



(12) 发明专利申请

(10) 申请公布号 CN 112650942 A

(43) 申请公布日 2021. 04. 13

(21) 申请号 202011513766.1

G06N 3/08 (2006.01)

(22) 申请日 2020.12.18

(71) 申请人 北京京东尚科信息技术有限公司
地址 100086 北京市海淀区知春路76号8层
申请人 北京京东世纪贸易有限公司

(72) 发明人 孙冰 陈蒙

(74) 专利代理机构 中科专利商标代理有限责任
公司 11021

代理人 王文思

(51) Int. Cl.

G06F 16/9536 (2019.01)

G06Q 30/06 (2012.01)

G06F 40/289 (2020.01)

G06F 40/30 (2020.01)

G06K 9/62 (2006.01)

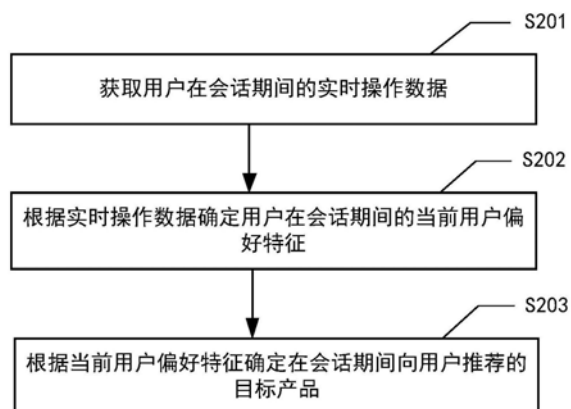
权利要求书2页 说明书13页 附图6页

(54) 发明名称

产品推荐方法、装置、计算机系统和计算机
可读存储介质

(57) 摘要

本公开提供了一种产品推荐方法、产品推荐装置、计算机系统和计算机可读存储介质。其中，产品推荐方法包括：获取用户在会话期间的实时操作数据；根据实时操作数据确定用户在会话期间的当前用户偏好特征；以及根据当前用户偏好特征确定在会话期间向用户推荐的目标产品。



1. 一种产品推荐方法,包括:

获取用户在会话期间的实时操作数据;

根据所述实时操作数据确定所述用户在所述会话期间的当前用户偏好特征;以及

根据所述当前用户偏好特征确定在所述会话期间向所述用户推荐的目标产品。

2. 根据权利要求1所述的方法,其中,根据所述实时操作数据确定所述用户在所述会话期间的当前用户偏好特征包括:

根据所述实时操作数据生成用户操作序列,其中,所述用户操作序列中包括按照操作时间顺序排列的一个或多个产品标识;

确定所述用户操作序列中的产品关键词;以及

将所述产品关键词作为所述用户在所述会话期间的当前用户偏好特征。

3. 根据权利要求1所述的方法,其中,所述实时操作数据包括用户客服对话数据,根据所述实时操作数据确定所述用户在所述会话期间的当前用户偏好特征还包括:

获取所述用户客服对话数据中的用户方输出数据,其中,所述用户方输出数据至少包括第一用户语句和第二用户语句,所述第一用户语句包括一条或多条语句;

确定所述第一用户语句中的产品关键词;

根据所述第一用户语句中的产品关键词确定所述第二用户语句中的产品关键词;以及

将所述第一用户语句中的产品关键词和所述第二用户语句中的产品关键词作为所述用户在所述会话期间的当前用户偏好特征。

4. 根据权利要求3所述的方法,其中,确定所述第一用户语句中的产品关键词包括:

确定所述第一用户语句中的多个初始产品关键词;

在所述多个初始产品关键词的数量大于第一预设值的情况下,确定每个所述初始产品关键词的输出时间;

根据每个所述初始产品关键词的输出时间,从所述多个初始产品关键词中筛选出先输出的所述第一预设值个初始产品关键词;以及

将所述先输出的所述第一预设值个初始产品关键词作为所述第一用户语句中的产品关键词。

5. 根据权利要求3所述的方法,其中,根据所述第一用户语句中的产品关键词确定所述第二用户语句中的产品关键词包括:

将所述第二用户语句进行分词处理,得到多个词;

根据所述第一用户语句中的产品关键词确定所述第二用户语句中的每个词的权重;以及

根据所述每个词的权重确定所述第二用户语句中的产品关键词。

6. 根据权利要求1所述的方法,其中,根据所述当前用户偏好特征确定在所述会话期间向所述用户推荐的目标产品包括:

获取所述会话期间的产品候选集;

根据特征属性确定所述产品候选集中的待推荐产品,其中,所述特征属性包括以下至少之一:用户属性、产品属性和由所述当前用户偏好特征得到的属性;

计算所述待推荐产品与所述当前用户偏好特征的相似度;

根据所述相似度对所述待推荐产品进行排序,得到排序结果;以及

将所述排序结果中的前第二预设值个所述待推荐产品作为在所述会话期间向所述用户推荐的目标产品。

7. 一种产品推荐装置, 包括:

获取模块, 用于获取用户在会话期间的实时操作数据;

第一确定模块, 用于根据所述实时操作数据确定所述用户在所述会话期间的当前用户偏好特征; 以及

第二确定模块, 用于根据所述当前用户偏好特征确定在所述会话期间向所述用户推荐的目标产品。

8. 根据权利要求7所述的装置, 其中, 所述第一确定模块包括:

生成单元, 用于根据所述实时操作数据生成用户操作序列, 其中, 所述用户操作序列中包括按照操作时间顺序排列的一个或多个产品标识;

第一确定单元, 用于确定所述用户操作序列中的产品关键词; 以及

第一定义单元, 用于将所述产品关键词作为所述用户在所述会话期间的当前用户偏好特征。

9. 一种计算机系统, 包括:

一个或多个处理器;

存储器, 用于存储一个或多个程序,

其中, 当所述一个或多个程序被所述一个或多个处理器执行时, 使得所述一个或多个处理器实现权利要求1至6中任一项所述的方法。

10. 一种计算机可读存储介质, 其上存储有可执行指令, 该指令被处理器执行时使处理器实现权利要求1至6中任一项所述的方法。

产品推荐方法、装置、计算机系统和计算机可读存储介质

技术领域

[0001] 本公开涉及计算机、互联网技术领域,更具体地,涉及一种产品推荐方法、产品推荐装置、计算机系统和计算机可读存储介质。

背景技术

[0002] 随着互联网技术的快速发展,电子商务已逐步深入日常生活。为了应对不同用户的喜好和需求,在电子商务领域通常会设计相关的产品推荐技术,以便可以针对性的为不同的用户推荐更符合自身偏好和需求的产品。

[0003] 在实现本公开构思的过程中,发明人发现相关技术中至少存在如下问题,现有的产品推荐技术均是基于历史数据确定用户偏好,得到的推荐结果更趋向于用户历史偏好,导致无法为用户较为准确的推荐符合用户偏好的产品。

发明内容

[0004] 有鉴于此,本公开提供了一种产品推荐方法、产品推荐装置、计算机系统和计算机可读存储介质。

[0005] 本公开的一个方面提供了一种产品推荐方法,包括:获取用户在会话期间的实时操作数据;根据所述实时操作数据确定所述用户在所述会话期间的当前用户偏好特征;以及根据所述当前用户偏好特征确定在所述会话期间向所述用户推荐的目标产品。

[0006] 根据本公开的实施例,根据所述实时操作数据确定所述用户在所述会话期间的当前用户偏好特征包括:根据所述实时操作数据生成用户操作序列,其中,所述用户操作序列中包括按照操作时间顺序排列的一个或多个产品标识;确定所述用户操作序列中的产品关键词;以及将所述产品关键词作为所述用户在所述会话期间的当前用户偏好特征。

[0007] 根据本公开的实施例,所述实时操作数据包括用户客服对话数据,根据所述实时操作数据确定所述用户在所述会话期间的当前用户偏好特征还包括:获取所述用户客服对话数据中的用户方输出数据,其中,所述用户方输出数据至少包括第一用户语句和第二用户语句,所述第一用户语句包括一条或多条语句;确定所述第一用户语句中的产品关键词;根据所述第一用户语句中的产品关键词确定所述第二用户语句中的产品关键词;以及将所述第一用户语句中的产品关键词和所述第二用户语句中的产品关键词作为所述用户在所述会话期间的当前用户偏好特征。

[0008] 根据本公开的实施例,确定所述第一用户语句中的产品关键词包括:确定所述第一用户语句中的多个初始产品关键词;在所述多个初始产品关键词的数量大于第一预设值的情况下,确定每个所述初始产品关键词的输出时间;根据每个所述初始产品关键词的输出时间,从所述多个初始产品关键词中筛选出先输出的所述第一预设值个初始产品关键词;以及将所述先输出的所述第一预设值个初始产品关键词作为所述第一用户语句中的产品关键词。

[0009] 根据本公开的实施例,根据所述第一用户语句中的产品关键词确定所述第二用户

语句中的产品关键词包括：将所述第二用户语句进行分词处理，得到多个词；根据所述第一用户语句中的产品关键词确定所述第二用户语句中的每个词的权重；以及根据所述每个词的权重确定所述第二用户语句中的产品关键词。

[0010] 根据本公开的实施例，根据所述当前用户偏好特征确定在所述会话期间向所述用户推荐的目标产品包括：获取所述会话期间的产品候选集；根据特征属性确定所述产品候选集中的待推荐产品，其中，所述特征属性包括以下至少之一：用户属性、产品属性和由所述当前用户偏好特征得到的属性；计算所述待推荐产品与所述当前用户偏好特征的相似度；根据所述相似度对所述待推荐产品进行排序，得到排序结果；以及将所述排序结果中的前第二预设值个所述待推荐产品作为在所述会话期间向所述用户推荐的目标产品。

[0011] 本公开的另一个方面提供了一种产品推荐装置，包括：获取模块，用于获取用户在会话期间的实时操作数据；第一确定模块，用于根据所述实时操作数据确定所述用户在所述会话期间的当前用户偏好特征；以及第二确定模块，用于根据所述当前用户偏好特征确定在所述会话期间向所述用户推荐的目标产品。

[0012] 根据本公开的实施例，所述第一确定模块包括：生成单元，用于根据所述实时操作数据生成用户操作序列，其中，所述用户操作序列中包括按照操作时间顺序排列的一个或多个产品标识；第一确定单元，用于确定所述用户操作序列中的产品关键词；以及第一定义单元，用于将所述产品关键词作为所述用户在所述会话期间的当前用户偏好特征。

[0013] 根据本公开的实施例，所述实时操作数据包括用户客服对话数据，所述第一确定模块还包括：第一获取单元，用于获取所述用户客服对话数据中的用户方输出数据，其中，所述用户方输出数据至少包括第一用户语句和第二用户语句，所述第一用户语句包括一条或多条语句；第二确定单元，用于确定所述第一用户语句中的产品关键词；第三确定单元，用于根据所述第一用户语句中的产品关键词确定所述第二用户语句中的产品关键词；以及第二定义单元，用于将所述第一用户语句中的产品关键词和所述第二用户语句中的产品关键词作为所述用户在所述会话期间的当前用户偏好特征。

[0014] 根据本公开的实施例，所述第二确定单元包括：第一确定子单元，用于确定所述第一用户语句中的多个初始产品关键词；第二确定子单元，用于在所述多个初始产品关键词的数量大于第一预设值的情况下，确定每个所述初始产品关键词的输出时间；筛选子单元，用于根据每个所述初始产品关键词的输出时间，从所述多个初始产品关键词中筛选出先输出的所述第一预设值个初始产品关键词；以及定义子单元，用于将所述先输出的所述第一预设值个初始产品关键词作为所述第一用户语句中的产品关键词。

[0015] 根据本公开的实施例，所述第三确定单元包括：分词处理单元，用于将所述第二用户语句进行分词处理，得到多个词；第三确定子单元，用于根据所述第一用户语句中的产品关键词确定所述第二用户语句中的每个词的权重；以及第四确定子单元，用于根据所述每个词的权重确定所述第二用户语句中的产品关键词。

[0016] 根据本公开的实施例，所述第二确定模块包括：第二获取单元，用于获取所述会话期间的产品候选集；第四确定单元，用于根据特征属性确定所述产品候选集中的待推荐产品，其中，所述特征属性包括以下至少之一：用户属性、产品属性和由所述当前用户偏好特征得到的属性；计算单元，用于计算所述待推荐产品与所述当前用户偏好特征的相似度；排序单元，用于根据所述相似度对所述待推荐产品进行排序，得到排序结果；以及第三定义单

元,用于将所述排序结果中的前第二预设值个所述待推荐产品作为在所述会话期间向所述用户推荐的目标产品。

[0017] 本公开的另一方面提供了一种计算机系统,包括:一个或多个处理器;存储器,用于存储一个或多个程序,其中,当所述一个或多个程序被所述一个或多个处理器执行时,使得所述一个或多个处理器实现如上所述的方法。

[0018] 本公开的另一方面提供了一种计算机可读存储介质,其上存储有计算机可执行指令,所述指令在被执行时用于实现如上所述的方法。

[0019] 本公开的另一方面提供了一种计算机程序,所述计算机程序包括计算机可执行指令,所述指令在被执行时用于实现如上所述的方法。

[0020] 根据本公开的实施例,通过采用了获取用户在会话期间的实时操作数据;根据实时操作数据确定用户在会话期间的当前用户偏好特征;以及根据当前用户偏好特征确定在会话期间向用户推荐的目标产品的技术手段,由于根据用户在会话中的实时操作数据确定了当前用户偏好特征,所以至少部分地克服了根据历史数据无法实时确定用户当前偏好的技术问题,进而达到了可以实时推荐符合用户当前偏好的产品技术效果。

附图说明

[0021] 通过以下参照附图对本公开实施例的描述,本公开的上述以及其他目的、特征和优点将更为清楚,在附图中:

[0022] 图1示意性示出了根据本公开实施例的可以应用产品推荐方法的示例性系统架构;

[0023] 图2示意性示出了根据本公开实施例的产品推荐方法的流程图;

[0024] 图3示意性示出了根据本公开实施例的基于用户点击浏览序列构建当前用户偏好特征的示意图;

[0025] 图4示意性示出了根据本公开实施例的基于用户客服对话构建当前用户偏好特征的示意图;

[0026] 图5示意性示出了根据本公开实施例的seq2seq模型的训练过程;

[0027] 图6示意性示出了根据本公开实施例的seq2seq模型的预测过程;

[0028] 图7示意性示出了根据本公开实施例的产品推荐系统的整体框架图;

[0029] 图8示意性示出了根据本公开的实施例的产品推荐装置的框图;以及

[0030] 图9示意性示出了根据本公开实施例的适于实现上述的产品推荐方法的计算机系统的框图。

具体实施方式

[0031] 以下,将参照附图来描述本公开的实施例。但是应该理解,这些描述只是示例性的,而并非要限制本公开的范围。在下面的详细描述中,为便于解释,阐述了许多具体的细节以提供对本公开实施例的全面理解。然而,明显地,一个或多个实施例在没有这些具体细节的情况下也可以被实施。此外,在以下说明中,省略了对公知结构和技术的描述,以避免不必要地混淆本公开的概念。

[0032] 在此使用的术语仅仅是为了描述具体实施例,而并非意在限制本公开。在此使用

的术语“包括”、“包含”等表明了所述特征、步骤、操作和/或部件的存在,但是并不排除存在或添加一个或多个其他特征、步骤、操作或部件。

[0033] 在此使用的所有术语(包括技术和科学术语)具有本领域技术人员通常所理解的含义,除非另外定义。应注意,这里使用的术语应解释为具有与本说明书的上下文相一致的含义,而不应以理想化或过于刻板的方式来解释。

[0034] 在使用类似于“A、B和C等中至少一个”这样的表述的情况下,一般来说应该按照本领域技术人员通常理解该表述的含义来予以解释(例如,“具有A、B和C中至少一个的系统”应包括但不限于单独具有A、单独具有B、单独具有C、具有A和B、具有A和C、具有B和C、和/或具有A、B、C的系统等)。在使用类似于“A、B或C等中至少一个”这样的表述的情况下,一般来说应该按照本领域技术人员通常理解该表述的含义来予以解释(例如,“具有A、B或C中至少一个的系统”应包括但不限于单独具有A、单独具有B、单独具有C、具有A和B、具有A和C、具有B和C、和/或具有A、B、C的系统等)。

[0035] 产品推荐技术的推荐算法大致可以归为四类:基于内容的推荐算法、协同过滤推荐算法、基于知识的推荐算法和混合推荐算法。

[0036] 但在实现本公开构思的过程中,发明人发现基于内容的推荐算法,是根据用户选择和评价的对象,推荐与用户历史喜欢的对象相似的物品,但是在用户为新用户的情况下,由于新用户尚未产生历史数据,导致基于内容的推荐算法无法对新用户进行推荐,此外,该算法在训练分类器的过程中需要大量数据。

[0037] 发明人还发现协同过滤推荐算法主要有两类,基于用户的协同过滤推荐算法和基于物品的协同过滤推荐算法。基于用户的协同过滤推荐算法的思想是,如果存在多个用户对某些物品的评分比较相似,则可以预测这些用户对其他物品的评分也相似,就可以根据这些用户的偏好确定推荐至其他相似用户的推荐物品;基于物品的协同过滤推荐算法则认为用户对不同物品(例如可以为球鞋和球衣)评分存在相似性,就可以根据用户对一个物品(如球鞋)的偏好确定推荐至该用户的另一个物品(如球衣)。但是对于协同过滤推荐算法,存在数据稀疏的问题,由于数据稀疏分布,会导致算法计算复杂度高,此外,该算法还存在无法为新用户推荐产品,以及新产品得不到推荐的问题。

[0038] 发明人还发现基于知识的推荐算法不是建立在用户需要和偏好基础上的推荐,而是利用针对特定领域制定规则来进行推荐,从而达到推荐的目的。基于知识的推荐算法可以根据用户的领域将用户的需求映射到产品,而且非产品属性也纳入了考虑范围,但是对用户和产品的知识(例如领域等)的挖掘很困难,而且适应性差,导致该算法无法很好地移植到别的产品上。

[0039] 组合推荐算法是将以上三种算法加以组合以避免或减少各自单独使用的不足,目前应用最广泛的是基于内容推荐和基于协同过滤推荐的组合。

[0040] 综合以上四种推荐算法,发明人还发现,由其构建的推荐系统存在一个共同的问题,即如果本次推荐的结果准确度不高,不是用户想要的产品,那么这种情况下推荐系统不能做到进一步纠正推荐结果,不能和用户继续交互以挖掘用户需求,提供最适合用户产品的功能。即当前的研究通常假设一个一次性的交互范式,其中用户偏好是根据过去历史观察到的行为来估计的,并且按照等级排序的推荐列表是用户交互的主要的且是单向的形式。尽管这种方法在各个领域都很常见,也很有用,但是它可能有一些潜在的限制。例如,在

许多应用程序场景中,当用户登录网站时,他们未必已经知道自己的偏好,即用户的首选项,用户可能是在点击浏览商品中确定其首选项,也可能是在和客服的对话中表达自己的偏好,此时,用户首选项无法从其过去的历史中可靠的估计出来。

[0041] 有鉴于此,本公开的实施例提供了一种产品推荐方法、产品推荐装置、计算机系统和计算机可读存储介质。该方法包括获取用户在会话期间的实时操作数据;根据实时操作数据确定用户在会话期间的当前用户偏好特征;以及根据当前用户偏好特征确定在会话期间向用户推荐的目标产品。

[0042] 图1示意性示出了根据本公开实施例的可以应用产品推荐方法的示例性系统架构100。需要注意的是,图1所示仅为可以应用本公开实施例的系统架构的示例,以帮助本领域技术人员理解本公开的技术内容,但并不意味着本公开实施例不可以用于其他设备、系统、环境或场景。

[0043] 如图1所示,根据该实施例的系统架构100可以包括终端设备101、102、103,网络104和服务器105。网络104用以在终端设备101、102、103和服务器105之间提供通信链路的介质。网络104可以包括各种连接类型,例如有线和/或无线通信链路等等。

[0044] 用户可以使用终端设备101、102、103通过网络104与服务器105交互,以接收或发送消息等。终端设备101、102、103上可以安装有各种通讯客户端应用,例如可以包括各种购物类应用和单子商务类应用等。

[0045] 终端设备101、102、103可以是具有显示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。

[0046] 服务器105可以是提供各种服务的服务器,例如对用户利用终端设备101、102、103所浏览的网站提供支持的后台管理服务器。后台管理服务器可以对接收到的用户请求等数据进行分析等处理,并将处理结果(例如根据用户请求获取或生成的网页、信息、或数据等)反馈给终端设备。

[0047] 需要说明的是,本公开实施例所提供的产品推荐方法一般可以由服务器105执行。相应地,本公开实施例所提供的产品推荐装置一般可以设置于服务器105中。本公开实施例所提供的产品推荐方法也可以由不同于服务器105且能够与终端设备101、102、103和/或服务器105通信的服务器或服务器集群执行。相应地,本公开实施例所提供的产品推荐装置也可以设置于不同于服务器105且能够与终端设备101、102、103和/或服务器105通信的服务器或服务器集群中。或者,本公开实施例所提供的产品推荐方法也可以由终端设备101、102或103执行,或者也可以由不同于终端设备101、102或103的其他终端设备执行。相应地,本公开实施例所提供的产品推荐装置也可以设置于终端设备101、102或103中,或设置于不同于终端设备101、102或103的其他终端设备中。

[0048] 例如,实时操作数据可以是原本作用于终端设备101、102、或103中的任意一个(例如,终端设备101,但不限于此)之上的操作数据,或者作用于外部设备上并可以导入到终端设备101中的操作数据。然后,终端设备101可以在本地执行本公开实施例所提供的产品推荐方法,或者将实时操作数据发送到其他终端设备、服务器、或服务器集群,并由接收该实时操作数据的其他终端设备、服务器、或服务器集群来执行本公开实施例所提供的产品推荐方法。

[0049] 应该理解,图1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。

[0050] 图2示意性示出了根据本公开实施例的产品推荐方法的流程图。

[0051] 如图2所示,该方法包括操作S201~S203。

[0052] 在操作S201,获取用户在会话期间的实时操作数据。

[0053] 根据本公开的实施例,上述会话例如可以是终端用户与交互系统进行通讯的整个过程,上述会话期间例如可以是会话的开始和结束之间的一段时间。其中,会话的开始例如可以是用户打开电商网站或应用,会话的结束例如可以是用户断开连接,或者长时间没响应,在上述会话期间用户可以进行各种与获取产品信息相关的操作,这些操作即构成上述实时操作数据。

[0054] 在操作S202,根据实时操作数据确定用户在会话期间的当前用户偏好特征。

[0055] 根据本公开的实施例,上述实时操作数据由于均与产品相关,可以基于这些相关的产品提取产品特征,并将这些产品特征作为上述当前用户偏好特征,以此来记录当前会话期间的用户偏好和需求。

[0056] 需要说明的是,上述实时操作数据例如可以包括多条,需要实时记录与每条实时操作数据相关的产品的特征,以得到完整的用户偏好和需求,保证后续推荐结果的准确度。

[0057] 在操作S203,根据当前用户偏好特征确定在会话期间向用户推荐的目标产品。

[0058] 根据本公开的实施例,结合操作S202中确定的当前会话期间的用户偏好和需求,从海量的产品候选集中筛选出符合该偏好和需求的部分产品作为目标产品并推荐至该用户。

[0059] 通过本公开的上述实施例,由于可以在会话期间实时获取用户的兴趣点或者偏好,并将该兴趣点或者偏好作为确定在该会话期间推荐至用户的目标产品的依据,达到了可以根据用户实时偏好动态的为用户推荐符合用户首选项产品的目的,解决了现有推荐算法无法从过去的历史中可靠的估计在当前会话期间的用户首选项的问题,从而可以极大的满足用户需求,提升用户满意度。

[0060] 根据本公开的实施例,上述当前用户偏好特征的来源例如可以包括两部分,第一部分例如是用户在会话期间点击浏览商品的序列,第二部分例如是用户在和客服对话过程中用户的话术。

[0061] 下面结合具体实施例,并参考图3~图4,对如何构建上述当前用户偏好特征做进一步详细说明。

[0062] 根据本公开的一实施例,上述当前用户偏好特征例如为基于用户点击浏览序列构建。在此种情况下,上述操作S202包括:根据实时操作数据生成用户操作序列,其中,用户操作序列中包括按照操作时间顺序排列的一个或多个产品标识;确定用户操作序列中的产品关键词;以及将产品关键词作为用户在会话期间的当前用户偏好特征。

[0063] 图3示意性示出了根据本公开实施例的基于用户点击浏览序列构建当前用户偏好特征的示意图。

[0064] 一般来说,当用户登录购物网站或者APP时,通常会先点击浏览自己想要购买的产品(例如可以为生活日用品、食品、电子产品或其他商品等,此处以商品为例),本实施例就是根据用户的点击行为,实时获取用户偏好。在现实场景下,由于用户首先看到的往往是商品的标题,而不是商品的ID,所以,本实施例中利用用户在会话期间点击浏览的序列中的商

品标题信息,构建一个商品标题的集合,通过分析该商品标题集合,提取关键词,能够根据该关键词实时获取用户在短期内的偏好。

[0065] 如图3所示,例如用户在当前会话期间点击浏览了多个序列,根据该多个序列确定了其对应的商品标题集合(详细内容如图3所示),最终通过分析该些商品标题集合提取出了在本次会话当前阶段的关键词为5G、华为、8GB和全网通,即将该关键词作为当前用户偏好特征,从而完成当前用户偏好特征的构建。其中,提取关键词的过程例如可以包括:首先进行分词、去停用词等处理,然后按照词频选择前k(可人为设定)个词,作为(商品)关键词。

[0066] 通过本公开的上述实施例,由于当前用户偏好特征是根据高频关键词来构建,避免了用户在实际场景中因受到电商平台广告推荐的影响意外点击并不符合用户首选项的产品(例如用户在浏览手机的过程中意外点击了一件上衣)而导致关键词提取不准确的问题,通过高频关键词构建的当前用户偏好特征可以充分代表此刻用户的部分偏好,从而可以更准确的为用户推荐符合当前用户首选项的产品。

[0067] 根据本公开的另一实施例,上述当前用户偏好特征例如为基于用户客服对话构建。在此种情况下,上述实时操作数据例如包括用户客服对话数据,上述操作S202包括:获取用户客服对话数据中的用户方输出数据,其中,用户方输出数据至少包括第一用户语句和第二用户语句,第一用户语句包括一条或多条语句;确定第一用户语句中的产品关键词;根据第一用户语句中的产品关键词确定第二用户语句中的产品关键词;以及将第一用户语句中的产品关键词和第二用户语句中的产品关键词作为用户在会话期间的当前用户偏好特征。

[0068] 图4示意性示出了根据本公开实施例的基于用户客服对话构建当前用户偏好特征的示意图。

[0069] 在电商真实场景会话中,一般情况下用户会先自行点击浏览一些商品,这些代表了用户在当前会话期间某一时刻的一部分偏好,接着用户会向客服发起咨询,在和客服的对话过程中,也会有商品ID序列或者体现用户偏好的文本数据信息,此时可从该文本数据信息中提取关键词,实时构建用户偏好。

[0070] 如图4所示,其中用户发出的链接表示当前浏览的商品,连接中的100012015170为商品ID,后面用户提到的“5G、超级快充”都将作为关键词提取出来,用以完成当前用户偏好特征的构建。

[0071] 通过本公开的上述实施例,由于在提取关键词的过程中还考虑到了用户客服对话中的文本数据信息,可进一步提取到例如在用户点击浏览序列的过程中不能得到的一部分特征,或可进一步增加对已有特征(关键词)的关注度,即可以进一步确定用户的偏好,从而能够更加准确的预判用户的当前首选项,使最终推荐结果更为符合用户需求。

[0072] 根据本公开的实施例,上述根据第一用户语句中的产品关键词确定第二用户语句中的产品关键词的过程包括:将第二用户语句进行分词处理,得到多个词;根据第一用户语句中的产品关键词确定第二用户语句中的每个词的权重;以及根据每个词的权重确定第二用户语句中的产品关键词。

[0073] 根据本公开的实施例,上述第一用户语句和第二用户语句中的产品关键词的生成均采用了seq2seq模型(sequence to sequence,一种序列到序列的自然语言生成模型),该模型的架构如图5~图6所示。

[0074] 图5示意性示出了根据本公开实施例的seq2seq模型的训练过程。

[0075] 如图5所示,seq2seq模型包含一个编码器(encoder,可以对输入数据进行编码表示,得到表征向量)和一个解码器(decoder,用于对编码结果进行解码,输出词概率分布),编码器负责对输入的文本数据按时序进行编码,通常都采用RNN模型(Recurrent Neural Network,循环神经网络,也可以替换为如LSTM、GRU等神经网络模型)。

[0076] 在训练过程中,首先准备好训练数据,训练数据例如可以为用户客服对话中的用户方输出数据(例如可以为上述的第一用户语句和第二用户语句,在这里简写为句子1、句子2、句子3、句子4等),人工从每个用户语句中提取出若干关键词,训练时将该些关键词送入seq2seq模型的编码器进行编码,seq2seq模型的解码器的输出结果就是对应的该些关键词。据此完成对seq2seq模型的训练。

[0077] 图6示意性示出了根据本公开实施例的seq2seq模型的预测过程。

[0078] 如图6所示,展示了seq2seq模型的预测过程,它和训练过程基本一致,区别主要在于预测过程中只需要输入实际用户客服对话中的用户方输出数据(即句子1、句子2、句子3、句子4等),模型的解码器会自动生成对应该些句子的一系列关键词。

[0079] 根据本公开的实施例,为了增加上下文之间的语义连贯性,每一次训练时,会把当前时刻的前N行文本一起送给seq2seq模型的编码器进行编码,以训练生成上述关键词,在该过程中,还采用了attention机制,通过根据上一个输入文本(例如可以为第一用户输入语句)输出的关键词赋予下一个输入文本(例如可以为第二用户输入语句)不同的权重,来确定该下一个输入文本的关键词。

[0080] 结合上述图4~图6的具体实施例,本实施例中,在训练阶段,在第一用户语句仅包括一条语句的情况下,第一用户语句例如为句子1,关键词1例如为句子1中对应的关键词,此时第二用户语句例如为句子2,关键词2例如可以包括关键词1和句子2对应的关键词;在第一用户语句包括多条语句的情况下,第一用户语句例如包括句子1、句子2和句子3,此时第二用户语句例如为句子4,关键词4例如可以包括关键词1、关键词2、关键词3和句子4对应的关键词。基于该准则训练完成的seq2seq模型,在其预测阶段的预测关键词1、2、3和4等遵循同样的准则。

[0081] 根据本公开的实施例,在上述第一用户语句对应的预测关键词较多或是该第一用户语句包括多条语句(如包括句子1、句子2、句子3等)而生产较多预测关键词的情况下,上述确定第一用户语句中的产品关键词的过程包括:确定第一用户语句中的多个初始产品关键词;在多个初始产品关键词的数量大于第一预设值的情况下,确定每个初始产品关键词的输出时间;根据每个初始产品关键词的输出时间,从多个初始产品关键词中筛选出先输出的第一预设值个初始产品关键词;以及将先输出的第一预设值个初始产品关键词作为第一用户语句中的产品关键词。

[0082] 在本实施例中,上述初始产品关键词例如为句子1、句子2和句子3对应的所有预测关键词,在该所有预测关键词的个数例如大于N(即上述第一预设值,可人为设定)的情况下,下一时刻解码(即对句子4进行预测)时会将当前句子(即句子4)、历史输出的前N个关键词(即句子1、句子2和句子3对应的所有预测关键词的前N个,其中,N值要和训练过程保持一致,可人为设定),一起送给解码器完成该下一时刻的预测输出。

[0083] 通过本公开的上述实施例,由于训练了具体模型对用户克服对话中的关键词进行

提取,可通过在会话期间的各种用户交互方式更为全面的确定用户实时偏好,有利于进一步提升用户满意度。

[0084] 根据本公开的上述实施例,上述操作S203包括:获取会话期间的产品候选集;根据特征属性确定产品候选集中的待推荐产品,其中,特征属性包括以下至少之一:用户属性、产品属性和由当前用户偏好特征得到的属性;计算待推荐产品与当前用户偏好特征的相似度;根据相似度对待推荐产品进行排序,得到排序结果;以及将排序结果中的前第二预设值个待推荐产品作为在会话期间向用户推荐的目标产品。

[0085] 图7示意性示出了根据本公开实施例的产品推荐系统的整体框架图。

[0086] 参见图7所示,本实施例中例如将产品(仍以商品为例)推荐系统分为召回阶段(采用召回算法)和排序阶段(采用排序模型)两个阶段。

[0087] 在召回阶段,采用召回算法并引入用户画像和特征工程,来实现从海量的商品库(即商品候选集)里,快速召回一部分用户潜在感兴趣的商品(即待推荐商品)。本实施例中,用户画像例如可以包括用户的人群属性、历史行为、兴趣内容和偏好倾向等多维度的用户属性;特征工程例如可以包括商品的通用属性、销量、上架时间、前述用户画像以及上述实时构建的当前用户偏好特征等属性;召回算法采用多路召回的策略,该多路召回例如可以包括userCF(基于用户的协同过滤算法)召回、热度召回、新品召回、搜索召回等。其中,userCF召回用于找出与当前用户相似的一批其他用户,然后从这些其他用户最近购买过的商品中过滤掉该当前用户购买的商品,剩余的作为召回候选集(即待推荐商品);热度召回用于根据商品的销量、点击量作为热度召回的来源,销量需要每天维护更新;新品召回用于根据商品上架时间,取一部分最近上新的商品作为召回候选集;搜索召回用于通过在商城网站搜索关键字,取topK(即前K个,此处K用于描述上述第二预设值)商品也作为待推荐商品的一部分。

[0088] 通过上述召回阶段的多种召回方式,无需专业的知识,可以做到自动推荐,容易发现用户的兴趣点,同时还解决了新物品的得不到推荐以及数据稀疏等问题。

[0089] 在排序阶段,接收上述召回阶段召回的商品候选集,同时结合上述用户画像和特征工程实现最终的排序。本实施例中,排序模型采用GBDT(Gradient Boosting Decision Tree,梯度提升决策树)模型。GBDT是一种基于迭代思路构造的决策树算法,该算法在实际问题中将生成多棵决策树,并将所有树的结果进行汇总来得到最终答案,该算法将决策树与集成思想进行了有效的结合,通过将弱学习器提升为强学习器的集成方法来提高预测精度。GBDT是一类泛化能力较强的学习算法。

[0090] 结合上述图3~图7的具体实施例,在生成了一系列代表用户偏好的关键词之后,将关键词作为实时构建的当前用户偏好特征指导对召回的待推荐商品进行排序。该排序过程例如包括:首先采用DSSM(Deep Structured Semantic Models,深度语义匹配模型)模型计算关键词和召回的待推荐商品的标题等的语义相似度,即相似度越高,代表越符合用户的当前偏好;之后,在构建好所有的特征(包括上述用户画像和特征工程中的所有属性)后,用GBDT排序,并取排序结果的topK个待推荐商品作为目标商品推荐给用户。

[0091] 通过本公开的上述实施例,由于将用户在当前会话期间的点击浏览序列和用户客服对话等一系列行为生成的关键词作为用户实时偏好,并通过计算关键词和待推荐商品的予以相似度构建当前用户偏好特征参与到GBDT的排序中,可以显着提高现有的例如电商等

推荐系统的推荐效果。

[0092] 图8示意性示出了根据本公开的实施例的产品推荐装置的框图。

[0093] 如图8所示,产品推荐装置800包括获取模块810、第一确定模块820和第二确定模块830。

[0094] 获取模块810,用于获取用户在会话期间的实时操作数据。

[0095] 第一确定模块820,用于根据实时操作数据确定用户在会话期间的当前用户偏好特征。

[0096] 第二确定模块830,用于根据当前用户偏好特征确定在会话期间向用户推荐的目标产品。

[0097] 根据本公开的实施例,上述第一确定模块包括生成单元、第一确定单元和第一定义单元。

[0098] 生成单元,用于根据实时操作数据生成用户操作序列,其中,用户操作序列中包括按照操作时间顺序排列的一个或多个产品标识。

[0099] 第一确定单元,用于确定用户操作序列中的产品关键词。

[0100] 第一定义单元,用于将产品关键词作为用户在会话期间的当前用户偏好特征。

[0101] 根据本公开的实施例,上述实时操作数据包括用户客服对话数据,上述第一确定模块还包括第一获取单元、第二确定单元、第三确定单元和第二定义单元。

[0102] 第一获取单元,用于获取用户客服对话数据中的用户方输出数据,其中,用户方输出数据至少包括第一用户语句和第二用户语句,第一用户语句包括一条或多条语句。

[0103] 第二确定单元,用于确定第一用户语句中的产品关键词。

[0104] 第三确定单元,用于根据第一用户语句中的产品关键词确定第二用户语句中的产品关键词。

[0105] 第二定义单元,用于将第一用户语句中的产品关键词和第二用户语句中的产品关键词作为用户在会话期间的当前用户偏好特征。

[0106] 根据本公开的实施例,上述第二确定单元包括第一确定子单元、第二确定子单元、筛选子单元和定义子单元。

[0107] 第一确定子单元,用于确定第一用户语句中的多个初始产品关键词。

[0108] 第二确定子单元,用于在多个初始产品关键词的数量大于第一预设值的情况下,确定每个初始产品关键词的输出时间。

[0109] 筛选子单元,用于根据每个初始产品关键词的输出时间,从多个初始产品关键词中筛选出先输出的第一预设值个初始产品关键词。

[0110] 定义子单元,用于将先输出的第一预设值个初始产品关键词作为第一用户语句中的产品关键词。

[0111] 根据本公开的实施例,上述第三确定单元包括分词处理单元、第三确定子单元和第四确定子单元。

[0112] 分词处理单元,用于将第二用户语句进行分词处理,得到多个词。

[0113] 第三确定子单元,用于根据第一用户语句中的产品关键词确定第二用户语句中的每个词的权重。

[0114] 第四确定子单元,用于根据每个词的权重确定第二用户语句中的产品关键词。

[0115] 根据本公开的实施例,上述第二确定模块包括第二获取单元、第四确定单元、计算单元、排序单元和第三定义单元。

[0116] 第二获取单元,用于获取会话期间的产品候选集。

[0117] 第四确定单元,用于根据特征属性确定产品候选集中的待推荐产品,其中,特征属性包括以下至少之一:用户属性、产品属性和由当前用户偏好特征得到的属性。

[0118] 计算单元,用于计算待推荐产品与当前用户偏好特征的相似度。

[0119] 排序单元,用于根据相似度对待推荐产品进行排序,得到排序结果。

[0120] 第三定义单元,用于将排序结果中的前第二预设值个待推荐产品作为在会话期间向用户推荐的目标产品。

[0121] 根据本公开的实施例的模块、单元、子单元中的任意多个、或其中任意多个的至少部分功能可以在一个模块中实现。根据本公开实施例的模块、单元、子单元中的任意一个或多个可以被拆分成多个模块来实现。根据本公开实施例的模块、单元、子单元中的任意一个或多个可以至少被部分地实现为硬件电路,例如现场可编程门阵列(FPGA)、可编程逻辑阵列(PLA)、片上系统、基板上的系统、封装上的系统、专用集成电路(ASIC),或可以通过对电路进行集成或封装的任何其他的合理方式的硬件或固件来实现,或以软件、硬件以及固件三种实现方式中任意一种或以其中任意几种的适当组合来实现。或者,根据本公开实施例的模块、单元、子单元中的一个或多个可以至少被部分地实现为计算机程序模块,当该计算机程序模块被运行时,可以执行相应的功能。

[0122] 例如,获取模块810、第一确定模块820和第二确定模块830中的任意多个可以合并在一个模块/单元/子单元中实现,或者其中的任意一个模块/单元/子单元可以被拆分成多个模块/单元/子单元。或者,这些模块/单元/子单元中的一个或多个模块/单元/子单元的至少部分功能可以与其他模块/单元/子单元的至少部分功能相结合,并在一个模块/单元/子单元中实现。根据本公开的实施例,获取模块810、第一确定模块820和第二确定模块830中的至少一个可以至少被部分地实现为硬件电路,例如现场可编程门阵列(FPGA)、可编程逻辑阵列(PLA)、片上系统、基板上的系统、封装上的系统、专用集成电路(ASIC),或可以通过对电路进行集成或封装的任何其他的合理方式等硬件或固件来实现,或以软件、硬件以及固件三种实现方式中任意一种或以其中任意几种的适当组合来实现。或者,获取模块810、第一确定模块820和第二确定模块830中的至少一个可以至少被部分地实现为计算机程序模块,当该计算机程序模块被运行时,可以执行相应的功能。

[0123] 需要说明的是,本公开的实施例中产品推荐装置部分与本公开的实施例中产品推荐方法部分是相对应的,产品推荐装置部分的描述具体参考产品推荐方法部分,在此不再赘述。

[0124] 图9示意性示出了根据本公开实施例的适于实现上述的产品推荐方法的计算机系统的框图。图9示出的计算机系统仅仅是一个示例,不应对本公开实施例的功能和使用范围带来任何限制。

[0125] 如图9所示,根据本公开实施例的计算机系统900包括处理器901,其可以根据存储在只读存储器(ROM) 902中的程序或者从存储部分908加载到随机访问存储器(RAM) 903中的程序而执行各种适当的动作和处理。处理器901例如可以包括通用微处理器(例如CPU)、指令集处理器和/或相关芯片组和/或专用微处理器(例如,专用集成电路(ASIC)),等等。处理

器901还可以包括用于缓存用途的板载存储器。处理器901可以包括用于执行根据本公开实施例的方法流程的不同动作的单一处理单元或者是多个处理单元。

[0126] 在RAM 903中,存储有系统900操作所需的各种程序和数据。处理器901、ROM 902以及RAM 903通过总线904彼此相连。处理器901通过执行ROM 902和/或RAM 903中的程序来执行根据本公开实施例的方法流程的各种操作。需要注意,所述程序也可以存储在除ROM 902和RAM 903以外的一个或多个存储器中。处理器901也可以通过执行存储在所述一个或多个存储器中的程序来执行根据本公开实施例的方法流程的各种操作。

[0127] 根据本公开的实施例,系统900还可以包括输入/输出(I/O)接口905,输入/输出(I/O)接口905也连接至总线904。系统900还可以包括连接至I/O接口905的以下部件中的一项或多项:包括键盘、鼠标等的输入部分906;包括诸如阴极射线管(CRT)、液晶显示器(LCD)等以及扬声器等的输出部分907;包括硬盘等的存储部分908;以及包括诸如LAN卡、调制解调器等的网络接口卡的通信部分909。通信部分909经由诸如因特网的网络执行通信处理。驱动器910也根据需要连接至I/O接口905。可拆卸介质911,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910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908。

[0128] 根据本公开的实施例,根据本公开实施例的方法流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存储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909从网络上被下载和安装,和/或从可拆卸介质911被安装。在该计算机程序被处理器901执行时,执行本公开实施例的系统中限定的上述功能。根据本公开的实施例,上文描述的系统、设备、装置、模块、单元等可以通过计算机程序模块来实现。

[0129] 本公开还提供了一种计算机可读存储介质,该计算机可读存储介质可以是上述实施例中描述的设备/装置/系统中所包含的;也可以是单独存在,而未装配入该设备/装置/系统中。上述计算机可读存储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执行时,实现根据本公开实施例的方法。

[0130] 根据本公开的实施例,计算机可读存储介质可以是非易失性的计算机可读存储介质。例如可以包括但不限于: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)、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。

[0131] 例如,根据本公开的实施例,计算机可读存储介质可以包括上文描述的ROM 902和/或RAM 903和/或ROM 902和RAM 903以外的一个或多个存储器。

[0132] 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。本领域技术人员可以理解,本公开的各个实施例和/或权利要求中记载的特征可以进行多种组合和/或结合,即使这样的组合或结合没有明确记载于本公开中。特别地,在不脱离本公开精神和教导的情况下,本公开的各个实施例和/或权利要求中记载的特征可以进行多种组合和/或结合。所有这些组合和/或结合均落入本公开的范围。

[0133] 以上对本公开的实施例进行了描述。但是,这些实施例仅仅是为了说明的目的,而并非为了限制本公开的范围。尽管在以上分别描述了各实施例,但是这并不意味着各个实施例中的措施不能有利地结合使用。本公开的范围由所附权利要求及其等同物限定。不脱离本公开的范围,本领域技术人员可以做出多种替代和修改,这些替代和修改都应落在本公开的范围之内。

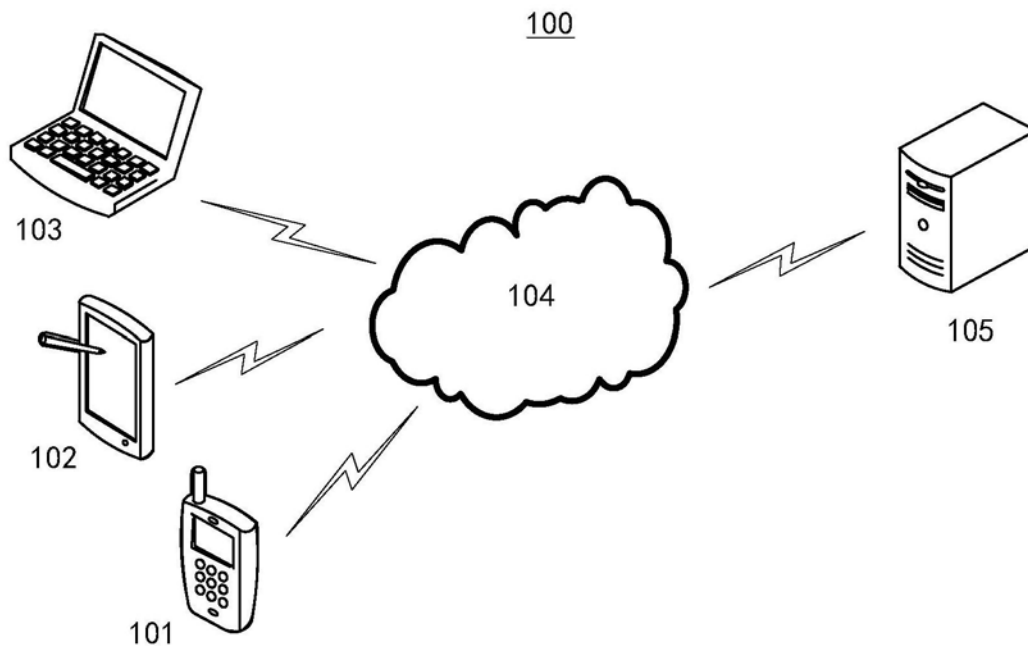


图1

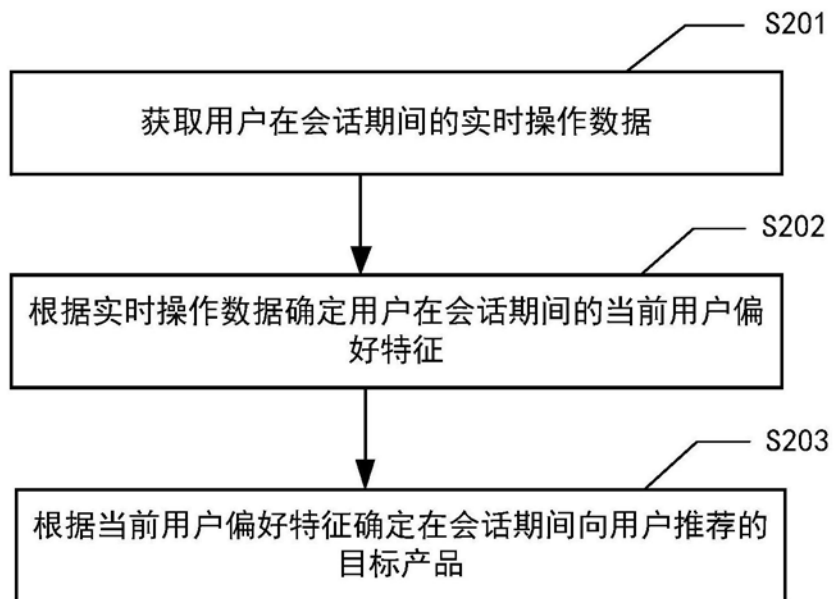


图2

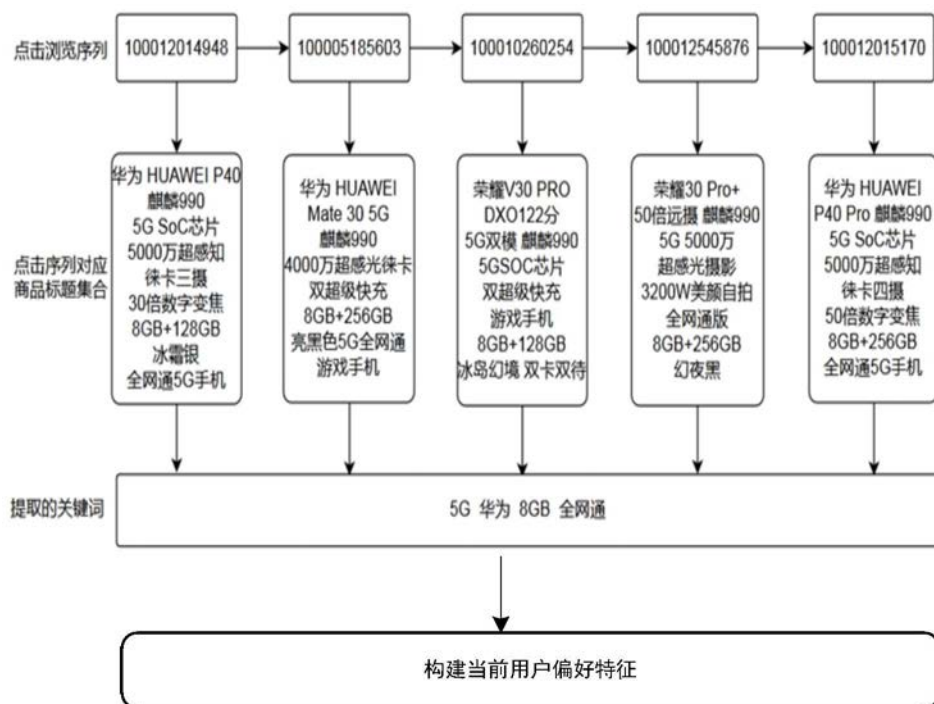


图3

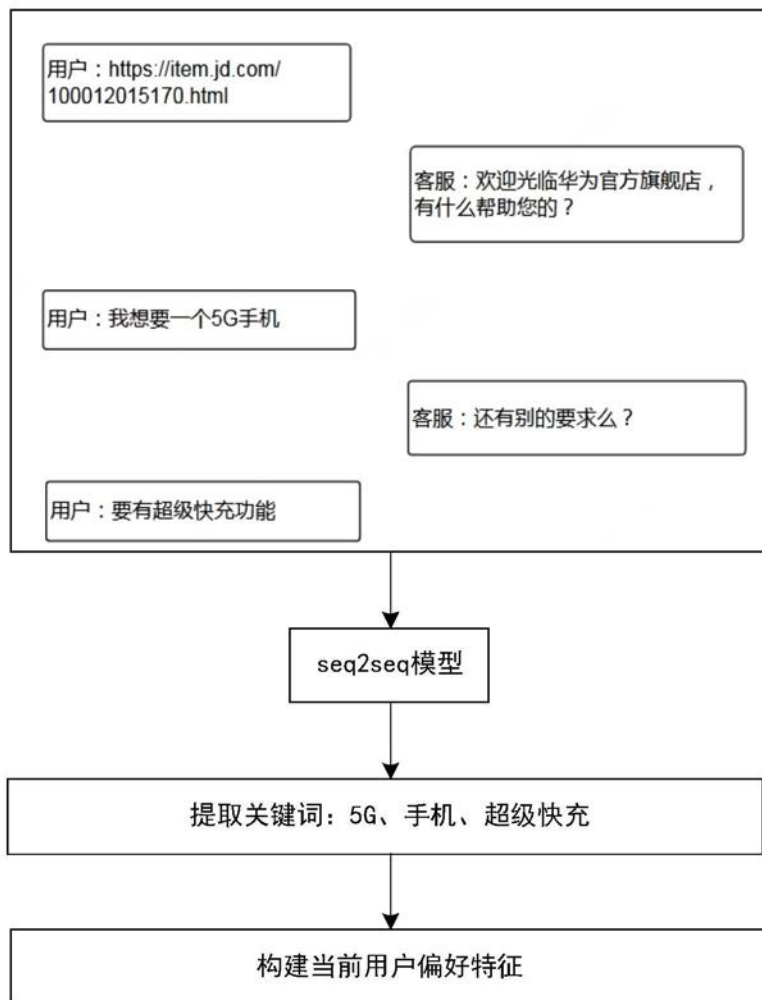


图4

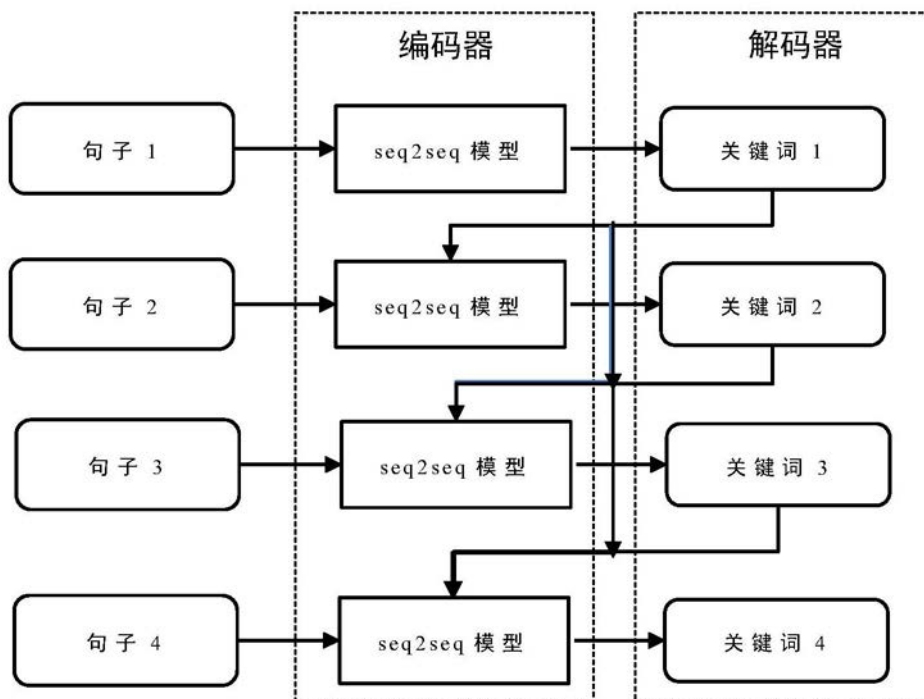


图5

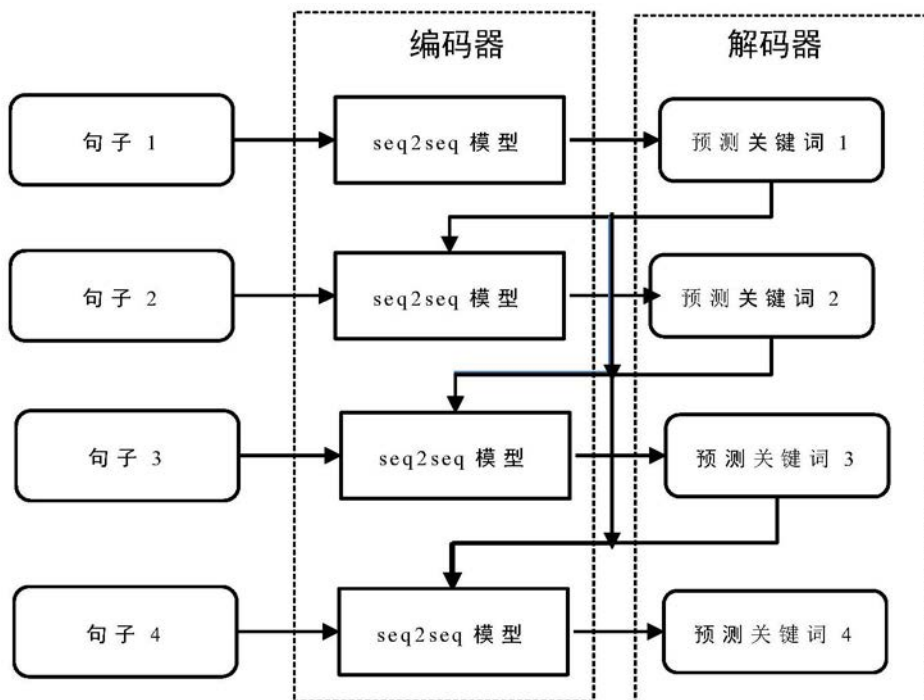


图6

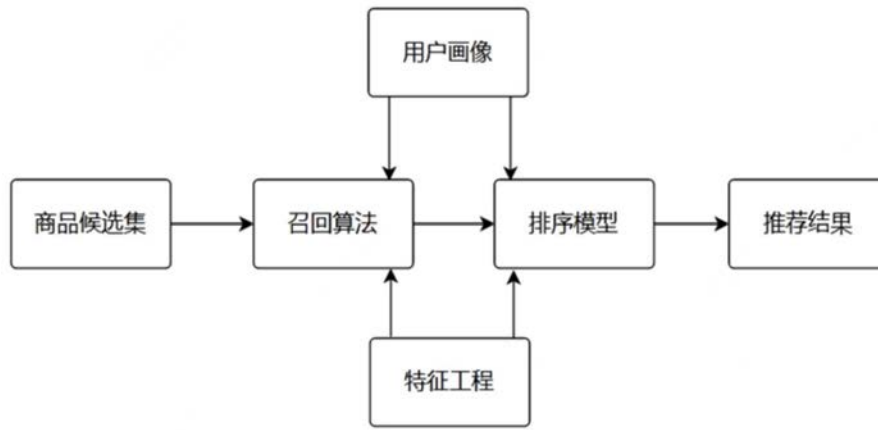


图7

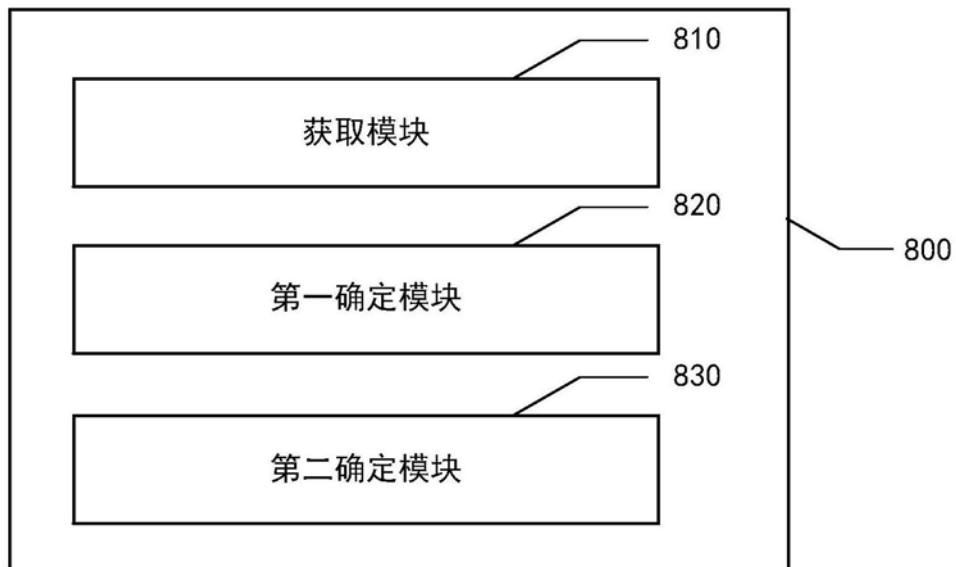


图8

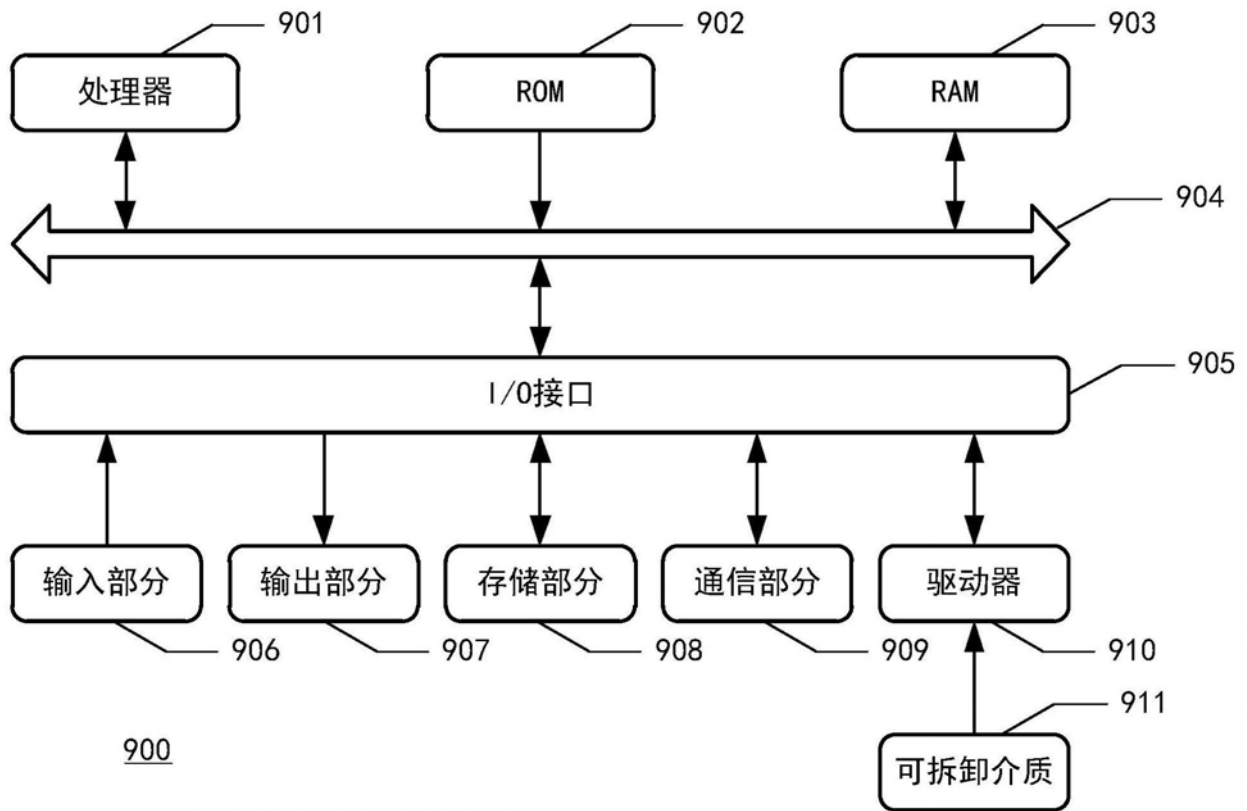


图9