



(12)发明专利申请

(10)申请公布号 CN 108765076 A

(43)申请公布日 2018.11.06

(21)申请号 201810509212.0

(22)申请日 2018.05.24

(71)申请人 蜜芽宝贝(北京)网络科技有限公司

地址 100000 北京市朝阳区太阳宫中路16
号院1号楼16层1606室

(72)发明人 蒲鹏 刘楠 郭春梅

(74)专利代理机构 北京超凡志成知识产权代理
事务所(普通合伙) 11371

代理人 徐丽

(51)Int.Cl.

G06Q 30/06(2012.01)

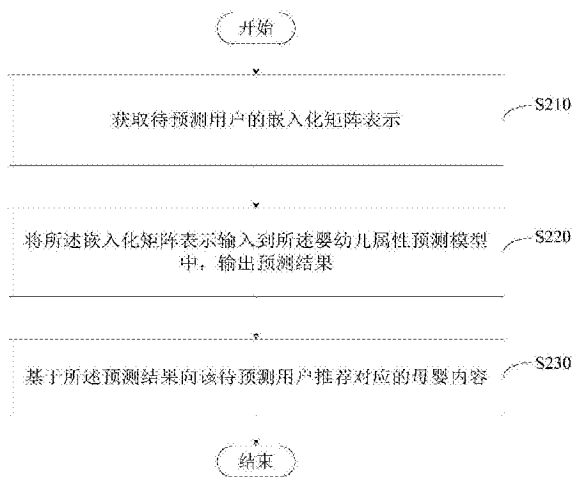
权利要求书3页 说明书10页 附图3页

(54)发明名称

母婴内容推荐方法、装置及可读存储介质

(57)摘要

本发明实施例提供一种母婴内容推荐方法、装置及可读存储介质。该方法包括：获取待预测用户的嵌入化矩阵表示，所述嵌入化矩阵表示中包括有与该待预测用户关联的商品数据；将所述嵌入化矩阵表示输入到婴幼儿属性预测模型中，输出预测结果，所述预测结果包括该待预测用户的至少一种婴幼儿属性；基于所述预测结果向该待预测用户推荐对应的母婴内容。由此，能够为用户提供匹配的社区内容，提升母婴电商用户向母婴社区的导流和对社区内容的兴趣度，进而提高社区内容的点击率，培养用户对内容社区的认知。



1. 一种母婴内容推荐方法,其特征在于,应用于服务器,所述服务器中配置有婴幼儿属性预测模型,所述方法包括:

获取待预测用户的嵌入化矩阵表示,所述嵌入化矩阵表示中包括有与该待预测用户关联的商品数据;

将所述嵌入化矩阵表示输入到所述婴幼儿属性预测模型中,输出预测结果,所述预测结果包括该待预测用户的至少一种婴幼儿属性;

基于所述预测结果向该待预测用户推荐对应的母婴内容。

2. 根据权利要求1所述的母婴内容推荐方法,其特征在于,在所述获取待预测用户的嵌入化矩阵的步骤之前,所述方法还包括:

配置所述婴幼儿属性预测模型,其中,所述婴幼儿属性预测模型包括输入层、卷积层、池化层、全连接层以及输出层;

所述配置所述婴幼儿属性预测模型的方式,包括:

获取训练样本,所述训练样本包括具有婴幼儿属性的多个用户;

对所述训练样本进行处理,得到待输入的嵌入化矩阵表示;

将所述嵌入化矩阵表示通过所述输入层输入到所述卷积层;

通过所述卷积层对所述嵌入化矩阵表示进行卷积运算,向所述池化层输出对应的特征图;

通过所述池化层将所述特征图转换为对应的多维向量后输出到所述全连接层,其中,所述全连接层的节点与所述池化层的节点连接;

所述全连接层随机丢弃预定比例的节点对所述特征图进行计算,并将计算结果输出至分段线性函数进行计算处理,向所述输出层输出计算处理结果;

所述输出层将所述计算处理结果输入到softmax分类函数中计算出每种婴幼儿属性类别的概率值;

在上述训练过程中,根据所述每种婴幼儿属性类别的概率值和对应用户的婴幼儿属性计算交叉熵损失值,并根据所述交叉熵损失值使用随机梯度下降算法更新所述婴幼儿属性预测模型各层的参数,直到所述交叉熵损失值满足训练终止条件时,停止训练,输出目标模型参数和计算图;

根据所述目标模型参数和计算图得到所述婴幼儿属性预测模型。

3. 根据权利要求2所述的母婴内容推荐方法,其特征在于,所述对所述训练样本进行处理,得到待输入的嵌入化矩阵表示的步骤,包括:

针对具有婴幼儿属性的每个用户,获取具有与该用户关联的商品数据,所述商品数据包括商品浏览数据、商品收藏数据、商品购买数据中的其中一种或者多种组合,所述商品浏览数据、商品收藏数据、商品购买数据分别包括商品名称和对应的商品属性;

对所述商品数据进行过滤,并分别对过滤后的商品数据中的商品名称和对应的商品属性进行数值化处理,生成字典文件,所述字典文件包括各个商品的名称编码和商品属性编码;

对所述字典文件进行处理后得到待输入的嵌入化矩阵表示。

4. 根据权利要求3所述的母婴内容推荐方法,其特征在于,所述对所述商品数据进行过滤的方式,包括:

将所述商品数据中商品名称和商品属性的出现次数小于预设次数的商品进行过滤。

5. 根据权利要求3所述的母婴内容推荐方法,其特征在於,所述对所述字典文件进行处理后得到待输入的嵌入化矩阵表示的步骤,包括:

统计所述字典文件中的商品名称数量和商品属性数量,生成统计结果;

根据所述统计结果生成所述婴幼儿属性预测模型的待输入长度;

从预先配置的词嵌入矩阵查找所述字典文件中每个商品名称和商品属性的嵌入向量;

基于所述待输入长度和查找到的每个商品名称和商品属性的嵌入向量生成待输入的嵌入化矩阵表示。

6. 根据权利要求1所述的母婴内容推荐方法,其特征在於,所述婴幼儿属性预测模型包括输入层、卷积层、池化层、全连接层以及输出层,所述将所述嵌入化矩阵表示输入到所述婴幼儿属性预测模型中,输出预测结果的步骤,包括:

将所述嵌入化矩阵表示通过所述输入层输入到所述卷积层;

通过所述卷积层对所述嵌入化矩阵表示进行卷积运算,向所述池化层输出对应的特征图;

通过所述池化层将所述特征图转换为对应的多维向量后输出到所述全连接层,其中,所述全连接层的节点与所述池化层的节点连接;

所述全连接层使用全部节点对所述特征图进行计算,并将计算结果输出至分段线性函数进行计算处理后,向所述输出层输出对应的计算处理结果;

所述输出层将所述计算处理结果输入到softmax分类函数中计算出每种婴幼儿属性类别的概率值;

选择概率值最高的婴幼儿属性作为预测结果。

7. 根据权利要求1所述的母婴内容推荐方法,其特征在於,所述基于所述预测结果向该待预测用户推荐对应的母婴内容的步骤,包括:

基于所述预测结果向该用户推荐匹配的UGC内容和PGC内容,其中,所述PGC内容关联有相应的婴幼儿属性选项;

基于所述预测结果获取对应的用户集合和用户集合中每个用户的商品数据,并根据所述对应的用户集合和用户集合中每个用户的商品数据向该用户推荐匹配的商品内容。

8. 根据权利要求7所述的母婴内容推荐方法,其特征在於,所述根据所述对应的用户集合和用户集合中每个用户的商品数据向该用户推荐匹配的商品内容的步骤,包括:

根据置信阈值对所述用户集合和所述用户集合中每个用户的商品数据进行分类,得到分类结果;

根据所述分类结果向该用户推荐匹配的商品内容。

9. 一种母婴内容推荐装置,其特征在於,应用于服务器,所述服务器中配置有婴幼儿属性预测模型,所述装置包括:

获取模块,用于获取待预测用户的嵌入化矩阵表示,所述嵌入化矩阵表示中包括有与该待预测用户关联的商品数据;

输入模块,用于将所述嵌入化矩阵表示输入到所述婴幼儿属性预测模型中,输出预测结果,所述预测结果包括该待预测用户的至少一种婴幼儿属性;

推荐模块,用于基于所述预测结果向该待预测用户推荐对应的母婴内容。

10. 一种可读存储介质,其特征在于,所述可读存储介质中存储有计算机程序,所述计算机程序被执行时实现权利要求1-8中任意一项所述的母婴内容推荐方法。

母婴内容推荐方法、装置及可读存储介质

技术领域

[0001] 本发明涉及母婴内容推荐领域,具体而言,涉及一种母婴内容推荐方法、装置及可读存储介质。

背景技术

[0002] 随着垂直母婴电商的发展,单纯提供电商的购买功能,越来越受限于品类的扩容和供应链的竞争力,而母婴用户群天然的具有高粘度的社区属性,通过发展母婴社区,可以提升用户活跃和APP停留时长,更好地反哺电商平台。内容+电商的模式逐渐成为垂直电商行业的标配,提高电商和内容社区之间互相导流的效率非常有价值。

[0003] 母婴电商用户在进行购物的过程中,对内容社区的认知和需求是偏弱的,用户的浏览行为主要以电商销售商品、口碑评论为主,搜索也是偏电商类SKU为主,对电商平台附加的社区内容缺乏认知。如何为用户提供匹配的社区内容,提升母婴电商用户向母婴社区的导流,是本领域技术人员亟待解决的技术问题。

发明内容

[0004] 为了克服现有技术中的上述不足,本发明的目的在于提供一种母婴内容推荐方法、装置及可读存储介质,能够为用户提供匹配的社区内容,提升母婴电商用户向母婴社区的导流和对社区内容的兴趣度,进而提高社区内容的点击率,培养用户对内容社区的认知。

[0005] 为了实现上述目的,本发明实施例采用的技术方案如下:

[0006] 第一方面,本发明实施例提供一种母婴内容推荐方法,应用于服务器,所述服务器中配置有婴幼儿属性预测模型,所述方法包括:

[0007] 获取待预测用户的嵌入化矩阵表示,所述嵌入化矩阵表示中包括有与该待预测用户关联的商品数据;

[0008] 将所述嵌入化矩阵表示输入到所述婴幼儿属性预测模型中,输出预测结果,所述预测结果包括该待预测用户的至少一种婴幼儿属性;

[0009] 基于所述预测结果向该待预测用户推荐对应的母婴内容。

[0010] 可选地,在所述获取待预测用户的嵌入化矩阵的步骤之前,所述方法还包括:

[0011] 配置所述婴幼儿属性预测模型,其中,所述婴幼儿属性预测模型包括输入层、卷积层、池化层、全连接层以及输出层;

[0012] 所述配置所述婴幼儿属性预测模型的方式,包括:

[0013] 获取训练样本,所述训练样本包括具有婴幼儿属性的多个用户;

[0014] 对所述训练样本进行处理,得到待输入的嵌入化矩阵表示;

[0015] 将所述嵌入化矩阵表示通过所述输入层输入到所述卷积层;

[0016] 通过所述卷积层对所述嵌入化矩阵表示进行卷积运算,向所述池化层输出对应的特征图;

[0017] 通过所述池化层将所述特征图转换为对应的多维向量后输出到所述全连接层,其

中,所述全连接层的节点与所述池化层的节点连接;

[0018] 所述全连接层随机丢弃预定比例的节点对所述特征图进行计算,并将计算结果输出至分段线性函数进行计算处理,向所述输出层输出计算处理结果;

[0019] 所述输出层将所述计算处理结果输入到softmax分类函数中计算出每种婴幼儿属性类别的概率值;

[0020] 在上述训练过程中,根据所述每种婴幼儿属性类别的概率值和对应用户的婴幼儿属性计算交叉熵损失值,并根据所述交叉熵损失值使用随机梯度下降算法更新所述婴幼儿属性预测模型各层参数,直到所述交叉熵损失值满足训练终止条件时,停止训练,输出目标模型参数和计算图;

[0021] 根据所述目标模型参数和计算图得到所述婴幼儿属性预测模型。

[0022] 可选地,所述对所述训练样本进行处理,得到待输入的嵌入化矩阵表示的步骤,包括:

[0023] 针对具有婴幼儿属性的每个用户,获取具有与该用户关联的商品数据,所述商品数据包括商品浏览数据、商品收藏数据、商品购买数据中的其中一种或者多种组合,所述商品浏览数据、商品收藏数据、商品购买数据分别包括商品名称和对应的商品属性;

[0024] 对所述商品数据进行过滤,并分别对过滤后的商品数据中的商品名称和对应的商品属性进行数值化处理,生成字典文件,所述字典文件包括各个商品的商品名称编码和商品属性编码;

[0025] 对所述字典文件进行处理后得到待输入的嵌入化矩阵表示。

[0026] 可选地,所述对所述商品数据进行过滤的方式,包括:

[0027] 将所述商品数据中商品名称和商品属性的出现次数小于预设次数的商品进行过滤。

[0028] 可选地,所述对所述字典文件进行处理后得到待输入的嵌入化矩阵表示的步骤,包括:

[0029] 统计所述字典文件中的商品名称数量和商品属性数量,生成统计结果;

[0030] 根据所述统计结果生成所述婴幼儿属性预测模型的待输入长度;

[0031] 从预先配置的词嵌入矩阵查找所述字典文件中每个商品名称和商品属性的嵌入向量;

[0032] 基于所述待输入长度和查找到的每个商品名称和商品属性的嵌入向量生成待输入的嵌入化矩阵表示。

[0033] 可选地,所述婴幼儿属性预测模型包括输入层、卷积层、池化层、全连接层以及输出层,所述将所述嵌入化矩阵表示输入到所述婴幼儿属性预测模型中,输出预测结果的步骤,包括:

[0034] 将所述嵌入化矩阵表示通过所述输入层输入到所述卷积层;

[0035] 通过所述卷积层对所述嵌入化矩阵表示进行卷积运算,向所述池化层输出对应的特征图;

[0036] 通过所述池化层将所述特征图转换为对应的多维向量后输出到所述全连接层,其中,所述全连接层的节点与所述池化层的节点连接;

[0037] 所述全连接层使用全部节点对所述特征图进行计算,并将计算结果输出至分段线

性函数进行计算处理后,向所述输出层输出对应的计算处理结果;

[0038] 所述输出层将所述计算处理结果输入到softmax分类函数中计算出每种婴幼儿属性类别的概率值;

[0039] 选择概率值最高的婴幼儿属性作为预测结果。

[0040] 可选地,所述基于所述预测结果向该待预测用户推荐对应的母婴内容的步骤,包括:

[0041] 基于所述预测结果向该用户推荐匹配的UGC内容和PGC内容,其中,所述PGC内容关联有相应的婴幼儿属性选项;

[0042] 基于所述预测结果获取对应的用户集合和用户集合中每个用户的商品数据,并根据所述对应的用户集合和用户集合中每个用户的商品数据向该用户推荐匹配的商品内容。

[0043] 可选地,所述根据所述对应的用户集合和用户集合中每个用户的商品数据向该用户推荐匹配的商品内容的步骤,包括:

[0044] 根据置信阈值对所述用户集合和所述用户集合中每个用户的商品数据进行分类,得到分类结果;

[0045] 根据所述分类结果向该用户推荐匹配的商品内容。

[0046] 第二方面,本发明实施例还提供一种母婴内容推荐装置,应用于服务器,所述服务器中配置有婴幼儿属性预测模型,所述装置包括:

[0047] 获取模块,用于获取待预测用户的嵌入化矩阵表示,所述嵌入化矩阵表示中包括有与该待预测用户关联的商品数据;

[0048] 输入模块,用于将所述嵌入化矩阵表示输入到所述婴幼儿属性预测模型中,输出预测结果,所述预测结果包括该待预测用户的至少一种婴幼儿属性;

[0049] 推荐模块,用于基于所述预测结果向该待预测用户推荐对应的母婴内容。

[0050] 第三方面,本发明实施例还提供一种可读存储介质,所述可读存储介质中存储有计算机程序,所述计算机程序被执行时实现上述的母婴内容推荐方法。

[0051] 相对于现有技术而言,本发明具有以下有益效果:

[0052] 本发明实施例提供的母婴内容推荐方法、装置及可读存储介质,通过获取待预测用户的嵌入化矩阵表示,并将所述嵌入化矩阵表示输入到婴幼儿属性预测模型中,输出预测结果,所述预测结果包括该待预测用户的至少一种婴幼儿属性,而后基于所述预测结果向该待预测用户推荐对应的母婴内容。由此,能够根据预测的用户婴幼儿属性,为用户提供匹配的社区内容,提升母婴电商用户向母婴社区的导流和对社区内容的兴趣度,进而提高社区内容的点击率,培养用户对内容社区的认知。

附图说明

[0053] 为了更清楚地说明本发明实施例的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,应当理解,以下附图仅示出了本发明的某些实施例,因此不应被看作是对范围的限定,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它相关的附图。

[0054] 图1为本发明实施例提供的母婴内容推荐方法的应用场景示意图;

[0055] 图2为本发明实施例提供的母婴内容推荐方法的一种流程示意图;

- [0056] 图3为本发明实施例提供的婴幼儿属性预测模型的结构图；
- [0057] 图4为本发明实施例提供的用于实现上述母婴内容推荐方法的服务器的一种结构框图；
- [0058] 图5为图4中所示的母婴内容推荐装置的功能模块图。
- [0059] 图标:100-服务器;110-总线;120-处理器;130-存储介质;140-总线接口;150-网络适配器;160-用户接口;200-母婴内容推荐装置;210-获取模块;220-输入模块;230-推荐模块;300-用户终端。

具体实施方式

[0060] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。通常在此处附图中描述和示出的本发明实施例的组件可以以各种不同的配置来布置和设计。

[0061] 因此,以下对在附图中提供的本发明的实施例的详细描述并非旨在限制要求保护的本发明的范围,而是仅仅表示本发明的选定实施例。基于本发明中的实施例,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它实施例,都属于本发明保护的范围。

[0062] 应注意到:相似的标号和字母在下面的附图中表示类似项,因此,一旦某一项在一个附图中被定义,则在随后的附图中不需要对其进行进一步定义和解释。同时,在本发明的描述中,术语“第一”、“第二”等仅用于区分描述,而不能理解为指示或暗示相对重要性。

[0063] 目前,随着垂直母婴电商的发展,单纯提供电商的购买功能,越来越受限于品类的扩容和供应链的竞争力,而母婴用户群天然的具有高粘度的社区属性,通过发展母婴社区,可以提升用户活跃和APP 停留时长,更好地反哺电商平台。内容+电商的模式逐渐成为垂直电商行业的标配,提高电商和内容社区之间互相导流的效率非常有价值。

[0064] 发明人在使用社区过程中发现,母婴社区的一大特性就是用户分群,不同婴幼儿属性的用户具有较大的购买和社区需求差异。例如,孕期的用户更关注孕妇装、防辐射服、叶酸等保健品类商品,在母婴社区中更关注孕期需要注意的事项,比如能不能吃某些食物、身体发生某些变化是否正常、孕检阶段需要注意的事项和关键指标解读。而 0-1岁婴幼儿的妈妈更关注母乳喂养、婴幼儿生长指标、新手妈妈必读知识等,购物也偏向于奶粉、纸尿裤等商品为主。1-3岁婴幼儿的妈妈更关注幼儿认知培养,会购买绘本、玩具、童装相关品类,社区方面更多偏晒娃、旅游、幼儿教育类相关知识为主。3-6岁婴幼儿的妈妈更关注大龄婴幼儿的童装、婴幼儿营养相关的商品,社区方面更关注大龄儿童教育、儿童活动相关内容。所以,发明人得出结论,对用户分群可以更好地帮助电商和社区进行精准流量分发,提升用户粘性。

[0065] 然而,发明人发现了现有技术中存在如下问题:母婴电商用户在进行购物的过程中,对社区的认知和需求是偏弱的,用户的浏览行为主要以电商销售商品、口碑评论为主,搜索也是偏电商类SKU为主,对电商平台附加的社区内容缺乏认知,如果只是提供通用的社区内容而不针对用户婴幼儿属性做精细化匹配,大部分用户对这部分社区内容的兴趣度并不高。鉴于此,发明人认为如果能在用户浏览电商商品、口碑的时候能够基于用户的浏览行为、收藏行为、购物行为刻画该户的精准用户画像、婴幼儿属性,提供精准的知识、话题等相

关社区内容,可以提高电商用户对社区内容的兴趣度,进而提高用户对社区内容的点击率,培养用户对内容社区的认知。

[0066] 具体地,母婴社区的标配是初始用户使用时需要设置婴幼儿属性信息,这样有利于将用户进行分群,提供更精准的内容。但是在电商为主的平台上,设置婴幼儿属性不是惯用做法,因为即使提供设置入口,用户填写意愿相对于母婴社区平台不强,母婴电商平台的社区因为缺乏流量入口,能够获取到婴幼儿属性信息则更少,而获取不到婴幼儿属性信息,就无法提供更精准的内容吸引用户浏览,形成一个负反馈的闭环。为了解决上述问题,发明人在研究中发现母婴电商平台上存在非常丰富的电商用户行为,因此可以通过已有婴幼儿属性的电商用户行为进行标注,通过深度学习模型加上新用户行为预估出没有填写信息的用户婴幼儿属性。同时,已有婴幼儿属性浏览的社区内容又可以形成资源类的样本标注,通过刻画这部分资源映射的婴幼儿属性,又可以精准分发这部分内容给适合的婴幼儿属性的用户,进而提升母婴电商用户向母婴社区的导流。由此,本发明提出下述实施例以解决现有技术中存在的诸多弊端。

[0067] 以上现有技术中的方案所存在的缺陷,均是发明人在经过实践并仔细研究后得出的结果,因此,上述问题的发现过程以及下文中本实用新型实施例针对上述问题所提出的解决方案,都应该是发明人在本发明过程中对本发明做出的贡献。

[0068] 请参阅图1,为本发明实施例提供的母婴内容推荐方法的一种应用场景示意图。本实施例中,所述应用场景包括至少一个用户终端 300以及与所述至少一个用户终端300通信连接的服务器100。

[0069] 本实施例中,所述用户终端300可以是个人电脑、智能手机、平板电脑、可穿戴设备等等,本实施例对此不作详细限制。

[0070] 根据本发明的一些实施例,所述用户终端300可以包括:包含应用处理部和射频/数位讯号处理器的处理装置;显示屏;可包含物理键、覆盖在显示幕上的触摸键或它们的组合的袖珍键盘;用户识别模组卡;可以包含ROM、RAM、快闪存储器或它们的任意组合的存储器装置;Wi-Fi和/或蓝牙接口;NFC芯片、用于无线充电的无线电能接收线圈、无线电话接口;带有关联电池的电源管理电路;USB接口和连接器;带有关联麦克风、扬声器和耳机插孔的音讯管理系统;以及各种诸如照相机、全球定位系统、加速器等的选择的附属部件。此外,在用户终端300上可以安装各种用户端应用,用户端应用可以用于允许使用用户终端300来传送适合于和其他设备操作的命令。这类应用可以从服务器100上下载并安装到用户终端300的存储器中,也可以预先已被安装在用户终端300上。在本发明实施例中,用户终端300上安装有母婴内容推荐应用,母婴内容推荐应用可以指点用户实现用户注册、用户登录、商品收藏、商品购买、母婴社区内容浏览、用户邀请等功能。

[0071] 本实施例中,所述服务器100应被理解为提供处理、资料库、通讯设施的业务点。举例而言,服务器100可以指具有相关通信和资料存储和资料库设施的单个的物理处理器,或它可以指联网或集聚的处理器、相关网路和存放装置的集合体,并且对软体和一个或多个资料库系统和支援服务器100所提供的服务的应用软体进行操作。服务器 100可以在配置或性能上差异很大,但是服务器100一般可以包括一个或多个中央处理单元和存储单元。服务器100还可以包括一个或多个大型存放区设备、一个或多个电源、一个或多个有线或无线网络组件、一个或多个输入/输出组件、或一个或多个作业系统,诸如, Windows Server、

Mac OS X、Unix、Linux、FreeBSD等等。

[0072] 进一步地,请参阅图2,为本发明实施例提供的母婴内容推荐方法的一种流程示意图,所述方法由图1中所示的服务器100执行。所应说明的是,本发明实施例提供的母婴内容推荐方法不以图2及以下所述的具体顺序为限制,该母婴内容推荐方法可以通过如下步骤实现:

[0073] 步骤S210,获取待预测用户的嵌入化矩阵表示。

[0074] 本实施例中,所述嵌入化矩阵表示中包括有与该待预测用户关联的商品数据,所述商品数据可表征该待预测用户的用户行为,具体可商品浏览数据、商品收藏数据、商品购买数据中的其中一种或者多种组合。例如,用户电商浏览的sku-id,sku属性id、收藏的sku-id,sku 属性id、购买的sku-id,sku属性id等等。所述sku-id可以理解为具体商品名称,所述sku属性id可以理解为该商品属性,例如对于婴幼儿衣服而言,sku属性id可包括颜色、尺码、样式等等。

[0075] 详细地,在对步骤S210进行进一步阐述之前,首先对所述婴幼儿属性预测模型的配置方式进行说明。请参阅图3,在本实施例中,所述婴幼儿属性预测模型可包括输入层、卷积层、池化层、全连接层以及输出层。

[0076] 首先,获取训练样本,所述训练样本包括具有婴幼儿属性的多个用户,所述婴幼儿属性可以根据实际场景进行定义和扩展,例如,所述婴幼儿属性可包括:婴幼儿的年龄、婴幼儿的性别、婴幼儿的星座、年龄划分段位可以为备孕期、怀孕期、0-1岁、1-3岁、3-6岁,婴幼儿的性别主要分为男、女,婴幼儿的星座主要分为十二星座等等属性,在此不作详细限制。

[0077] 接着,对所述训练样本进行处理,得到待输入的嵌入化矩阵表示。下面结合图3对该嵌入化矩阵表示的生成过程中进行说明,值得说明的是,上述待预测用户的嵌入化矩阵表示的生成也可以参照下述方式实现。

[0078] 首先,针对具有婴幼儿属性的每个用户,获取具有与该用户关联的商品数据。所述商品数据包括商品浏览数据、商品收藏数据、商品购买数据中的其中一种或者多种组合。所述商品浏览数据、商品收藏数据、商品购买数据分别包括商品名称sku-id和对应的商品属性sku 属性id。

[0079] 接着,对所述商品数据进行过滤,并分别对过滤后的商品数据中的商品名称和对应的商品属性进行数值化处理,也即,对sku-id和 sku属性id进行统一编码,对sku-id、sku属性id,用其在字典文件中的编号代替,生成sku-id、sku属性id编码,将其数值化,生成字典文件D,因此所述字典文件包括各个商品的名称编码和商品属性编码。可选地,对所述商品数据进行过滤的方式可以是将所述商品数据中商品名称和商品属性的出现次数小于预设次数(例如两次)的商品进行过滤,从而可以提高训练样本的准确性。

[0080] 而后,对所述字典文件进行处理后得到待输入的嵌入化矩阵表示。作为一种实施方式,可以通过统计所述字典文件中的商品名称数量和商品属性数量,生成统计结果,并根据所述统计结果生成所述婴幼儿属性预测模型的待输入长度。例如,可以统计所有用户每周访问的去重的sku数、sku属性数,选取去重后的sku数加上sku属性数的平均长度作为所述婴幼儿属性预测模型的待输入长度L,长度大于 L的可以作截断处理,长度小于L的标题用0补齐。而后,随机初始化维度为 $D \times S$ 的词嵌入矩阵E,从该词嵌入矩阵E中查找所述字典

文件D中每个商品名称和商品属性的嵌入向量,也即,对于每个 sku-id、sku属性id通过对应的编码查找对应词嵌入矩阵E中的长度为S的嵌入向量。最后,基于所述待输入长度和查找到的每个商品名称和商品属性的嵌入向量组合成维度为L×S的矩阵X,该矩阵X也即所述待输入的嵌入化矩阵表示,该用户的婴幼儿属性作为训练标签。

[0081] 然后,将所述嵌入化矩阵表示X作为输入通过所述输入层输入到所述卷积层,同时将该用户的婴幼儿属性作为训练标签。

[0082] 接着,通过所述卷积层对所述嵌入化矩阵表示进行卷积运算,向所述池化层输出对应的特征图。作为一种实施方式,可以通过所述卷积层将所述输入层用128个宽度为5的和128个宽度为2的一维卷积核进行步长为1的卷积运算,其中,不同宽度的卷积核分别负责提取不同范围大小的局部特征,在进行卷积运算时需要将所述嵌入化矩阵表示进行相应的零填充使得输出的特征图的长度和输入的嵌入化矩阵表示的长度相等。在卷积运算后,可以得到两个L×128的特征图,将两个特征图简单拼接后组成L×256的特征图。

[0083] 接着,通过所述池化层将所述特征图转换为对应的多维向量后输出到所述全连接层。由此,通过最大化池化运算将L×256的特征图转换为256维的向量,池化运算降低了特征图的维度,提高了网络的泛化能力。其中,所述全连接层的节点与所述池化层的节点连接。

[0084] 接着,所述全连接层随机丢弃预定比例(例如50%)的节点对所述特征图进行计算,并将计算结果输出至分段线性函数(ReLU)进行计算处理,向所述输出层输出计算处理结果。其中,所述分段线性函数(ReLU)的计算公式为 $f(x) = \max(0, x)$,其中,x是全连接层的输出,f(x)为分段线性函数的输出。

[0085] 接着,所述输出层将所述计算处理结果输入到softmax分类函数中计算出每种婴

幼儿属性类别的概率值,其中,所述softmax的计算公式为

$$\sigma(z_j) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$$

[0086] 在上述训练过程中,根据所述每种婴幼儿属性类别的概率值和对应用户的婴幼儿属性计算交叉熵损失值(cross-entropy cost),并根据所述交叉熵损失值使用随机梯度下降算法更新所述婴幼儿属性预测模型各层参数,直到所述交叉熵损失值满足训练终止条件时,停止训练,输出目标模型参数和计算图。由此,根据所述目标模型参数和计算图可以得到所述婴幼儿属性预测模型。

[0087] 在上述婴幼儿属性预测模型配置完成后,在获取到待预测用户的嵌入化矩阵表示事,首先加载婴幼儿属性预测模型的计算图和目标模型参数。

[0088] 步骤S220,将所述嵌入化矩阵表示输入到所述婴幼儿属性预测模型中,输出预测结果。

[0089] 在实施时,首先将所述嵌入化矩阵表示通过所述输入层输入到所述卷积层,而后通过所述卷积层对所述嵌入化矩阵表示进行卷积运算,向所述池化层输出对应的特征图,并通过所述池化层将所述特征图转换为对应的多维向量后输出到所述全连接层。而后所述全连接层使用全部节点对所述特征图进行计算,并将计算结果输出至分段线性函数进行计算处理后,向所述输出层输出对应的计算处理结果。所述输出层将所述计算处理结果输入到softmax分类函数中计算出每种婴幼儿属性类别的概率值,并选择概率值最高的婴幼儿

属性作为预测结果,所述预测结果包括该待预测用户的至少一种婴幼儿属性。

[0090] 上述过程的具体操作方法可参照上述婴幼儿属性预测模型的配置方式中相应的详细描述,在此不再重复赘述。

[0091] 由此,本实施例将没有婴幼儿属性的用户访问、收藏、购买的 sku-id、sku属性id使用婴幼儿属性预测模型进行分类分值预测得到结果集合 $P\{p_1, p_2, \dots, p_j\}$,将上述结果集合作为用户婴幼儿属性的最终分类。例如 p_1 表示性别属性为女, p_2 表示属性为0-1岁,那么用户婴幼儿属性刻画为0-1岁的女宝宝。

[0092] 步骤S230,基于所述预测结果向该待预测用户推荐对应的母婴内容。

[0093] 本实施例中,母婴内容可以分为两种形式:生产类内容和消费类内容,生产类内容也即UGC内容和PGC内容,消费类内容也即用户主动浏览、点赞、收藏的内容。作为一种实施方式,在得到预测结果后,可基于所述预测结果向该用户推荐匹配的UGC内容和PGC内容。

[0094] UGC(User Generated Content)内容指的是用户原创内容,也即其他具有婴幼儿属性的用户在母婴社区发布相应的UGC内容时,如果该用户的婴幼儿属性与上述待预测用户的婴幼儿属性相似,服务器100可以将该用户发布的UGC内容推送给该待预测用户的用户终端300。

[0095] PGC(Professional Generated Content)内容指的是专业生产内容。所述PGC内容关联有相应的婴幼儿属性选项,也即,相关专家在创作PGC内容时可以勾选匹配的婴幼儿属性,如果勾选的婴幼儿属性与上述待预测用户的婴幼儿属性相似,服务器100可以将所述PGC内容推送给该待预测用户的用户终端300。由此,可以为该待预测用户推送精准的UGC内容和PGC内容。

[0096] 而对消费类内容而言,可以基于所述预测结果获取对应的用户集合和用户集合中每个用户的商品数据,并根据所述对应的用户集合和用户集合中每个用户的商品数据向该用户推荐匹配的商品内容。例如,可以获取消费过内容的所有用户集合 $U\{u_1, u_2 \dots u_n\}$ 的婴幼儿属性 $P\{p_{u1}, p_{u2} \dots p_{un}\}$,根据置信阈值 μ 得到 p_{ui} 的可信分类结果集合 $C\{c|c>\mu\}$,然后根据分类结果向该用户推荐匹配的商品内容。

[0097] 由此,通过上述步骤,本实施例能够根据预测的用户婴幼儿属性,为用户提供匹配的社区内容,提升母婴电商用户向母婴社区的导流和对社区内容的兴趣度,进而提高社区内容的点击率,培养用户对内容社区的认知。

[0098] 请参阅图4,为本发明较佳实施例提供的服务器100的一种结构示意图。如图4所示,所述服务器100可以由总线110作一般性的总线体系结构来实现。根据服务器100的具体应用和整体设计约束条件,总线110可以包括任意数量的互连总线和桥接。总线110将各种电路连接在一起,这些电路包括处理器120、存储介质130和总线接口140。可选地,服务器100可以使用总线接口140将网络适配器150等经由总线110连接。网络适配器150可用于实现无线通信网络中物理层的信号处理功能,并通过天线实现射频信号的发送和接收。用户接口160可以连接外部设备,例如:键盘、显示器、鼠标或者操纵杆等。总线110还可以连接各种其它电路,如定时源、外围设备、电压调节器或者功率管理电路等,这些电路是本领域所熟知的,因此不再详述。

[0099] 可以替换的,服务器100也可配置成通用处理系统,例如通称为芯片,该通用处理系统包括:提供处理功能的一个或多个微处理器,以及提供存储介质130的至少一部分的外

部存储器,所有这些都通过外部总线体系结构与其它支持电路连接在一起。

[0100] 可替换的,服务器100可以使用下述来实现:具有处理器120、总线接口140、用户接口160的ASIC(专用集成电路);以及集成在单个芯片中的存储介质130的至少一部分,或者,服务器100可以使用下述来实现:一个或多个FPGA(现场可编程门阵列)、PLD(可编程逻辑器件)、控制器、状态机、门逻辑、分立硬件部件、任何其它适合的电路、或者能够执行本发明通篇所描述的各种功能的电路的任意组合。

[0101] 其中,处理器120负责管理总线110和一般处理(包括执行存储在存储介质130上的软件)。处理器120可以使用一个或多个通用处理器和/或专用处理器来实现。处理器120的例子包括微处理器、微控制器、DSP处理器和能够执行软件的其它电路。应当将软件广义地解释为表示指令、数据或其任意组合,而不论是将其称作为软件、固件、中间件、微代码、硬件描述语言还是其它。

[0102] 在图4中存储介质130被示为与处理器120分离,然而,本领域技术人员很容易明白,存储介质130或其任意部分可位于服务器100 之外。举例来说,存储介质130可以包括传输线、用数据调制的载波波形、和/或与无线节点分离开的计算机制品,这些介质均可以由处理器120通过总线接口140来访问。可替换地,存储介质130或其任意部分可以集成到处理器120中,例如,可以是高速缓存和/或通用寄存器。

[0103] 所述处理器120可执行上述实施例,具体地,所述存储介质130 中可以存储有所述母婴内容推荐装置200,所述处理器120可以用于执行所述母婴内容推荐装置200。

[0104] 进一步地,请参阅图5,所述母婴内容推荐装置200可包括:

[0105] 获取模块210,用于获取待预测用户的嵌入化矩阵表示,所述嵌入化矩阵表示中包括有与该待预测用户关联的商品数据;

[0106] 输入模块220,用于将所述嵌入化矩阵表示输入到所述婴幼儿属性预测模型中,输出预测结果,所述预测结果包括该待预测用户的至少一种婴幼儿属性;

[0107] 推荐模块230,用于基于所述预测结果向该待预测用户推荐对应的母婴内容。

[0108] 可以理解的是,本实施例中的各功能模块的具体操作方法可参照上述方法实施例中相应步骤的详细描述,在此不再重复赘述。

[0109] 本发明实施例还提供一种可读存储介质,所述可读存储介质中存储有计算机程序,所述计算机程序被执行时实现上述的母婴内容推荐方法。

[0110] 综上所述,本发明实施例提供的母婴内容推荐方法、装置及可读存储介质,通过获取待预测用户的嵌入化矩阵表示,并将所述嵌入化矩阵表示输入到婴幼儿属性预测模型中,输出预测结果,所述预测结果包括该待预测用户的至少一种婴幼儿属性,而后基于所述预测结果向该待预测用户推荐对应的母婴内容。由此,能够根据预测的用户婴幼儿属性,为用户提供匹配的社区内容,提升母婴电商用户向母婴社区的导流和对社区内容的兴趣度,进而提高社区内容的点击率,培养用户对内容社区的认知。

[0111] 在本发明所提供的实施例中,应该理解到,所揭露的装置和方法,也可以通过其它的方式实现。以上所描述的装置和方法实施例仅仅是示意性的,例如,附图中的流程图和框图显示了根据本发明的多个实施例的系统、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,所述模块、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方式中,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,两个连续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,框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合,可以用执行规定的功能或动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。

[0112] 另外,在本发明各个实施例中的各功能模块可以集成在一起形成一个独立的部分,也可以是各个模块单独存在,也可以两个或两个以上模块集成形成一个独立的部分。

[0113] 可以替换的,可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时,可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时,全部或部分地产生按照本发明实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中,或者从一个计算机可读存储介质向另一个计算机可读存储介质传输,例如,所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线(例如同轴电缆、光纤、数字用户线(DSL))或无线(例如红外、无线、微波等)方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质,(例如,软盘、硬盘、磁带)、光介质(例如,DVD)、或者半导体介质(例如固态硬盘 Solid State Disk(SSD))等

[0114] 需要说明的是,在本文中,术语“包括”、“包含”或者其任何其它变体意在涵盖非排它性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其它要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0115] 对于本领域技术人员而言,显然本发明不限于上述示范性实施例的细节,而且在不背离本发明的精神或基本特征的情况下,能够以其它的具体形式实现本发明。因此,无论从哪一点来看,均应将实施例看作是示范性的,而且是非限制性的,本发明的范围由所附权利要求而不是上述说明限定,因此旨在将落在权利要求的等同要件的含义和范围内的所有变化囊括在本发明内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。

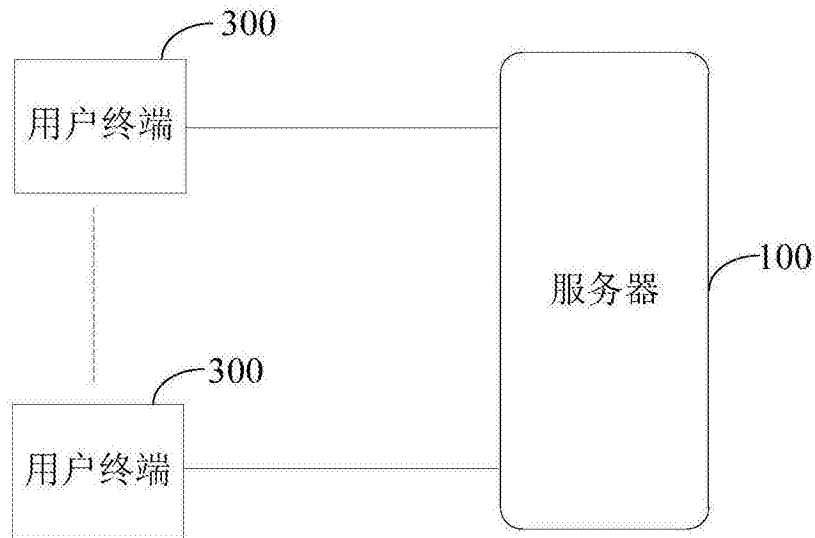


图1

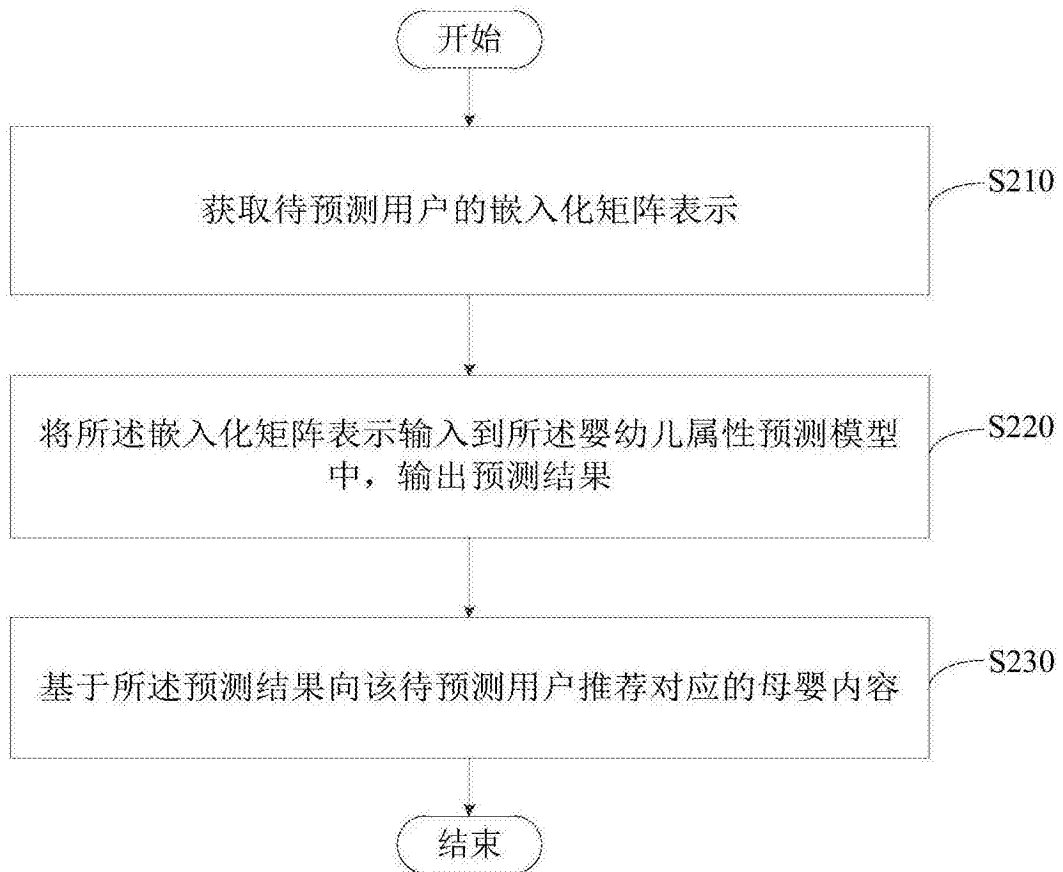


图2

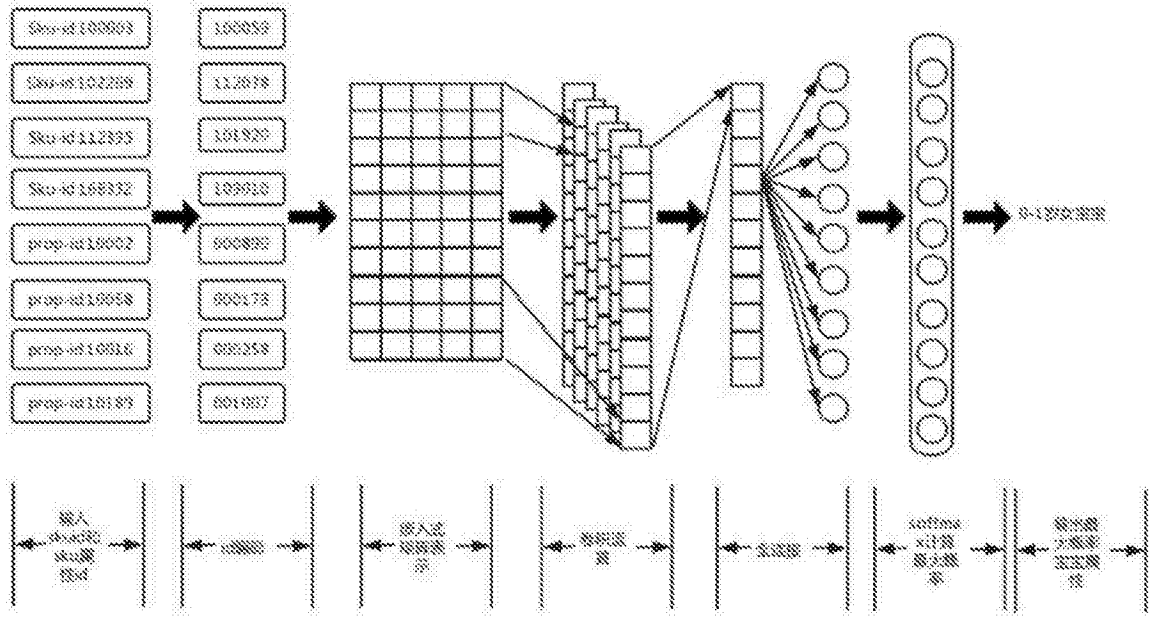


图3

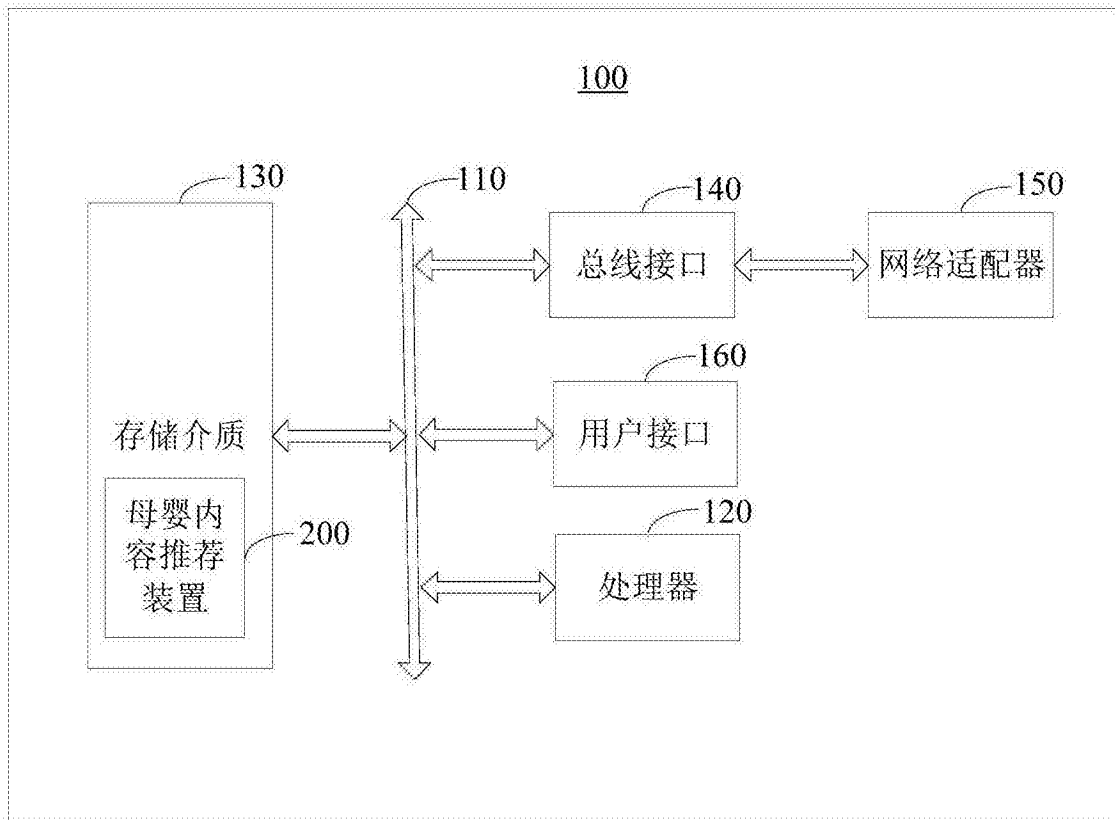


图4

200

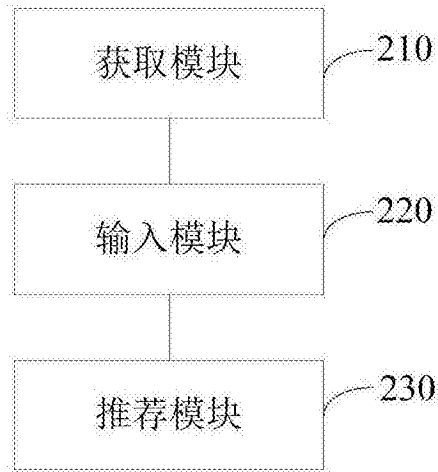


图5