐模式,确定为候选推荐模式;其中,所述第一设定数量为大于1的整数;

根据所述推荐信息的内容特征、展示位置特征、所述对象特征、以及所述候选推荐模式的素材特征进行预测处理,并

对得到的预测触发率和预测转化率进行融合处理,得到第一评分;

根据所述第一评分,对所述候选推荐模式进行筛选处理,得到一个推荐模式以作为筛选出的推荐模式。

3. 根据权利要求2所述的推荐信息推荐方法,其特征在于,所述根据所述第一评分,对所述候选推荐模式进行筛选处理,得到一个推荐模式以作为筛选出的推荐模式,包括:

将所述候选推荐模式中曝光量小于或等于曝光量阈值的推荐模式,添加至第一集合,并

将所述候选推荐模式中曝光量大于所述曝光量阈值的推荐模式,添加至第二集合;

根据选择概率对所述第一集合和所述第二集合进行选择处理;其中,所述第一集合的选择概率与所述第二集合的选择概率的加和为1;

当选择的所述第一集合为非空集合、或者所述第二集合为空集合时,对所述第一集合内的推荐模式的第一评分进行归一化处理,得到归一化概率,并

根据所述归一化概率,对所述第一集合内的推荐模式进行选择处理,得到一个推荐模式以作为筛选出的推荐模式;

当选择的所述第二集合为非空集合、或者所述第一集合为空集合时,将所述第二集合内的、第一评分最高的推荐模式,作为筛选出的推荐模式。

4. 根据权利要求2所述的推荐信息推荐方法,其特征在于,所述根据应用有筛选出的推荐模式的所述推荐信息,执行针对所述对象的推荐操作,包括:

在多个所述推荐信息的、筛选出的推荐模式的第一评分的降序排序中,将排序在前的第二设定数量的推荐信息,确定为目标推荐信息;其中,所述第二设定数量为大于0的整数;

根据应用有筛选出的推荐模式的所述目标推荐信息,执行针对所述对象的推荐操作。

5. 根据权利要求2所述的推荐信息推荐方法,其特征在于,还包括:

通过人工智能模型,对样本推荐信息在历史推荐记录中的内容特征、展示位置特征、历史对象特征以及应用的推荐模式的素材特征进行预测处理,得到预测结果;

根据所述预测结果与历史推荐结果之间的差异,更新所述人工智能模型的权重参数;

其中,所述预测结果为预测触发率及预测转化率中的任意一种;更新后的所述人工智能模型用于对所述推荐信息的内容特征、展示位置特征、所述对象特征、以及所述候选推荐模式的素材特征进行预测处理。

6. 根据权利要求1所述的推荐信息推荐方法,其特征在于,所述确定历史推荐记录中的历史对象特征与历史推荐结果之间的关联关系,包括:

初始化所述推荐模式对应的第一关联参数和第二关联参数;

将所述历史推荐记录中的历史对象特征与所述历史对象特征的转置进行乘积处理,并将乘积结果与所述第一关联参数进行求和处理,以根据求和结果更新所述第一关联参数;

将所述历史推荐记录中的历史对象特征与历史推荐结果进行乘积处理,并将乘积结果与所述第二关联参数进行求和处理,以根据求和结果更新所述第二关联参数;

将更新后的所述第一关联参数的逆矩阵、与更新后的所述第二关联参数进行乘积处理,得到第三关联参数;

其中,更新后的所述第一关联参数、以及所述第三关联参数用于表示所述关联关系。

7. 根据权利要求6所述的推荐信息推荐方法,其特征在于,所述根据所述对象特征、以及所述推荐模式对应的关联关系进行预测处理,得到所述推荐模式的预测推荐结果的置信区间界限,包括:

将所述对象特征的转置、与所述第三关联参数进行乘积处理,得到所述推荐模式的预测推荐结果;

将所述对象特征、所述对象特征的转置、与更新后的所述第一关联参数的逆矩阵进行乘积处理,并

将乘积结果的平方根与设定系数进行乘积处理,得到所述预测推荐结果的置信区间宽度;

将所述预测推荐结果与所述置信区间宽度进行求和处理,得到置信区间界限。

8. 根据权利要求6所述的推荐信息推荐方法,其特征在于,所述将所述历史推荐记录中的历史对象特征与所述历史对象特征的转置进行乘积处理之前,还包括:

对多个所述历史对象特征进行聚类处理,得到多个中心向量;

根据多个所述中心向量,对所述历史推荐记录中的历史对象特征进行更新处理。

9. 根据权利要求8所述的推荐信息推荐方法,其特征在于,所述根据多个所述中心向量,对所述历史推荐记录中的历史对象特征进行更新处理,包括:

执行以下任意一种处理:

确定所述历史推荐记录中的历史对象特征、与每个所述中心向量之间的相似度,并根据得到的多个相似度,构建更新后的所述历史对象特征;

确定所述历史推荐记录中的历史对象特征在每个所述中心向量上的投影分量,并以所述投影分量作为权重,对多个所述中心向量进行加权处理,得到更新后的所述历史对象特征。

10. 根据权利要求1所述的推荐信息推荐方法,其特征在于,

所述确定历史推荐记录中的历史对象特征与历史推荐结果之间的关联关系,包括:

将所述历史推荐记录中的历史对象特征与展示位置特征进行拼接处理,得到第一拼接特征,并

确定所述第一拼接特征与历史推荐结果之间的关联关系;

所述根据所述对象特征、以及所述推荐模式对应的关联关系进行预测处理,得到所述推荐模式的预测推荐结果的置信区间界限,包括:

将所述对象特征与所述推荐信息的展示位置特征进行拼接处理,得到第二拼接特征,并

根据所述第二拼接特征、以及所述推荐模式对应的关联关系进行预测处理,得到所述推荐模式的预测推荐结果的置信区间界限。

11. 根据权利要求1至10任一项所述的推荐信息推荐方法,其特征在于,

所述获取符合所述对象特征的推荐信息之前,还包括:

针对数据库中的每个推荐信息,根据所述推荐信息的历史推荐记录确定所述推荐信息的触发率及转化率;

将所述推荐信息的触发率及转化率进行融合处理,得到所述推荐信息的第二评分;

根据所述第二评分对多个所述推荐信息进行筛选处理;

所述获取符合所述对象特征的推荐信息,包括:

获取筛选出的推荐信息的定向推荐特征;

当所述定向推荐特征与所述对象特征匹配成功时,将所述定向推荐特征对应的推荐信息,确定为符合所述对象特征的推荐信息。

12. 根据权利要求1至10任一项所述的推荐信息推荐方法,其特征在于,所述针对应用所述推荐模式的所述推荐信息,确定历史推荐记录中的历史对象特征与历史推荐结果之间的关联关系之前,还包括:

根据所述对象特征和所述推荐信息的内容特征进行预测处理,并

对得到的预测触发率和预测转化率进行融合处理,得到第三评分;

根据所述第三评分,对多个所述推荐信息进行筛选处理。

13. 一种基于人工智能的推荐信息推荐装置,其特征在于,包括:

获取模块,用于获取待接收推荐信息的对象的对象特征,并获取符合所述对象特征的推荐信息;其中,所述推荐信息包括多个推荐模式;

关系确定模块,用于针对应用所述推荐模式的所述推荐信息,确定历史推荐记录中的历史对象特征与历史推荐结果之间的关联关系;

预测模块,用于根据所述对象特征、以及所述推荐模式对应的关联关系进行预测处理,得到所述推荐模式的预测推荐结果的置信区间界限;

筛选模块,用于根据所述置信区间界限对多个所述推荐模式进行筛选处理,并

根据应用有筛选出的推荐模式的所述推荐信息,执行针对所述对象的推荐操作。

14. 一种电子设备,其特征在于,包括:

存储器,用于存储可执行指令;

处理器,用于执行所述存储器中存储的可执行指令时,实现权利要求1至12任一项所述的基于人工智能的推荐信息推荐方法。

15. 一种计算机可读存储介质,其特征在于,存储有可执行指令,用于被处理器执行时,实现权利要求1至12任一项所述的基于人工智能的推荐信息推荐方法。

## 基于人工智能的推荐信息推荐方法、装置及电子设备

### 技术领域

[0001] 本申请涉及人工智能和大数据技术,尤其涉及一种基于人工智能的推荐信息推荐方法、装置、电子设备及计算机可读存储介质。

### 背景技术

[0002] 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。

[0003] 智能推荐是人工智能的一个重要分支,也涉及到云技术的大数据处理,主要研究针对一个特定的对象,如何将合适的推荐信息推荐至该对象。对于包括多个推荐模式的推荐信息,如包括多个创意的广告来说,在相关技术提供的方案中,通常是应用随机策略,即一个推荐信息的所有推荐模式按照均等概率进行推荐。但是,该方案容易导致推荐效果不佳,推荐资源的利用率低。

### 发明内容

[0004] 本申请实施例提供一种基于人工智能的推荐信息推荐方法、装置、电子设备及计算机可读存储介质,能够优化推荐效果,提升推荐资源的利用率。

[0005] 本申请实施例的技术方案是这样实现的:

本申请实施例提供一种基于人工智能的推荐信息推荐方法,包括:

获取待接收推荐信息的对象的对象特征,并获取符合所述对象特征的推荐信息;其中,所述推荐信息包括多个推荐模式;

针对应用所述推荐模式的所述推荐信息,确定历史推荐记录中的历史对象特征与历史推荐结果之间的关联关系;

根据所述对象特征、以及所述推荐模式对应的关联关系进行预测处理,得到所述推荐模式的预测推荐结果的置信区间界限;

根据所述置信区间界限对多个所述推荐模式进行筛选处理,并

根据应用有筛选出的推荐模式的所述推荐信息,执行针对所述对象的推荐操作。

[0006] 在上述方案中,所述根据所述置信区间界限对多个所述推荐模式进行筛选处理,包括:

将多个所述推荐模式中置信区间界限最大的一个推荐模式,作为筛选出的推荐模式。

[0007] 在上述方案中,还包括:执行以下任意一种处理:

获取与推荐信息的多个待填充部分分别对应的一个素材,并

将每个所述素材填充至对应的所述待填充部分中,得到所述推荐信息的一个推荐模式;

获取与推荐信息的多个待填充部分分别对应的多个素材,并

在每个所述待填充部分对应的多个素材中进行随机选择,将选择的素材填充至对应的所述待填充部分中,得到所述推荐信息的一个推荐模式。

[0008] 在上述方案中,所述历史推荐结果的类型包括:

第一推荐结果,用于表示应用所述推荐模式的所述推荐信息已曝光、且未触发;

第二推荐结果,用于表示应用所述推荐模式的所述推荐信息已触发、且未转化;

第三推荐结果,用于表示应用所述推荐模式的所述推荐信息已转化。

[0009] 本申请实施例提供一种基于人工智能的推荐信息推荐装置,包括:

获取模块,用于获取待接收推荐信息的对象的对象特征,并获取符合所述对象特征的推荐信息;其中,所述推荐信息包括多个推荐模式;

关系确定模块,用于针对应用所述推荐模式的所述推荐信息,确定历史推荐记录中的历史对象特征与历史推荐结果之间的关联关系;

预测模块,用于根据所述对象特征、以及所述推荐模式对应的关联关系进行预测处理,得到所述推荐模式的预测推荐结果的置信区间界限;

筛选模块,用于根据所述置信区间界限对多个所述推荐模式进行筛选处理,并

根据应用有筛选出的推荐模式的所述推荐信息,执行针对所述对象的推荐操作。

[0010] 本申请实施例提供一种电子设备,包括:

存储器,用于存储可执行指令;

处理器,用于执行所述存储器中存储的可执行指令时,实现本申请实施例提供的基于人工智能的推荐信息推荐方法。

[0011] 本申请实施例提供一种计算机可读存储介质,存储有可执行指令,用于引起处理器执行时,实现本申请实施例提供的基于人工智能的推荐信息推荐方法。

[0012] 本申请实施例具有以下有益效果:

根据历史对象特征与历史推荐结果之间的关联关系、以及待接收推荐信息的对象的对象特征进行预测处理,得到推荐模式的预测推荐结果的置信区间界限,进而根据置信区间界限对多个推荐模式进行筛选处理,如此,通过引入对象特征,加强了推荐的针对性和精准性,使得最终筛选出的推荐模式能够更加符合对象的需求,提升了推荐资源的利用率。

## 附图说明

[0013] 图1是本申请实施例提供的基于人工智能的推荐信息推荐系统的一个可选的架构示意图;

图2是本申请实施例提供的终端设备的一个可选的架构示意图;

图3A是本申请实施例提供的基于人工智能的推荐信息推荐方法的一个可选的流程示意图;

图3B是本申请实施例提供的基于人工智能的推荐信息推荐方法的一个可选的流程示意图;

图3C是本申请实施例提供的基于人工智能的推荐信息推荐方法的一个可选的流程示意图;

图3D是本申请实施例提供的基于人工智能的推荐信息推荐方法的一个可选的流程示意图;

图4是本申请实施例提供的上传素材的一个可选的界面示意图；  
图5是本申请实施例提供的上传素材的一个可选的界面示意图；  
图6是本申请实施例提供的推荐效果数据的一个可选的示意图；  
图7是本申请实施例提供的广告推荐系统的一个可选的架构示意图；  
图8是本申请实施例提供的聚类处理的一个可选的流程示意图；  
图9是本申请实施例提供的广告推荐系统的一个可选的架构示意图；  
图10是本申请实施例提供的创意筛选的一个可选的示意图；  
图11是本申请实施例提供的粗排和精排的一个可选的示意图。

### 具体实施方式

[0014] 为了使本申请的目的、技术方案和优点更加清楚，下面将结合附图对本申请作进一步地详细描述，所描述的实施例不应视为对本申请的限制，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其它实施例，都属于本申请保护的范围。

[0015] 在以下的描述中，涉及到“一些实施例”，其描述了所有可能实施例的子集，但是可以理解，“一些实施例”可以是所有可能实施例的相同子集或不同子集，并且可以在不冲突的情况下相互结合。

[0016] 在以下的描述中，所涉及的术语“第一\第二\第三”仅仅是是区别类似的对象，不代表针对对象的特定排序，可以理解地，“第一\第二\第三”在允许的情况下可以互换特定的顺序或先后次序，以使这里描述的本申请实施例能够以除了在这里图示或描述的以外的顺序实施。在以下的描述中，所涉及的术语“多个”是指至少两个。

[0017] 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。本文中所使用的术语只是为了描述本申请实施例的目的，不是旨在限制本申请。

[0018] 对本申请实施例进行进一步详细说明之前，对本申请实施例中涉及的名词和术语进行说明，本申请实施例中涉及的名词和术语适用于如下的解释。

[0019] 1) 推荐信息：用于推荐特定的内容，如商品、音乐或游戏等。在本申请实施例中，推荐信息可以是广告，但并不限于广告的形式。

[0020] 2) 对象：用于接收推荐信息，对象可以是某个实际用户，也可以是某个虚拟账号，例如在电商平台中注册的用户账号。

[0021] 3) 推荐模式：指推荐信息展现时应用的展现形式，对于一个推荐信息来说，其可能包括多个推荐模式。例如，推荐信息的推荐模式可以包括文字模式、图文模式及视频模式等，对于文字模式来说，还可根据所添加的文字的不同，扩展出多种文字模式，以此类推。在多创意(Multi Creative, MC)广告中，一个广告包括多个创意，其中，创意是指为了达到广告目的，对广告主题、内容和表现形式所进行的创造性表达和构想，创意即对应上文的推荐模式。

[0022] 4) 推荐结果：将应用某个推荐模式的推荐信息推荐至对象后，接收到的反馈结果，在本申请实施例中，推荐结果可包括曝光且未触发、触发且未转化、以及已转化三种类型。其中，曝光是指推荐信息已进行展示；触发是指接收到对象执行的触发操作，例如接收到对象点击某个商品的广告的操作，跳转至该商品的详情页面；转化是指在接收到触发操作的

基础上,接收到与推荐信息对应的、设定的转化操作,例如,在该商品的详情页面中,接收到对象将该商品加入购物车并进行支付的操作。

[0023] 5) 置信区间 (Confidence Interval): 包括预测推荐结果的一个数值区间,真值 (实际值) 落入该数值区间的概率大于设定的概率阈值。置信区间界限是指置信区间的上限或下限。

[0024] 6) 人工智能模型: 基于人工智能原理构建的模型,本申请实施例对人工智能模型的类型不做限定,例如可以是神经网络模型。

[0025] 7) 数据库 (Database): 以一定方式储存在一起、能与多个用户共享、具有尽可能小的冗余度、与应用程序彼此独立的数据集合,用户可以对数据库中的数据执行新增、查询、更新及删除等操作。

[0026] 对于推荐信息的多个推荐模式的筛选,相关技术提供了随机策略的方案,即一个推荐信息的多个推荐模式按照均等概率进行推荐,但是,在应用随机策略时,会导致在劣质的推荐模式上浪费较多曝光机会,推荐资源的利用率低,推荐信息的推荐效果也不佳。

[0027] 本申请实施例提供一种基于人工智能的推荐信息推荐方法、装置、电子设备和计算机可读存储介质,能够优化推荐效果,提升推荐资源的利用率。下面说明本申请实施例提供的电子设备的示例性应用,本申请实施例提供的电子设备可以实施为笔记本电脑,平板电脑,台式计算机,机顶盒,移动设备(例如,移动电话,便携式音乐播放器,个人数字助理,专用消息设备,便携式游戏设备)等各种类型的终端设备,也可以实施为服务器。

[0028] 电子设备通过运行本申请实施例提供的推荐方案,能够筛选出效果最好的推荐模式并执行推荐操作,加强推荐资源的利用率,即提高电子设备自身的推荐性能,适用于多种推荐场景。例如,电子设备可以是电商平台的服务器,电商平台中某个商品的广告包括多个创意,通过将应用有筛选出的创意的广告,推荐至电商平台的用户账号,提升该用户账号根据该广告购买该商品的概率,即提升触发率和转化率;又例如,电子设备可以是终端设备,其上运行有办公软件或游戏软件,推荐信息是办公软件或游戏软件的教程,包括多个推荐模式(例如文字模式、图文模式及视频模式等),电子设备通过将应用有筛选出的推荐模式的教程,显示于办公软件或游戏软件的界面中,提升用户根据教程进行学习的概率,便于用户快速上手,即提升触发率和转化率。

[0029] 参见图1,图1是本申请实施例提供的基于人工智能的推荐信息推荐系统100的一个可选的架构示意图,终端设备400通过网络300连接服务器200,服务器200连接数据库500,其中,网络300可以是广域网或者局域网,又或者是二者的组合。

[0030] 在一些实施例中,以电子设备是终端设备为例,本申请实施例提供的基于人工智能的推荐信息推荐方法可以由终端设备实现。例如,终端设备400在获取到应用推荐模式的推荐信息的历史推荐记录后,确定历史推荐记录中的历史对象特征与历史推荐结果之间的关联关系,并根据推荐模式对应的关联关系、以及待接收推荐信息的对象的对象特征进行预测处理,得到推荐模式的预测推荐结果的置信区间界限。如此,终端设备400可根据得到的置信区间界限,对多个推荐模式进行筛选处理,并将应用有筛选出的推荐模式的推荐信息,呈现于图形界面410中。其中,历史推荐记录可以预先存储在终端设备400本地,也可以是终端设备400向外界(如数据库500)获取到的,包括多个推荐模式的推荐信息同理。

[0031] 在一些实施例中,以电子设备是服务器为例,本申请实施例提供的基于人工智能

的推荐信息推荐方法,也可以由服务器实现。例如,服务器200从数据库500中获取包括多个推荐模式的推荐信息、以及各个推荐模式对应的历史推荐记录,根据历史推荐记录确定各个推荐模式对应的关联关系。然后,服务器200根据推荐模式对应的关联关系、以及从终端设备400获取到的对象特征进行预测处理,得到推荐模式的预测推荐结果的置信区间界限。如此,服务器200可根据得到的置信区间界限对多个推荐模式进行筛选处理,并将应用有筛选出的推荐模式的推荐信息,发送至终端设备400。值得说明的是,本申请实施例对推荐信息以及历史推荐记录的存储位置不做限定,例如可以是数据库500、服务器200的分布式文件系统或者区块链等位置。

[0032] 终端设备400用于在图形界面410中,显示推荐过程中的各种结果和最终结果。在图1中,示例性地示出了应用有筛选出的推荐模式的推荐信息,其包括图片1及文案1。

[0033] 在一些实施例中,服务器200可以是独立的物理服务器,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,其中,云服务可以是推荐服务,供终端设备400进行调用,以根据终端设备400发送的对象特征,将应用有筛选出的推荐模式的推荐信息,发送至终端设备400。终端设备400可以是智能手机、平板电脑、笔记本电脑、台式计算机、智能音箱及智能手表等,但并不局限于此。终端设备以及服务器可以通过有线或无线通信方式进行直接或间接地连接,本申请实施例中不做限制。

[0034] 以本申请实施例提供的电子设备是终端设备为例说明,可以理解的,对于电子设备是服务器的情况,图2中示出的结构中的部分(例如用户接口、呈现模块和输入处理模块)可以缺省。参见图2,图2是本申请实施例提供的终端设备400的结构示意图,图2所示的终端设备400包括:至少一个处理器410、存储器450、至少一个网络接口420和用户接口430。终端设备400中的各个组件通过总线系统440耦合在一起。可理解,总线系统440用于实现这些组件之间的连接通信。总线系统440除包括数据总线之外,还包括电源总线、控制总线和状态信号总线。但是为了清楚说明起见,在图2中将各种总线都标为总线系统440。

[0035] 处理器410可以是一种集成电路芯片,具有信号的处理能力,例如通用处理器、数字信号处理器(DSP, Digital Signal Processor),或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等,其中,通用处理器可以是微处理器或者任何常规的处理器等。

[0036] 用户接口430包括使得能够呈现媒体内容的一个或多个输出装置431,包括一个或多个扬声器和/或一个或多个视觉显示屏。用户接口430还包括一个或多个输入装置432,包括有助于用户输入的用户接口部件,比如键盘、鼠标、麦克风、触屏显示屏、摄像头、其他输入按钮和控件。

[0037] 存储器450可以是可移除的,不可移除的或其组合。示例性的硬件设备包括固态存储器,硬盘驱动器,光盘驱动器等。存储器450可选地包括在物理位置上远离处理器410的一个或多个存储设备。

[0038] 存储器450包括易失性存储器或非易失性存储器,也可包括易失性和非易失性存储器两者。非易失性存储器可以是只读存储器(ROM, Read Only Memory),易失性存储器可以是随机存取存储器(RAM, Random Access Memory)。本申请实施例描述的存储器450旨在

包括任意适合类型的存储器。

[0039] 在一些实施例中,存储器450能够存储数据以支持各种操作,这些数据的示例包括程序、模块和数据结构或者其子集或超集,下面示例性说明。

[0040] 操作系统451,包括用于处理各种基本系统服务和执行硬件相关任务的系统程序,例如框架层、核心库层、驱动层等,用于实现各种基础业务以及处理基于硬件的任务;

网络通信模块452,用于经由一个或多个(有线或无线)网络接口420到达其他计算设备,示例性的网络接口420包括:蓝牙、无线相容性认证(WiFi)、和通用串行总线(USB, Universal Serial Bus)等;

呈现模块453,用于经由一个或多个与用户接口430相关联的输出装置431(例如,显示屏、扬声器等)使得能够呈现信息(例如,用于操作外围设备和显示内容和信息的用户接口);

输入处理模块454,用于对一个或多个来自一个或多个输入装置432之一的一个或多个用户输入或互动进行检测以及翻译所检测的输入或互动。

[0041] 在一些实施例中,本申请实施例提供的装置可以采用软件方式实现,图2示出了存储在存储器450中的基于人工智能的推荐信息推荐装置455,其可以是程序和插件等形式的软件,包括以下软件模块:获取模块4551、关系确定模块4552、预测模块4553及筛选模块4554,这些模块是逻辑上的,因此根据所实现的功能可以进行任意的组合或进一步拆分。将在下文中说明各个模块的功能。

[0042] 在另一些实施例中,本申请实施例提供的装置可以采用硬件方式实现,作为示例,本申请实施例提供的基于人工智能的推荐信息推荐装置可以是采用硬件译码处理器形式的处理器,其被编程以执行本申请实施例提供的基于人工智能的推荐信息推荐方法,例如,硬件译码处理器形式的处理器可以采用一个或多个应用专用集成电路(ASIC, Application Specific Integrated Circuit)、DSP、可编程逻辑器件(PLD, Programmable Logic Device)、复杂可编程逻辑器件(CPLD, Complex Programmable Logic Device)、现场可编程门阵列(FPGA, Field-Programmable Gate Array)或其他电子元件。

[0043] 将结合本申请实施例提供的电子设备的示例性应用和实施,说明本申请实施例提供的基于人工智能的推荐信息推荐方法。

[0044] 参见图3A,图3A是本申请实施例提供的基于人工智能的推荐信息推荐方法的一个可选的流程示意图,将结合图3A示出的步骤进行说明。

[0045] 在步骤101中,获取待接收推荐信息的对象的对象特征,并获取符合对象特征的推荐信息;其中,推荐信息包括多个推荐模式。

[0046] 这里,对象特征用于描述待接收推荐信息的对象的情况,以对象为用户账号的情况举例,则对象特征包括但不限于持有用户账号的用户的年龄、性别、居住城市、婚恋状况、消费水平、教育程度及工作状态。在获取到对象特征之后,获取符合对象特征的推荐信息,例如在数据库的多个推荐信息中,查找出符合对象特征的推荐信息,其中,推荐信息包括多个推荐模式。

[0047] 在一些实施例中,在步骤101之前,还包括:执行以下任意一种处理:获取与推荐信息的多个待填充部分分别对应的一个素材,并将每个素材填充至对应的待填充部分中,得到推荐信息的一个推荐模式;获取与推荐信息的多个待填充部分分别对应的多个素材,并

在每个待填充部分对应的多个素材中进行随机选择,将选择的素材填充至对应的待填充部分中,得到推荐信息的一个推荐模式。

[0048] 本申请实施例提供了两种方式,以得到推荐信息。第一种方式是,获取与推荐信息的多个待填充部分分别对应的一个素材,并将每个素材填充至对应的待填充部分中,得到推荐信息的一个推荐模式。以推荐信息为广告的情况举例,广告的待填充部分包括图片部分和文案部分,针对图片部分,获取到广告主上传的图片素材1;针对文案部分,获取到广告主上传的文案素材1。然后,将图片素材1填充至图片部分,将文案素材1填充至文案部分,得到广告的一个创意。广告主可以通过上传素材,继续制作广告的下一个创意。

[0049] 第二种方式是,获取与推荐信息的多个待填充部分分别对应的多个素材,并在每个待填充部分对应的多个素材中进行随机选择,将选择的素材填充至对应的待填充部分中,得到推荐信息的一个推荐模式。例如,针对某广告的图片部分,获取到广告主上传的图片素材1和图片素材2;针对该广告的文案部分,获取到广告主上传的文案素材1和文案素材2。然后,将随机选择出的图片素材1填充至图片部分,将随机选择出的文案素材1填充至文案部分。值得说明的是,这里可以穷尽不同待填充部分对应的素材的组合方式,直至得到所有可能出现的推荐模式,例如得到(图片素材1,文案素材1)、(图片素材1,文案素材2)、(图片素材2,文案素材1)以及(图片素材2,文案素材2)这4个推荐模式,如此,可以增加得到的推荐模式的数量。通过上述两种方式,提升了生成推荐信息的灵活性。

[0050] 在步骤102中,针对应用推荐模式的推荐信息,确定历史推荐记录中的历史对象特征与历史推荐结果之间的关联关系。

[0051] 这里,历史推荐记录包括推荐信息应用的推荐模式、执行的推荐操作针对的历史对象的历史对象特征以及历史推荐结果。例如,推荐信息包括推荐模式1和推荐模式2,在数据库中总共获取到推荐信息的20条历史推荐记录,其中,前10条历史推荐记录中应用的均是推荐模式1,后10条历史推荐记录中应用的均是推荐模式2。则确定前10条历史推荐记录中的历史对象特征与历史推荐结果之间的关联关系,作为推荐模式1对应的关联关系;确定后10条历史推荐记录中的历史对象特征与历史推荐结果之间的关联关系,作为推荐模式2对应的关联关系。本申请实施例对关联关系的确定方式不做限定,例如可通过人工智能模型来确定历史对象特征与历史推荐结果之间的关联关系。

[0052] 在一些实施例中,历史推荐结果的类型包括:第一推荐结果,用于表示应用推荐模式的推荐信息已曝光、且未触发;第二推荐结果,用于表示应用推荐模式的推荐信息已触发、且未转化;第三推荐结果,用于表示应用推荐模式的推荐信息已转化。

[0053] 在本申请实施例中,可以预先设定推荐信息对应的触发操作及转化操作,例如,对于某个商品的广告来说,设定对应的触发操作是对该广告的点击操作,对应的转化操作是对该商品的付费操作。如此,可以将推荐结果(如历史推荐结果)分为第一推荐结果、第二推荐结果和第三推荐结果三类,其中,第一推荐结果用于表示应用推荐模式的推荐信息已经进行了显示,但未接收到对应的触发操作;第二推荐结果用于表示已接收到与应用推荐模式的推荐信息对应的触发操作,但未接收到对应的转化操作;第三推荐结果用于表示已接收到与应用推荐模式的推荐信息对应的转化操作。当然,这并不构成对本申请实施例的限定,例如可以将推荐结果划分为第一推荐结果和第二推荐结果两类。此外,为了便于确定关联关系,可以对推荐结果进行数值化,例如将第一推荐结果数值化为0,将第二推荐结果数

值化为1,将第三推荐结果数值化为2,其中,在转化操作可能涉及到多个转化量(例如用户购买了n个商品,那么转化量就是n)的情况下,可以将第三推荐结果数值化为1+转化量。通过上述的分类方式,能够提升确定关联关系、以及基于关联关系进行预测处理的准确性。

[0054] 在步骤103中,根据对象特征、以及推荐模式对应的关联关系进行预测处理,得到推荐模式的预测推荐结果的置信区间界限。

[0055] 在通过步骤102确定出推荐信息的每个推荐模式对应的关联关系后,针对每个推荐模式,根据待接收推荐信息的对象的对象特征、以及推荐模式对应的关联关系进行预测处理,得到该推荐模式的预测推荐结果的置信区间界限。其中,置信区间界限可以是预测推荐结果的置信区间的上限或下限。

[0056] 在一些实施例中,可以通过这样的方式来实现上述的确定历史推荐记录中的历史对象特征与历史推荐结果之间的关联关系:将历史推荐记录中的历史对象特征与展示位置特征进行拼接处理,得到第一拼接特征,并确定第一拼接特征与历史推荐结果之间的关联关系;可以通过这样的方式来实现上述的根据对象特征、以及推荐模式对应的关联关系进行预测处理,得到推荐模式的预测推荐结果的置信区间界限:将对象特征与推荐信息的展示位置特征进行拼接处理,得到第二拼接特征,并根据第二拼接特征、以及推荐模式对应的关联关系进行预测处理,得到推荐模式的预测推荐结果的置信区间界限。

[0057] 除了对象特征之外,在本申请实施例中,还可引入展示位置特征。例如,针对应用有推荐模式的推荐信息,将历史推荐记录中的历史对象特征与展示位置特征进行拼接处理,为了便于区分,将得到的拼接特征命名为第一拼接特征,然后,确定第一拼接特征与历史推荐结果之间的关联关系,作为该推荐模式对应的关联关系。

[0058] 得到推荐模式对应的关联关系后,将待接收推荐信息的对象的对象特征、与推荐信息的展示位置特征进行拼接处理,得到第二拼接特征,并根据第二拼接特征、以及推荐模式对应的关联关系进行预测处理,得到推荐模式的预测推荐结果的置信区间界限。由于展示位置与推荐结果存在一定的关联,故通过引入展示位置特征,能够提升确定出的关联关系以及置信区间界限的准确性。

[0059] 在步骤104中,根据置信区间界限对多个推荐模式进行筛选处理,并根据应用有筛选出的推荐模式的推荐信息,执行针对对象的推荐操作。

[0060] 对于推荐信息包括的每个推荐模式来说,都可得到一个置信区间界限,这里,根据置信区间界限对多个推荐模式进行筛选处理,得到一个筛选出的推荐模式,其中,筛选处理倾向于筛选出置信区间界限更大的推荐模式。然后,根据应用有筛选出的推荐模式的推荐信息,执行针对对象的推荐操作,本申请实施例对推荐操作并不做限定,例如可以呈现于对象当前打开的页面中,又例如以短信、邮件或即时消息等方式推荐给对象。

[0061] 在一些实施例中,可以通过这样的方式来实现上述的根据置信区间界限对多个推荐模式进行筛选处理:将多个推荐模式中置信区间界限最大的一个推荐模式,作为筛选出的推荐模式。

[0062] 本申请实施例提供了筛选处理的一种示例,即可以在推荐信息的多个推荐模式中,将置信区间界限最大的一个推荐模式,确定为筛选出的推荐模式。

[0063] 如图3A所示,本申请实施例通过引入对象特征,并通过计算出的置信区间界限对多个推荐模式进行筛选处理,使得最终筛选出的推荐模式能够更加符合对象的实际需求,

加强了推荐过程中的针对性和精准性,提升了推荐资源的利用率。

[0064] 在一些实施例中,参见图3B,图3B是本申请实施例提供的基于人工智能的推荐信息推荐方法的一个可选的流程示意图,图3A示出的步骤104可以通过步骤201至步骤204实现,将结合各步骤进行说明。

[0065] 在步骤201中,在多个推荐模式的置信区间界限的降序排序中,将排序在前的第一设定数量的推荐模式,确定为候选推荐模式;其中,第一设定数量为大于1的整数。

[0066] 除了将置信区间界限最大的一个推荐模式,确定为筛选出的推荐模式之外,本申请实施例还提供了筛选处理的另一种方式。首先,对推荐信息包括的多个推荐模式的置信区间界限进行降序排序,并确定出排序在前的第一设定数量的推荐模式,为了便于区分,将确定出的推荐模式命名为候选推荐模式。其中,第一设定数量为大于1的整数,可根据实际应用场景进行设定,如设定为5。然后,再对多个候选推荐模式进行进一步精选。

[0067] 在步骤202中,根据推荐信息的内容特征、展示位置特征、对象特征、以及候选推荐模式的素材特征进行预测处理,并对得到的预测触发率和预测转化率进行融合处理,得到第一评分。

[0068] 这里,获取推荐信息所推荐内容的内容特征,以推荐信息为某商品的广告进行举例,则内容特征包括但不限于该商品的种类、浏览量、购买量、好评数量及差评数量;同时,获取推荐信息对应的展示位置的展示位置特征,例如展示位置可以是广告预期展示的广告位,展示位置特征可以包括广告位所在的页面的ID以及广告位在页面中的具体位置的描述参数;此外,还获取候选推荐模式中所用素材的素材特征,例如可以包括所用图片的尺寸及所用文案的文本量等。当然,这并不构成对本申请实施例的限定,根据实际应用场景的不同,可对内容特征、展示位置特征以及素材特征进行针对性调整。

[0069] 针对得到的每个候选推荐模式,根据推荐信息的内容特征、展示位置特征、对象特征、以及候选推荐模式的素材特征进行预测处理,得到该候选推荐模式对应的预测触发率和预测转化率。然后,对候选推荐模式对应的预测触发率和预测转化率进行融合处理,得到该候选推荐模式的第一评分。例如,将预测触发率和预测转化率进行乘积处理,得到第一评分;又例如,在推荐信息为广告的情况下,可以将广告的出价、候选推荐模式对应的预测触发率以及预测转化率进行乘积处理,得到该候选推荐模式的第一评分。

[0070] 在一些实施例中,步骤202之前,还包括:通过人工智能模型,对样本推荐信息在历史推荐记录中的内容特征、展示位置特征、历史对象特征以及应用的推荐模式的素材特征进行预测处理,得到预测结果;根据预测结果与历史推荐结果之间的差异,更新人工智能模型的权重参数;其中,预测结果为预测触发率及预测转化率中的任意一种;更新后的人工智能模型用于对推荐信息的内容特征、展示位置特征、对象特征、以及候选推荐模式的素材特征进行预测处理。

[0071] 在本申请实施例中,可以通过人工智能模型进行预测处理,得到预测触发率和预测转化率。首先,对人工智能模型进行训练,例如,通过人工智能模型,对样本推荐信息在历史推荐记录中的内容特征、展示位置特征、历史对象特征以及应用的推荐模式的素材特征进行预测处理,得到预测结果。其中,样本推荐信息可以是步骤101中确定出的符合对象特征的推荐信息,也可以是其他的推荐信息。在样本推荐信息的历史推荐记录中,还包括基于应用推荐模式的样本推荐信息执行推荐操作后,得到的历史推荐结果。

[0072] 值得说明的是,预测结果是预测触发率及预测转化率中的任意一种,即可以针对预测触发率和预测转化率,分别训练一个人工智能模型,为了便于理解,以训练预测触发率对应的人工智能模型进行说明。在通过人工智能模型,对样本推荐信息的某条历史推荐记录进行预测处理,得到预测触发率后,确定预测触发率与该历史推荐记录中的历史推荐结果之间的差异(即损失值),根据该差异在人工智能模型中进行反向传播,并在反向传播的过程中,沿梯度下降方向更新人工智能模型的权重参数。其中,对确定差异的方式不做限定,例如可通过交叉熵损失函数,来确定预测触发率与历史推荐结果之间的差异。当满足设定的迭代条件时,确定人工智能模型更新完成,更新后的人工智能模型用于对推荐信息的内容特征、展示位置特征、对象特征、以及候选推荐模式的素材特征进行预测处理,得到候选推荐模式对应的预测触发率,其中,迭代条件如设定的迭代次数,或者设定的精确度阈值等。训练预测转化率对应的人工智能模型同理,在此不做赘述。通过上述方式,提升了预测处理得到的预测触发率和预测转化率的精确率。

[0073] 在步骤203中,根据第一评分,对候选推荐模式进行筛选处理,得到一个推荐模式以作为筛选出的推荐模式。

[0074] 对于推荐信息的每个候选推荐模式,通过步骤202均可得到对应的第一评分。然后,根据第一评分对多个候选推荐模式进行筛选处理,最终得到一个推荐模式以作为筛选出的推荐模式,例如,将第一评分最大的候选推荐模式,作为筛选出的推荐模式,当然这并不构成对本申请实施例的限定。

[0075] 在一些实施例中,可以通过这样的方式来实现上述的根据第一评分,对候选推荐模式进行筛选处理,得到一个推荐模式以作为筛选出的推荐模式:将候选推荐模式中曝光量小于或等于曝光量阈值的推荐模式,添加至第一集合,并将候选推荐模式中曝光量大于曝光量阈值的推荐模式,添加至第二集合;根据选择概率对第一集合和第二集合进行选择处理;其中,第一集合的选择概率与第二集合的选择概率的加和为1;当选择的第一集合为非空集合、或者第二集合为空集合时,对第一集合内的推荐模式的第一评分进行归一化处理,得到归一化概率,并根据归一化概率,对第一集合内的推荐模式进行选择处理,得到一个推荐模式以作为筛选出的推荐模式;当选择的第二集合为非空集合、或者第一集合为空集合时,将第二集合内的、第一评分最高的推荐模式,作为筛选出的推荐模式。

[0076] 本申请实施例提供了根据第一评分对多个候选推荐模式进行筛选处理的另一种示例。首先,针对推荐信息的每个候选推荐模式,获取候选推荐模式的曝光量,即应用该候选推荐模式的推荐信息在历史上的曝光量。将多个候选推荐模式中曝光量小于或等于曝光量阈值的推荐模式,添加至第一集合,同时将多个候选推荐模式中曝光量大于曝光量阈值的推荐模式,添加至第二集合,曝光量阈值可根据实际应用场景进行设定,如设定为5000次。然后,根据设定的选择概率,对第一集合和第二集合进行选择处理,其中,第一集合的选择概率与第二集合的选择概率的加和为1,例如设定第一集合的选择概率为0.2,第二集合的选择概率为0.8。

[0077] 经过选择处理后,若选择的第一集合为非空集合、或者第二集合为空集合(无论选择的是什么集合),对第一集合内的推荐模式的第一评分进行归一化处理,例如进行softmax归一化处理,得到归一化概率。然后,根据归一化概率,对第一集合内的推荐模式进行选择处理,得到一个推荐模式以作为筛选出的推荐模式。

[0078] 经过选择处理后,若选择的第二集合为非空集合、或者第一集合为空集合(无论选择的是什么集合),将第二集合内的、第一评分最高的推荐模式,作为筛选出的推荐模式。通过上述的划分第一集合和第二集合,并进行选择的方式,能够挖掘出潜在(曝光量小于曝光量阈值)的优质推荐模式,适用于冷启动(历史推荐记录的数量较少)等场景。

[0079] 在步骤204中,根据应用有筛选出的推荐模式的推荐信息,执行针对对象的推荐操作。

[0080] 在一些实施例中,可以通过这样的方式来实现上述的根据应用有筛选出的推荐模式的推荐信息,执行针对对象的推荐操作:在多个推荐信息的、筛选出的推荐模式的第一评分的降序排序中,将排序在前的第二设定数量的推荐信息,确定为目标推荐信息;其中,第二设定数量为大于0的整数;根据应用有筛选出的推荐模式的目标推荐信息,执行针对对象的推荐操作。

[0081] 由于步骤101中确定出的符合对象特征的推荐信息可能存在多个,故针对每个推荐信息,得到筛选出的推荐模式后,根据筛选出的推荐模式的第一评分,对多个推荐信息进行降序排序,并将排序在前的第二设定数量的推荐信息,确定为目标推荐信息,其中,第二设定数量为大于0的整数,例如可设定为1。然后,根据应用有筛选出的推荐模式的目标推荐信息,执行针对对象的推荐操作。通过上述方式,能够基于第一评分,对符合对象特征的推荐信息进行进一步筛选,从而进一步提升了推荐效果。

[0082] 如图3B所示,本申请实施例在通过置信区间界限确定出候选推荐模式的基础上,根据第一评分对候选推荐模式进行精选,进一步提升了筛选处理的精度和最终的推荐效果。

[0083] 在一些实施例中,参见图3C,图3C是本申请实施例提供的基于人工智能的推荐信息推荐方法的一个可选的流程示意图,图3A示出的步骤102可以通过步骤301至步骤304实现,将结合各步骤进行说明。

[0084] 在步骤301中,针对推荐信息包括的每个推荐模式,初始化推荐模式对应的第一关联参数和第二关联参数。

[0085] 这里,对于推荐信息包括的多个推荐模式来说,均需要确定对应的关联关系,为了便于理解,以推荐信息包括推荐模式A、B和C,且确定推荐模式A的关联关系为例,进行说明。首先,针对推荐模式A,初始化推荐模式A对应的第一关联参数和第二关联参数,例如,将第一关联参数初始化为单位矩阵,将第二关联参数初始化为零向量,其中,单位矩阵和零向量的维度,均与历史对象特征的维度相同。

[0086] 在步骤302中,将历史推荐记录中的历史对象特征与历史对象特征的转置进行乘积处理,并将乘积结果与第一关联参数进行求和处理,以根据求和结果更新第一关联参数。

[0087] 获取应用推荐模式A的推荐信息的历史推荐记录,并根据历史推荐记录中的历史对象特征,更新推荐模式A对应的第一关联参数。例如,将历史推荐记录中的历史对象特征与历史对象特征的转置进行乘积处理,并将乘积结果与第一关联参数进行求和处理,以根据求和结果更新第一关联参数。

[0088] 在一些实施例中,步骤302之前,还包括:对多个历史对象特征进行聚类处理,得到多个中心向量;根据多个中心向量,对历史推荐记录中的历史对象特征进行更新处理。

[0089] 由于历史对象特征的维度可能较高,稀疏性较强,故在本申请实施例中,可以对历

史对象特征进行更新处理。例如,可以从数据库中获取所有历史推荐记录中的历史对象特征,并对多个历史对象特征进行聚类处理,得到多个中心向量,其中,多个历史对象特征的格式相同,例如都包括年龄、性别及居住城市。此外,在进行聚类处理之前,可以对多个历史对象特征进行特征工程处理,如离散化处理,以性别举例,可以将其离散为两类,对应二维向量,如果是男,则性别对应的向量为 $[1,0]$ ;如果是女,则性别对应的向量 $[0,1]$ 。

[0090] 根据聚类得到的多个中心向量,对应用推荐模式A的推荐信息的历史推荐记录中的历史对象特征进行更新处理。本申请实施例对更新处理的方式不做限定,例如一种方式是对历史对象特征进行降维,另一种方式是在维度不变的基础上,增加历史对象特征中的信息量。然后,可将更新后的历史对象特征用于步骤302及步骤303中,以对推荐模式A对应的第一关联参数和第二关联参数进行更新。通过上述方式,提升了历史对象特征的质量,便于根据历史对象特征,对第一关联参数和第二关联参数进行有效更新。

[0091] 值得说明的是,在后续的步骤305和步骤306中,所用的对象特征也可以是经过离散化处理,并根据多个中心向量进行更新处理后得到的。

[0092] 在一些实施例中,可以通过这样的方式来实现上述的对多个历史对象特征进行聚类处理,得到多个中心向量:对多个历史对象特征进行随机选取,得到第三设定数量的中心向量,并根据每个中心向量创建对应的向量簇;其中,第三设定数量为大于1的整数;针对每个历史对象特征,将历史对象特征添加至距离最近的中心向量所在的向量簇;确定每个向量簇的质心,以作为向量簇的新的中心向量,并确定向量簇的中心向量偏移距离;当向量簇的中心向量偏移距离大于或等于距离阈值时,针对每个历史对象特征,将历史对象特征添加至距离最近的中心向量所在的向量簇,直至向量簇的中心向量偏移距离小于距离阈值。

[0093] 本申请实施例提供了聚类处理的一种示例,首先,对多个历史对象特征进行随机选取,得到第三设定数量的中心向量,并根据得到的每一个中心向量,创建对应的向量簇,其中,第三设定数量为大于1的整数,例如可设定为5。然后,针对每个历史对象特征(包括被选取为中心向量的历史对象特征),将历史对象特征添加至距离最近的中心向量所在的向量簇,直至所有历史对象特征都被添加至向量簇中。

[0094] 对于每个向量簇,计算向量簇的质心,例如对向量簇内的所有历史对象特征进行平均处理,得到质心,以作为该向量簇的新的中心向量。然后,确定向量簇的新的中心向量与原有的中心向量之间的距离,以作为中心向量偏移距离。当中心向量偏移距离大于或等于距离阈值时,针对每个历史对象特征,将历史对象特征添加至距离最近的中心向量(指新的中心向量)所在的向量簇,直至向量簇的中心向量偏移距离小于距离阈值为止。停止迭代后,即可根据多个向量簇的中心向量,对历史对象特征进行更新处理。本申请实施例对计算向量之间距离的方式不做限定,例如距离可以是欧式距离或曼哈顿距离等。通过上述方式,能够对多个历史对象特征进行准确归类,提升得到的中心向量的准确性。

[0095] 在一些实施例中,可以通过这样的方式来实现上述的根据多个中心向量,对历史推荐记录中的历史对象特征进行更新处理:执行以下任意一种处理:确定历史推荐记录中的历史对象特征、与每个中心向量之间的相似度,并根据得到的多个相似度,构建更新后的历史对象特征;确定历史推荐记录中的历史对象特征在每个中心向量上的投影分量,并以投影分量作为权重,对多个中心向量进行加权处理,得到更新后的历史对象特征。

[0096] 本申请实施例提供了更新处理的两种方式。第一种方式是,确定历史推荐记录中

的历史对象特征、与每个中心向量之间的相似度,本申请实施例对相似度的类型不做限定,例如可以是余弦相似度等。然后,根据得到的多个相似度,构建更新后的历史对象特征,例如,可以直接将多个相似度作为更新后的历史对象特征中的数值;又例如,确定最大的相似度对应的中心向量,将该中心向量对应的类别数值更新为1,将其他中心向量对应的类别数值更新为0,然后将这些类别数值作为更新后的历史对象特征中的数值,该更新后的历史对象特征的维度与中心向量的数量相同。如此,能够对历史对象特征进行有效降维。

[0097] 举例来说,进行聚类处理后得到5个中心向量,历史推荐记录中的历史对象特征、与5个中心向量之间的相似度分别为0.3(即30%)、0.6、0.2、0.7和0.4,则更新后的历史对象特征可以是[0.3,0.6,0.2,0.7,0.4],也可以是[0,0,0,1,0]。

[0098] 第二种方式是,确定历史推荐记录中的历史对象特征在每个中心向量上的投影分量,并以投影分量作为权重,对多个中心向量进行加权处理,如进行加权求和,得到更新后的历史对象特征。如此,能够增加历史对象特征中的信息量。通过上述方式,提升了更新处理的灵活性,根据实际应用场景可选用任一种更新处理的方式。

[0099] 在步骤303中,将历史推荐记录中的历史对象特征与历史推荐结果进行乘积处理,并将乘积结果与第二关联参数进行求和处理,以根据求和结果更新第二关联参数。

[0100] 根据历史推荐记录中的历史对象特征及历史推荐结果,更新推荐模式A对应的第二关联参数。例如,将历史推荐记录中的历史对象特征与历史推荐结果进行乘积处理,并将乘积结果与第二关联参数进行求和处理,以根据求和结果更新第二关联参数。

[0101] 值得说明的是,在应用推荐模式A的推荐信息的历史推荐记录包括多条时,可以多次执行步骤302和步骤303,以对第一关联参数和第二关联参数进行多次更新。

[0102] 在步骤304中,将更新后的第一关联参数的逆矩阵、与更新后的第二关联参数进行乘积处理,得到第三关联参数;其中,更新后的第一关联参数、以及第三关联参数用于表示关联关系。

[0103] 在根据应用推荐模式A的推荐信息的历史推荐记录,完成对推荐模式A对应的第一关联参数和第二关联参数的更新后,将更新后的第一关联参数的逆矩阵、与更新后的第二关联参数进行乘积处理,得到第三关联参数。其中,更新后的第一关联参数、以及第三关联参数用于表示推荐模式A对应的关联关系,第二关联参数相当于一个中间参数,用于对第三关联参数进行更新。同理,也可以对推荐模式B和C对应的参数(指第一关联参数、第二关联参数及第三关联参数)进行更新,在此不做赘述。

[0104] 在图3C中,图3A示出的步骤103可通过步骤305至步骤307实现。

[0105] 在步骤305中,将对象特征的转置、与第三关联参数进行乘积处理,得到推荐模式的预测推荐结果。

[0106] 同样以推荐模式A为例,将待接收推荐信息的对象的对象特征、与推荐模式A对应的第三关联参数进行乘积处理,得到推荐模式A的预测推荐结果。

[0107] 在步骤306中,将对象特征、对象特征的转置、与更新后的第一关联参数的逆矩阵进行乘积处理,并将乘积结果的平方根与设定系数进行乘积处理,得到预测推荐结果的置信区间宽度。

[0108] 同时,将对象特征、对象特征的转置、与推荐模式A对应的更新后的第一关联参数的逆矩阵进行乘积处理,并将乘积结果的平方根与设定系数进行乘积处理,得到预测推荐

结果的置信区间宽度。其中,设定系数可根据实际应用场景进行设定,设定系数越大,表示越倾向于探索(Explore),即挖掘潜在的推荐模式;设定系数越小,表示越倾向于利用(Exploit),即利用已有的经验。

[0109] 值得说明的是,在引入展示位置特征的情况下,可以将步骤302和步骤303中的历史对象特征替换为第一拼接特征,可以将步骤305和步骤306中的对象特征替换为第二拼接特征。

[0110] 在步骤307中,将预测推荐结果与置信区间宽度进行求和处理,得到置信区间界限。

[0111] 将推荐模式A对应的预测推荐结果与置信区间宽度进行求和处理,得到推荐模式A的置信区间界限,这里的置信区间界限即为置信区间的上界。同理,可以得到推荐模式B和C的置信区间界限,然后,便可根据置信区间界限,对推荐模式A、B和C进行筛选处理。

[0112] 如图3C所示,本申请实施例通过历史推荐记录来更新第一关联参数、第二关联参数及第三关联参数,进而能够预测出更加准确的置信区间界限,提升了后续推荐的精度。

[0113] 在一些实施例中,参见图3D,图3D是本申请实施例提供的基于人工智能的推荐信息推荐方法的一个可选的流程示意图,基于图3A,在步骤101之前,还可以在步骤401中,针对数据库中的每个推荐信息,根据推荐信息的历史推荐记录确定推荐信息的触发率及转化率。

[0114] 在本申请实施例中,可以对数据库中的推荐信息进行离线筛选,以筛选出更具价值的推荐信息。首先,针对数据库中的每个推荐信息,根据推荐信息的多条历史推荐记录,确定该推荐信息的触发率及转化率,其中,触发率=触发量/曝光量,转化率=转化量/触发量。

[0115] 在步骤402中,将推荐信息的触发率及转化率进行融合处理,得到推荐信息的第二评分。

[0116] 例如,将推荐信息的触发率及转化率进行乘积处理,得到该推荐信息的第二评分;又例如,当推荐信息为广告时,将广告的出价、触发率及转化率进行乘积处理,得到该广告的第二评分。

[0117] 在步骤403中,根据第二评分对多个推荐信息进行筛选处理。

[0118] 例如,在多个推荐信息的第二评分的降序排序中,将排序在前的第四设定数量的推荐信息,作为筛选出的推荐信息,其中,第四设定数量为大于1的整数。

[0119] 在图3D中,图3A示出的步骤101可通过步骤404至步骤406实现。

[0120] 在步骤404中,获取待接收推荐信息的对象的对象特征。

[0121] 在存在待接收推荐信息的对象时,获取该对象的对象特征。

[0122] 在步骤405中,获取筛选出的推荐信息的定向推荐特征。

[0123] 同时,获取通过步骤403筛选出的推荐信息的定向推荐特征,其中,推荐信息的定向推荐特征可以预先设定,根据推荐信息所推荐内容的特点而定,例如,推荐信息是某个商品的广告,该商品的目标客户是年龄段20~40岁的男性,则可设定对应的定向推荐特征包括年龄段20~40岁、男性。

[0124] 在步骤406中,当定向推荐特征与对象特征匹配成功时,将定向推荐特征对应的推荐信息,确定为符合对象特征的推荐信息。

[0125] 当对象特征与定向推荐特征相同、或落入定向推荐特征的范围时,确定定向推荐特征与对象特征匹配成功。以上述例子再次举例,若对象特征中的年龄是30岁,性别是男性,则确定定向推荐特征与对象特征匹配成功;若对象特征中的年龄是30岁,性别是女性,则确定定向推荐特征与对象特征匹配失败。然后,将匹配成功的定向推荐特征对应的推荐信息,确定为符合对象特征的推荐信息。

[0126] 在图3D中,步骤406之后,还在步骤407中,根据对象特征、以及符合对象特征的推荐信息的内容特征进行预测处理,并对得到的预测触发率和预测转化率进行融合处理,得到第三评分。

[0127] 在确定出符合对象特征的推荐信息的基础上,还可进行进一步筛选,例如,根据对象特征、以及符合对象特征的推荐信息的内容特征进行预测处理,并对得到的预测触发率和预测转化率进行融合处理,得到该推荐信息的第三评分。其中,融合处理的方式与步骤402同理,此处不再赘述。

[0128] 值得说明的是,可以通过人工智能模型来对对象特征和内容特征进行预测处理,在此之前,需要对人工智能模型进行训练,例如,根据样本推荐信息在历史推荐记录中的内容特征、历史对象特征以及历史推荐结果,对人工智能模型的权重参数进行更新。

[0129] 在步骤408中,根据第三评分,对符合对象特征的多个推荐信息进行筛选处理。

[0130] 例如,在多个推荐信息的第三评分的降序排序中,筛选出排序在前的第五设定数量的推荐信息,以执行后续的步骤102。其中,第五设定数量为大于1的整数,并且第五设定数量小于第四设定数量。

[0131] 如图3D所示,本申请实施例通过对数据库中的推荐信息进行离线筛选,另外在得到符合对象特征的推荐信息后,根据预测的第三评分进行进一步筛选,从而能够得到更有推荐价值的推荐信息,提升推荐资源的利用率。

[0132] 下面,将说明本申请实施例在一个实际的应用场景中的示例性应用。为了便于理解,以推荐信息为广告的情况进行说明。本申请实施例适用于多创意广告,不同于普通广告只允许广告主上传一个创意,在多创意广告中,允许广告主在一个广告下上传多个创意,其中,创意即对应上文的推荐模式。在构建多创意广告时,可以获取广告主上传的、与广告的不同待填充部分分别对应的多个素材,并将不同待填充部分对应的素材进行组合,得到创意;也可以获取广告主上传的、已经制作完成的创意。

[0133] 本申请实施例提供了如图4所示的上传素材的界面示意图,在图4中,示出了广告的待填充部分41(创意图片)、待填充部分42(创意文案)以及待填充部分43(按钮文案)。针对每个待填充部分,可以获取广告主上传的多个素材,例如针对待填充部分41,获取到创意图片素材1和创意图片素材2;针对待填充部分42,获取到创意文案素材1和创意文案素材2;针对待填充部分43,获取到按钮文案素材1和按钮文案素材2。则将不同待填充部分对应的素材进行组合后,可以得到创意包括(创意图片素材1,创意文案素材1,按钮文案素材1)、(创意图片素材1,创意文案素材1,按钮文案素材2)、……。

[0134] 本申请实施例还提供了如图5所示的上传素材的界面示意图,在图5中,广告主已经在线下确定好了各个创意所用的素材,因此,针对待填充部分51、待填充部分52和待填充部分53,均只获取广告主上传的一个素材,以形成广告的一个创意,如图5中示出的创意1。

[0135] 在形成多创意广告后,即可将多创意广告推荐至用户,即进行线上展示。在线上展

示的过程,在广告的多个创意中筛选出最佳的创意,并将应用最佳的创意的广告,进行曝光。通过构建多创意广告,能够避免频繁创建多个相似的广告,提升投放效率以及探索优质创意的效率,并且,能够为不同的用户展示最合适的创意,实现个性化、精准化的创意推荐,从而丰富广告的曝光人群,提升推荐效果。

[0136] 本申请实施例提供了如图6所示的投放多创意广告之后的推荐效果示意图,图6中示出了创意的多个维度的数据,包括所属广告、创意形式、曝光量、点击量(对应上文的触发量)、点击率、点击均价及花费,当然,这并不构成对本申请实施例的限定,例如还可显示转化量、转化率、回报投入比(Return On Investment,ROI)等。通过该推荐效果示意图,能够有助于广告主对不同创意的优劣有直观的认识,帮助广告主挑选出其中整体表现优秀的创意,以在后续过程中进行更多次推荐。

[0137] 除了手动监控的方式外,在本申请实施例中,还提供了自动筛选创意的方式。作为示例,提供了如图7所示的广告推荐系统的框架示意图,在图7中,示出了数据反馈模块、线下模块、粗排模块和精排模块,将对各模块进行详细说明:

1) 数据反馈模块:负责收集历史上进行广告推荐时,广告粒度和创意粒度的推荐效果数据,具体如曝光、点击及转化(即上文的历史推荐结果)等数据;此外,还收集用户特征(对应上文的历史对象特征),包括但不限于年龄、性别、居住城市、婚恋状况、消费水平、教育程度及工作状态等不同类型的特征。以某个广告包括创意A和创意B举例,那么,广告粒度的推荐效果数据由创意A和创意B的推荐效果数据综合得到;创意粒度的数据就仅包括某个创意,如创意A的推荐效果数据。收集的这些数据即对应上文的历史推荐记录,这些数据能够为后续模块如线下模块提供数据源头。

[0138] 2) 线下模块:包括用户特征聚类模块和关联参数训练模块,以下进行分别说明:

① 用户特征聚类模块:在数据反馈模块中,已经获取到历史上每一次广告推荐时,对应的用户特征及相应的推荐效果数据(曝光、点击及转化)。如图8所示,在用户特征聚类模块中,会进行特征工程处理、用户特征聚类,最终输出中心向量供线上使用。

[0139] 在特征工程部分,对用户特征进行离散化,例如将年龄离散成10个区间,对应一个10维的向量,若用户特征中的年龄符合第一个区间,将得到对应的向量为 $[1,0,0,0,0,0,0,0,0,0]$ 。又例如,将性别离散成2类,对应2维向量,若用户特征中的性别是男,则得到对应的向量为 $[1,0]$ ;若是女,则对应的向量为 $[0,1]$ 。对于用户特征中其他类型的特征,以此类推。将用户特征进行离散化后,可以表示为:

$$U = [u_1, u_2, u_3, \dots, u_n]$$

其中, $u_i$ 表示将用户特征内各个类型的特征进行离散化后,得到的向量,例如 $u_1$ 、 $u_2$ 和 $u_3$ 分别是年龄对应的离散化向量、性别对应的离散化向量和居住城市对应的离散化向量。

[0140] 经过特征过程处理后,得到用户特征 $U$ ,然后对用户特征进行聚类处理。以得到的用户特征包括 $U_1, U_2, U_3, \dots, U_N$ 举例,在聚类处理的第一个步骤中,在所有用户特征中随机选取 $K$ 个用户特征,作为中心向量,记为 $X_1, X_2, \dots, X_K$ ,其中, $K$ 为大于0的整数,例如可取5。根据选取出的每个中心向量,创建对应的向量簇。

[0141] 在聚类处理的第二个步骤中,执行循环操作:

{对于每一个用户特征,根据以下公式计算其所属的向量簇:

$$C_i = \underset{j}{\operatorname{argmin}} \|U_i - X_j\|^2$$

即,将第*i*个用户特征,添加至与其距离最小的中心向量(表示为 $X_j$ )所在的向量簇(表示为 $C_i$ );

对于每一个向量簇,计算该向量簇的质心,并将计算出的质心作为新的中心向量:

$$X_j = \frac{\sum_{i=1}^N 1\{C_i = j\} U_i}{\sum_{i=1}^N 1\{C_i = j\}}$$

当中心向量 $X_j$ 的偏移距离大于或等于距离阈值 $\epsilon$ 时,继续执行循环,直至各个向量簇的中心向量 $X_j$ 的偏移距离均小于 $\epsilon$ ;

其中, $\epsilon$ 为大于0的数,如可取0.01}

完成聚类处理后,可以得到5个中心向量,以供其他模块使用。

[0142] ②关联参数训练模块:在该模块中,首先,根据聚类处理得到的中心向量,对用户特征 $U_i$ 进行更新处理,得到上下文特征 $x_i$ (对应上文的更新后的历史对象特征),其中,更新处理的方式在后文进行阐述。

[0143] 然后,设定创意 $a$ 带来的收益同上下文特征 $x_i$ 成线性关系,线性因子为 $\theta_a$ ,则收益 $E$ 可以表示为:

$$E[r_{t,a} | x_{t,a}] = x_{t,a} \theta_a = x_i \theta_a$$

其中, $t$ 表示第*t*次广告推荐, $x_{t,a}$ 即为第*t*次推荐应用创意 $a$ 的广告时对应的上下文特征,此处即为上下文特征 $x_i$ 。

[0144] 定义 $D_a (m \times d)$ 为第*t*次广告推荐时,创意 $a$ 对应的历史观察用户矩阵,包括之前*m*次推荐创意 $a$ 时的上下文特征(维度为*d*),同时定义历史收益向量 $c_a (\mathbb{R}^m)$ ,包括之前*m*次推荐创意 $a$ 时的收益 $r$ (对应上文的历史推荐结果)。此处收益 $r$ 定义为用户的点击转化效果,当曝光无点击时, $r$ 的值为0;当点击无转化时, $r$ 的值为1;当已转化时, $r$ 的值为1+转化量。

[0145] 在上述定义的基础上,利用岭回归(Ridge Regression),可以得到 $\theta_a$ 的估计值:

$$\hat{\theta}_a = (D_a^T D_a + I_d)^{-1} D_a^T c_a$$

其中, $I_d$ 是维度*d*×*d*的单位矩阵。进而,可以保证以下公式成立的概率大于 $1 - \delta$ :

$$\left| x_{t,a}^T \hat{\theta}_a - E[r_{t,a} | x_{t,a}] \right| \leq \alpha \sqrt{x_{t,a}^T (D_a^T D_a + I_d)^{-1} x_{t,a}}$$

其中,  $\alpha=1+\sqrt{\ln(2/\delta)/2}$ ,  $\alpha$  即对应上文的设定系数。 $\delta$  为大于0的数, 如可取0.1。

[0146] 进而, 可以得到创意 $a$ 的预测收益(对应上文的预测推荐结果)的置信区间界限为:

$$x_{t,a}^T \hat{\theta}_a + \alpha \sqrt{x_{t,a}^T (D_a^T D_a + I_d)^{-1} x_{t,a}}$$

令  $A_a = D_a^T D_a + I_d$ ,  $b_a = D_a^T c_a$ , 则可以得到:  $\hat{\theta}_a \leftarrow A_a^{-1} b_a$ 。其中,  $A_a$  对应上文的第一关联参数,  $b_a$  对应上文的第二关联参数,  $\hat{\theta}_a$  对应上文的第三关联参数。

[0147] 因此, 在线下模块中, 只需更新每个创意对应的关联参数  $A_a$  和  $b_a$  即可。

[0148] 在更新过程中, 首先对关联参数  $A_a$  和  $b_a$  进行初始化:

$$A_a \leftarrow I_d, \quad b_a \leftarrow \mathbf{0}_{d \times 1}$$

其中,  $\mathbf{0}_{d \times 1}$  表示 $d$ 维的0向量。

[0149] 在第 $t$ 次广告推荐时, 得到广告所应用的创意 $a$ 、相应的上下文特征  $x_{t,a}$  以及收益  $r_t$ , 对关联参数的更新公式如下:

$$A_{at} \leftarrow A_{at} + x_{t,a_t} x_{t,a_t}^T, \quad b_{at} \leftarrow b_{at} + r_t x_{t,a_t}$$

$$\hat{\theta}_a \leftarrow A_a^{-1} b_a$$

在线下模块中, 完成关联参数的更新后, 将更新后的关联参数传递至线上, 供粗排模块使用。

[0150] 3) 粗排模块。

[0151] 在阐述线上的粗排模块之前, 先对广告推荐系统进行阐述。本申请实施例提供了如图9所示的广告推荐系统的架构图, 在图9中, 全量广告即是指数据库(广告库)包括的所有广告。在图9示出的离线排序环节中, 计算数据库中每个广告的点击率(Click Through Rate, CTR)和转化率(Conversion Rate, CVR), 进而计算广告的千次展示收益(effective Cost Per Mile, eCPM), 例如  $eCPM1 = bid \times CTR \times CVR$ , 其中,  $bid$  即为该广告的出价, 这里计算出的  $eCPM1$  对应上文的第二评分。根据  $eCPM1$  对多个广告进行降序排序, 取出其中靠前的广告(例如上文的第四设定数量的广告)进入在线索引池。在图9中的检索(召回)环节中, 根据线上请求带有的用户特征(对应上文的对象特征), 从在线索引池中检索出符合用户特征的广告, 以进入粗排模块。

[0152] 在粗排模块中, 首先针对进入的每个广告, 结合该广告的广告特征(对应上文的内容特征)以及线上请求中的用户特征, 进行预测处理, 得到预测点击率和预测转化率, 将预测点击率表示为  $liteCTR$ , 将预测转化率表示为  $liteCVR$ , 可以得到广告的  $eCPM2 = bid \times$

liteCTR×liteCVR,这里得到的eCPM2对应上文的第三评分。在本申请实施例中,可以通过人工智能模型,如深度双塔网络模型,来进行预测处理,得到liteCTR和liteCVR。

[0153] 然后,根据得到的eCPM2对多个广告进行降序排序,取出其中靠前的广告(例如上文的第五设定数量的广告)进行进一步处理。接下来,继续说明精排模块包括的用户特征更新模块和创意筛选模块:

①用户特征更新模块:当接收到线上请求,即接收到待进行广告推荐的用户的用户特征 $U$ 时,根据用户特征聚类模块传递过来的5个中心向量(即包括 $X_1, X_2, \dots, X_5$ ),对用户特征 $U$ 进行更新处理,得到上下文特征 $x$ 。

[0154] 本申请实施例提供了更新处理的两种方式,一种方式是:

$$user\ class = \underset{i}{argmax} Similarity(U, X_i)$$

$$x = [\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5]$$

$$\lambda_i = \begin{cases} 1, & i = user\ class \\ 0, & i \neq user\ class \end{cases}$$

其中, $Similarity()$ 为相似度函数,在最终得到的上下文特征 $x$ 中,只有一个数值为1,即与用户特征 $U$ 相似度最大的中心向量对应的数值,其他数值均为0。

[0155] 另一种方式是,确定用户特征 $U$ 在每个中心向量上的投影分量 $a_i$ ,将投影分量作为相应的中心向量的系数,对5个中心向量进行加权矢量和,得到上下文特征,公式为:

$$x = sum(X_i * a_i), \text{ where } i = 1 \text{ to } 5$$

值得说明的是,这里的对用户特征 $U$ 进行更新处理的方式,同样适用于关联参数训练模块中。此外,本申请实施例中还可引入广告位特征(对应上文的展示位置特征),对广告位特征进行同样的聚类处理和更新处理后,将更新后的广告位特征、与更新后的用户特征,拼接为上下文特征。

[0156] ②创意筛选模块:在该模块中,获取关联参数训练模块传递过来的关联参数 $A_a$ 和 $\hat{\theta}_a$ ,并针对根据eCPM2筛选出的广告包括的每个创意,利用公式

$x_{t,a}^T \hat{\theta}_a + \alpha \sqrt{x_{t,a}^T (D_a^T D_a + I_d)^{-1} x_{t,a}}$  计算创意的预测收益的置信区间界限,这里的 $x_{t,a}$ 表示用户特征更新模块中得到的上下文特征 $x$ 。

[0157] 针对每个广告,根据置信区间界限对广告包括的多个创意进行降序排序,并筛选出排序在前(例如上文的第一设定数量)的创意,进入精排模块,例如,筛选出5个创意以进入精排模块。

[0158] 4)精排模块:包括预测模块及创意筛选模块,以下进行分别说明。

[0159] ①预测模块:针对粗排模块传递过来的广告包括的每个创意,进行预测处理,得到预测点击率和预测转化率,为了便于区分,将这里得到的预测点击率和预测转化率分别表

示为pCTR和pCVR。由于进入精排模块的广告数量和创意数量较少,因此在pCTR和pCVR的预测过程中,可以使用较多的特征,例如广告特征、创意特征(对应上文的素材特征)、用户特征及广告位特征,预测精度较之粗排模块也更高。

[0160] ②创意筛选模块:本申请实施例提供了如图10所示的创意筛选的示意图,以对某一个广告包括的5个创意进行筛选进行说明,首先,根据创意的曝光量,将曝光量小于或等于5000次(对应上文的曝光量阈值)的创意,如创意1、创意2和创意3,添加至新创意集合101(对应上文的第一集合);将曝光量大于5000次的创意,如创意4和创意5,添加至老创意集合102(对应上文的第二集合)。

[0161] 同时,计算每个创意的 $eCPM3 = bid \times pCTR \times pCVR$ ,这里得到的eCPM3对应上文的第一评分。对于一个广告的不同创意而言,bid都是相同的,故在这里,也可以计算创意的 $eCPM3 = pCTR \times pCVR$ 。

[0162] 然后,概率性地将当前流量定为探索流量或者利用流量,设定命中探索流量的概率和命中利用流量的概率的加和为1,例如命中探索流量的概率是0.2,命中利用流量的概率是0.8。

[0163] 若当前流量命中探索流量,则在新创意集合101中进行筛选。由于新创意集合101中创意的曝光不充分,pCTR和pCVR的精度相较于老创意集合102中创意来说,误差更大,故在新创意集合101中,对各创意的eCPM3进行softmax归一化,并按照得到的归一化概率来选择出一个创意,如此,能够在一定程度上给予eCPM3较小的新创意一些探索的机会。其中,softmax归一化的公式可以为:

$$\sigma(z_j) = \frac{e^{z_j}}{\sum_{s=1}^S e^{z_s}} \quad \text{for } s = 1, \dots, S$$

其中, $z_j$ 表示新创意集合101中的第j个创意的eCPM3, $\sigma(z_j)$ 表示第j个创意的归一化概率,S表示新创意集合101包括的创意的数量。

[0164] 若当前流量命中利用流量,则在老创意集合102中进行筛选。由于老创意集合102中创意的曝光已经足够充分,pCTR和pCVR的精度相对较高,则直接根据不同创意的eCPM3进行排序,筛选出eCPM3最大的创意。

[0165] 此外,若当前流量命中探索流量、且新创意集合101为空集合时,则在老创意集合102中筛选出eCPM3最大的创意;若当前流量命中利用流量、且老创意集合102为空集合时,则在新创意集合101中按照归一化概率进行选择。

[0166] 如图11所示,针对进入精排模块的每个广告,在筛选出每个广告对应的一个创意后,根据筛选出的广告的广告的eCPM3,对多个广告进行降序排序,将排序在前(例如上文的第二设定数量)的广告,作为目标广告(对应上文的推荐信息),在图11中,以第二设定数量为1的情况进行示例说明。然后,将应用有筛选出的创意的目标广告,推荐至用户,完成整个广告推荐流程。

[0167] 本申请实施例在粗排模块和精排模块中均使用了探索利用的优选策略方法,在创意累计的推荐效果数据不足时进行探索,在推荐效果数据充足时进行利用,优选其中最佳创意,适用于冷启动等场景,也有效避免了对劣质创意进行曝光;同时,由于在粗排模块和

精排模块中均考虑了上下文信息,即引入了用户特征,提升了广告推荐的精准化和个性化,从而提高了多创意广告的整体长期收益;另外,对于引入的用户特征,通过聚类及更新处理的方式,解决了用户特征稀疏性较高的问题。

[0168] 下面继续说明本申请实施例提供的基于人工智能的推荐信息推荐装置455实施为软件模块的示例性结构,在一些实施例中,如图2所示,存储在存储器450的基于人工智能的推荐信息推荐装置455中的软件模块可以包括:获取模块4551,用于获取待接收推荐信息的对象的对象特征,并获取符合对象特征的推荐信息;其中,推荐信息包括多个推荐模式;关系确定模块4552,用于针对应用推荐模式的推荐信息,确定历史推荐记录中的历史对象特征与历史推荐结果之间的关联关系;预测模块4553,用于根据对象特征、以及推荐模式对应的关联关系进行预测处理,得到推荐模式的预测推荐结果的置信区间界限;筛选模块4554,用于根据置信区间界限对多个推荐模式进行筛选处理,并根据应用有筛选出的推荐模式的推荐信息,执行针对对象的推荐操作。

[0169] 在一些实施例中,筛选模块4554,还用于:在多个推荐模式的置信区间界限的降序排序中,将排序在前的第一设定数量的推荐模式,确定为候选推荐模式;其中,第一设定数量为大于1的整数;根据推荐信息的内容特征、展示位置特征、对象特征、以及候选推荐模式的素材特征进行预测处理,并对得到的预测触发率和预测转化率进行融合处理,得到第一评分;根据第一评分,对候选推荐模式进行筛选处理,得到一个推荐模式以作为筛选出的推荐模式。

[0170] 在一些实施例中,筛选模块4554,还用于:将候选推荐模式中曝光量小于或等于曝光量阈值的推荐模式,添加至第一集合,并将候选推荐模式中曝光量大于曝光量阈值的推荐模式,添加至第二集合;根据选择概率对第一集合和第二集合进行选择处理;其中,第一集合的选择概率与第二集合的选择概率的加和为1;当选择的第一集合为非空集合、或者第二集合为空集合时,对第一集合内的推荐模式的第一评分进行归一化处理,得到归一化概率,并根据归一化概率,对第一集合内的推荐模式进行选择处理,得到一个推荐模式以作为筛选出的推荐模式;当选择的第二集合为非空集合、或者第一集合为空集合时,将第二集合内的、第一评分最高的推荐模式,作为筛选出的推荐模式。

[0171] 在一些实施例中,筛选模块4554,还用于:在多个推荐信息的、筛选出的推荐模式的第一评分的降序排序中,将排序在前的第二设定数量的推荐信息,确定为目标推荐信息;其中,第二设定数量为大于0的整数;根据应用有筛选出的推荐模式的目标推荐信息,执行针对对象的推荐操作。

[0172] 在一些实施例中,基于人工智能的推荐信息推荐装置455还包括:模型预测模块,用于通过人工智能模型,对样本推荐信息在历史推荐记录中的内容特征、展示位置特征、历史对象特征以及应用的推荐模式的素材特征进行预测处理,得到预测结果;模型更新模块,用于根据预测结果与历史推荐结果之间的差异,更新人工智能模型的权重参数;其中,预测结果为预测触发率及预测转化率中的任意一种;更新后的人工智能模型用于对推荐信息的内容特征、展示位置特征、对象特征、以及候选推荐模式的素材特征进行预测处理。

[0173] 在一些实施例中,筛选模块4554,还用于:将多个推荐模式中置信区间界限最大的一个推荐模式,作为筛选出的推荐模式。

[0174] 在一些实施例中,关系确定模块4552,还用于:初始化推荐模式对应的第一关联参

数和第二关联参数;将历史推荐记录中的历史对象特征与历史对象特征的转置进行乘积处理,并将乘积结果与第一关联参数进行求和处理,以根据求和结果更新第一关联参数;将历史推荐记录中的历史对象特征与历史推荐结果进行乘积处理,并将乘积结果与第二关联参数进行求和处理,以根据求和结果更新第二关联参数;将更新后的第一关联参数的逆矩阵、与更新后的第二关联参数进行乘积处理,得到第三关联参数;其中,更新后的第一关联参数、以及第三关联参数用于表示关联关系。

[0175] 在一些实施例中,预测模块4553,还用于:将对象特征的转置、与第三关联参数进行乘积处理,得到推荐模式的预测推荐结果;将对象特征、对象特征的转置、与更新后的第一关联参数的逆矩阵进行乘积处理,并将乘积结果的平方根与设定系数进行乘积处理,得到预测推荐结果的置信区间宽度;将预测推荐结果与置信区间宽度进行求和处理,得到置信区间界限。

[0176] 在一些实施例中,基于人工智能的推荐信息推荐装置455还包括:聚类模块,用于对多个历史对象特征进行聚类处理,得到多个中心向量;特征更新模块,用于根据多个中心向量,对历史推荐记录中的历史对象特征进行更新处理。

[0177] 在一些实施例中,聚类模块还用于:对多个历史对象特征进行随机选取,得到第三设定数量的中心向量,并根据每个中心向量创建对应的向量簇;其中,第三设定数量为大于1的整数;针对每个历史对象特征,将历史对象特征添加至距离最近的中心向量所在的向量簇;确定每个向量簇的质心,以作为向量簇的新的中心向量,并确定向量簇的中心向量偏移距离;当向量簇的中心向量偏移距离大于或等于距离阈值时,针对每个历史对象特征,将历史对象特征添加至距离最近的中心向量所在的向量簇,直至向量簇的中心向量偏移距离小于距离阈值。

[0178] 在一些实施例中,特征更新模块还用于:执行以下任意一种处理:确定历史推荐记录中的历史对象特征、与每个中心向量之间的相似度,并根据得到的多个相似度,构建更新后的历史对象特征;确定历史推荐记录中的历史对象特征在每个中心向量上的投影分量,并以投影分量作为权重,对多个中心向量进行加权处理,得到更新后的历史对象特征。

[0179] 在一些实施例中,关系确定模块4552,还用于:将历史推荐记录中的历史对象特征与展示位置特征进行拼接处理,得到第一拼接特征,并确定第一拼接特征与历史推荐结果之间的关联关系;预测模块4553,还用于:将对象特征与推荐信息的展示位置特征进行拼接处理,得到第二拼接特征,并根据第二拼接特征、以及推荐模式对应的关联关系进行预测处理,得到推荐模式的预测推荐结果的置信区间界限。

[0180] 在一些实施例中,基于人工智能的推荐信息推荐装置455还包括:推荐效果确定模块,用于针对数据库中的每个推荐信息,根据推荐信息的历史推荐记录确定推荐信息的触发率及转化率;第一融合模块,用于将推荐信息的触发率及转化率进行融合处理,得到推荐信息的第二评分;离线筛选模块,用于根据第二评分对多个推荐信息进行筛选处理。

[0181] 在一些实施例中,获取模块4551,还用于:获取筛选出的推荐信息的定向推荐特征;当定向推荐特征与对象特征匹配成功时,将定向推荐特征对应的推荐信息,确定为符合对象特征的推荐信息。

[0182] 在一些实施例中,基于人工智能的推荐信息推荐装置455还包括:第二融合模块,用于根据对象特征和推荐信息的内容特征进行预测处理,并对得到的预测触发率和预测转

化率进行融合处理,得到第三评分;前置筛选模块,用于根据第三评分,对多个推荐信息进行筛选处理。

[0183] 在一些实施例中,基于人工智能的推荐信息推荐装置455还包括:推荐模式获取模块,用于执行以下任意一种处理:获取与推荐信息的多个待填充部分分别对应的一个素材,并将每个素材填充至对应的待填充部分中,得到推荐信息的一个推荐模式;获取与推荐信息的多个待填充部分分别对应的多个素材,并在每个待填充部分对应的多个素材中进行随机选择,将选择的素材填充至对应的待填充部分中,得到推荐信息的一个推荐模式。

[0184] 在一些实施例中,历史推荐结果的类型包括:第一推荐结果,用于表示应用推荐模式的推荐信息已曝光、且未触发;第二推荐结果,用于表示应用推荐模式的推荐信息已触发、且未转化;第三推荐结果,用于表示应用推荐模式的推荐信息已转化。

[0185] 本申请实施例提供了一种计算机程序产品或计算机程序,该计算机程序产品或计算机程序包括计算机指令,该计算机指令存储在计算机可读存储介质中。计算机设备的处理器从计算机可读存储介质读取该计算机指令,处理器执行该计算机指令,使得该计算机设备执行本申请实施例上述的基于人工智能的推荐信息推荐方法。

[0186] 本申请实施例提供一种存储有可执行指令的计算机可读存储介质,其中存储有可执行指令,当可执行指令被处理器执行时,将引起处理器执行本申请实施例提供的方法,例如,如图3A、图3B、图3C或图3D示出的基于人工智能的推荐信息推荐方法。值得说明的是,计算机包括终端设备和服务器在内的各种计算设备。

[0187] 在一些实施例中,计算机可读存储介质可以是FRAM、ROM、PROM、EPROM、EEPROM、闪存、磁表面存储器、光盘、或CD-ROM等存储器;也可以是包括上述存储器之一或任意组合的各种设备。

[0188] 在一些实施例中,可执行指令可以采用程序、软件、软件模块、脚本或代码的形式,按任意形式的编程语言(包括编译或解释语言,或者声明性或过程性语言)来编写,并且其可按任意形式部署,包括被部署为独立的程序或者被部署为模块、组件、子例程或者适合在计算环境中使用的其它单元。

[0189] 作为示例,可执行指令可以但不一定对应于文件系统中的文件,可以可被存储在保存其它程序或数据的文件的一部分,例如,存储在超文本标记语言(HTML,Hyper Text Markup Language)文档中的一个或多个脚本中,存储在专用于所讨论的程序的单个文件中,或者,存储在多个协同文件(例如,存储一个或多个模块、子程序或代码部分的文件)中。

[0190] 作为示例,可执行指令可被部署为在一个计算设备上执行,或者在位于一个地点的多个计算设备上执行,又或者,在分布在多个地点且通过通信网络互连的多个计算设备上执行。

[0191] 综上,通过本申请实施例能够实现以下技术效果:

1)在推荐过程中引入对象特征,并根据确定出的置信区间界限来筛选推荐模式,提升了推荐的针对性和精准性,有效地避免了推荐资源的浪费;在此基础上,还可引入展示位置特征,进一步提升推荐效果。

[0192] 2)提供了构建推荐模式的两种方式,适用于不同的场景,提升了灵活性。

[0193] 3)在根据置信区间界限确定出候选推荐模式的基础上,根据预测出的第一评分进行精选,进一步提升了推荐效果;在精选时,可以使用探索利用策略,便于挖掘出潜在的优

质推荐模式,适用于冷启动等场景。

[0194] 4)在引入对象特征的基础上,可对对象特征进行更新,例如通过确定对象特征与多个中心向量之间的相似度,从而对对象特征进行降维,降低其稀疏性;又例如,根据对象特征在多个中心向量上的投影向量,对多个中心向量进行加权,从而提升对象特征包含的信息量。

[0195] 5)可以预先对数据库中的推荐信息进行离线筛选,另外,在得到符合对象特征的推荐信息后,可以根据预测的第三评分进行进一步筛选,从而能够得到更有推荐价值的推荐信息,提升推荐资源的利用率。

[0196] 以上,仅为本申请的实施例而已,并非用于限定本申请的保护范围。凡在本申请的精神和范围之内所作的任何修改、等同替换和改进等,均包含在本申请的保护范围之内。

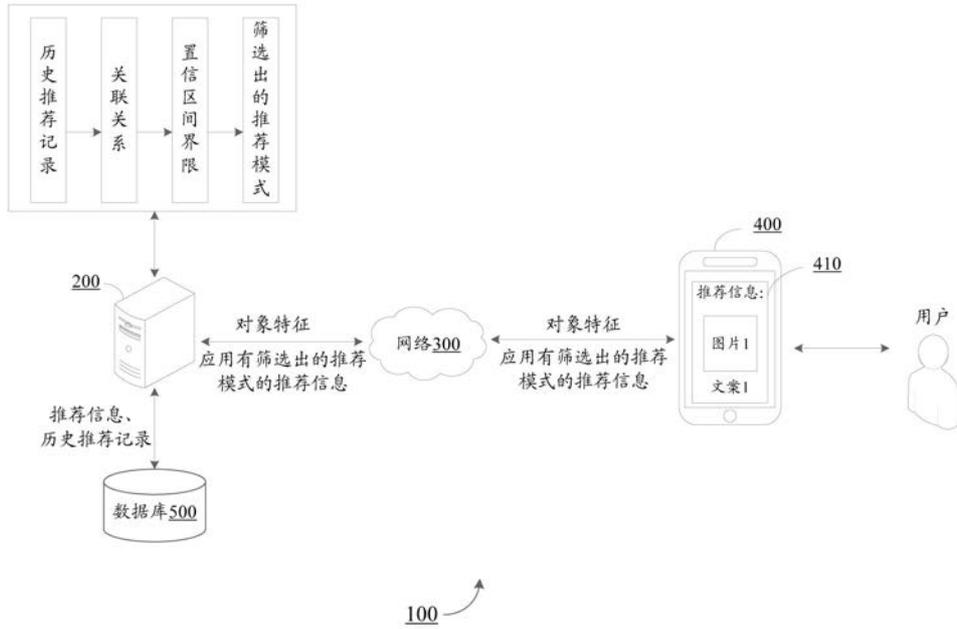


图1

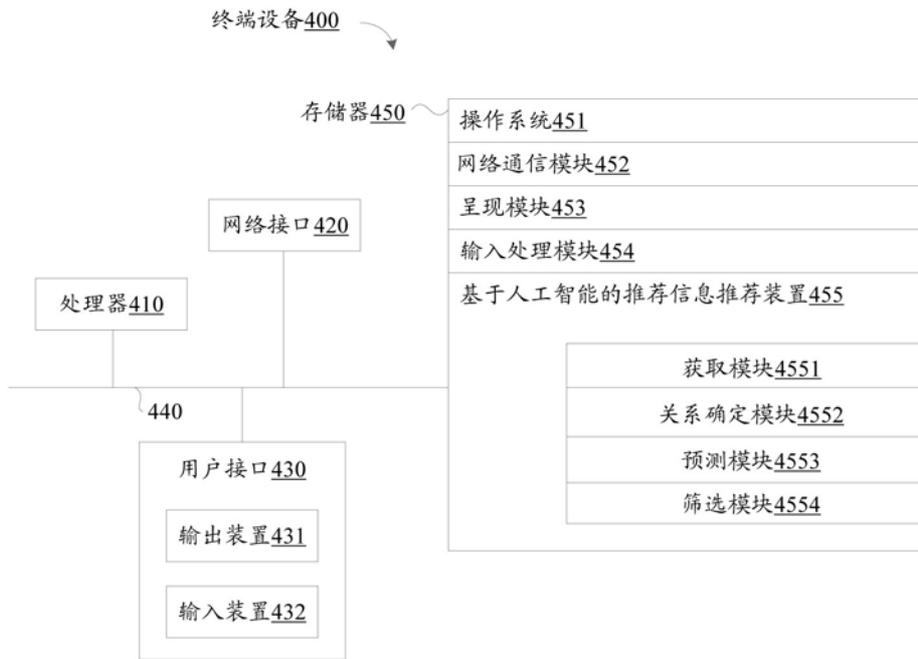


图2

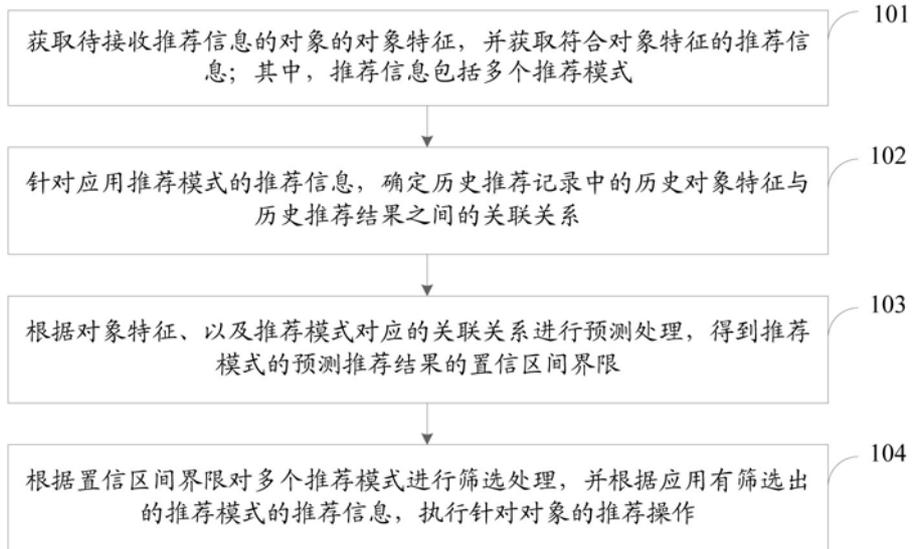


图3A

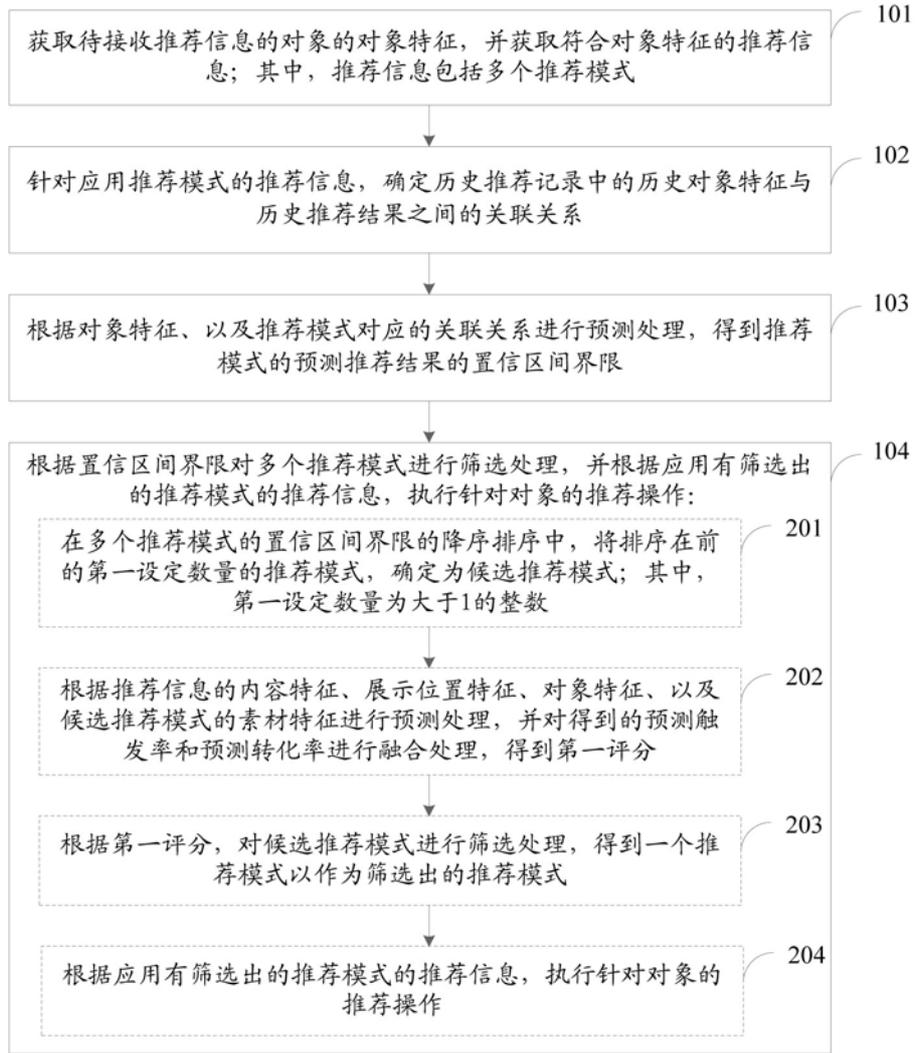


图3B

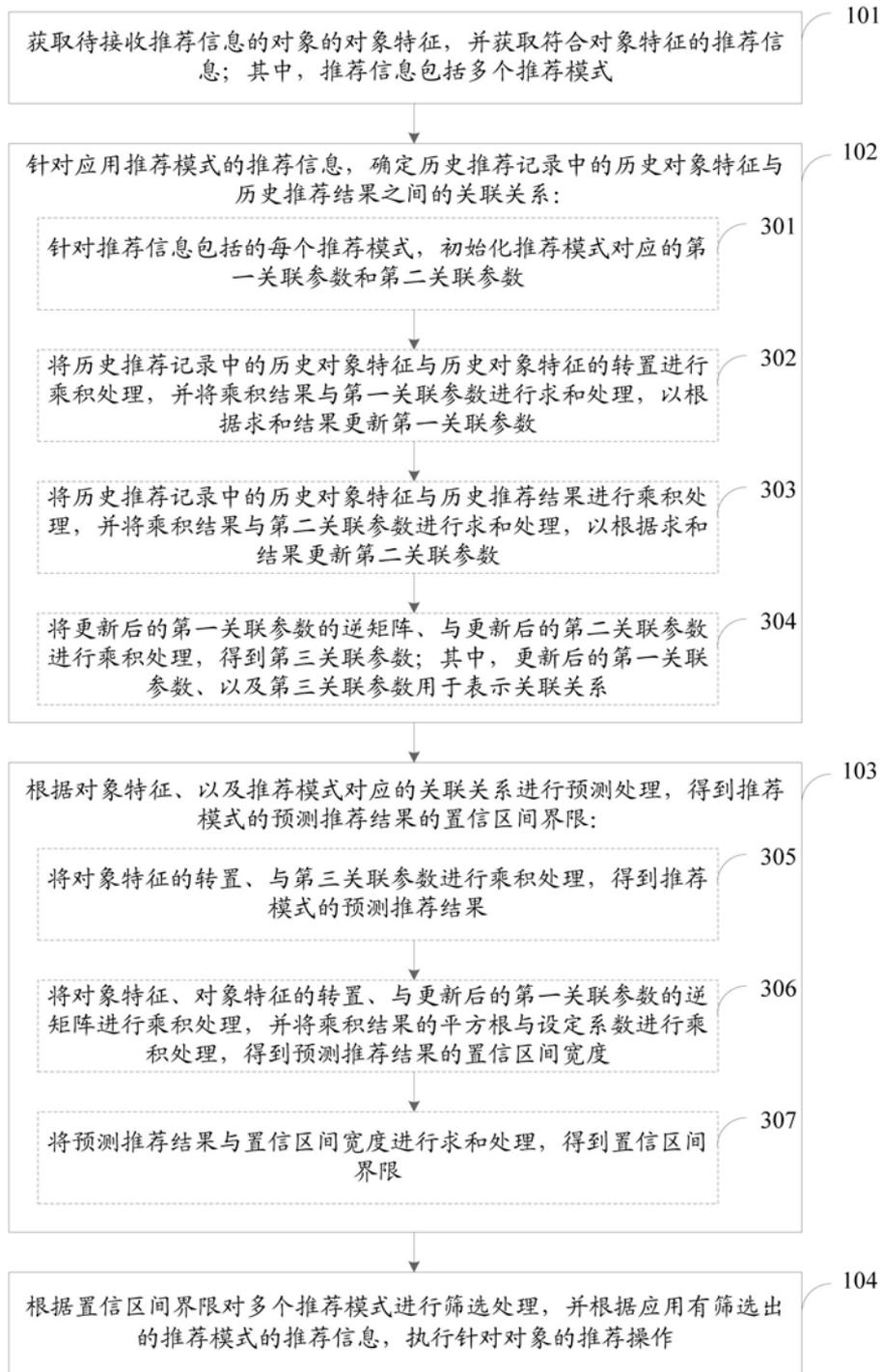


图3C

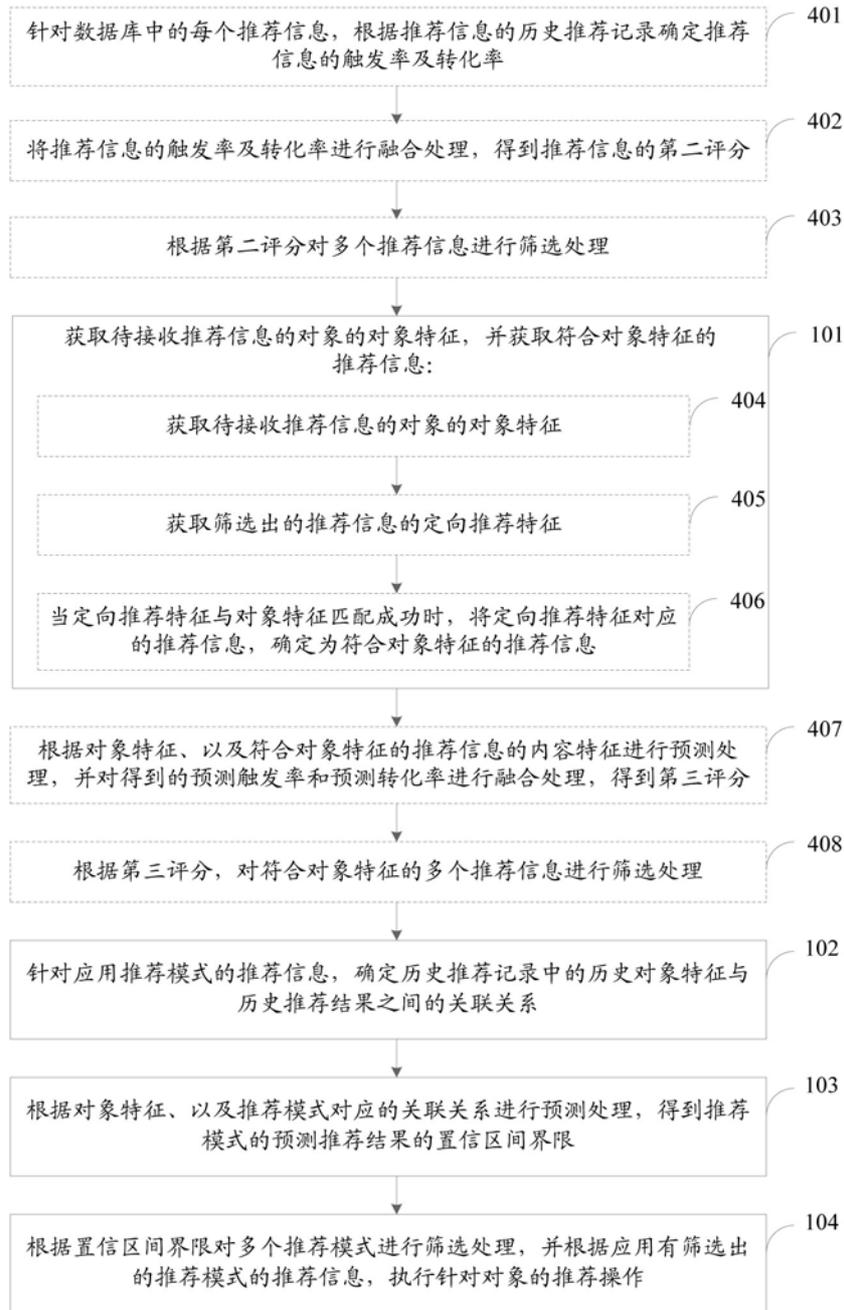


图3D

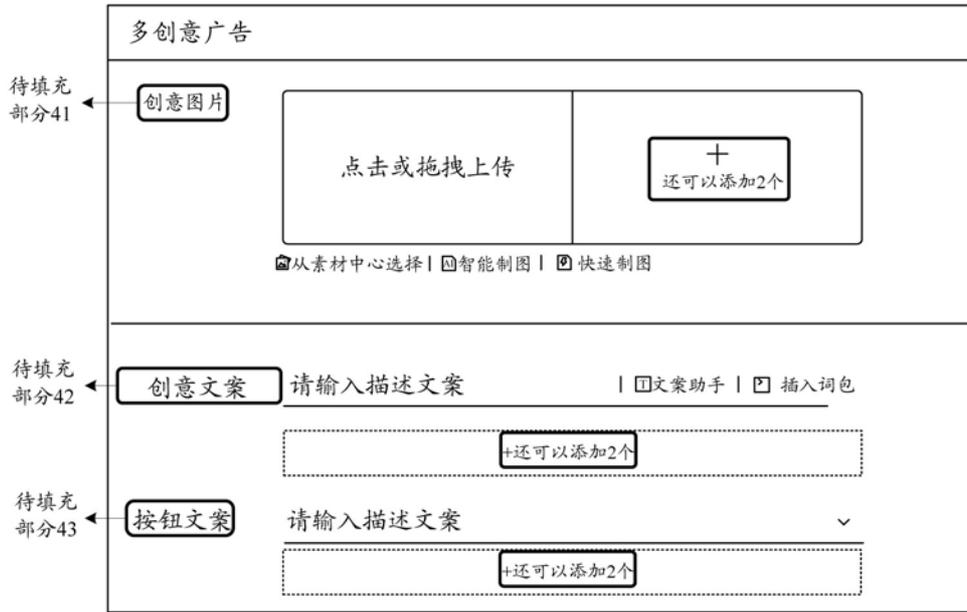


图4

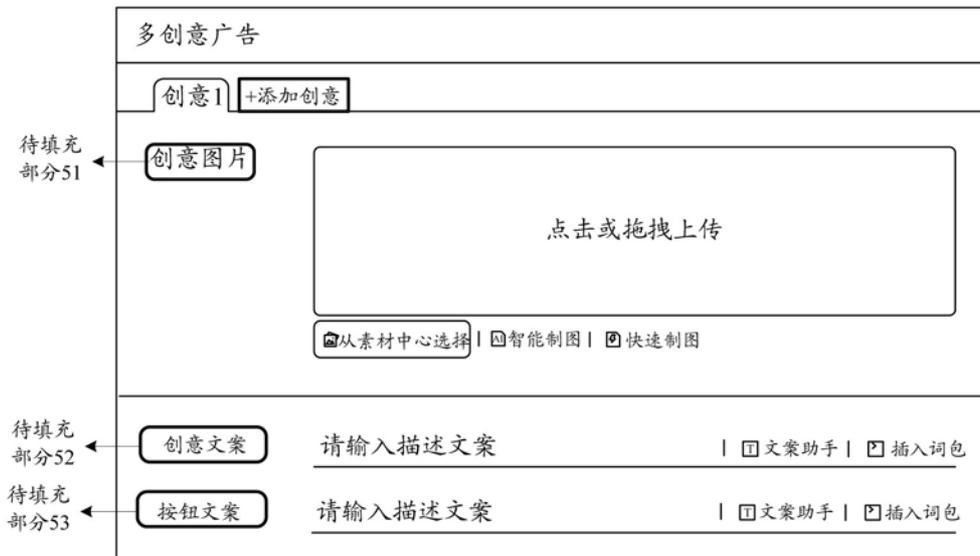


图5

<input type="checkbox"/>	广告包括的创意	所属广告	创意形式	曝光量	点击量	点击率	点击均价	花费
<input type="checkbox"/>	 创意1	ID: 13-01	横版大图16:9	13241	41	0.31%	0.12	5.03
<input type="checkbox"/>	 创意2	ID: 13-01	横版大图16:9	3716	3	0.08%	0.14	0.43
<input type="checkbox"/>	 创意3	ID: 13-01	横版大图16:9	7134	16	0.22%	0.15	2.34

图6

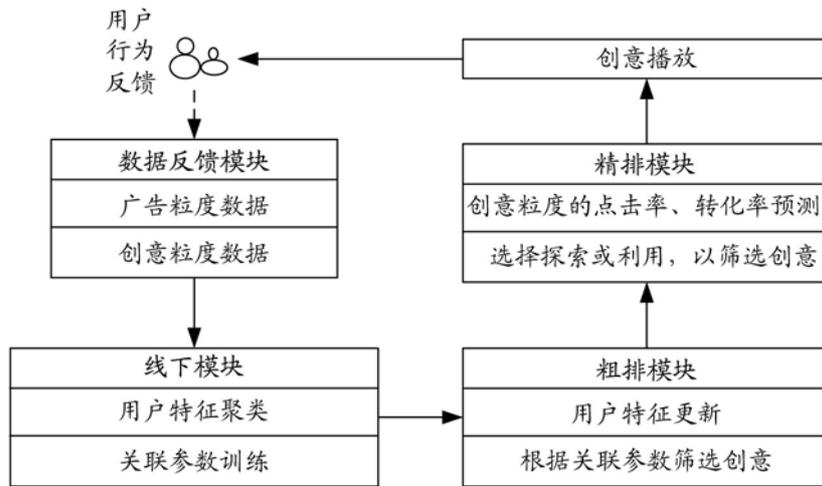


图7

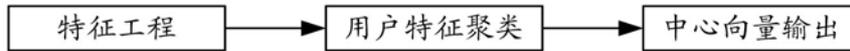


图8

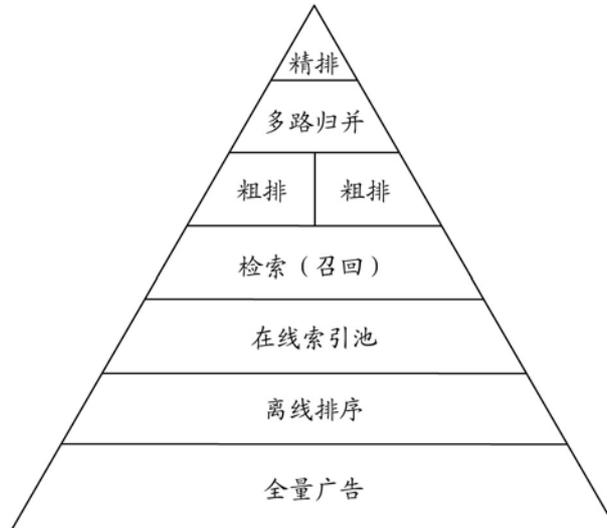


图9

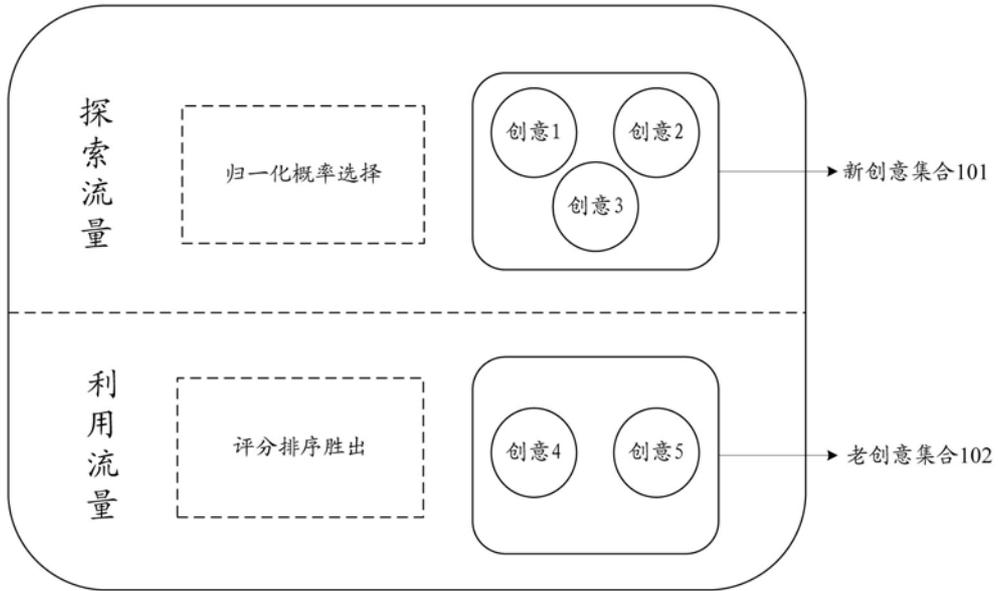


图10

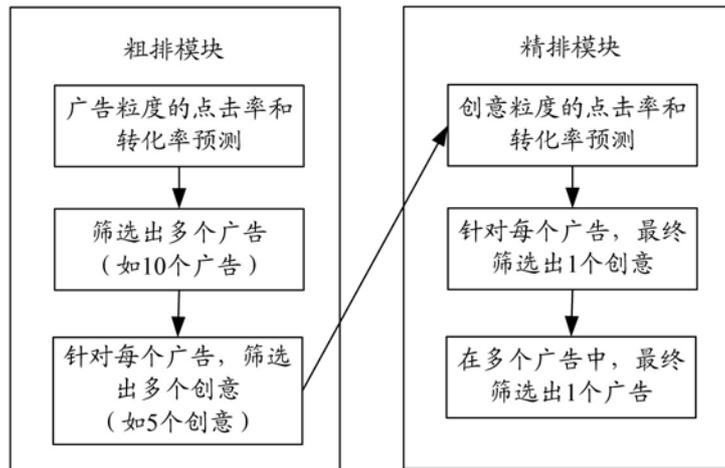


图11