(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 113763096 B
(45)授权公告日 2025.01.17

(21)申请号 202011466850.2
(22)申请日 2020.12.14
(65)同一申请的已公布的文献号
申请公布号 CN 113763096 A
(43)申请公布日 2021.12.07
(73)专利权人 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
专利权人 北京京东世纪贸易有限公司
(72)发明人 潘博 陈蒙
(74)专利代理机构 中科专利商标代理有限责任公司 11021
专利代理师 王文思
(51)Int.Cl.
*G06Q 30/0601*(2023.01)
*G06F 16/2457*(2019.01)

(56)对比文件
CN 109285030 A,2019.01.29

审查员 刘辉

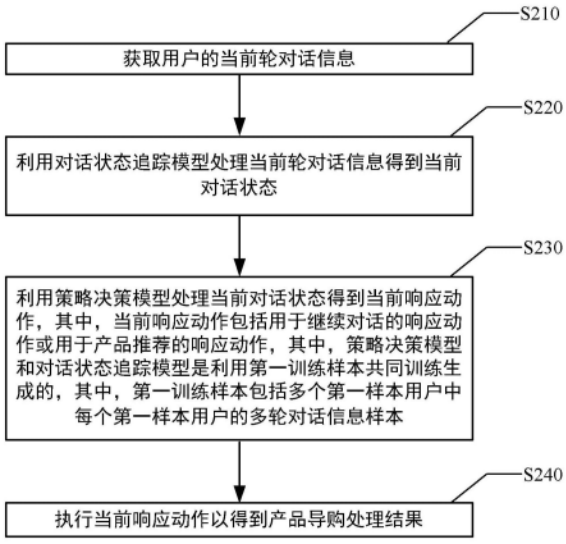
权利要求书3页 说明书15页 附图8页

(54)发明名称

产品导购处理方法、装置、电子设备、介质及程序产品

(57)摘要

本公开实施例提供了产品导购处理方法、装置、电子设备、介质及程序产品。该方法包括:获取用户的当前轮对话信息;利用对话状态追踪模型处理当前轮对话信息得到当前对话状态,当前对话状态包括多个产品属性中每个产品属性的当前概率分布,每个当前概率分布用于表征与产品属性对应的一个或多个产品属性值中每个产品属性值的当前概率;利用策略决策模型处理当前对话状态得到当前响应动作,当前响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作,策略决策模型和对话状态追踪模型是利用第一训练样本共同训练生成的,第一训练样本包括多个第一样本用户中每个第一样本用户的多轮对话信息样本;执行当前响应动作以得到产品导购处理结果。




CN 113763096 B

1.一种产品导购处理方法,包括:

获取用户的当前轮对话信息;

利用对话状态追踪模型处理所述当前轮对话信息得到当前对话状态,其中,所述当前对话状态包括多个产品属性中每个产品属性的当前概率分布,其中,每个所述当前概率分布用于表征与所述产品属性对应的一个或多个产品属性值中每个所述产品属性值的当前概率;

利用策略决策模型处理所述当前对话状态得到当前响应动作,其中,所述当前响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作,其中,所述策略决策模型和所述对话状态追踪模型是利用第一训练样本,共同训练策略网络模型和多个循环神经网络模型生成的,其中,所述策略决策模型与所述策略网络模型相对应,所述对话状态追踪模型与所述多个循环神经网络模型相对应,所述策略网络模型和所述多个循环神经网络模型的网络参数是根据相同的目标函数输出值进行调整的,所述第一训练样本包括多个第一样本用户中每个第一样本用户的多轮对话信息样本;以及

执行所述当前响应动作以得到产品导购处理结果。

2.根据权利要求1所述的方法,其中,所述利用策略决策模型处理所述当前对话状态得到当前响应动作,包括:

利用策略决策模型处理所述当前对话状态得到当前向量,其中,所述当前向量包括当前推荐概率和所述多个产品属性中每个产品属性的当前预测概率;以及

根据所述当前向量,确定所述当前响应动作。

3.根据权利要求2所述的方法,其中,所述根据所述当前向量,确定所述当前响应动作,包括:

从所述当前推荐概率和多个所述当前预测概率中确定最大概率;

在确定所述最大概率为所述当前预测概率的情况下,确定所述当前响应动作为用于继续对话的响应动作;以及

在确定所述最大概率为所述当前推荐概率的情况下,确定所述当前响应动作为用于产品推荐的响应动作。

4.根据权利要求3所述的方法,其中,所述当前响应动作包括用于继续对话的响应动作;

所述执行所述响应动作以得到产品导购处理结果,包括:

确定目标产品属性,其中,所述目标产品属性为与所述最大概率对应的产品属性;以及

将所述目标产品属性确定为与所述用户继续对话的回复信息。

5.根据权利要求1或3所述的方法,其中,所述当前响应动作包括用于产品推荐的响应动作;

所述执行所述响应动作以得到产品导购处理结果,包括:

获取所述用户的特征信息和候选产品的特征信息;以及

利用产品推荐模型处理所述用户的特征信息、所述候选产品的特征信息和所述当前对话状态,得到推荐产品,其中,所述产品推荐模型是利用第二训练样本训练生成的,其中,所述第二训练样本包括多个第二样本用户中每个第二样本用户的特征信息、对话状态样本和真实兴趣产品,以及,多个样本产品中每个样本产品的特征信息。

6.根据权利要求5所述的方法,其中,所述产品推荐模型是利用第二训练样本训练生成的,包括:

利用所述第二训练样本训练因子分解机模型,得到所述产品推荐模型。

7.根据权利要求1所述的方法,其中,所述策略决策模型和所述对话状态追踪模型是利用第一训练样本共同训练生成的,包括:

针对每个所述第一样本用户,将所述多轮对话信息样本中每轮对话信息样本分别输入多个循环神经网络模型,输出多个产品属性中每个产品属性的预测概率分布,其中,每个所述循环神经网络模型用于输出对应的所述产品属性的预测概率分布;

针对每轮所述对话信息样本,根据多个产品属性中每个产品属性的预测概率分布生成本轮对话状态;

将本轮对话状态输入策略网络模型中输出与本轮对话对应的本轮向量,其中,所述本轮向量包括本轮推荐概率和所述多个产品属性中每个产品属性的本轮预测概率;

根据所述本轮向量,确定与本轮对话对应的响应动作;

执行所述与本轮对话对应的响应动作,得到用户针对所述响应动作的反馈结果,其中,所述与本轮对话对应的响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作;

确定与所述反馈结果对应的反馈得分;

将多轮对话的响应概率和反馈得分输入目标函数中,输出目标函数的输出值,其中,所述响应概率为与所述响应动作对应的概率,所述响应概率为所述本轮推荐概率或所述本轮预测概率;

根据所述输出值调整所述多个循环神经网络模型和所述策略网络模型的网络参数,直至所述目标函数收敛;以及

将训练后得到的多个循环神经网络模型作为所述对话状态追踪模型,将训练后得到的策略网络模型作为所述策略决策模型。

8.根据权利要求7所述的方法,其中,所述确定与所述反馈结果对应的反馈得分,包括:

在确定所述反馈结果为购买推荐产品的情况下,确定与所述反馈结果对应的反馈得分为第一奖励得分;

在确定所述反馈结果为浏览推荐产品且未购买所述产品的情况下,确定与所述反馈结果对应的反馈得分为第二奖励得分,其中,所述第一奖励得分大于所述第二奖励得分;

在确定所述反馈结果为继续对话或未浏览推荐产品的情况下,确定与所述反馈结果对应的反馈得分为第一惩罚得分;

在确定所述反馈结果为终止对话的情况下,确定与所述反馈结果对应的反馈得分为第二惩罚得分,其中,所述第一惩罚得分大于所述第二惩罚得分。

9.一种产品导购处理装置,包括:

获取模块,用于获取用户的当前轮对话信息;

第一处理模块,用于利用对话状态追踪模型处理所述当前轮对话信息得到当前对话状态,其中,所述当前对话状态包括多个产品属性中每个产品属性的当前概率分布,其中,每个所述当前概率分布用于表征与所述产品属性对应的一个或多个产品属性值中每个所述产品属性值的概率;

第二处理模块,用于利用策略决策模型处理所述当前对话状态得到当前响应动作,其中,所述当前响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作,其中,所述策略决策模型和所述对话状态追踪模型是利用第一训练样本,共同训练策略网络模型和多个循环神经网络模型生成的,其中,所述策略决策模型与所述策略网络模型相对应,所述对话状态追踪模型与所述多个循环神经网络模型相对应,所述策略网络模型和所述多个循环神经网络模型的网络参数是根据相同的目标函数输出值进行调整的,所述第一训练样本包括多个第一样本用户中每个第一样本用户的多轮对话信息样本;以及

执行模块,用于执行所述当前响应动作以得到产品导购处理结果。

10.一种电子设备,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1~8中任一项所述的方法。

11.一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1~8中任一项所述的方法。

12.一种计算机程序产品,所述计算机程序产品包括计算机程序,所述计算机程序被处理器执行时用于实现权利要求1~8中任一项所述的方法。

# 产品导购处理方法、装置、电子设备、介质及程序产品

## 技术领域

[0001] 本公开实施例涉及计算机技术领域，更具体地，涉及一种产品导购处理方法、装置、电子设备、介质及程序产品。

## 背景技术

[0002] 随着计算机技术的不断发展，提供了通过产品导购的方式向用户推荐产品，以提高销售转换率。

[0003] 在相关技术中，通常采用基于搜索对话方式或推荐方式实现产品导购。其中，基于搜索对话方式是通过与用户之间的多轮对话信息，预测用户的购买意图，并根据预测出的购买意图，向用户推荐产品。基于推荐方式是根据用户的历史行为记录，向用户推荐产品。

[0004] 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：采用相关技术实现产品导购的销售转换率不高。

## 发明内容

[0005] 有鉴于此，本公开实施例提供了一种产品导购处理方法、装置、电子设备、介质及程序产品。

[0006] 本公开实施例的一个方面提供了一种产品导购处理方法，包括：获取用户的当前轮对话信息；利用对话状态追踪模型处理上述当前轮对话信息得到当前对话状态，其中，上述当前对话状态包括多个产品属性中每个产品属性的当前概率分布，其中，每个上述当前概率分布用于表征与上述产品属性对应的一个或多个产品属性值中每个上述产品属性值的当前概率；利用策略决策模型处理上述当前对话状态得到当前响应动作，其中，上述当前响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作，其中，上述策略决策模型和上述对话状态追踪模型是利用第一训练样本共同训练生成的，其中，上述第一训练样本包括多个第一样本用户中每个第一样本用户的多轮对话信息样本；以及，执行上述当前响应动作以得到产品导购处理结果。

[0007] 本公开实施例的另一个方面提供了一种产品导购处理装置，包括：获取模块，用于获取用户的当前轮对话信息；第一处理模块，用于利用对话状态追踪模型处理上述当前轮对话信息得到当前对话状态，其中，上述当前对话状态包括多个产品属性中每个产品属性的当前概率分布，其中，每个上述当前概率分布用于表征与上述产品属性对应的一个或多个产品属性值中每个上述产品属性值的概率；第二处理模块，用于利用策略决策模型处理上述当前对话状态得到当前响应动作，其中，上述当前响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作，其中，上述策略决策模型和上述对话状态追踪模型是利用第一训练样本共同训练生成的，其中，上述第一训练样本包括多个第一样本用户中每个第一样本用户的多轮对话信息样本；以及，执行模块，用于执行上述当前响应动作以得到产品导购处理结果。

[0008] 本公开实施例的另一个方面提供了一种电子设备，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,当上述一个或多个程序被上述一个或多个处理器执行时,使得上述一个或多个处理器实现如上所述的方法。

[0009] 本公开实施例的另一个方面提供了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如上所述的方法。

[0010] 本公开实施例的另一个方面提供了一种计算机程序产品,上述计算机程序产品包括计算机程序,上述计算机程序被处理器执行时用于实现如上所述的方法。

[0011] 根据本公开的实施例,通过获取用户的当前轮对话信息,利用对话状态追踪模型处理当前轮对话信息得到当前对话状态,当前对话状态包括多个产品属性中每个产品属性的当前概率分布,利用策略决策模型处理当前对话状态得到当前响应动作,当前响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作,策略决策模型和对话状态追踪模型是利用第一训练样本共同训练生成的,并执行当前响应动作以得到产品导购处理结果。由于利用策略决策模型处理基于对话状态追踪模型得到的当前对话状态,能够确定当前响应动作是需要与用户继续交流以明确用户需求的响应动作,还是进行产品推荐的响应动作,实现了在合适的时间向用户推荐合适的产品,因而,至少部分地克服了采用相关技术实现产品导购的销售转换率不高的技术问题,提高了产品导购的销售转换率。

## 附图说明

[0012] 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:

[0013] 图1示意性示出了根据本公开实施例的可以应用产品导购方法的示例性系统架构;

[0014] 图2示意性示出了根据本公开实施例的一种产品导购处理方法的流程图;

[0015] 图3示意性示出了根据本公开实施例的一种当前轮对话信息的示意图;

[0016] 图4示意性示出了根据本公开实施例的一种用于继续对话的响应动作的示意图;

[0017] 图5示意性示出了根据本公开实施例的一种用于产品推荐的响应动作的示意图;

[0018] 图6示意性示出了根据本公开实施例的另一种产品导购处理方法的流程图;

[0019] 图7示意性示出了根据本公开实施例的一种产品导购处理方法的应用示意图;

[0020] 图8示意性示出了根据本公开的实施例的一种产品导购处理装置的框图;以及

[0021] 图9示意性示出了根据本公开实施例的适于实现产品导购处理方法的框图。

## 具体实施方式

[0022] 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。

[0023] 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。

[0024] 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。

[0025] 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)。

[0026] 在相关技术中,发明人发现由于基于搜索对话方式没有考虑用户的历史行为记录,而通常根据用户的历史行为记录,可以发现用户的潜在兴趣点,因此,可能会错失一些潜在兴趣点比较高的产品,进而降低了销售转换率。由于基于推荐方式无法获取到用户的当前需求信息,因此,可能会导致推荐的产品并不符合用户的需求,进而降低了销售转换率。

[0027] 在实现本公开构思的过程中,发明人发现提高销售转换率的方式可以为在合适的时间向用户推荐合适的产品。其中,所谓在合适的时间向用户推荐产品即为在充分了解用户的当前需求信息后,再进行产品推荐。所谓推荐合适的产品即为推荐既符合用户的当前需求信息,也包括用户的潜在兴趣点的产品。发明人还发现基于搜索对话方式可以获取用户的当前需求信息,基于推荐方式可以获取用户的潜在兴趣点,由此,可以将两种方式结合起来使用,但还发现直接将两种方式结合起来并不能解决问题,这是由于实现向用户推荐合适的产品的前提是获取到了足够能够用于向用户进行产品推荐的当前需求信息,如果还没有获取到足够的当前需求信息,便向用户进行产品推荐,则推荐的产品可能与用户的实际需求不符,相反的,如果已经获取到足够的当前需求信息,可以向用户进行产品推荐而没有进行产品推荐反而继续与用户进行对话以继续获取当前需求信息,则会造成不必要的资源消耗,可能还会引起用户的反感,因此,在与用户的对话过存在着在当前时间节点上,是需要继续基于搜索对话方式继续与用户对话获取用户的当前需求信息,以进一步明确用户的当前需求,还是需要基于推荐方式向用户推荐合适的产品的问题。

[0028] 为了解决两种方式结合过程中所存在的问题,发明人提出了一种基于策略决策模型将两种实现方式结合使用的技术方案,即本公开的实施例提供了一种产品导购处理方法、产品导购处理装置以及能够应用该方法的电子设备。该方法包括预测过程和执行过程。在预测过程中,获取用户的当前轮对话信息,利用对话状态追踪模块处理当前轮对话信息得到当前对话状态,当前对话状态包括多个产品属性中每个产品属性的当前概率分布,每个当前概率分布用于表征与产品属性对应的一个或多个产品属性值中每个产品属性值的当前概率。利用策略决策模块处理当前对话状态得到当前响应动作,当前响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作,策略决策模型和对话状态追踪模型是利用第一训练样本共同训练生成的,第一训练样本包括多个第一样本用户中每个第一样本用户的多轮对话信息样本。在执行过程,执行当前响应动作以得到产品导购处理结果。

[0029] 图1示意性示出了根据本公开实施例的可以应用产品导购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。

[0030] 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。

[0031] 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、邮箱客户端和/或社交平台软件等(仅为示例)。

[0032] 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。

[0033] 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(仅为示例)。后台管理服务器可以对接收到的用户的导购对话请求进行分析等处理,并将当前响应动作反馈给终端设备。

[0034] 需要说明的是,本公开实施例所提供的产品导购处理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导购处理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导购处理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导购处理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导购处理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导购处理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。

[0035] 例如,对话状态追踪模型和策略决策模型可以原本存储在终端设备101、102、或103中的任意一个(例如,终端设备101,但不限于此)之中,或者存储在外部存储设备上并可以导入到终端设备101中。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导购处理方法,或者将对话状态追踪模型和策略决策模型发送到其他终端设备、服务器、或服务器集群,并由接收该对话状态追踪模型和策略决策模型的其他终端设备、服务器、或服务器集群来执行本公开实施例所提供的产品导购处理方法。

[0036] 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、网络和服务器。

[0037] 图2示意性示出了根据本公开实施例的一种产品导购处理方法的流程图。

[0038] 如图2所示,该方法包括操作S210～S240。

[0039] 在操作S210,获取用户的当前轮对话信息。

[0040] 根据本公开的实施例,可以获取与用户进行对话产生的当前轮对话信息,其中,用户可以利用语音、文字和图像中的至少一种方式进行对话。当前轮对话信息可以包括针对产品的多个产品属性,当前轮对话信息可以包括多组对话。

[0041] 示例性的,如用户通过触发应用程序上的对话请求控件进入对话界面,并在文本框中输入针对产品的产品需求信息,响应用户输入的产品需求信息,生成用户的当前轮对

话信息。图3示意性示出了根据本公开实施例的一种当前轮对话信息的示意图。图3中用户需要获得针对豆浆机的产品推荐。用户的当前轮对话信息可以包括与用户进行的两组对话。

[0042] 在操作S220,利用对话状态追踪模型处理当前轮对话信息得到当前对话状态,其中,当前对话状态包括多个产品属性中每个产品属性的当前概率分布,其中,每个当前概率分布用于表征与产品属性对应的一个或多个产品属性值中每个产品属性值的当前概率。

[0043] 根据本公开的实施例,对话状态追踪模型可以基于循环神经网络模型训练生成。由于当前轮对话信息可以包括多个产品属性,因此,每个产品属性可以具有对应的循环神经网络模型。

[0044] 在获得用户的当前轮对话信息后,可以利用对话状态追踪模型处理当前轮对话信息,得到当前轮对话状态。其中,当前轮对话状态可以包括多个产品属性中每个产品属性的当前概率分布,每个产品属性可以对应一个或多个产品属性值,每个产品属性值具有对应的当前概率,产品属性值的概率可以表征当产品属性的产品属性值为该产品属性值的概率。

[0045] 示例性的,如产品属性为颜色,与颜色对应的产品属性值包括红色、黄色和蓝色。与红色对应的当前概率为0.3,与黄色对应的当前概率为0.1,与蓝色对应的当前概率为0.6。颜色的当前概率分布为[0.3 0.1 0.6]。

[0046] 在操作S230,利用策略决策模型处理当前对话状态得到当前响应动作,其中,当前响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作,其中,策略决策模型和对话状态追踪模型是利用第一训练样本共同训练生成的,其中,第一训练样本包括多个第一样本用户中每个第一样本用户的多轮对话信息样本。

[0047] 在操作S240,执行当前响应动作以得到产品导购处理结果。

[0048] 根据本公开的实施例,策略决策模型可以用于根据当前对话状态,确定当前响应动作。策略决策模型和对话状态追踪模型可以是利用第一训练样本联合训练生成的,即策略决策模型和对话状态追踪模型训练完成的条件相同。策略决策模型可以是策略网络模型。

[0049] 根据本公开的实施例,用于继续对话的响应动作可以表征需要继续与用户进行对话,以进一步明确用户的当前需求信息。用于产品推荐的响应动作可以表征向用户推荐产品。其中,可以利用策略决策模型处理当前对话状态得到当前向量,当前向量包括推荐概率和多个产品属性中每个产品属性的预测概率,并根据当前向量,确定当前响应动作,其中,推荐概率可以用于表征进行产品推荐的概率。在根据当前向量,确定当前响应动作,可以包括根据推荐概率和多个预测概率,确定最大概率,如果最大概率为预测概率,则可以确定当前响应动作为用于继续对话的响应动作。如果最大概率为推荐概率,则可以确定当前响应动作为用于产品推荐的响应动作。

[0050] 根据本公开的实施例,如果当前响应动作为用于继续对话的响应动作,则可以将最大概率对应的产品属性确定为目标产品属性,将目标产品属性确定为与用户继续对话的回复信息。如果当前响应动作为用于产品推荐的响应动作,则可以获取用户的特征信息和候选产品的特征信息,利用产品推荐模型处理用户的特征信息、候选产品的特征信息和当前对话状态,得到推荐产品。

[0051] 根据本公开实施例的技术方案，通过获取用户的当前轮对话信息，利用对话状态追踪模型处理当前轮对话信息得到当前对话状态，当前对话状态包括多个产品属性中每个产品属性的当前概率分布，利用策略决策模型处理当前对话状态得到当前响应动作，当前响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作，策略决策模型和对话状态追踪模型是利用第一训练样本共同训练生成的，并执行当前响应动作以得到产品导购处理结果。由于利用策略决策模型处理基于对话状态追踪模型得到的当前对话状态，能够确定当前响应动作是需要与用户继续交流以明确用户需求的响应动作，还是进行产品推荐的响应动作，实现了在合适的时间向用户推荐合适的产品，因而，至少部分地克服了采用相关技术实现产品导购的销售转换率不高的技术问题，提高了产品导购的销售转换率。

[0052] 根据本公开的实施例，利用策略决策模型处理当前对话状态得到当前响应动作，可以包括如下操作。

[0053] 利用策略决策模型处理当前对话状态得到当前向量，其中，当前向量包括当前推荐概率和多个产品属性中每个产品属性的当前预测概率。根据当前向量，确定当前响应动作。

[0054] 根据本公开的实施例，当前推荐概率可以用于表征在当前轮进行产品推荐的概率。每个产品属性的当前预测概率可以用于表征在当前轮中该产品属性的预测概率。可以根据当前推荐概率和多个当前预测概率，确定最大概率，并根据最大概率，确定当前响应动作。

[0055] 示例性的，如产品属性的数量包括L个，每个产品属性的当前预测概率可以表示为$B_j$，其中，$j\in\{1,2,\ldots\ldots,L-1,L\}$。当前推荐概率可以表示为$B_0$。当前向量可以表示为$[B_0,B_1,B_2,\ldots\ldots,B_{L-1},B_L]$。

[0056] 根据本公开的实施例，根据当前向量，确定当前响应动作，可以包括如下操作。

[0057] 从当前推荐概率和多个当前预测概率中确定最大概率。在确定最大概率为当前预测概率的情况下，确定当前响应动作为用于继续对话的响应动作。在确定最大概率为当前推荐概率的情况下，确定当前响应动作为用于产品推荐的响应动作。

[0058] 根据本公开的实施例，将当前推荐概率和多个产品属性中每个产品属性的当前预测概率进行比较，确定出最大概率。如果最大概率为当前推荐概率，则可以确定当前响应动作即为用于产品推荐的响应动作。如果最大概率为多个当前预测概率中的某个当前预测概率，则可以确定当前响应动作即为用于继续对话的响应动作。

[0059] 根据本公开的实施例，如果最大概率的数量为大于一个，则可以根据实际情况进行确定。示例性的，如最大概率包括当前推荐概率和多个当前预测概率中的一个或多个，则可以将产品推荐的响应动作确定为当前响应动作。

[0060] 根据本公开的实施例，当前响应动作包括用于继续对话的响应动作。执行响应动作以得到产品导购处理结果，可以包括如下操作。

[0061] 确定目标产品属性，其中，目标产品属性为与最大概率对应的产品属性。将目标产品属性确定为与用户继续对话的回复信息。

[0062] 根据本公开的实施例，如果最大概率为当前预测概率，则可以将与该当前预测概率对应的产品属性确定为目标产品属性，并将目标产品属性作为与用户继续对话的回复信息，即回复用户的信息为目标产品属性。图4示意性示出了根据本公开实施例的一种用于继

续对话的响应动作的示意图。在图3的基础上,图4中确定出的目标产品属性为免洗,根据目标产品属性,确定与用户继续对话的回复信息为需要免洗吗?

[0063] 根据本公开的实施例,如果最大概率为当前预测概率,并且为最大概率的当前预测概率的数量为大于一个,则可以从多个当前预测概率中任选一个当前预测概率,将与该当前预测概率对应的产品属性作为目标产品属性。

[0064] 根据本公开的实施例,当前响应动作包括用于产品推荐的响应动作。

[0065] 执行响应动作以得到产品导购处理结果,可以包括如下操作。

[0066] 获取用户的特征信息和候选产品的特征信息。利用产品推荐模型处理用户的特征信息、候选产品的特征信息和当前对话状态,得到推荐产品,其中,产品推荐模型是利用第二训练样本训练生成的,其中,第二训练样本包括多个第二样本用户中每个第二样本用户的特征信息、对话状态样本和真实兴趣产品,以及,多个样本产品中每个样本产品的特征信息。

[0067] 根据本公开的实施例,如果当前响应动作为用于产品推荐的响应动作,则可以获取用户的特征信息和候选产品的特征信息,将用户的特征信息、候选产品的特征信息和当前对话状态进行向量化处理,得到用户的特征向量、候选产品的特征向量和当前对话状态向量,将用户的特征向量、候选产品的特征向量和当前对话状态向量进行拼接,得到目标特征向量。利用产品推荐模型处理目标特征向量,得到推荐产品。图5示意性示出了根据本公开实施例的一种用于产品推荐的响应动作的示意图。在图3的基础上,图5采用向用户发送相关产品的网络链接的方式实现产品推荐。

[0068] 根据本公开的实施例,产品推荐模型是利用第二训练样本训练生成的,可以包括如下操作。

[0069] 利用第二训练样本训练因子分解机模型,得到产品推荐模型。

[0070] 根据本公开的实施例,获取多个第二样本用户中每个第二样本用户的对话状态样本可以包括:针对每个第二样本用户,获取与该第二样本用户对应的对话信息样本,将对话信息样本输入对话状态追踪模型,输出对话状态样本,对话状态样本包括多个产品属性中每个产品属性的概率分布。

[0071] 根据本公开的实施例,将每个第二样本用户的特征信息和每个样本产品的特征信息进行向量化处理,得到每个第二样本用户的特征向量和每个样本产品的产品特征向量,将每个第二样本用户的特征向量、每个样本产品的特征信息和每个第二样本用户的对话状态样本进行拼接,得到多个训练特征向量,利用多个训练特征向量训练因子分解机(Factorization Machine,FM)模型,得到产品推荐模型。即将多个训练特征向量输入因子分解机模型,输出每个第二样本用户的预测兴趣产品。将每个第二样本用户的预测兴趣产品和真实兴趣产品输入因子分解机模型的损失函数,输出损失函数的输出值。根据损失函数的输出值调整因子分解机模型的模型参数,直至损失函数收敛,并将训练后的因子分解机模型作为产品推荐模型。其中,可以采用随机梯度递减方法、交替最小二乘方法或马尔科夫链蒙特卡罗推理方法训练因子分解机模型。

[0072] 示例性的,如第二训练样本包括U个第二样本用户,I个样本产品。第二样本用户m的特征向量可以表示为$u_m$={0,0,.....,1,.....,0},其中,$u_m$中与第二样本用户m对应的维度用1表示。样本产品n的特征向量可以表示为$i_n$={0,0,.....,1,.....,0},其中,$i_n$中

与样本产品n对应的维度用1表示。第二样本用户m在第t轮的对话状态样本可以用$s_{tm}$表示。训练特征向量$x_{mnt}$可以表示为$x_{mnt}=u_m \oplus i_n \oplus s_{tm}$。

[0073] 根据本公开的实施例,策略决策模型和对话状态追踪模型是利用第一训练样本共同训练生成的,可以包括如下操作。

[0074] 针对每个第一样本用户,将多轮对话信息样本中每轮对话信息样本分别输入多个循环神经网络模型,输出多个产品属性中每个产品属性的预测概率分布,其中,每个循环神经网络模型用于输出对应的产品属性的预测概率分布。针对每轮对话信息样本,根据多个产品属性中每个产品属性的预测概率分布生成本轮对话状态。将本轮对话状态输入策略网络模型中输出与本轮对话对应的本轮向量,其中,本轮向量包括本轮推荐概率和多个产品属性中每个产品属性的本轮预测概率。根据本轮向量,确定与本轮对话对应的响应动作。执行与本轮对话对应的响应动作,得到用户针对响应动作的反馈结果,其中,与本轮对话对应的响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作。确定与反馈结果对应的反馈得分。将多轮对话的响应概率和反馈得分输入目标函数中,输出目标函数的输出值,其中,响应概率为与响应动作对应的概率,响应概率为本轮推荐概率或本轮预测概率。根据输出值调整多个循环神经网络模型和策略网络模型的网络参数,直至目标函数收敛。将训练后得到的多个循环神经网络模型作为对话状态追踪模型,将训练后得到的策略网络模型作为策略决策模型。

[0075] 根据本公开的实施例,可以通过策略决策模型将基于搜索对话方式和基于推荐方式实现的产品导购结合起来,其中,策略决策模型和对话状态追踪模型可以是利用强化学习训练策略网络模型和多个循环神经网络模型得到的,多个循环神经网络模型的输出是训练策略网络模型的输入。循环神经网络模型可以包括长短期记忆神经网络模型。

[0076] 根据本公开的实施例,标准的强化学习的要素可以包括动作、状态、奖励和策略。强化学习是在一个无人指导的环境中,智能体与环境相互作用,智能体每执行一个动作,就会收到一个来自环境的反馈,即是这个动作对应的状态和奖励。强化学习的目标即是通过动作与状态的不断更新,在环境中寻求一个最优化的策略以积累出最大化的奖励,即目标函数的输出值收敛,即使得目标函数的输出值最大。强化学习可以包括策略梯度算法。

[0077] 根据本公开的实施例,可以策略决策模型和对话状态追踪模型可以是利用策略梯度算法训练策略网络模型和多个循环神经网络模型得到的。策略梯度算法是直接根据当前的状态来选择动作,其是回合更新,即在回合中记录每一步的状态、动作和奖励,在一个回合完成以后,策略网络模型预测回合中每步状态的动作,与已执行的动作做交叉熵,如果一个动作得到的的奖励多,则该动作出现的概率增加,如果一个动作得到的奖励少,则该动作出现的概率减小。

[0078] 策略网络模型可以为包括两个全连接层的神经网络模型,神经网络模型的输入为本轮对话状态,即包括多个长短期记忆神经网络模型的对话状态追踪模型的输出,神经网络模型的输出为本轮向量,即本轮推荐概率和多个产品属性中每个产品属性的本轮预测概率。通常神经网络模型是根据损失函数来进行网络参数的更新的,而从严格意义上讲,策略梯度算法是没有误差的,这是由于其输出只是选择的概率。为了进行网络参数的更新,可以根据目标函数的输出值来实现,其中,目标函数可以理解为从初始轮对话开始到当前轮对话累积奖励的期望。目标函数的输入参数可以包括响应概率和反馈得分。反馈得分可以为

与反馈结果对应的得分，反馈结果可以是执行与本轮对话对应的响应动作得到的用户针对响应动作的反馈结果，响应动作可以根据本轮向量确定。响应动作可以根据本轮向量确定，可以包括从本轮推荐概率和多个本轮预测概率中确定最大概率，在确定最大概率为本轮预测概率的情况下，确定本轮响应动作为用于继续对话的响应动作，在确定最大概率为本轮推荐概率的情况下，确定本轮响应动作为用于产品推荐的响应动作。

[0079] 在获得从初始轮对话到当前轮对话的多轮对话中每轮对话的响应概率和反馈得分后，可以将多个响应概率和多个反馈得分输入目标函数中，输出目标函数的输出值，可以为将多轮对话中每轮对话的响应概率和反馈得分相乘，得到本轮乘积结果，将多个本轮乘积结果相加，得到目标函数的输出值。在此基础上，由于越远的奖励对当前的状态影响越小，因此，可以增加一个折扣参数来影响衰减。由此，本轮乘积结果可以通过如下方式确定：将多轮对话中每轮对话的响应概率、反馈得分和本轮的折扣因子相乘，得到本轮乘积结果。在获得目标函数的输出值后，可以根据输出值调整多个循环神经网络模型和策略网络模型的网络参数，直至目标函数收敛，并将训练后得到的多个循环神经网络模型作为对话状态追踪模型，将训练后得到的神经网络模型作为策略决策模型。

[0080] 根据本公开的实施例，第t轮的本轮向量可以通过如下公式(1)确定。

[0081] $$S_t = f_{1t} \oplus f_{2t} \oplus \ldots\ldots \oplus f_{jt} \oplus \ldots\ldots \oplus f_{G-1t} \oplus f_{Gt} \quad (1)$$

[0082] 其中，$S_t$表示第t轮的本轮向量。$f_{jt}$表示第t轮的对话状态中的产品属性j的概率分布，$j \in \{1,2,\ldots\ldots,L-1,L\}$，L表示产品属性的数量。

[0083] 目标函数可以通过如下公式(2)确定。

[0084] $$\eta(\theta)=E_{\pi}\left[\sum_{t=0}^{T}\gamma^{t}R_{t}\right]=\sum_{t=0}^{T}\gamma^{t}R_{t}P_{\pi(A_t|S_t)}\left(A_t|S_t\right) \quad (2)$$

[0085] 其中，$\eta(\theta)$表示目标函数，$\theta$表示学习参数，即神经网络模型和多个长短期记忆神经网络模型的网络参数。$\pi$表示策略，t表示轮次，T表示总轮数，$t \in \{1,2,\ldots\ldots,T-1,T\}$，$A_t$表示第t轮的动作，$S_t$表示第t轮的状态。$R_t$表示第t的反馈得分。J表示折扣参数，$J \in [0,1]$。$P_{\pi(A_t|S_t)}\left(A_t|S_t\right)$表示在状态$S_t$下，根据策略$\pi(A_t|S_t)$确定的第t轮的响应动作的响应概率。

[0086] 根据本公开的实施例，通过利用强化学习训练得到的策略决策模型能够将基于搜索对话方式和基于推荐方式实现的产品导购结合起来，解决了在当前时间节点，是需要继续基于搜索对话方式与用户继续对话获取用户的当前需求信息，以进一步明确用户的当前需求，还是需要基于推荐方式向用户推荐合适的产品的问题。

[0087] 根据本公开的实施例，确定与反馈结果对应的反馈得分，可以包括如下操作。

[0088] 在确定反馈结果为购买推荐产品的情况下，确定与反馈结果对应的反馈得分为第一奖励得分。在确定反馈结果为浏览推荐产品且未购买产品的情况下，确定与反馈结果对应的反馈得分为第二奖励得分，其中，第一奖励得分大于第二奖励得分。在确定反馈结果为继续对话或未浏览推荐产品的情况下，确定与反馈结果对应的反馈得分为第一惩罚得分。在确定反馈结果为终止对话的情况下，确定与反馈结果对应的反馈得分为第二惩罚得分，其中，第一惩罚得分大于第二惩罚得分。

[0089] 根据本公开的实施例，根据反馈结果确定对应的反馈得分可以分为如下三种情况。

[0090] 情况一、推荐下的激励。如果执行本轮响应动作后，用户针对本轮响应动作的反馈结果为购买推荐产品或浏览推荐产品，则可以认为本轮响应动作得到了用户的良好反馈，模型可以通过强化学习受到激励。如果确定反馈结果为购买推荐产品，则可以确定与反馈结果对应的反馈得分为第一奖励得分。如果确定反馈结果为浏览推荐产品且未购买产品，则可以确定与反馈结果对应的反馈得分为第二奖励得分，其中，奖励得分可以为正数，第一奖励得分大于第二奖励得分。示例性的，如在第t轮对话中，用户购买了推荐产品，则第一奖励得分可以为$R_t=5$。在第t轮对话中，用户浏览了但没有购买推荐产品，则第二奖励得分可以为$R_t=1$。

[0091] 情况二、对话轮次增加下的惩罚。如果反馈结果为继续对话，则可以说明当前没有推荐产品，使得对话轮次增加。在此情况下，可以认为在当前轮对话中，系统没有明确用户对于产品的需求，模型会通过强化学习受到惩罚。或者，如果当前有推荐产品，反馈结果为未浏览推荐产品，则也可以认为在当前轮对话中，系统没有明确用户对于产品的需求，模型会通过强化学习受到惩罚。示例性的，如当前处于第t轮，$t=5$且系统没有推荐产品，或者，系统有推荐产品且用户未浏览推荐产品，则第一惩罚得分可以为$R_t=1$。

[0092] 情况三、终止对话下的惩罚。如果反馈结果为终止对话，则可以认为用户对系统的体验不佳，模型会通过强化学习受到惩罚。示例性的，如在第t轮对话中，用户主动终止了对话，则第二惩罚得分可以为$R_t=3$。

[0093] 图6示意性示出了根据本公开实施例的另一种产品导购处理方法的流程图。

[0094] 如图6所示，该方法包括操作S601～S611。

[0095] 在操作S601，获取用户的当前轮对话信息。

[0096] 在操作S602，利用对话状态追踪模型处理当前轮对话信息得到当前对话状态。

[0097] 在操作S603，利用策略决策模型处理当前对话状态得到当前向量，其中，当前向量包括当前推荐概率和多个产品属性中每个产品属性的当前预测概率。

[0098] 在操作S604，从当前推荐概率和多个当前预测概率中确定最大概率。

[0099] 在操作S605，确定最大概率是否为推荐概率；若是，则执行操作S606；若否，则执行操作S607。

[0100] 在操作S606，确定当前响应动作为用于产品推荐的响应动作，并执行操作S608。

[0101] 在操作S607，确定当前响应动作为用于继续对话的响应动作，并执行操作S610。

[0102] 在操作S608，获取用户的特征信息和候选产品的特征信息。

[0103] 在操作S609，利用产品推荐模型处理用户的特征信息、候选产品的特征信息和当前对话状态，得到推荐产品。

[0104] 在操作S610，确定目标产品属性，其中，目标产品属性为与最大概率对应的产品属性。

[0105] 在操作S611，将目标产品属性确定为与用户继续对话的回复信息。

[0106] 根据本公开的实施例，图7示意性示出了根据本公开实施例的一种产品导购处理方法的应用示意图。

[0107] 图7中对话状态追踪模型的输出可以作为策略决策模型和产品推荐模型的输入。其中，对话状态追踪模型可以包括k个长短期记忆神经网络模型，k可以是正整数。

[0108] 根据本公开实施例的技术方案，由于利用策略决策模型处理基于对话状态追踪模

型得到的当前对话状态,能够确定当前响应动作是需要与用户继续交流以明确用户需求的响应动作,还是调用产品推荐模型进行产品推荐的响应动作。其中,基于对话状态追踪模型可以获得反映用户的当前需求信息的当前对话状态,产品推荐模型进行产品推荐时结合了反映用户的历史行为记录的用户的特征信息和反映用户的当前需求信息的当前对话状态,使得推荐的产品既符合用户的当前需求,也体现了用户的潜在兴趣点,实现了在合适的时间向用户推荐合适的产品,因而,至少部分地克服了采用相关技术实现产品导购的销售转换率不高的技术问题,提高了产品导购的销售转换率。

[0109] 图8示意性示出了根据本公开的实施例的一种产品导购处理装置的框图。

[0110] 如图8所示,产品导购处理装置800包括获取模块810、第一处理模块820、第二处理模块830和执行模块840。

[0111] 获取模块810、第一处理模块820、第二处理模块830和执行模块840通信连接。

[0112] 获取模块810,用于获取用户的当前轮对话信息。

[0113] 第一处理模块820,用于利用对话状态追踪模型处理当前轮对话信息得到当前对话状态,其中,当前对话状态包括多个产品属性中每个产品属性的当前概率分布,其中,每个当前概率分布用于表征与产品属性对应的一个或多个产品属性值中每个产品属性值的概率。

[0114] 第二处理模块830,用于利用策略决策模型处理当前对话状态得到当前响应动作,其中,当前响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作,其中,策略决策模型和对话状态追踪模型是利用第一训练样本共同训练生成的,其中,第一训练样本包括多个第一样本用户中每个第一样本用户的多轮对话信息样本。

[0115] 执行模块840,用于执行当前响应动作以得到产品导购处理结果。

[0116] 根据本公开实施例的技术方案,通过获取用户的当前轮对话信息,利用对话状态追踪模型处理当前轮对话信息得到当前对话状态,当前对话状态包括多个产品属性中每个产品属性的当前概率分布,利用策略决策模型处理当前对话状态得到当前响应动作,当前响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作,策略决策模型和对话状态追踪模型是利用第一训练样本共同训练生成的,并执行当前响应动作以得到产品导购处理结果。由于利用策略决策模型处理基于对话状态追踪模型得到的当前对话状态,能够确定当前响应动作是需要与用户继续交流以明确用户需求的响应动作,还是进行产品推荐的响应动作,实现了在合适的时间向用户推荐合适的产品,因而,至少部分地克服了采用相关技术实现产品导购的销售转换率不高的技术问题,提高了产品导购的销售转换率。

[0117] 根据本公开的实施例,第二处理模块830可以包括第一处理子模块和第一确定子模块。

[0118] 第一处理子模块,用于利用策略决策模型处理当前对话状态得到当前向量,其中,当前向量包括当前推荐概率和多个产品属性中每个产品属性的当前预测概率。

[0119] 第一确定子模块,用于根据当前向量,确定当前响应动作。

[0120] 根据本公开的实施例,第一确定子模块可以包括第一确定单元、第二确定单元和第三确定单元。

[0121] 第一确定单元,用于从当前推荐概率和多个当前预测概率中确定最大概率。

[0122] 第二确定单元,用于在确定最大概率为当前预测概率的情况下,确定当前响应动

作为用于继续对话的响应动作。

[0123] 第三确定单元,用于在确定最大概率为当前推荐概率的情况下,确定当前响应动作为用于产品推荐的响应动作。

[0124] 根据本公开的实施例,当前响应动作包括用于继续对话的响应动作。

[0125] 执行模块840可以包括第二确定子模块和第三确定子模块。

[0126] 第二确定子模块,用于确定目标产品属性,其中,目标产品属性为与最大概率对应的产品属性。

[0127] 第三确定子模块,用于将目标产品属性确定为与用户继续对话的回复信息。

[0128] 根据本公开的的实施例,当前响应动作包括用于产品推荐的响应动作。

[0129] 执行模块840可以包括获取子模块和第二处理子模块。

[0130] 获取子模块,用于获取用户的特征信息和候选产品的特征信息。

[0131] 第二处理子模块,用于利用产品推荐模型处理用户的特征信息、候选产品的特征信息和当前对话状态,得到推荐产品,其中,产品推荐模型是利用第二训练样本训练生成的,其中,第二训练样本包括多个第二样本用户中每个第二样本用户的特征信息、对话状态样本和真实兴趣产品,以及,多个样本产品中每个样本产品的特征信息。

[0132] 根据本公开的实施例,产品推荐模型是利用第二训练样本训练生成的,可以包括如下操作。

[0133] 利用第二训练样本训练因子分解机模型,得到产品推荐模型。

[0134] 根据本公开的实施例,策略决策模型和对话状态追踪模型是利用第一训练样本共同训练生成的,可以包括如下操作。

[0135] 针对每个第一样本用户,将多轮对话信息样本中每轮对话信息样本分别输入多个循环神经网络模型,输出多个产品属性中每个产品属性的预测概率分布,其中,每个循环神经网络模型用于输出对应的产品属性的预测概率分布。针对每轮对话信息样本,根据多个产品属性中每个产品属性的预测概率分布生成本轮对话状态。将本轮对话状态输入策略网络模型中输出与本轮对话对应的本轮向量,其中,本轮向量包括本轮推荐概率和多个产品属性中每个产品属性的本轮预测概率。根据本轮向量,确定与本轮对话对应的响应动作。执行与本轮对话对应的响应动作,得到用户针对响应动作的反馈结果,其中,与本轮对话对应的响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作。确定与反馈结果对应的反馈得分。将多轮对话的响应概率和反馈得分输入目标函数中,输出目标函数的输出值,其中,响应概率为与响应动作对应的概率,响应概率为本轮推荐概率或本轮预测概率。根据输出值调整多个循环神经网络模型和策略网络模型的网络参数,直至目标函数收敛。将训练后得到的多个循环神经网络模型作为对话状态追踪模型,将训练后得到的策略网络模型作为策略决策模型。

[0136] 根据本公开的实施例,确定与反馈结果对应的反馈得分,可以包括如下操作。

[0137] 在确定反馈结果为购买推荐产品的情况下,确定与反馈结果对应的反馈得分为第一奖励得分。在确定反馈结果为浏览推荐产品且未购买产品的情况下,确定与反馈结果对应的反馈得分为第二奖励得分,其中,第一奖励得分大于第二奖励得分。在确定反馈结果为继续对话或未浏览推荐产品的情况下,确定与反馈结果对应的反馈得分为第一惩罚得分。在确定反馈结果为终止对话的情况下,确定与反馈结果对应的反馈得分为第二惩罚得分,

其中,第一惩罚得分大于第二惩罚得分。

[0138] 根据本公开的实施例的模块、子模块、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、子模块、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、子模块、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(Field Programmable Gate Array,FPGA)、可编程逻辑阵列(Programmable Logic Arrays,PLA)、片上系统、基板上的系统、封装上的系统、专用集成电路(Application Specific Integrated Circuit,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、子模块、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。

[0139] 例如,获取模块810、第一处理模块820、第二处理模块830和执行模块840中的任意多个可以合并在一个模块/子模块/单元中实现,或者其中的任意一个模块/子模块/单元可以被拆分成多个模块/子模块/单元。或者,这些模块/子模块/单元中的一个或多个模块/子模块/单元的至少部分功能可以与其他模块/子模块/单元的至少部分功能相结合,并在一个模块/子模块/单元中实现。根据本公开的实施例,获取模块810、第一处理模块820、第二处理模块830和执行模块840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、第一处理模块820、第二处理模块830和执行模块840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。

[0140] 需要说明的是,本公开的实施例中产品导购处理装置部分与本公开的实施例中产品导购处理方法部分是相对应的,产品导购处理装置部分的描述具体参考产品导购处理方法部分,在此不再赘述。

[0141] 图9示意性示出了根据本公开实施例的适于实现上文描述的方法的电子设备的框图。图9示出的电子设备仅仅是一个示例,不应对本公开实施例的功能和使用范围带来任何限制。

[0142] 如图9所示,根据本公开实施例的电子设备900包括处理器901,其可以根据存储在只读存储器(Read-Only Memory,ROM)902中的程序或者从存储部分908加载到随机访问存储器(Random Access Memory,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。

[0143] 在RAM 903中,存储有电子设备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

ROM902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。

[0144] 根据本公开的实施例,电子设备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。电子设备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(Liquid Crystal Display,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。

[0145] 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、子模块、单元等可以通过计算机程序模块来实现。

[0146] 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。

[0147] 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((Erasable Programmable Read Only Memory,EPROM)或闪存)、便携式紧凑磁盘只读存储器(Computer Disc Read-Only Memory,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。

[0148] 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。

[0149] 根据本公开实施例的技术方案,通过获取用户的当前轮对话信息,利用对话状态追踪模型处理当前轮对话信息得到当前对话状态,当前对话状态包括多个产品属性中每个产品属性的当前概率分布,利用策略决策模型处理当前对话状态得到当前响应动作,当前响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作,策略决策模型和对话状态追踪模型是利用第一训练样本共同训练生成的,并执行当前响应动作以得到产品导购处理结果。由于利用策略决策模型处理基于对话状态追踪模型得到的当前对话状态,能够确定当前响应动作是需要与用户继续交流以明确用户需求的响应动作,还是进行产品推荐的响应动作,实现了在合适的时间向用户推荐合适的产品,因而,至少部分地克服了采用相关技术实现产品导购的销售转换率不高的技术问题,提高了产品导购的销售转换率。

[0150] 本公开的实施例还包括一种计算机程序产品,其包括计算机程序,该计算机程序包含用于执行本公开实施例所提供的方法的程序代码,当计算机程序产品在电子设备上运行时,该程序代码用于使电子设备实现本公开实施例所提供的产品导购处理方法。

[0151] 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/装置中限定的上述功能。根据本公开的实施例,上文描述的系统、装置、模块、单元等可以通过计算机程序模块来实现。

[0152] 在一种实施例中,该计算机程序可以依托于光存储器件、磁存储器件等有形存储介质。在另一种实施例中,该计算机程序也可以在网络介质上以信号的形式进行传输、分发,并通过通信部分909被下载和安装,和/或从可拆卸介质911被安装。该计算机程序包含的程序代码可以用任何适当的网络介质传输,包括但不限于:无线、有线等等,或者上述的任意合适的组合。

[0153] 根据本公开的实施例,可以以一种或多种程序设计语言的任意组合来编写用于执行本公开实施例提供的计算机程序的程序代码,具体地,可以利用高级过程和/或面向对象的编程语言、和/或汇编/机器语言来实施这些计算程序。程序设计语言包括但不限于诸如Java,C++,python,“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算设备上执行、部分地在用户设备上执行、部分在远程计算设备上执行、或者完全在远程计算设备或服务器上执行。在涉及远程计算设备的情形中,远程计算设备可以通过任意种类的网络,包括局域网(Local Area Network,LAN)或广域网(Wide Area Networks,WAN),连接到用户计算设备,或者,可以连接到外部计算设备(例如利用因特网服务提供商来通过因特网连接)。

[0154] 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。也要注意的是,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。

[0155] 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。

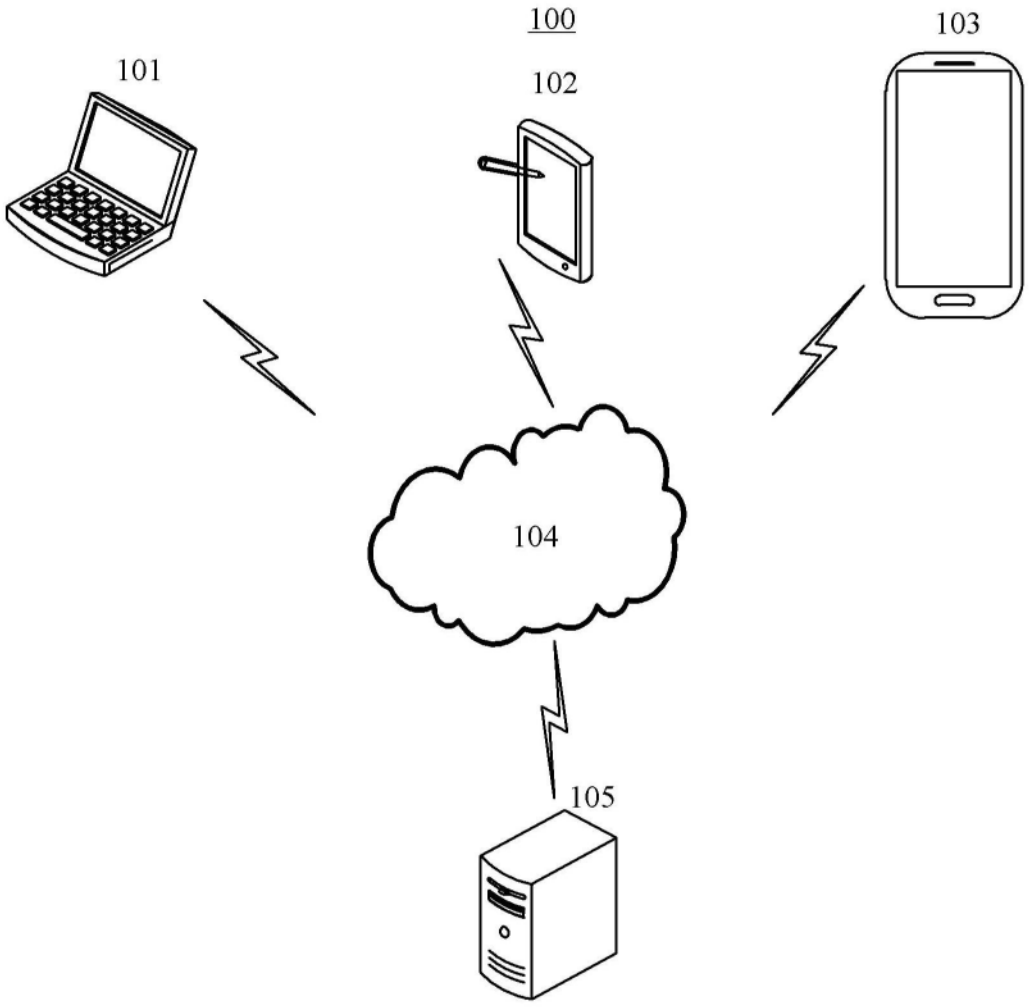



图1

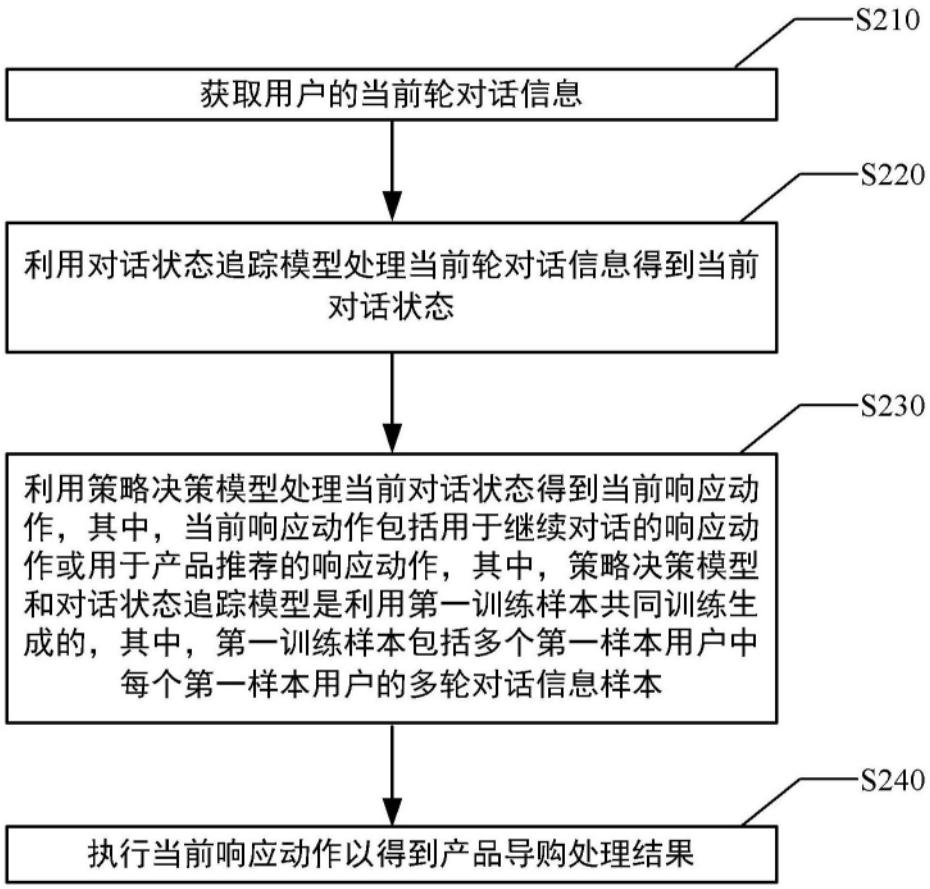

S210
获取用户的当前轮对话信息
S220
利用对话状态追踪模型处理当前轮对话信息得到当前对话状态
S230
利用策略决策模型处理当前对话状态得到当前响应动作，其中，当前响应动作包括用于继续对话的响应动作或用于产品推荐的响应动作，其中，策略决策模型和对话状态追踪模型是利用第一训练样本共同训练生成的，其中，第一训练样本包括多个第一样本用户中每个第一样本用户的多轮对话信息样本
S240
执行当前响应动作以得到产品导购处理结果

图2

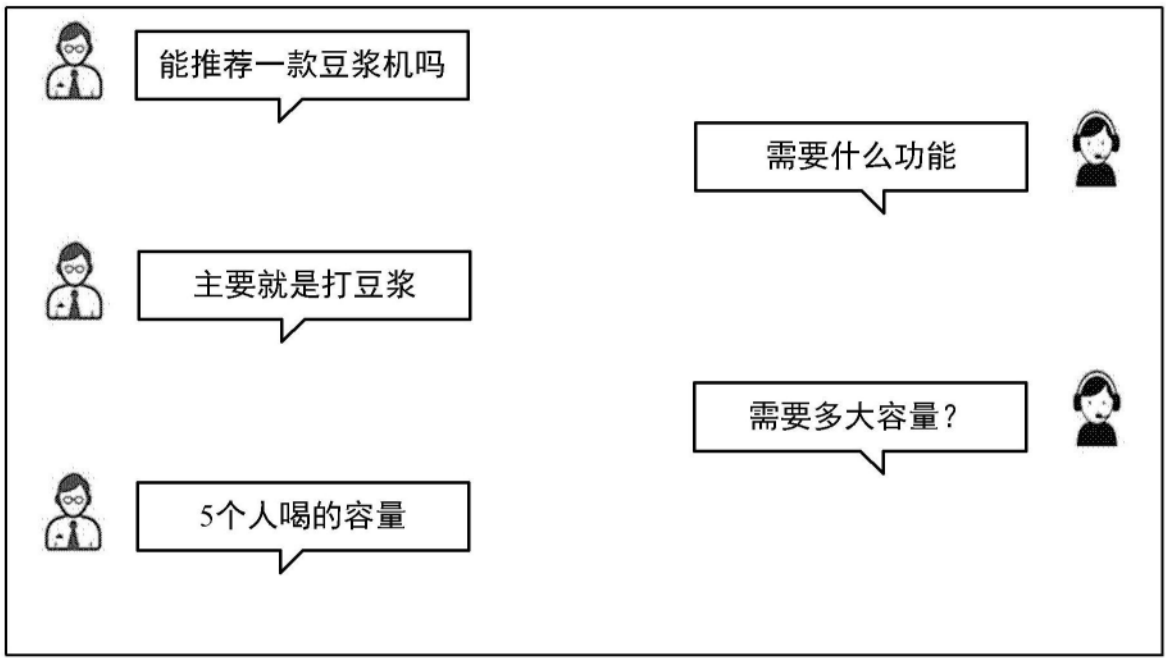

能推荐一款豆浆机吗
需要什么功能
主要就是打豆浆
需要多大容量？
5个人喝的容量

图3

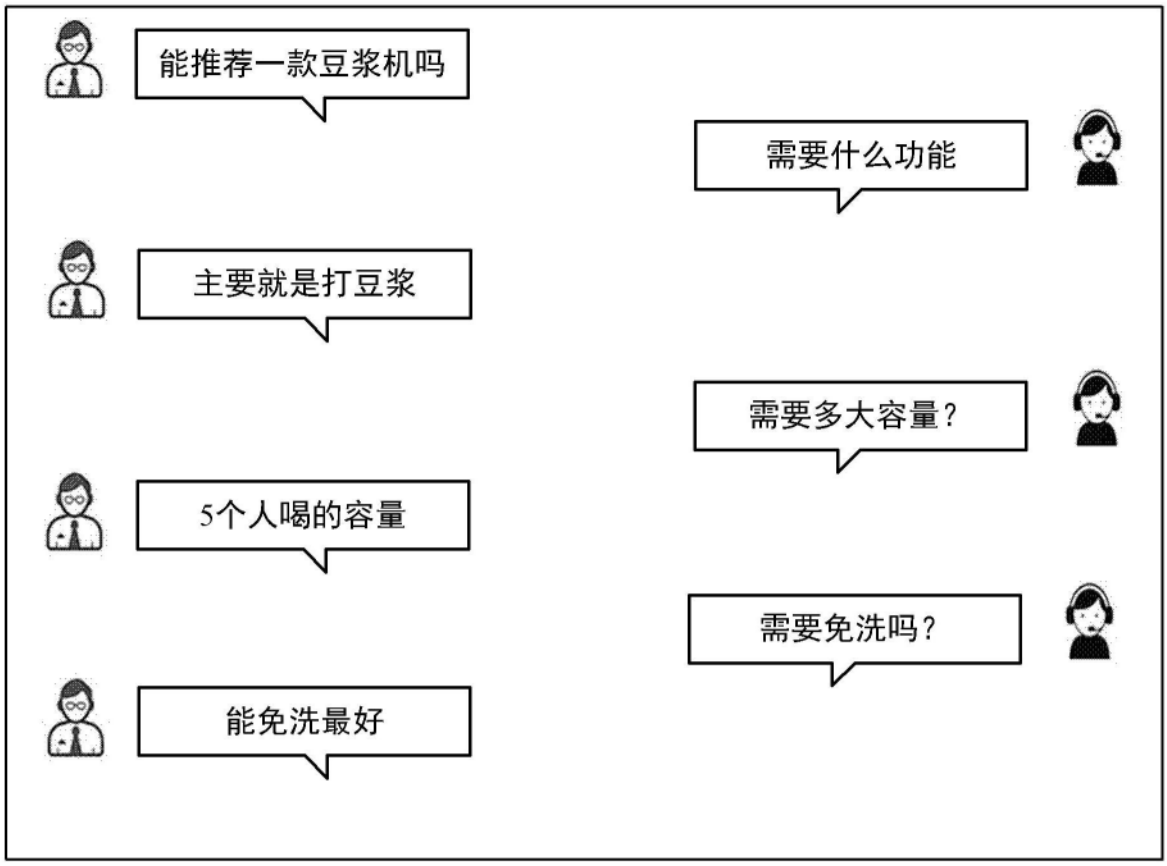



图4

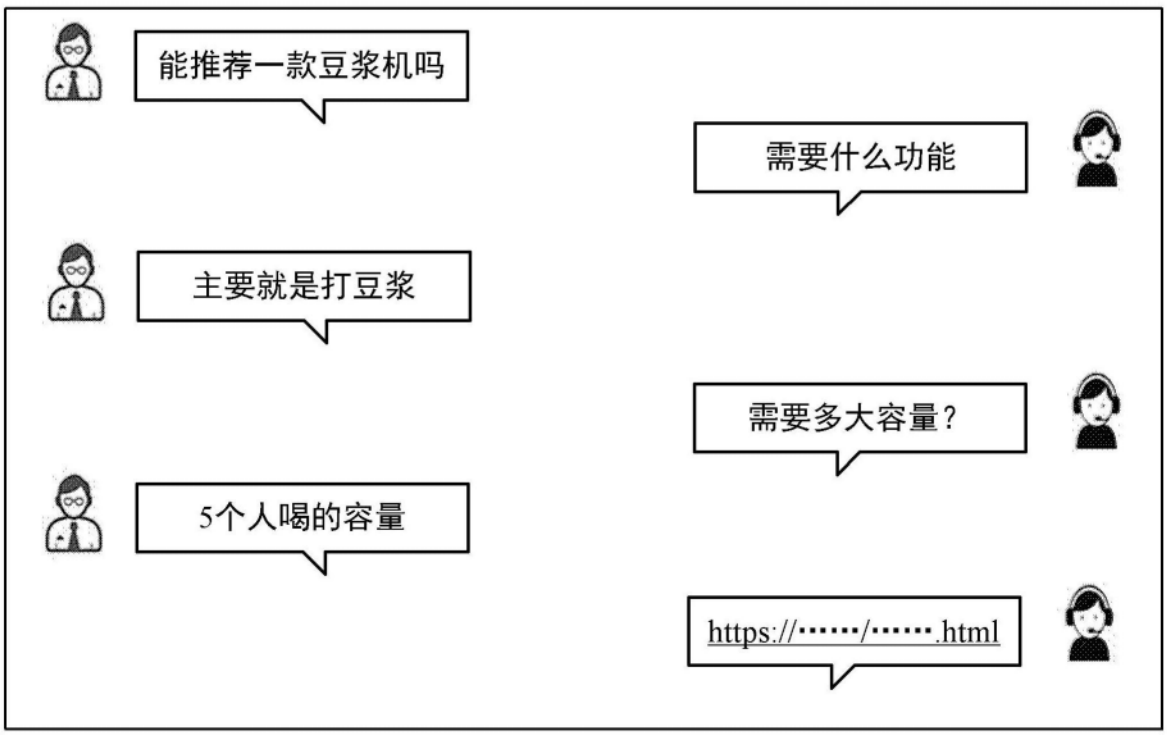



图5

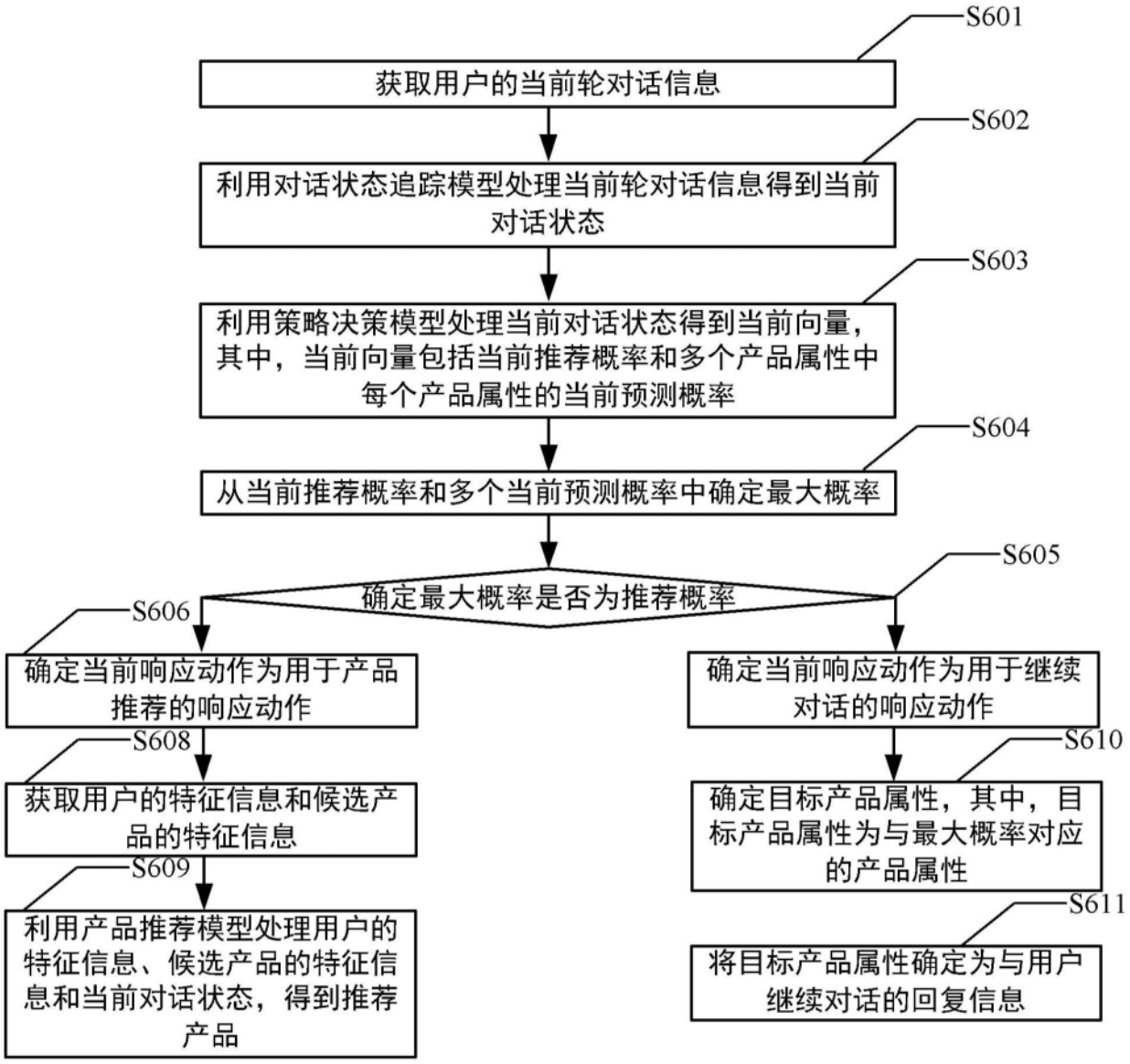

S601
获取用户的当前轮对话信息
S602
利用对话状态追踪模型处理当前轮对话信息得到当前对话状态
S603
利用策略决策模型处理当前对话状态得到当前向量，其中，当前向量包括当前推荐概率和多个产品属性中每个产品属性的当前预测概率
S604
从当前推荐概率和多个当前预测概率中确定最大概率
S605
确定最大概率是否为推荐概率
S606
确定当前响应动作为用于产品推荐的响应动作
S608
获取用户的特征信息和候选产品的特征信息
S609
利用产品推荐模型处理用户的特征信息、候选产品的特征信息和当前对话状态，得到推荐产品
确定当前响应动作为用于继续对话的响应动作
S610
确定目标产品属性，其中，目标产品属性为与最大概率对应的产品属性
S611
将目标产品属性确定为与用户继续对话的回复信息

图6

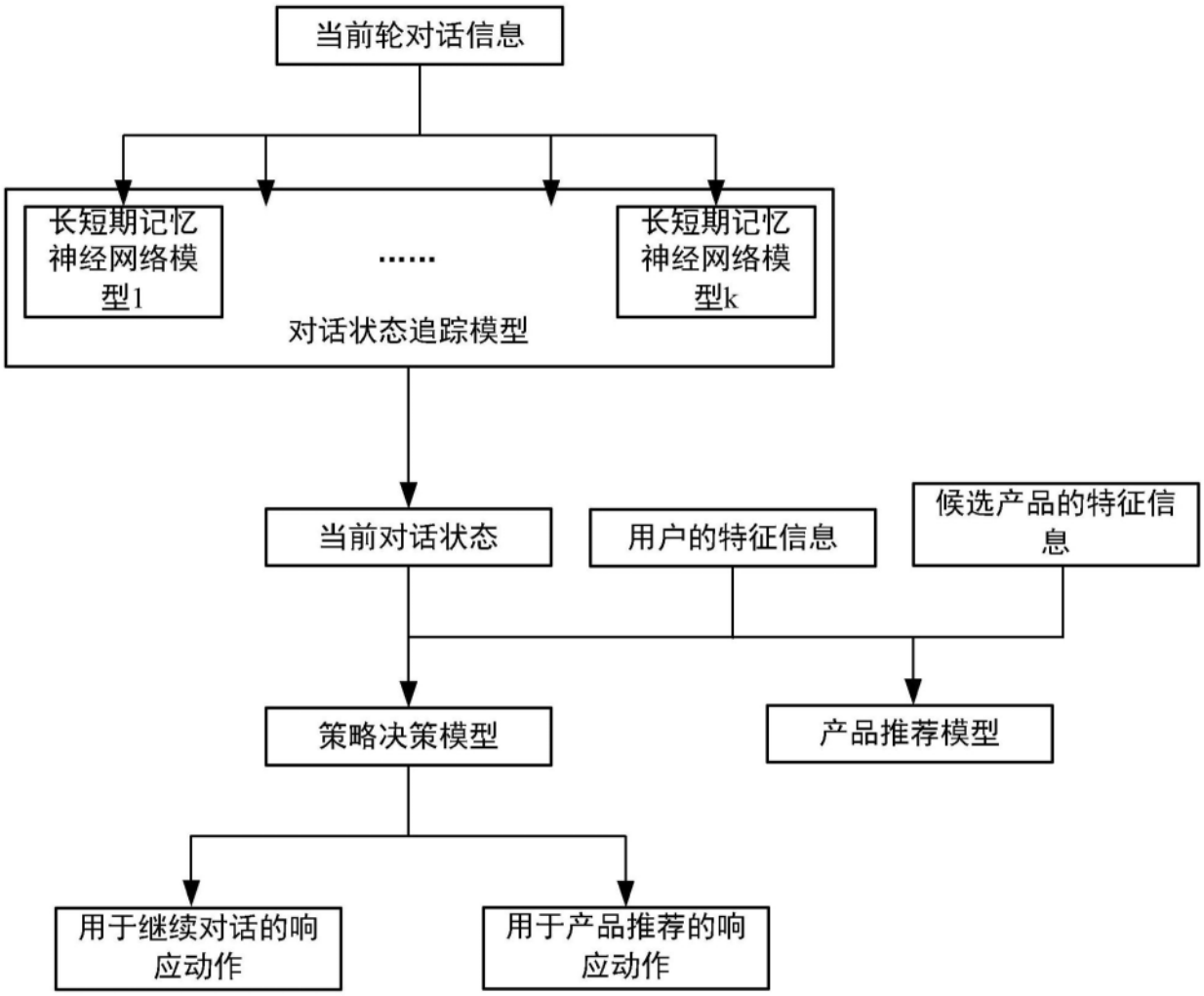

当前轮对话信息
长短期记忆神经网络模型1
.....
长短期记忆神经网络模型k
对话状态追踪模型
当前对话状态
用户的特征信息
候选产品的特征信息
策略决策模型
产品推荐模型
用于继续对话的响应动作
用于产品推荐的响应动作

图7

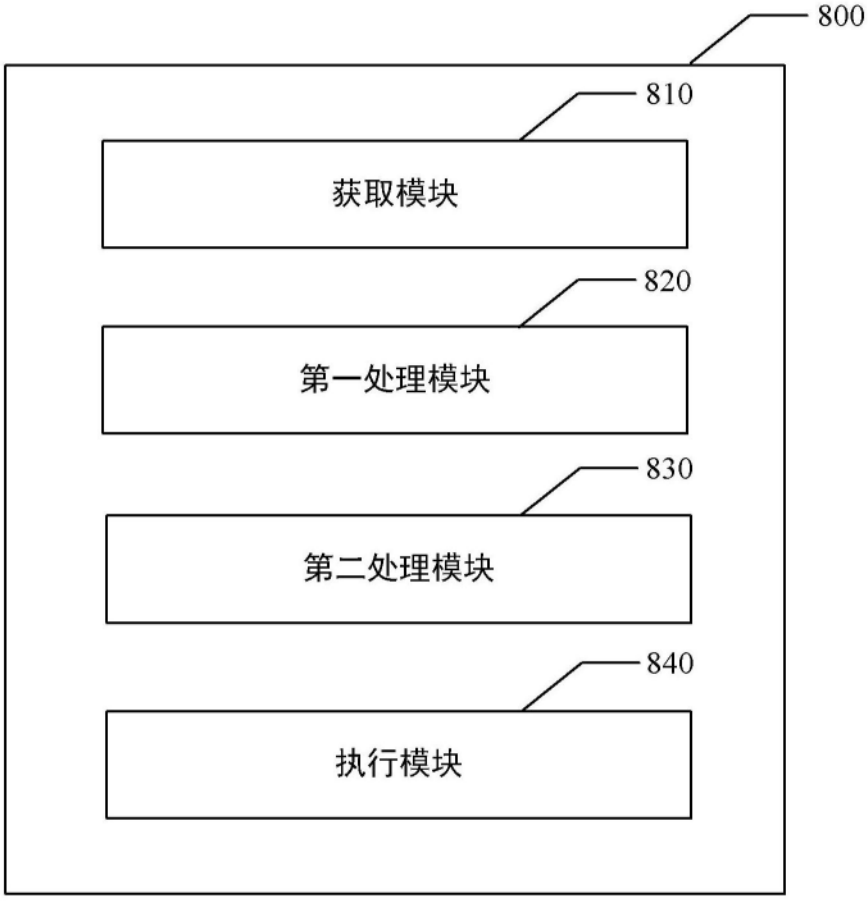



图8

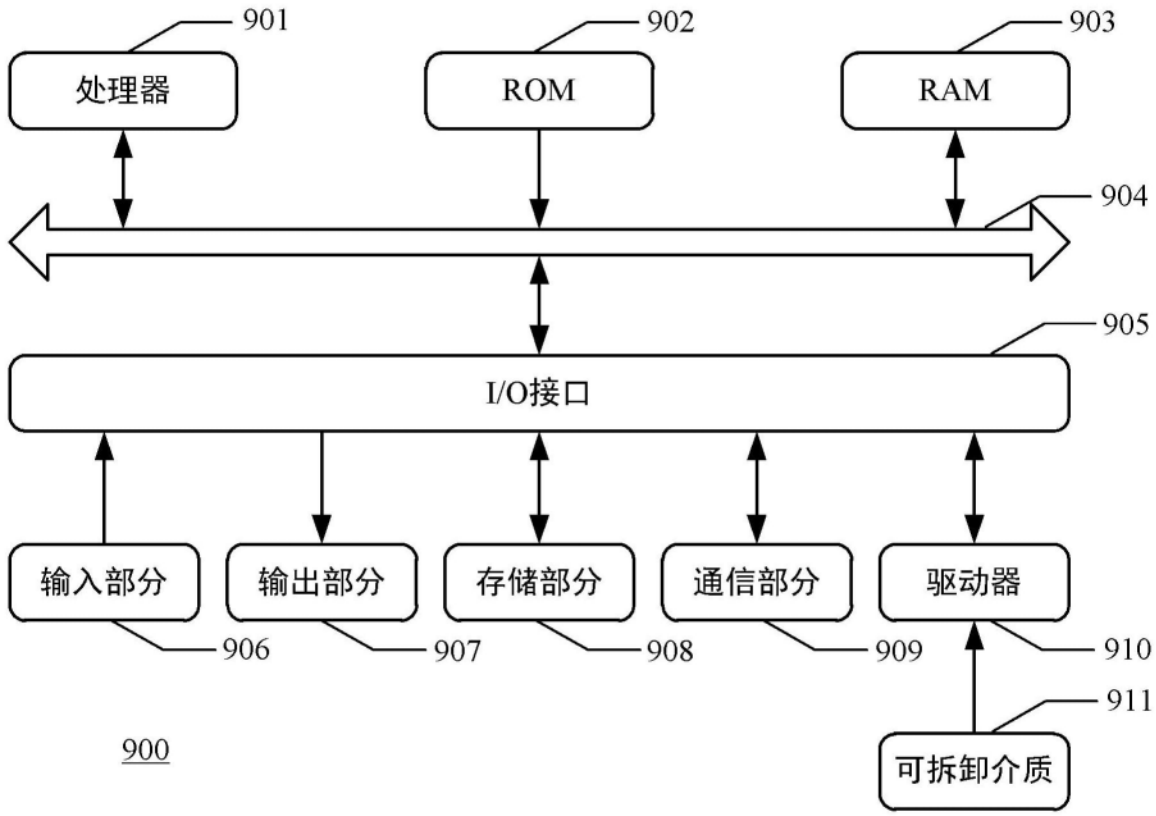

901
处理器
902
ROM
903
RAM
904
905
I/O接口
输入部分
输出部分
存储部分
通信部分
驱动器
906
907
908
909
910
911
900
可拆卸介质

图9