



(12)发明专利申请

(10)申请公布号 CN 112581203 A

(43)申请公布日 2021.03.30

(21)申请号 201910941459.4

(22)申请日 2019.09.30

(71)申请人 微软技术许可有限责任公司

地址 美国华盛顿州

(72)发明人 吴先超

(74)专利代理机构 永新专利商标代理有限公司

72002

代理人 张立达

(51)Int.Cl.

G06Q 30/06(2012.01)

G06F 16/332(2019.01)

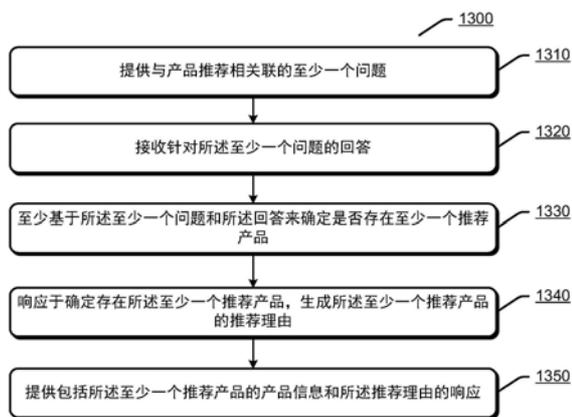
权利要求书3页 说明书19页 附图12页

(54)发明名称

在会话中提供解释性产品推荐

(57)摘要

本公开提供了用于在会话中提供解释性产品推荐的方法和装置。可以提供与产品推荐相关联的至少一个问题。可以接收针对所述至少一个问题的回答。可以至少基于所述至少一个问题和所述回答来确定是否存在至少一个推荐产品。可以响应于确定存在所述至少一个推荐产品,生成所述至少一个推荐产品的推荐理由。可以提供包括所述至少一个推荐产品的产品信息和所述推荐理由的响应。



1. 一种用于在会话中提供解释性产品推荐的方法,包括:  
提供与产品推荐相关联的至少一个问题;  
接收针对所述至少一个问题的回答;  
至少基于所述至少一个问题和所述回答来确定是否存在至少一个推荐产品;  
响应于确定存在所述至少一个推荐产品,生成所述至少一个推荐产品的推荐理由;以  
及  
提供包括所述至少一个推荐产品的产品信息和所述推荐理由的响应。
2. 如权利要求1所述的方法,其中,所述确定是否存在至少一个推荐产品包括:  
至少基于所述至少一个问题和所述回答来对多个候选产品执行产品排序;以及  
基于所述产品排序的结果来确定所述多个候选产品中是否存在满足预定条件的至少  
一个候选产品。
3. 如权利要求2所述的方法,其中,所述执行产品排序包括:  
至少基于所述至少一个问题和所述回答来更新候选产品评估状态,所述候选产品评估  
状态包括所述多个候选产品中每一个候选产品的期望概率。
4. 如权利要求3所述的方法,其中,所述预定条件包括:  
候选产品的期望概率高于阈值。
5. 如权利要求1所述的方法,还包括:  
将所述至少一个问题和所述回答添加到所述会话的会话记录中,所述会话记录包括在  
所述会话中与所述产品推荐相关联的历史问题和历史回答。
6. 如权利要求5所述的方法,其中,所述生成推荐理由包括:  
确定所述会话记录中的每个历史问题针对所述至少一个推荐产品所产生的期望概率  
提升;  
选择产生最大期望概率提升的历史问题;以及  
至少基于所选择的历史问题和相应回答来生成所述推荐理由。
7. 如权利要求5所述的方法,其中,所述生成推荐理由包括:  
通过推荐理由生成模型,基于由所述历史回答所选择的属性、所述推荐产品的属性以  
及所述推荐产品的描述中的至少一个来生成所述推荐理由。
8. 如权利要求2所述的方法,还包括:  
响应于确定不存在所述至少一个推荐产品,至少基于所述产品排序的结果来对多个候  
选问题执行问题排序;以及  
基于所述问题排序的结果来选择将要提供的下一个问题。
9. 如权利要求8所述的方法,其中,所述问题排序是通过基于熵的排序或者基于策略的  
增强学习排序来执行的。
10. 如权利要求8所述的方法,其中,所述多个候选问题是至少基于所述多个候选产品  
的属性来预先确定的。
11. 如权利要求1所述的方法,其中,所述至少一个问题包括一个或多个选项,并且所述  
回答包括针对所述一个或多个选项的选择。
12. 如权利要求1所述的方法,其中,所述回答包括自然语言语句,并且所述确定是否存  
在至少一个推荐产品包括:

确定与所述自然语言语句相对应的一个或多个相关问题和对应回答;以及至少基于所述一个或多个相关问题和对应回答来确定是否存在所述至少一个推荐产品。

13. 一种用于在会话中提供解释性产品推荐的装置,包括:  
问题提供模块,用于提供与产品推荐相关联的至少一个问题;  
回答接收模块,用于接收针对所述至少一个问题的回答;  
推荐产品确定模块,用于至少基于所述至少一个问题和所述回答来确定是否存在至少一个推荐产品;

推荐理由生成模块,用于响应于确定存在所述至少一个推荐产品,生成所述至少一个推荐产品的推荐理由;以及

响应提供模块,用于提供包括所述至少一个推荐产品的产品信息和所述推荐理由的响应。

14. 如权利要求13所述的装置,其中,所述推荐产品确定模块用于:  
至少基于所述至少一个问题和所述回答来对多个候选产品执行产品排序;以及  
基于所述产品排序的结果来确定所述多个候选产品中是否存在满足预定条件的至少一个候选产品。

15. 如权利要求14所述的装置,其中,所述执行产品排序包括:  
至少基于所述至少一个问题和所述回答来更新候选产品评估状态,所述候选产品评估状态包括所述多个候选产品中每一个候选产品的期望概率。

16. 如权利要求13所述的装置,还包括:  
会话记录添加模块,用于将所述至少一个问题和所述回答添加到所述会话的会话记录中,所述会话记录包括在所述会话中与所述产品推荐相关联的历史问题和历史回答。

17. 如权利要求16所述的装置,其中,所述推荐理由生成模块用于:  
确定所述会话记录中的每个历史问题针对所述至少一个推荐产品所产生的期望概率提升;

选择产生最大期望概率提升的历史问题;以及  
至少基于所选择的历史问题和相应回答来生成所述推荐理由。

18. 如权利要求16所述的装置,其中,所述推荐理由生成模块用于:  
通过推荐理由生成模型,基于由所述历史回答所选择的属性、所述推荐产品的属性以及所述推荐产品的描述中的至少一个来生成所述推荐理由。

19. 如权利要求14所述的装置,还包括问题选择模块,用于:  
响应于确定不存在所述至少一个推荐产品,至少基于所述产品排序的结果来对多个候选问题执行问题排序;以及

基于所述问题排序的结果来选择将要提供的下一个问题。

20. 一种用于在会话中提供解释性产品推荐的装置,包括:  
至少一个处理器;以及  
存储器,其存储计算机可执行指令,当所述计算机可执行指令被执行时使所述至少一个处理器:

提供与产品推荐相关联的至少一个问题,

接收针对所述至少一个问题的回答，  
至少基于所述至少一个问题和所述回答来确定是否存在至少一个推荐产品，  
响应于确定存在所述至少一个推荐产品，生成所述至少一个推荐产品的推荐理由，以  
及  
提供包括所述至少一个推荐产品的产品信息和所述推荐理由的响应。

## 在会话中提供解释性产品推荐

### 背景技术

[0001] 人工智能(AI)聊天机器人变得越来越流行,并且正在越来越多的场景中得到应用。聊天机器人被设计用于模拟人类话语,并且可以通过文本、语音、图像等与用户进行聊天。通常,聊天机器人可以识别由用户输入的消息内的语言内容或对消息应用自然语言处理,并进而向用户提供针对该消息的响应。

### 发明内容

[0002] 提供本发明内容以便介绍一组概念,这组概念将在以下的具体实施方式中做进一步描述。本发明内容并非旨在标识所保护主题的关键特征或必要特征,也不旨在用于限制所保护主题的范围。

[0003] 本公开的实施例提出了用于在会话中提供解释性产品推荐的方法和装置。可以提供与产品推荐相关联的至少一个问题。可以接收针对所述至少一个问题的回答。可以至少基于所述至少一个问题和所述回答来确定是否存在至少一个推荐产品。可以响应于确定存在所述至少一个推荐产品,生成所述至少一个推荐产品的推荐理由。可以提供包括所述至少一个推荐产品的产品信息和所述推荐理由的响应。

[0004] 应当注意,以上一个或多个方面包括以下详细描述以及权利要求中具体指出的特征。下面的说明书及附图详细提出了所述一个或多个方面的某些说明性特征。这些特征仅仅指示可以实施各个方面的原理的多种方式,并且本公开旨在包括所有这些方面和其等同变换。

### 附图说明

[0005] 以下将结合附图描述所公开的多个方面,这些附图被提供用以说明而非限制所公开的多个方面。

[0006] 图1示出了根据实施例的部署聊天机器人的示例性网络架构。

[0007] 图2示出了根据实施例的示例性聊天机器人系统。

[0008] 图3示出了根据实施例的在候选产品集合与候选问题集合之间的映射。

[0009] 图4示出了根据实施例的用于提供解释性产品推荐的示例性总体过程。

[0010] 图5示出了根据实施例的用于提供解释性产品推荐的示例性具体过程。

[0011] 图6示出了根据实施例的用于训练推荐理由生成模型的示例性过程。

[0012] 图7示出了根据实施例的用于生成推荐理由的示例性过程。

[0013] 图8至图12示出了根据实施例的示例性聊天窗口。

[0014] 图13示出了根据实施例的用于在会话中提供解释性产品推荐的示例性方法的流程图。

[0015] 图14示出了根据实施例的用于在会话中提供解释性产品推荐的示例性装置。

[0016] 图15示出了根据实施例的用于在会话中提供解释性产品推荐的示例性装置。

## 具体实施方式

[0017] 现在将参考多种示例性实施方式来讨论本公开。应当理解,这些实施方式的讨论仅仅用于使得本领域技术人员能够更好地理解并从而实施本公开的实施例,而并非教导对本公开的范围的任何限制。

[0018] 通常,聊天机器人可以在与用户的会话中进行自动聊天。在本文中,“会话”可以指两个聊天参与者之间的时间连续对话,并且可以包括对话中的消息和响应。“消息”可以指由用户输入的任何信息,例如来自用户的查询、用户对聊天机器人的问题的回答、用户的意见等。术语“消息”和术语“查询”也可以互换使用。“响应”可以指由聊天机器人提供的任何信息,例如聊天机器人对用户的问题的答案、聊天机器人的评论、聊天机器人提出的问题等。

[0019] 在一些应用场景中,聊天机器人可以在与用户的会话中向用户提供产品推荐。在本文中,产品可以包括商品、服务等。然而,由聊天机器人提供产品推荐将面临多种挑战。在一方面,需要准备大规模标记语料库来对机器学习模型进行训练,以便用于捕获用户以自然语言表达的意图或需求。用户的意图或需求指示了用户对于推荐产品的产品属性的喜好。产品属性可以包括产品的各种参数、配置、特性等。在另一方面,在聊天机器人能够提供的产品与用户的需求之间存在信息不对称。例如,用户不知道聊天机器人能够提供哪些产品,并且不知道如何找到所期望的产品。用户需要在发送给聊天机器人的消息中包括描述产品属性的关键词,然而,聊天机器人未必能够使用这些关键词高效地找到推荐产品或者根本并不存在与这些关键词对应的产品。在又一方面,由于用户可能以低效的且低信息性的方式来与聊天机器人进行聊天,因此,聊天机器人需要花费较长的时间来逐步地收集用户需求。

[0020] 本公开的实施例提出了在聊天机器人与用户的会话期间,以高效的且准确的方式来提供解释性产品推荐。聊天机器人可以基于本公开的实施例所提出的学习-解释(LTE: learning-to-explain)架构来提供解释性产品推荐。该LTE架构可以在与用户的多轮会话中动态地提供一系列与产品推荐相关联的问题并且收集用户对这些问题的回答,并且可以至少从用户的回答中学习新的知识以便确定出推荐产品并且给出用于解释为什么推荐该产品的推荐理由。该LTE架构可以通过相对较短的会话来从大量的候选产品中筛选出推荐产品。

[0021] 向用户提供的问题可以针对各种产品属性。可选地,可以向问题附加指示不同产品属性的选项,从而用户可以在回答中直接选择所期望的选项。可选地,如果用户采用自然语言语句来回答问题或者发送消息,则可以对该自然语言语句进行解析,以便识别用户所期望的产品属性。

[0022] 每次从用户接收到对当前问题的回答时,LTE架构可以对多个候选产品执行产品排序。包括用户所选择的属性的候选产品将被排序地更高。通过产品排序,每个候选产品将具有对应的期望概率,其指示在该轮对话后该候选产品被用户所期望的可能性。可以针对每一轮会话来计算候选产品的期望概率。从而,随着会话的进行,每个候选产品的期望概率也会不断更新。可选地,在产品排序中也可以为每个候选产品计算权重,该权重可以被映射为期望概率。由于在候选产品的权重和期望概率之间具有特定的映射关系,因此两者在本文中相互替代地使用,或者被统称为概率权重。

[0023] 在对候选产品进行了产品排序之后,如果没有候选产品满足预定条件,则LTE架构可以进一步对多个候选问题执行问题排序。该预定条件可以包括例如候选产品的期望概率高于阈值等。这些候选问题可以是基于所述多个候选产品的属性来直接地或间接地预先确定的。例如,可以通过考虑之前提供过的问题以及用户的回答、候选产品的产品排序等,来对多个候选问题进行排序。该LTE架构可以通过基于熵的排序和/或基于策略的增强学习(reinforcement learning)排序来执行问题排序。通过问题排序,每个候选问题将具有对应的信息增益。可以从经排序的候选问题中选择具有例如最大信息增益的候选问题作为提供给用户的下一个问题。此处,最大信息增益可以指示该候选问题具有最高的区分度,从而能够以最快的速度来尽可能多地缩小后续将要考虑的候选产品的范围。可选地,在产品排序中也可以为每个候选问题计算权重,该权重可以被映射为信息增益。由于在候选问题的权重和信息增益之间具有特定的映射关系,因此两者在本文中相互替代地使用。

[0024] 如果在对候选产品进行了产品排序之后,存在满足预定条件的至少一个候选产品,则该LTE架构可以将该候选产品确定为推荐产品,并且向用户提供关于该推荐产品的产品信息。产品信息可以包括关于产品的任何信息,例如名称、编号、价格等。进一步地,该LTE架构可以确定该推荐产品的推荐理由。在一个方面,该LTE架构可以从之前提供给用户的问题中选择贡献最大的至少一个问题及相应的用户回答来生成推荐理由。例如,可以从提供给用户的历史问题中选择针对该推荐产品产生最大期望概率提升的问题,并且以该选择的问题及用户的回答为参考来生成推荐理由。在另一个方面,该LTE架构也可以通过预先训练的推荐理由生成模型来生成推荐理由。

[0025] 在一方面,本公开的实施例可以使用候选产品的产品属性列表来预先生成问题,并且可以从用户的回答中直接识别出用户所选择的属性或者通过将用户的回答与产品属性列表的比较来识别出用户所选择的属性。因此,本公开的实施例不需要用于捕获用户意图或需求的大规模标记语料库。在另一方面,由于本公开的实施例可以向用户提供附加有选项的问题,并且用户可以通过简单地选择选项来回答问题,因此聊天机器人可以有效地引导会话,以避免由在可提供的产品与用户需求之间的信息不对称所引起的问题并且提高产品推荐效率。在又一方面,根据本公开实施例的LTE架构能够在会话中实时地学习新的知识并且相应地引导下一轮会话,从而可以准确地理解用户意图或需求并且有效地提高用户的参与度、兴趣度等。在又一方面,本公开的实施例可以向用户提供推荐产品的推荐理由,从而可以增强用户对于推荐产品的关注度和信任度,并进而提高用户预订该推荐产品的可能性。

[0026] 根据本公开实施例的解释性产品推荐可以被应用在多种场景中。在一些场景中,本公开的实施例可以用于商品推荐,例如,礼物推荐等。在一些场景中,本公开的实施例可以用于服务推荐,例如面向任务的酒店、餐厅预定推荐等。

[0027] 图1示出了根据实施例的部署聊天机器人的示例性网络架构100。

[0028] 在图1中,网络110被应用于在终端设备120和聊天机器人服务器130之间进行互连。

[0029] 网络110可以是能够对网络实体进行互连的任何类型的网络。网络110可以是单个网络或各种网络的组合。在覆盖范围方面,网络110可以是局域网(LAN)、广域网(WAN)等。在承载介质方面,网络110可以是有线网络、无线网络等。在数据交换技术方面,网络110可以

是电路交换网络、分组交换网络等。

[0030] 终端设备120可以是能够连接到网络110、访问网络110上的服务器或网站、处理数据或信号等的任何类型的电子计算设备。例如,终端设备120可以是台式计算机、笔记本电脑、平板电脑、智能电话、AI终端等。尽管在图1中仅示出了一个终端设备,但是应当理解,可以有不同数量的终端设备连接到网络110。

[0031] 在一种实施方式中,终端设备120可以由用户使用。终端设备120可以包括可为用户提供自动聊天服务的聊天机器人客户端122。在一些情况下,聊天机器人客户端122可以与聊天机器人服务器130进行交互。例如,聊天机器人客户端122可以将用户输入的消息传送到聊天机器人服务器130,并且从聊天机器人服务器130接收与消息相关联的响应。然而,应当理解,在其它情况下,聊天机器人客户端122也可以本地地生成对用户输入的消息的响应,而不是与聊天机器人服务器130进行交互。

[0032] 聊天机器人服务器130可以连接到或包含聊天机器人数据库140。聊天机器人数据库140可以包括可由聊天机器人服务器130用于生成响应的信息。在一种实施方式中,聊天机器人服务器130还可以连接到产品数据库150。产品数据库150可以包括关于多个候选产品的各种产品信息,例如产品名称、产品属性等。产品信息可以是例如由产品提供方预先提供的或者从网络上抓取的。尽管产品数据库150被示为独立于聊天机器人数据库140,但是也可以将产品数据库150包括到聊天机器人数据库140中。当将一个或多个候选产品确定为推荐产品时,可以将与这些推荐产品相关联的产品信息提供给用户。

[0033] 应当理解,图1中所示的所有网络实体都是示例性的,根据具体的应用需求,网络架构100中可以涉及任何其它网络实体。

[0034] 图2示出了根据实施例的示例性聊天机器人系统200。

[0035] 聊天机器人系统200可以包括用于呈现聊天窗口的用户界面(UI) 210。聊天窗口可以由聊天机器人用于与用户进行交互。

[0036] 聊天机器人系统200可以包括核心处理模块220。核心处理模块220被配置用于通过与聊天机器人系统200的其它模块的协作,在聊天机器人的运行期间提供处理能力。

[0037] 核心处理模块220可以获得由用户在聊天窗口中输入的消息,这些消息被存储在消息队列232中。消息可以采用各种多媒体形式,例如文本、语音、图像、视频等。

[0038] 核心处理模块220可以用先进先出的方式处理消息队列232中的消息。核心处理模块220可以调用应用程序接口(API) 模块240中的处理单元来处理各种形式的消息。API模块240可以包括文本处理单元242、语音处理单元244、图像处理单元246等。

[0039] 对于文本消息,文本处理单元242可以对文本消息执行文本理解,并且核心处理模块220可以进一步确定文本响应。

[0040] 对于语音消息,语音处理单元244可以对语音消息执行语音到文本转换以获得文本语句,文本处理单元242可以对所获得的文本语句执行文本理解,并且核心处理模块220可以进一步确定文本响应。如果确定以语音提供响应,则语音处理单元244可以对文本响应执行文本到语音转换以生成相应的语音响应。

[0041] 对于图像消息,图像处理单元246可以对图像消息执行图像识别以生成相应的文本,并且核心处理模块220可以进一步确定文本响应。在一些情况下,图像处理单元246也可以用于基于文本响应来获得图像响应。

[0042] 此外,尽管在图2中未示出,API模块240还可以包括任何其它处理单元。例如,API模块240可以包括视频处理单元,该视频处理单元用于与核心处理模块220协作以处理视频消息并确定响应。

[0043] 核心处理模块220可以通过数据库250来确定响应。数据库250可以包括可由核心处理模块220访问以用于确定响应的各种信息。

[0044] 数据库250可以包括纯聊天索引集251。纯聊天索引集251可以包括索引项,索引项被准备用于聊天机器人和用户之间的自由聊天,并且可以是利用来自例如社交网络的数据来建立的。

[0045] 数据库250可以包括至少一个候选产品集合252。候选产品集合252可以包括候选产品列表、每个候选产品的属性等。在一种实施方式中,可以针对每一种产品类别建立一个候选产品集合,该集合包括属于该类别的多个候选产品。可以基于不同的层级或标准来设置产品类别,并且将候选产品划分到对应的一个或多个类别。例如,“礼物”类别可以包括巧克力、首饰、鲜花、工艺品等候选产品,“电子商品”类别可以包括手机、计算机、电视、耳机等候选产品,等等。此外,候选产品集合252中的候选产品也可以包括例如具体品牌、具体型号的候选产品,例如XX品牌YY型号的手机Z、酒店K等。以候选产品“手机Z”为例,其属性可以包括例如4G+5G移动网络、6寸屏幕、指纹识别功能、等等。候选产品集合252中所包括的各种信息可以是例如由产品提供方预先提供的或者从网络上抓取的。

[0046] 数据库250可以包括候选问题集合253。候选问题集合253可以包括至少基于候选产品集合252中的候选产品的属性所确定的候选问题列表。每个候选问题可以针对一项或多项产品属性,以便当用户提供针对该候选问题的回答时,可以基于该回答来确定用户所期望的属性并且识别出哪些候选产品具有所期望的属性。在一种实施方式中,每个候选问题还可以附加有一个或多个选项,这些选项对应于该候选问题所针对的产品属性。

[0047] 在一种实施方式中,候选产品集合252和候选问题集合253可以通过产品属性而链接在一起。图3示出了根据实施例的在候选产品集合与候选问题集合之间的映射300。候选产品集合 $P_i$ 包括多个候选产品 $\{p_1^i, p_2^i, \dots, p_M^i\}$ ,其中, $i$ 表示这些候选产品的类别, $M$ 为候选产品的数量。候选产品集合 $P_i$ 还包括每个候选产品的多个属性。例如,候选产品 $p_1^i$ 的属性包括 $\{v_{11}^i, v_{12}^i, \dots, v_{1N}^i\}$ ,其中, $N$ 为属性的数量。候选问题集合 $Q_i$ 包括多个候选问题 $\{q_1^i, q_2^i, \dots, q_N^i\}$ ,其中, $i$ 表示该候选问题集合 $Q_i$ 所针对的产品类别, $N$ 为候选问题的数量。候选问题集合 $Q_i$ 还可以包括不同候选产品对于每个候选问题的参考回答。例如,对于候选问题 $q_1^i$ ,候选产品 $p_1^i$ 的参考回答为 $v_{11}^i$ ,候选产品 $p_2^i$ 的参考回答为 $v_{21}^i$ ,等等。因此,图3中由候选产品的属性所构成的矩阵也可以被称为参考回答矩阵 $D_i$ 。参考回答矩阵 $D_i$ 将候选产品集合 $P_i$ 与候选问题集合 $Q_i$ 链接在一起。 $D_i$ 中的每个元素可以被表示为 $v_{mn}^i$ ,其中, $m$ 是候选产品索引, $n$ 是候选问题索引。应当理解,尽管图3中示出候选问题集合具有 $N$ 个候选问题,但是候选问题集合也可以具有任意数量的候选问题。尽管图3中示出每个候选产品具有 $N$ 个属性,但是不同的候选产品也可以具有不同数量的属性。

[0048] 应当理解,本公开的实施例并不局限于任何特定的基于候选产品的属性来构造候选问题的方式。作为示例,假设候选产品1为耳机X,其包括属性“入耳式”,候选产品2为耳机

Y,其包括属性“包耳式”,则可以构造候选问题“你喜欢入耳式的耳机还是包耳式的耳机?”。针对该候选问题,候选产品1的参考答案为“入耳式”,而候选产品2的参考答案为“包耳式”。如果用户的回答为“入耳式”,则可以确定候选产品1具有用户所选择的属性,相应地,候选产品1可以具有比候选产品2更高的排名。

[0049] 数据库250可以包括会话记录254。会话记录254可以包括在聊天机器人与用户的会话中,聊天机器人提供过的与产品推荐相关联的历史问题以及来自用户的相应历史回答。

[0050] 数据库250可以包括候选产品评估状态255。候选产品评估状态255可以包括在每一轮会话后针对每个候选产品所评估的期望概率或权重。

[0051] 聊天机器人系统200可以包括模块集合260,模块集合260是可以由核心处理模块220操作以生成或获得响应的功能模块的集合。

[0052] 模块集合260可以包括产品排序模块261。每次从用户接收到对当前问题的回答时,产品排序模块261可以重新计算每个候选产品的期望概率或权重,并且相应地对候选产品进行排序。所计算的期望概率或权重可以用于更新候选产品评估状态。当一个或多个候选产品满足预定条件时,可以将该一个或多个候选产品确定为将要提供给用户的推荐产品。

[0053] 模块集合260可以包括问题排序模块262。在提供下一个问题之前,问题排序模块262可以计算每个候选问题的权重,并且相应地对候选问题进行排序。问题排序可以至少基于产品排序的结果。例如,问题排序可以考虑当前候选产品评估状态中所包括的每个候选产品的当前期望概率或权重。此外,问题排序还可以考虑会话记录等。可以选择排名最高的候选问题,或者从排名最高的多个候选问题中随机选择一个候选问题,来作为将要提供给用户的下一个问题。

[0054] 应当理解,尽管产品排序模块261和问题排序模块262被示为单独的模块,但是这两个模块也可以合并在一起,从而可以通过执行一个统一的过程而实现产品排序和问题排序两者。

[0055] 模块集合260可以包括推荐理由生成模块263。推荐理由生成模块263可以通过各种方式来生成所确定的推荐产品的推荐理由。在一种方式中,推荐理由生成模块263可以从之前提供的问题中选择针对该推荐产品产生最大期望概率提升的问题,并且以该选择的问题及用户的回答为参考来生成推荐理由。在另一种方式中,推荐理由生成模块263可以通过预先训练的推荐理由生成模型来生成推荐理由。

[0056] 模块集合260可以包括语句解析模块264。当用户采用自然语言语句来回答问题或者发送消息时,语句解析模块264可以对该自然语言语句进行解析,以便识别用户所期望的产品属性。

[0057] 模块集合260可以包括响应提供模块265。响应提供模块265可以被配置用于提供或传递对用户的消息的响应。在一些实施方式中,响应提供模块265所提供的响应可以包括所确定的推荐产品的产品信息、推荐理由等。

[0058] 核心处理模块220可以将所确定的响应提供给响应队列或响应高速缓存234。例如,响应高速缓存234可以确保能够以适当的时序来显示响应序列。假设对于一个消息,由核心处理模块220确定了不少于两个响应,则对响应的延迟设置可能是必要的。例如,

如果用户输入的消息是“你吃早餐了吗?”,则可能确定出两个响应,例如,第一响应是“是的,我吃了面包”,第二响应是“你呢?还觉得饿吗?”。在这种情况下,通过响应高速缓存234,聊天机器人可以确保立即向用户提供第一响应。进而,聊天机器人可以确保以例如1或2秒的时间延迟来提供第二响应,使得第二响应将在第一响应之后的1或2秒被提供给用户。由此,响应高速缓存234可以管理将要发送的响应以及针对每个响应的适当时序。

[0059] 可以将响应队列或响应高速缓存234中的响应进一步传送到UI 210,以便可以在聊天窗口中将响应显示给用户。

[0060] 应当理解,图2中的聊天机器人系统200中所示的所有单元都是示例性的,并且根据具体的应用需求,在聊天机器人系统200中可以省略任何所示的单元并且可以涉及任何其它单元。

[0061] 图4示出了根据实施例的用于提供解释性产品推荐的示例性总体过程400。

[0062] 在410处,可以接收来自用户的消息。该消息可能指示用户想要获得产品推荐的意图。例如,该消息可以为“为我推荐个礼物”。

[0063] 在420处,可以至少基于在410处所接收的消息,确定将要推荐的产品所属于的产品类别。例如,当所接收的消息为“为我推荐个礼物”,则可以确定产品类别为“礼物”。例如,当所接收的消息为“给我推荐个电子商品”时,可以确定产品类别为“电子商品”。

[0064] 在430处,可以基于本公开实施例所提出的LTE架构来进行与用户的多轮会话。在该多轮会话中,可以动态地提供与在所确定的产品类别下的产品推荐相关联的问题并且收集用户对这些问题的回答,并且确定推荐产品以及推荐理由。后面将结合图5来具体讨论。

[0065] 在450处,可以向用户提供推荐产品的产品信息和推荐理由以作为响应。

[0066] 图5示出了根据实施例的用于提供解释性产品推荐的示例性具体过程500。过程500示出了根据本公开实施例的LTE架构的示例性工作过程。在执行过程500之前,已经确定了用户想要获得产品推荐的意图,并且确定了产品类别。因此,过程500将被执行用于在该产品类别下提供产品推荐。应当理解,过程500可以被迭代地执行,直到确定出推荐产品。

[0067] 在502处,可以向用户提供与产品推荐相关联的问题。该问题可以是针对产品属性的。可选地,可以与该问题一起向用户提供多个选项,以使用户可以从这些选项中选择回答。

[0068] 在504处,可以从用户处接收针对所提供问题的回答。

[0069] 在问题附加有选项的情况下,用户的回答可能是对所附加的选项中的一个或多个选项的直接选择。因此,可以在506处识别用户针对这些选项所做出的选择,以便确定用户所选择的产品属性。例如,如果问题附加了三个选项,并且用户的回答为第二个选项的索引“2”或者包括与第二个选项的内容相关联的表述,则可以确定第2个选项所指示的产品属性是用户所期望的。

[0070] 在问题没有附加选项的情况下,用户的回答可能采用自然语言语句的形式。可以在508处对自然语言语句执行解析。该解析可以采用任何已有的意图-槽(intent-slot)解析技术,以便从自然语言语句中检测出槽和对应的值。可以将这些值用作关键词,以从候选问题集合中检索出相关问题和对应回答。所检索出的相关问题可以包括提供给用户的问题以及其它问题。例如,假设提供给用户的问题是“你喜欢待在安静的地方么?”并且用户的回答是自然语言语句“我喜欢安静,但我也喜欢跑步”,则至少可以从该自然语言语句中检测

出关键词“安静”以及“跑步”。安静可以是对相关问题“你喜欢待在安静的地方么?”的回答,而“跑步”可以是对另一相关问题“你有什么爱好?”的对应回答。从而,从该自然语言语句中至少获得了用户对于两个相关问题的两个回答。

[0071] 在510处,可以将<问题,回答>对添加到会话记录512中。问题可以是所提供的问题或相关问题,回答可以是用户所选择的选项的内容或者相关问题的对应回答。会话记录512中所包括的问题和回答可以随着会话的进行而不断更新。

[0072] 在514处,可以至少基于所提供的问题和用户的回答来对多个候选产品516执行产品排序。通过产品排序,可以更新每个候选产品的期望概率,并且将所计算的期望概率包括到候选产品评估状态518中。在一种实施方式中,候选产品评估状态518可以采用向量的形式,该向量的维度对应于候选产品的数量,每个维度的值对应于一个候选产品的期望概率。此外,如前所述,也可以使用候选产品的权重来代替期望概率,从而候选产品评估状态518中包括了每个候选产品的权重。

[0073] 在520处,可以基于产品排序的结果来确定多个候选产品中是否存在满足预定条件的至少一个候选产品。该预定条件可以指示一个候选产品是否达到了被确定为推荐产品的条件。例如,预定条件可以为候选产品的期望概率或权重高于阈值,其中该阈值可以是预先经验性设定的。

[0074] 如果在520处确定不存在满足预定条件的候选产品,则过程500可以进一步向用户提供问题。在522处,可以对多个候选问题524执行问题排序,以便确定将提供给用户的下一个问题。问题排序可以至少基于产品排序的结果,例如,当前候选产品评估状态中所包括的每个候选产品的当前期望概率或权重。问题排序还可以基于会话记录512中的历史问题和历史回答等。在问题排序中,可以计算出每个候选问题的权重或信息增益,并且根据权重或信息增益来对候选问题进行排序。可以采用各种方式来执行问题排序,例如,通过基于熵的排序522-1、通过基于策略的增强学习排序522-2等,后面将对此具体讨论。

[0075] 在526处,可以基于候选问题的权重来选择将要提供给用户的下一个问题。例如,可以选择排名最高的候选问题作为下一个问题,可以从排名最高的多个候选问题中随机选择一个候选问题作为下一个问题,等等。

[0076] 在选择了一个问题之后,过程500迭代地返回到502以将所选择的下一个问题提供给用户,并接着执行后续步骤。

[0077] 如果在520处确定存在满足预定条件的至少一个候选产品,则可以在528处将该至少一个候选产品确定为是将要提供给用户的推荐产品。

[0078] 在530处,可以确定该推荐产品的推荐理由。可以采用各种方式来确定推荐理由。

[0079] 在一种方式中,可以从之前提供给用户的问题中选择贡献最大的至少一个问题及相应的用户回答来生成推荐理由。例如,可以从提供给用户历史问题中选择针对该推荐产品产生最大期望概率提升的问题,并且以该选择的问题及用户的回答为参考来生成推荐理由。可以基于预先定义的各种规则来基于所选择的问题及回答构建推荐理由。推荐理由可以包括对所选择的问题及回答中的至少一个的简单重复。推荐理由可以包括对所选择的问题及回答中的至少一个的经变换的表达。例如,推荐理由中可以包括从回答“喜欢KFC的汉堡”所变换得到的表述“喜欢快餐食品”。推荐理由可以包括对所选择的问题及回答中的至少一个的概括表述。例如,可以通过任何自然语言处理技术来将这些问题和回答的内容进

行语义概括。推荐理由可以包括一些在自由聊天中常用的词语或短语,以使得句子表述更为自然。例如,在推荐理由中添加“我给出以上推荐的理由是…”、“考虑到…,我决定推荐…”等表述。

[0080] 在另一种方式中,可以预先训练一个推荐理由生成模型,以基于由会话记录中的历史回答所选择的属性、推荐产品的属性、推荐产品的描述等中的至少一个来生成推荐理由,后面将具体讨论。

[0081] 在532处,可以向用户提供包括推荐产品的产品信息和推荐理由的响应。产品信息可以从例如图1的产品数据库150或者图2的候选产品集合252中提取的。

[0082] 应当理解,图5中的过程500所包括的所有处理步骤及其顺序都是示例性的,根据具体的应用需求,可以对过程500中的处理步骤进行任何添加、删除或替换。例如,在一种实施方式中,在聊天机器人接收到用户发送的自然语言消息并且该消息并非针对聊天机器人提供的任何与产品推荐相关联的问题的情况下,可以直接在508处对该自然语言消息执行解析,以便确定相对应的相关问题和对应回答,并进而执行后续处理。此外,在一种实施方式中,可以在520处定义或添加其它用于确定推荐产品的判断条件。例如,可以预先定义向用户提供的问题数量的阈值。如果确定在过程500中已经向用户提供了超过阈值数量的问题,则可以直接在528处确定推荐产品,该推荐产品可以是例如当前具有最高期望概率或权重的候选产品。

[0083] 下面将讨论通过基于熵的排序来选择问题的方法,其可以用于对候选问题进行排序并进而从候选问题中选择将要提供给用户的下一个问题。

[0084] 下一个问题可以被选择为使得该问题能够尽可能多地剔除以较低可能性作为推荐产品的那些候选产品,而无论用户对该问题给出怎样的回答。可以选择下一个问题,该问题能够将候选产品划分为两个具有相近大小或相近权重的子集。例如,在问题的回答为二元回答时,例如“是”或“否”,候选产品可以被划分为参考回答为“是”的子集和参考回答为“否”的子集。这两个子集可以具有相等或近似数量的候选产品,或者具有相等或近似累加权重的候选产品。例如,在问题的回答具有两个以上选项时,例如“便宜”、“中等”和“昂贵”,候选产品可以被划分为参考回答分别是“便宜”、“中等”和“昂贵”的三个子集,这些子集具有相近大小或相近权重。在一种实施方式中,在通过一轮会话确定了用户当前所选择的产品属性后,可以将符合用户当前所选择属性的那些候选产品作为确定下一轮会话的问题时所参考的候选产品集合。通过不断地执行上述过程,可以持续地缩小用于确定下一个问题的候选产品集合的规模。

[0085] 在基于熵的排序中,对于产品类别 $i$ ,可以初始地为每个候选产品 $p_m^i$ , ( $1 \leq m \leq M$ ) 分配一个先验概率权重 $w(\cdot)$ 。该权重可以是参考搜索引擎上针对该候选产品的搜索频率或者电子商务网站上针对该候选产品的点评分值或订购频率来设定的。然后,可以将 $w(\cdot)$  归一化为:

$$[0086] \quad w'(p_m^i) = \frac{w(p_m^i)}{\sum_{m=1}^M w(p_m^i)} \quad \text{公式 (1)}$$

[0087] 对于一个候选产品 $p_m^i$ ,可以计算其对选择候选问题 $q_n$ 的贡献 $Y_{m,n}^l$ :

$$[0088] \quad Y_{m,n}^l = \frac{f_{m,n}^l + \alpha I(v_{m,n}^{i,l} = \text{是})}{\sum_{l=1}^{L_n} \{f_{m,n}^l + \alpha I(v_{m,n}^{i,l} = \text{是})\}} \quad \text{公式 (2)}$$

[0089] 其中,  $f_{m,n}^l$  表示在历史数据中在用户们选择了候选问题  $q_n$  的选项  $l$  后最终选择候选产品  $p_m^i$  的频率。历史数据可以是通过大量用户的使用信息进行收集而获得的, 从而  $f_{m,n}^l$  反映了大量用户的历史使用信息。 $I(\cdot)$  是指示符函数, 其当  $\{v_{m,n}^{i,l} = \text{是}\}$  成立时返回 1, 否则返回 0, 其中,  $\{v_{m,n}^{i,l} = \text{是}\}$  表示选项  $l$  是候选产品  $p_m^i$  针对候选问题  $q_n$  的参考回答。参数  $\alpha$  用于在历史使用信息与参考回答矩阵  $D_i$  中的参考回答之间进行平衡。例如, 当  $\alpha$  为 0 时, 仅考虑历史使用信息, 而忽略参考回答。当  $\alpha$  被设置为极大的值时, 较大程度地考虑参考回答而忽略历史使用信息。此外, 也可以将  $\alpha$  扩展为时间衰减函数  $\alpha(t)$ , 其中  $t$  表示时间。

[0090] 在一种实施方式中, 可以将负香农熵用于选项的多变量伯努利分布, 并且计算参数  $M_{m,n}$ :

$$[0091] \quad M_{m,n} = \sum_{l=1}^{L_n} Y_{m,n}^l \log_2 Y_{m,n}^l \quad \text{公式 (3)}$$

[0092] 然后, 可以将候选问题  $q_n$  的权重  $w(q_n)$  计算为:

$$[0093] \quad w(q_n) = \sum_{m=1}^M w'(p_m^i) \times M_{m,n} \quad \text{公式 (4)}$$

[0094] 可以针对候选问题集合  $Q_i$  中的每一个候选问题执行上述过程, 以便计算出每个候选问题的权重。可以基于权重来对候选问题进行排序, 并且从经过排序的候选问题中选择将要提供给用户的下一个问题  $q_{n_t}$ 。

[0095] 基于用户针对问题  $q_{n_t}$  的回答  $a$ , 可以更新每个候选产品  $p_m^i$  的权重:

$$[0096] \quad w(p_m^i) = w'(p_m^i) \times \sum_{l=1}^{L_{n_t}} I(v_{m,n_t}^{i,l} = a) \times w(q_{n_t}) \quad \text{公式 (5)}$$

[0097] 在用户的回答  $a$  不与问题  $q_{n_t}$  的任何选项相匹配的情况下,  $w(p_m^i)$  将为 0。在这种情况下, 可以通过将  $I(\cdot)$  的返回值设置为极小的值, 例如 0.01, 而不是 0, 以避免完全舍弃该候选产品。

[0098] 可以将所计算的每个候选产品的权重添加到候选产品评估状态中。如果某个候选产品满足预定条件, 例如该候选产品的权重或期望概率超过阈值, 则可以将该候选产品确定为推荐产品。

[0099] 如果没有候选产品满足预定条件, 可以迭代地再次执行上述基于熵的排序。例如, 可以再次通过公式 (1) 来将  $w(p_m^i)$  归一化为  $w'(p_m^i)$ , 然后继续执行后续的处理。

[0100] 可以用脚标  $t$  和  $t+1$  分别表示当前阶段和下一阶段, 从而得到表示在当前阶段使用的候选产品权重的  $w'_t(p_m^i)$  以及表示在当前阶段之后将在下一阶段使用的候选产品权重的  $w'_{t+1}(p_m^i)$ 。然后, 可以计算由向用户提供的当前问题  $q_{n_t}$  以及用户回答  $a$  所引起的候选产品  $p_m^i$  的权重提升或期望概率提升  $g_{t+1}(p_m^i, q_{n_t}, a)$ :

$$[0101] \quad g_{t+1}(p_m^i, q_{n_t}, a) = w'_{t+1}(p_m^i) - w'_t(p_m^i) \quad \text{公式 (6)}$$

[0102] 在一种实施方式中,在将一个候选产品确定为推荐产品后,可以选择引起该推荐产品的最大权重提升或最大期望概率提升的问题及对应用户回答来生成该推荐产品的推荐理由。例如,假设第一个问题导致该推荐产品的期望概率从0变为0.2,第二个问题导致期望概率从0.2变为0.6,第三个问题导致期望概率从0.6变为0.9,其中0.9的期望概率超出了阈值0.8。由于第一个问题引起的期望概率提升为0.2,第二个问题引起的期望概率提升为0.4,第三个问题引起的期望概率提升为0.3,因此可以选择引起最大期望概率提升(即,0.4)的第二个问题及对应用户回答来生成推荐理由。此外,如前所述,也可以通过推荐理由生成模型来生成推荐理由。

[0103] 通过执行上述基于熵的排序,在一个方面,可以在会话期间不断地更新候选问题的权重,以便选择将要提供给用户的下一个问题,在另一个方面,可以不断地更新候选产品权重或期望概率,以便可以确定推荐产品。

[0104] 应当理解,以上所有公式都是示例性的,其仅用于说明示例性的处理过程,本公开的实施例并不局限于任何上述具体公式。例如,替代公式(3)和公式(4),也可以通过下面的过程来计算候选问题 $q_n$ 的权重。

[0105] 在通过公式(2)计算了 $Y_{m,n}^l$ 后,可以接着计算表示选项1在候选问题 $q_n$ 中的重要性的 $Y_n^l$ :

$$[0106] \quad Y_n^l = \sum_{m=1}^M w'(p_m^i) \times Y_{m,n}^l \quad \text{公式 (7)}$$

[0107] 然后,利用 $Y_n^l$ 的负方差来计算问题 $q_n$ 的权重 $w(q_n)$ :

$$[0108] \quad w(q_n) = - \sum_{l=1}^{L_n} (Y_n^l - \frac{\sum_{l=1}^{L_n} Y_n^l}{L_n})^2 \quad \text{公式 (8)}$$

[0109] 此外,应当理解,在以上描述基于熵的排序过程中,除了 $p_m^i$ 以外,为了简化而省略了大部分变量的类别索引 $i$ 。

[0110] 下面将讨论通过基于策略的增强学习排序来选择问题的方法,其可以用于对候选问题进行排序并进而从候选问题中选择将要提供给用户的下一个问题。

[0111] 基于策略的增强学习算法可以被用于在允许用户对单一属性问题做出回答的限制下预测特定实体(例如,名人等)。单一属性问题可以指目的在于获取二元回答(例如,是、否等)的问题。本公开的实施例将该算法调整为适合于多选项问题以用于解释性产品推荐。例如,该算法可以被调整为适合于附加有多个选项的问题的面向任务的场景。用户可以选择这些选项中的任意子集。这可以被视为是在若干个单一属性问题上的跨越。进一步地,在用户以涉及对多个问题的回答的自然语言语句来表达需求时,可以实现在与用户的单轮对话中对多个单一属性问题或组合问题的跨越。

[0112] 问题排序可以被归纳为由五元组 $\{S, A, P, R, \gamma\}$ 表示的有限马尔可夫决策过程(MDP)。S是连续的候选产品评估状态空间,S中的每一个状态 $s$ 表示存储了候选产品的期望概率的向量。 $A = \{q_1, q_2, \dots, q_n\}$ 是候选问题的集合。 $P(S_{t+1} = s' | S_t = s, A_t = q)$ 是状态转移概

率矩阵。 $R(s, q)$  表示回馈函数或回馈网络。 $\gamma \in [0, 1]$  是衰减系数以用于对长期返回值进行折算。在基于策略的增强学习算法中, 在每一个时间步骤  $t$ , 聊天机器人可以根据策略函数或策略网络  $\pi_\theta(q_t | s)$  来在当前候选问题评估状态  $s$  下提供候选问题  $q_t$ 。在提供了候选问题  $q_t$  并且接收到用户对  $q_t$  的回答后, 可以产生回馈分值  $r_{t+1}$  并且将候选问题评估状态  $s$  更新为  $s'$ 。可以将四元组  $\{s, q_t, r_{t+1}, s'\}$  用作增强学习过程中的情景 (episode)。可以将时间步骤  $t$  的长期回馈  $R_t$  定义为:  $R_t = \sum_{k=0}^{T-t-1} \gamma^k r_{t+k+1}$ 。

[0113] 可以使得候选问题评估状态  $s_t$  保持跟踪候选产品  $\{p_m^i\}$  在时间步骤  $t$  处的置信度, 例如期望概率。例如,  $s_t \in \mathbb{R}^M, \forall s_{t,m} \geq 0$ , 并且  $\sum_{m=1}^M s_{t,m} = 1$ 。此处,  $s_{t,m}$  表示在时间步骤  $t$  处用户期望产品  $p_m^i$  的置信度。初始地, 类似于上述基于熵的排序,  $s_0$  可以采用候选产品的先验期望概率。

[0114] 给定候选产品集合  $P_i = \{p_m^i\}$  以及候选问题集合  $Q = \{q_n\}$ , 可以计算用户的回答在 each 问题  $q_n$  的多个选项上的归一化置信度。即, 可以将候选产品评估状态的转移定义为:

$$[0115] \quad s_{t+1} = s_t \odot \beta \quad \text{公式 (9)}$$

[0116] 此处,  $\odot$  是点积运算符,  $\beta$  取决于用户对于在时间步骤  $t$  处所选择的问题  $q_t$  的回答  $x_t$ , 其中,  $q_t$  在候选问题集合  $Q = \{q_n\}$  中具有索引  $n_t$ 。当用户针对当前问题  $q_t$  的回答选择了选项  $\{l\} \in [1, \dots, L_{n_t}]$  时, 可以定义  $\beta = [Y_{1,n_t}^{\{l\}}, \dots, Y_{M,n_t}^{\{l\}}]$ , 并且  $Y_{m,n_t}^{\{l\}} = \sum_{l \in \{l\}} Y_{m,n_t}^l$  采用与公式 (2) 相类似的定义。通过这种方式, 可以基于用户对于时间步骤  $t$  处的问题  $q_t$  的回答  $\{l\}$  而将候选产品  $p_m^i$  的置信度  $s_{t,m}$  更新为  $s_{t+1,m}$ 。

[0117] 为了使得基于策略的增强学习算法能够将之前提供过的问题作为确定下一个问题的前提条件并且使用用户的历史选择来对候选问题进行排序, 本公开的实施例提出了使用基于神经网络的LTE回馈网络, 其采用四元组  $\{s_t, Q_t, \mathcal{L}_t, p_m^i\}$  作为输入, 并且输出下一步骤的回馈  $r_{t+1}$ 。LTE回馈网络采用具有sigmoid输出的最大层感知 (MLP), 以便在训练期间学习到适当的立即回馈。下面的表1示出了示例性的训练过程。

联合训练 LTE 回馈网络 $R$ 、策略函数 $\pi_\theta$ 以及值网络 $V$	
	1.1 $\mathcal{M} \leftarrow \emptyset$ // 初始化情景记忆
	1.2 随机初始化 $R$ 、 $V$ 和 $\pi_\theta$
	1.3 对于循环次数 1 到 $Z$ , 执行:
	1.4 $p_m^i \leftarrow \{p_m^i\}$ 中的加权选择, $S_1 \leftarrow \{\}$ , $S_2 \leftarrow \{\}$
	1.5 对于 $[1, T]$ 中的 $t$ , 执行:
	1.6 $q_t \leftarrow \pi_\theta(\cdot   s_{t-1})$ // 从候选问题中选择问题
	1.7 $s_t \leftarrow s_{t-1} \odot \beta$ // 更新候选产品评估状态
	1.8 $(s_t, Q_t, \mathcal{L}_t) \rightarrow S_1$
	1.9 如果 $\operatorname{argmax}(s_t) = p_m^i$ , 则终止 // 找到推荐产品, 停止会话
[0118]	1.10 如果 $\operatorname{argmax}(s_t) \neq p_m^i$ , 则 $r_T \leftarrow \kappa$ , 否则 $r_T \leftarrow -\kappa$
	1.11 对于 $S_1$ 中的 $(s_t, Q_t, \mathcal{L}_t)$ , 执行:
	1.12 $r_{t+1} \leftarrow R(s_t, Q_t, \mathcal{L}_t, p_m^i)$ // 计算回馈
	1.13 $(s_t, Q_t, \mathcal{L}_t, r_{t+1}) \rightarrow S_2$
	1.14 对于 $S_2$ 中的 $(s_t, Q_t, \mathcal{L}_t, r_{t+1})$ , 执行:
	1.15 $r'_{t+1} \leftarrow \operatorname{sigmoid}(\sum_{k=0}^{T-t-1} \gamma^k r_{t+k+1})$ // 累积后续的回馈
	1.16 利用 $(s_t, r'_{t+1})$ 来更新 $V$
	1.17 $v_{t+1} \leftarrow V(s_t)$
	1.18 $(s_t, Q_t, \mathcal{L}_t, p_m^i, r'_{t+1} - v_{t+1}) \rightarrow \mathcal{M}$
	1.19 如果 $ \mathcal{M}  > K_1$ , 则:
	1.20 利用 $\mathcal{M}$ 中的小批量数据 (mini-batch) 来更新 $R$
	1.21 利用 $\mathcal{M}$ 中的小批量数据来更新 $\pi_\theta$

[0119] 表1

[0120] 在表1所示的过程中, 对问题  $Q_t$  及对应回答  $\mathcal{L}_t$  进行嵌入, 并且将所得到的嵌入向量与  $s_t$  级联, 以用于在例如平方差损耗函数下对回馈网络  $R$  进行训练。回馈网络进一步被用于训练策略函数以对候选问题进行排序并选择下一个问题。策略函数可以是在例如交叉熵损耗函数下使用增强算法来训练的。

[0121] 值网络  $V$  可以用于对当前状态  $s_t$  的优良度进行评分。值网络可以估计当前状态自身有多优良以在情景中被选择。值网络可以使用例如平方差损耗函数, 并且采用累积的回馈  $r'_{t+1}$  作为参考分数。在更新后, 新估计的分值  $v_{t+1}$  被从  $r'_{t+1}$  中减去, 以进一步用于分别更新  $R$  和  $\pi_\theta$ 。

[0122] 在表1中包括四个循环。第一个循环是从1.3到1.21, 其将循环次数控制在  $Z$  以内。第二个循环是从1.5到1.9, 其应用策略函数来选择问题, 并且更新候选产品评估状态。在一种实施方式中, 可以限制一个候选问题在一个会话期间只能被选择并使用一次。将第二个循环中获得的结果存储在  $S_1$  中, 以便在后续步骤中使用。第三个循环是从1.11到1.13, 其应用回馈网络  $R$  以获得立即回馈。第四个循环是从1.14到1.21, 其通过从情景记忆  $\mathcal{M}$  中挑取小批量数据来更新策略函数和回馈网络中的参数。在一种实施方式中, 策略函数和回馈网络可以采用具有例如3层隐藏层的MLP, 并且利用基于ADAM优化器的算法。

[0123] 在应用基于策略的增强学习排序的阶段, 当前问题  $q_t$  以及用户的回答对于所有候选产品  $\{p_m^i\}$  所引起的  $t+1$  步骤的期望概率提升为:

$$[0124] \quad g_{t+1}(\{p_m^i\}, q_{n_t}, a) = s_{t+1} - s_t \quad \text{公式 (10)}$$

[0125] 由此,可以通过将每个候选产品的期望概率与预定条件进行比较来确定出推荐产品。

[0126] 应当理解,在以上关于基于熵的排序和基于策略的增强学习排序的讨论中所涉及的所有公式、变量、参数等都是示例性的。根据具体的应用需求,可以对这些公式、变量、参数等进行任何形式的删除、替换和添加。本公开的实施例并不局限于以上讨论的任何细节。

[0127] 图6示出了根据实施例的用于训练推荐理由生成模型的示例性过程600。

[0128] 可以从网络上,例如电子商务网站,收集用于训练推荐理由生成模型的训练数据。可以首先识别或指定用于生成训练数据的产品集合610。对于产品集合610中的每个产品,可以进而从提供该产品的网站上获得该产品的类别信息640。通常,产品的类别信息可以包括一系列不同层级的类别。例如,对于产品“三文鱼”,其可能对应于“食物”、“海鲜”、“鱼”等不同层级的多个类别。

[0129] 在一些情况下,在用户从电子商务网站上购买了产品后,其可能会提供针对该产品的点评(review),这些点评可能含有采用自然语言的对购买该产品的解释性理由。因此,可以收集针对产品集合610中的产品的点评620。可选地,过程600可以对点评620执行过滤以获得经过滤的点评622。例如,可以对点评执行情感分析,以便滤除负面的点评而保留正面的点评。此外,例如,可以采用预先定义的表达模式来检测点评的有效性,并且滤除表达中包括过少的词语或者过多的重复字符的无效的点评。应当理解,在下文中涉及到的术语“点评”可以广泛地指点评620和经过滤的点评622中的任意一者或两者。过程600可以从点评620或经过滤的点评622中提取产品的属性信息650。例如,对于点评“这些鞋子棒极了!超级柔软,减震,而且很轻”,可以提取出关于产品“鞋子”的属性,例如“柔软”、“减震”、“轻”等。

[0130] 通常,电子商务网站上还可能提供对产品的描述,例如,产品的特性、参数等等。这些描述通常明确地包括以采用自然语言表达的产品的各种属性。因此,可以收集针对产品集合610中的产品的描述630。可选地,过程600可以对产品的描述630执行概括以获得产品描述概要632。在一些情况下,产品的描述可能较长,因此可以仅将描述中的主要内容用于后续的训练。可以使用已有的非监督文本排序算法来对描述执行概括。应当理解,在下文中涉及到的术语“产品的描述”可以广泛地指产品的描述630和产品描述概要632中的任意一者或两者。过程600可以从产品的描述630或产品描述概要632中提取产品的属性信息650。

[0131] 通过上述的数据收集过程,可以形成采用<属性+描述,点评>对形式的训练数据集660。每个<属性+描述,点评>数据对与一个特定产品相关联。可以使用训练数据集660来对推荐理由生成模型670进行训练。可以将训练数据中的“属性+描述”作为模型的输入,而将“点评”作为模型的输出。推荐理由生成模型670可以采用变换器(transformer)架构,在该架构中,编码部分和解码部分都可以采用具有位置编码的自注意力机制来用于序列依存学习。在训练中,编码部分可以处理产品的属性和描述,而解码部分可以处理不同购买者对该产品的先前点评。经训练的推荐理由生成模型670可以生成采用自然语言的类似于点评的推荐理由。

[0132] 应当理解,过程600中的所有步骤都是示例性的,根据具体的应用需求和设计,可

以对过程600做出任何形式的改变。例如,当从网络上针对某个产品所收集到的点评的数量较少时,可以使用该产品的描述来代替点评以用于构建解释性理由。例如,训练数据可以采用<属性,描述>的形式,从而,在训练推荐理由生成模型时,可以将“属性”作为模型的输入,而将“描述”作为模型的输出。

[0133] 图7示出了根据实施例的用于生成推荐理由的示例性过程700。如前所述,在确定了推荐产品之后,可以进而确定该推荐产品的推荐理由,以便在响应中提供给用户。在过程700中使用推荐理由生成模型710来生成推荐产品的推荐理由,该模型710可以是通过图6的过程600来预先训练的。

[0134] 会话记录702中可以包括会话中由聊天机器人提供的历史问题和由用户提供历史回答。用户的历史回答可以指示用户所选择或期望的产品属性704。可以将用户选择的关于推荐产品的属性704提供给推荐理由生成模型710以作为输入。

[0135] 可以通过例如图6中的过程600来获得推荐产品的属性706,并且提供给推荐理由生成模型710以作为输入。此外,还可以通过例如图6中的过程600来获得推荐产品的描述708,并且提供给推荐理由生成模型710以作为输入。

[0136] 推荐理由生成模型710可以根据用户选择的属性704、推荐产品的属性706和推荐产品的描述708中的至少一个来生成推荐产品的推荐理由720。可选地,对于针对该推荐产品产生最大期望概率提升的问题,可以在生成推荐理由的过程中对由该问题的回答所指示的用户选择的属性给予较高的权重。

[0137] 图8示出了根据实施例的示例性聊天窗口800。

[0138] 在接收到用户提供的消息“帮我找个礼物”后,聊天机器人可以确定用户想要获得产品推荐,并且产品类别为“礼物”。接着,聊天机器人可以通过多轮会话向用户提供多个问题并接收用户的回答。这些问题可以是根据以上讨论的本公开的实施例来动态地依次确定的。每个问题都附加了选项,相应地,用户的回答包括对选项的明确选择。最后,聊天机器人向用户提供响应“根据你对问题6的选择‘河边’,我推荐你买钓鱼竿”。该响应包括推荐产品的产品信息“钓鱼竿”,其中,推荐产品可以是根据以上讨论的本公开的实施例来确定的。该响应还包括推荐理由“根据你对问题6的选择‘河边’”,其中,推荐理由可以是根据以上讨论的本公开的实施例来生成的,例如,基于对该推荐产品产生最大期望概率提升的问题6及对应回答来生成的。

[0139] 图9示出了根据实施例的示例性聊天窗口900。

[0140] 在接收到用户提供的消息“帮我推荐个电子商品作为礼物”后,聊天机器人可以确定用户想要获得产品推荐,并且产品类别为“礼物”和“电子商品”。接着,聊天机器人可以通过多轮会话向用户提供多个问题并接收用户的回答。最后,聊天机器人向用户提供响应“考虑到你喜欢‘跑步’,我推荐你买智能手环”。该响应包括推荐产品的产品信息“智能手环”,以及推荐理由“考虑到你喜欢‘跑步’”。推荐产品和推荐理由可以是根据以上讨论的本公开的实施例来确定的,其中,推荐理由可以是基于对该推荐产品产生最大期望概率提升的问题5及对应回答来生成的。

[0141] 通过图8与图9的比较可以看出,聊天机器人可以至少根据用户对于每个问题的回答而动态地确定下一个问题。

[0142] 图10示出了根据实施例的示例性聊天窗口1000。在图10的会话中,聊天机器人可

以向用户提供涉及酒店预定服务的产品推荐。问题1至问题2附加了选项以供用户选择。问题3至问题6没有附加选项,并且可以通过例如图5的步骤508对用户的采用自然语言语句的回答进行解析来确定用户所期望的产品属性。进一步地,可以通过解析来确定与用户的回答相对应的相关问题和对应回答,以用于确定下一个问题。

[0143] 图11示出了根据实施例的示例性聊天窗口1100。在图11的会话中,聊天机器人可以向用户提供涉及礼物的产品推荐。第一个问题和第二个问题附加有选项,并且用户的回答采用了自然语言语句的形式。可以通过例如图5的步骤506对用户的回答进行识别以确定用户所选择的选项,或者通过例如图5的步骤508对用户的回答进行解析以确定用户所期望的产品属性并且确定相关问题和对应回答。在聊天机器人提供的响应中包括了两个推荐理由。第一个推荐理由“我是基于关键词‘安静’和‘瑜伽/阅读’来推荐的”可以是基于对推荐产品产生最大期望概率提升的第一个问题和第二个问题及对应回答来生成的。第二个推荐理由“我希望这些礼物能帮助她在更安静的环境里享受生活”可以是通过推荐理由生成模型来生成的。

[0144] 图12示出了根据实施例的示例性聊天窗口1200。在图12的会话中,用户主动发送了采用自然语言语句的消息“我可以告诉你她喜欢安静过的地方、音乐和瑜伽”。可以通过例如图5的步骤508对该消息进行解析,以确定一组相关问题和对应回答,例如问题“她是否喜欢安静?”和对应回答“喜欢安静”,问题“她的爱好?”和对应回答“音乐”、“瑜伽”,等等。上述的“安静”、“音乐”、“瑜伽”等都可以被视为用户所期望的产品属性,并进而用于执行后续的产品排序、问题排序等。

[0145] 图13示出了根据实施例的用于在会话中提供解释性产品推荐的示例性方法1300的流程图。

[0146] 在1310处,可以提供与产品推荐相关联的至少一个问题。

[0147] 在1320处,可以接收针对所述至少一个问题的回答。

[0148] 在1330处,可以至少基于所述至少一个问题和所述回答来确定是否存在至少一个推荐产品。

[0149] 在1340处,可以响应于确定存在所述至少一个推荐产品,生成所述至少一个推荐产品的推荐理由。

[0150] 在1350处,可以提供包括所述至少一个推荐产品的产品信息和所述推荐理由的响应。

[0151] 在一种实施方式中,所述确定是否存在至少一个推荐产品可以包括:至少基于所述至少一个问题和所述回答来对多个候选产品执行产品排序;以及基于所述产品排序的结果来确定所述多个候选产品中是否存在满足预定条件的至少一个候选产品。

[0152] 在一种实施方式中,所述执行产品排序可以包括:至少基于所述至少一个问题和所述回答来更新候选产品评估状态,所述候选产品评估状态包括所述多个候选产品中每一个候选产品的期望概率。

[0153] 在一种实施方式中,所述预定条件可以包括:候选产品的期望概率高于阈值。

[0154] 在一种实施方式中,方法1300还可以包括:将所述至少一个问题和所述回答添加到所述会话的会话记录中,所述会话记录包括在所述会话中与所述产品推荐相关联的历史问题和历史回答。

[0155] 在一种实施方式中,所述生成推荐理由可以包括:确定所述会话记录中的每个历史问题针对所述至少一个推荐产品所产生的期望概率提升;选择产生最大期望概率提升的历史问题;以及至少基于所选择的历史问题和相应回答来生成所述推荐理由。

[0156] 在一种实施方式中,所述生成推荐理由可以包括:通过推荐理由生成模型,基于由所述历史回答所选择的属性、所述推荐产品的属性以及所述推荐产品的描述中的至少一个来生成所述推荐理由。

[0157] 在一种实施方式中,方法1300还可以包括:响应于确定不存在所述至少一个推荐产品,至少基于所述产品排序的结果来对多个候选问题执行问题排序;以及基于所述问题排序的结果来选择将要提供的下一个问题。

[0158] 在一种实施方式中,所述问题排序可以是通过基于熵的排序或者基于策略的增强学习排序来执行的。

[0159] 在一种实施方式中,所述多个候选问题可以是至少基于所述多个候选产品的属性来预先确定的。

[0160] 在一种实施方式中,所述至少一个问题可以包括一个或多个选项,并且所述回答可以包括针对所述一个或多个选项的选择。

[0161] 在一种实施方式中,所述回答可以包括自然语言语句。所述确定是否存在至少一个推荐产品可以包括:确定与所述自然语言语句相对应的一个或多个相关问题和对应回答;以及至少基于所述一个或多个相关问题和对应回答来确定是否存在所述至少一个推荐产品。

[0162] 应当理解,方法1300还可以包括根据上述本公开实施例的用于在会话中提供解释性产品推荐的任何步骤/过程。

[0163] 图14示出了根据实施例的用于在会话中提供解释性产品推荐的示例性装置1400。

[0164] 装置1400可以包括:问题提供模块1410,用于提供与产品推荐相关联的至少一个问题;回答接收模块1420,用于接收针对所述至少一个问题的回答;推荐产品确定模块1430,用于至少基于所述至少一个问题和所述回答来确定是否存在至少一个推荐产品;推荐理由生成模块1440,用于响应于确定存在所述至少一个推荐产品,生成所述至少一个推荐产品的推荐理由;以及响应提供模块1450,用于提供包括所述至少一个推荐产品的产品信息和所述推荐理由的响应。

[0165] 在一种实施方式中,所述推荐产品确定模块1430可以用于:至少基于所述至少一个问题和所述回答来对多个候选产品执行产品排序;以及基于所述产品排序的结果来确定所述多个候选产品中是否存在满足预定条件的至少一个候选产品。所述执行产品排序可以包括:至少基于所述至少一个问题和所述回答来更新候选产品评估状态,所述候选产品评估状态包括所述多个候选产品中每一个候选产品的期望概率。

[0166] 在一种实施方式中,装置1400还可以包括:会话记录添加模块,用于将所述至少一个问题和所述回答添加到所述会话的会话记录中,所述会话记录包括在所述会话中与所述产品推荐相关联的历史问题和历史回答。

[0167] 在一种实施方式中,所述推荐理由生成模块1440可以用于:确定所述会话记录中的每个历史问题针对所述至少一个推荐产品所产生的期望概率提升;选择产生最大期望概率提升的历史问题;以及至少基于所选择的历史问题和相应回答来生成所述推荐理由。

[0168] 在一种实施方式中,所述推荐理由生成模块1440可以用于:通过推荐理由生成模型,基于由所述历史回答所选择的属性、所述推荐产品的属性以及所述推荐产品的描述中的至少一个来生成所述推荐理由。

[0169] 在一种实施方式中,装置1400还可以包括问题选择模块,其用于:响应于确定不存在所述至少一个推荐产品,至少基于所述产品排序的结果来对多个候选问题执行问题排序;以及基于所述问题排序的结果来选择将要提供的下一个问题。

[0170] 此外,装置1400还可以包括根据上述本公开实施例的被配置用于在会话中提供解释性产品推荐的任何其它模块。

[0171] 图15示出了根据实施例的用于在会话中提供解释性产品推荐的示例性装置1500。

[0172] 装置1500可以包括至少一个处理器1510以及存储了计算机可执行指令的存储器1520。当执行所述计算机可执行指令时,处理器1510可以:提供与产品推荐相关联的至少一个问题;接收针对所述至少一个问题的回答;至少基于所述至少一个问题和所述回答来确定是否存在至少一个推荐产品;响应于确定存在所述至少一个推荐产品,生成所述至少一个推荐产品的推荐理由;以及提供包括所述至少一个推荐产品的产品信息和所述推荐理由的响应。此外,处理器1510还可以执行根据上述本公开实施例的用于在会话中提供解释性产品推荐的任何其它处理。

[0173] 本公开的实施例可以实施在非暂时性计算机可读介质中。该非暂时性计算机可读介质可以包括指令,当所述指令被执行时,使得一个或多个处理器执行根据上述本公开实施例的用于在会话中提供解释性产品推荐的方法的任何操作。

[0174] 应当理解,以上描述的方法中的所有操作都仅仅是示例性的,本公开并不限制于方法中的任何操作或这些操作的顺序,而是应当涵盖在相同或相似构思下的所有其它等同变换。

[0175] 还应当理解,以上描述的装置中的所有模块都可以通过各种方式来实施。这些模块可以被实施为硬件、软件、或其组合。此外,这些模块中的任何模块可以在功能上被进一步划分成子模块或组合在一起。

[0176] 已经结合各种装置和方法描述了处理器。这些处理器可以使用电子硬件、计算机软件或其任意组合来实施。这些处理器是实施为硬件还是软件将取决于具体的应用以及施加在系统上的总体设计约束。作为示例,本公开中给出的处理器、处理器的任意部分、或者处理器的任意组合可以实施为微处理器、微控制器、数字信号处理器(DSP)、现场可编程门阵列(FPGA)、可编程逻辑器件(PLD)、状态机、门逻辑、分立硬件电路、以及配置用于执行在本公开中描述的各种功能的其它适合的处理部件。本公开给出的处理器、处理器的任意部分、或者处理器的任意组合的功能可以实施为由微处理器、微控制器、DSP或其它适合的平台所执行的软件。

[0177] 软件应当被广泛地视为表示指令、指令集、代码、代码段、程序代码、程序、子程序、软件模块、应用、软件应用、软件包、例程、子例程、对象、运行线程、过程、函数等。软件可以驻留在计算机可读介质中。计算机可读介质可以包括例如存储器,存储器可以例如为磁性存储设备(如,硬盘、软盘、磁条)、光盘、智能卡、闪存设备、随机存取存储器(RAM)、只读存储器(ROM)、可编程ROM(PROM)、可擦除PROM(EPROM)、电可擦除PROM(EEPROM)、寄存器或者可移动盘。尽管在本公开给出的多个方面中将存储器示出为是与处理器分离的,但是存储器也

可以位于处理器内部(如,缓存或寄存器)。

[0178] 以上描述被提供用于使得本领域任何技术人员可以实施本文所描述的各个方面。这些方面的各种修改对于本领域技术人员是显而易见的,本文限定的一般性原理可以应用于其它方面。因此,权利要求并非旨在被局限于本文示出的方面。关于本领域技术人员已知或即将获知的、对本公开所描述各个方面的元素的所有结构和功能上的等同变换,都将被包含到本文中,并且旨在由权利要求所覆盖。

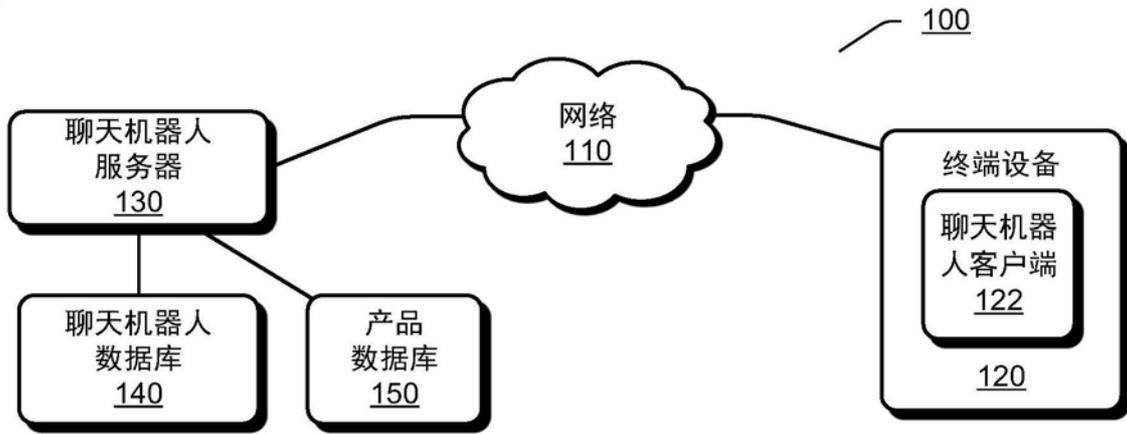


图1

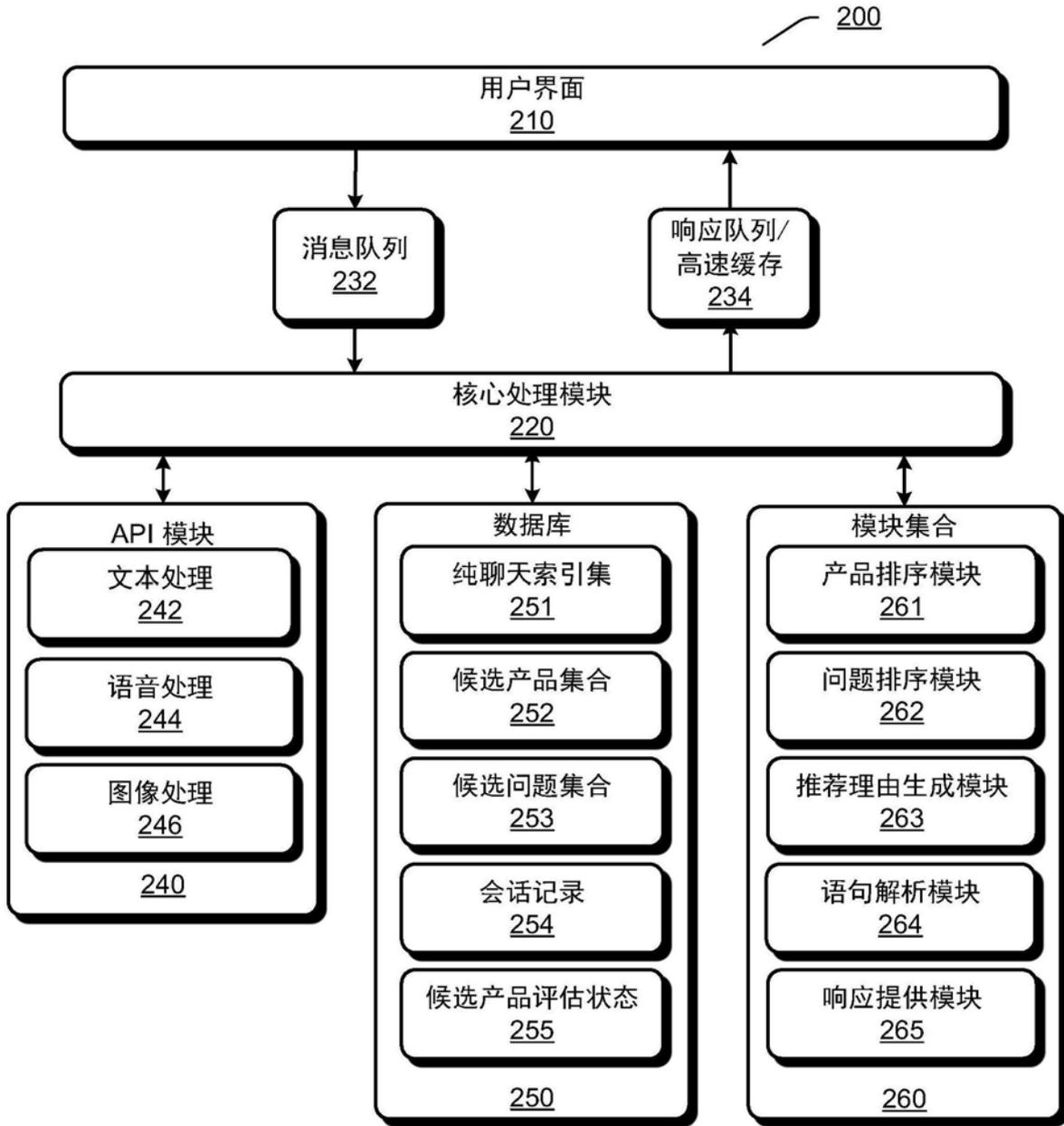


图2

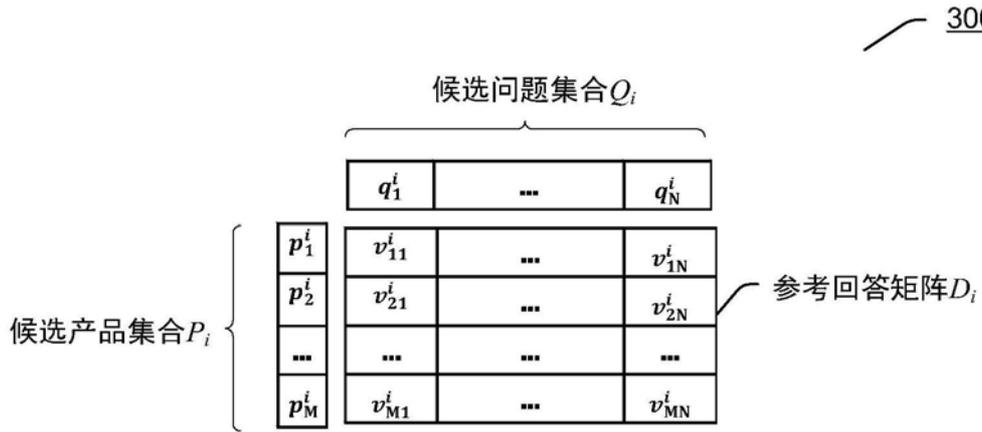


图3

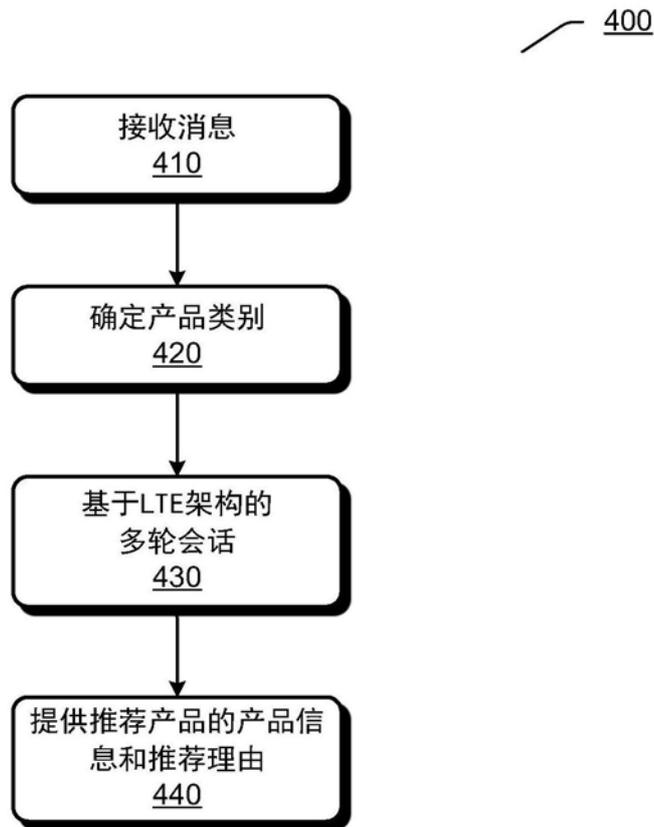


图4

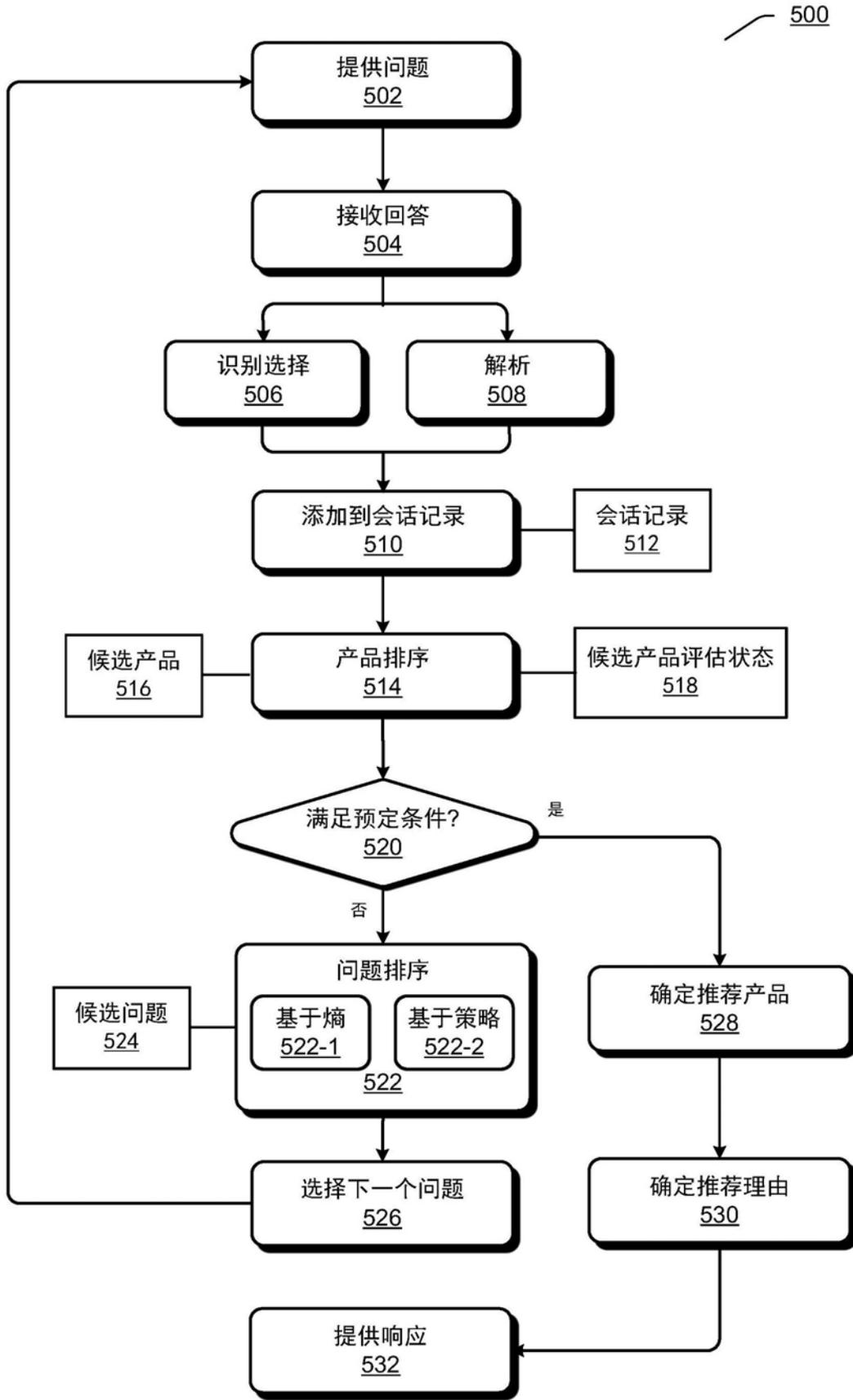


图5

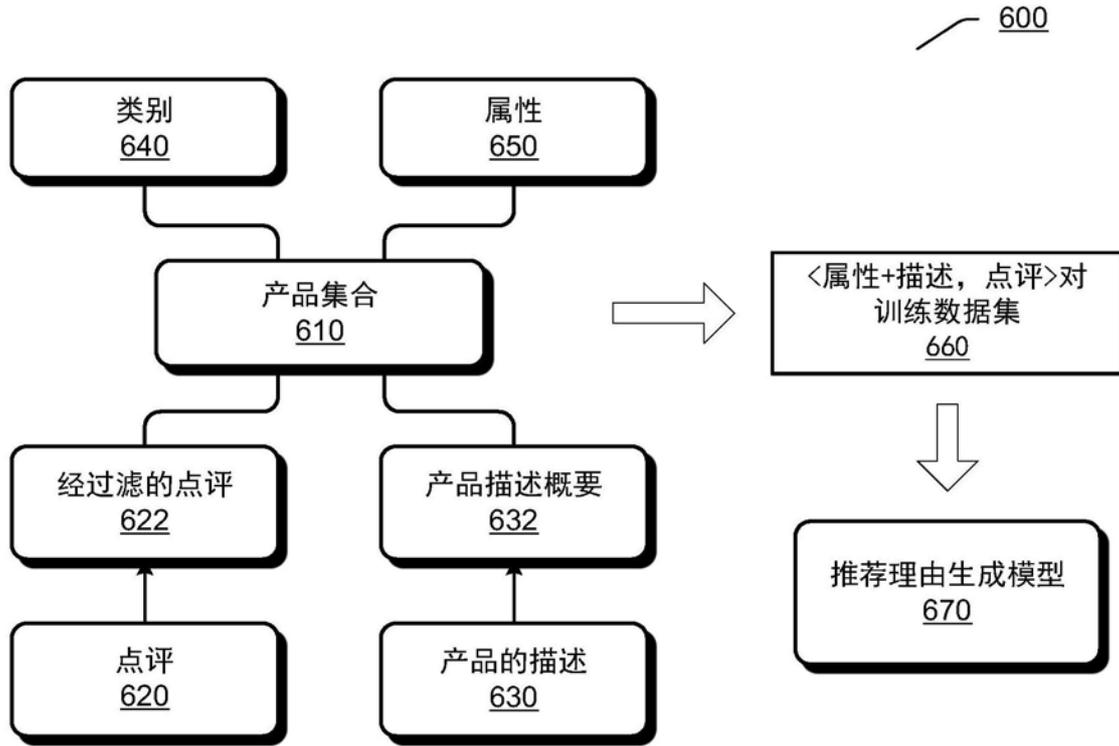


图6

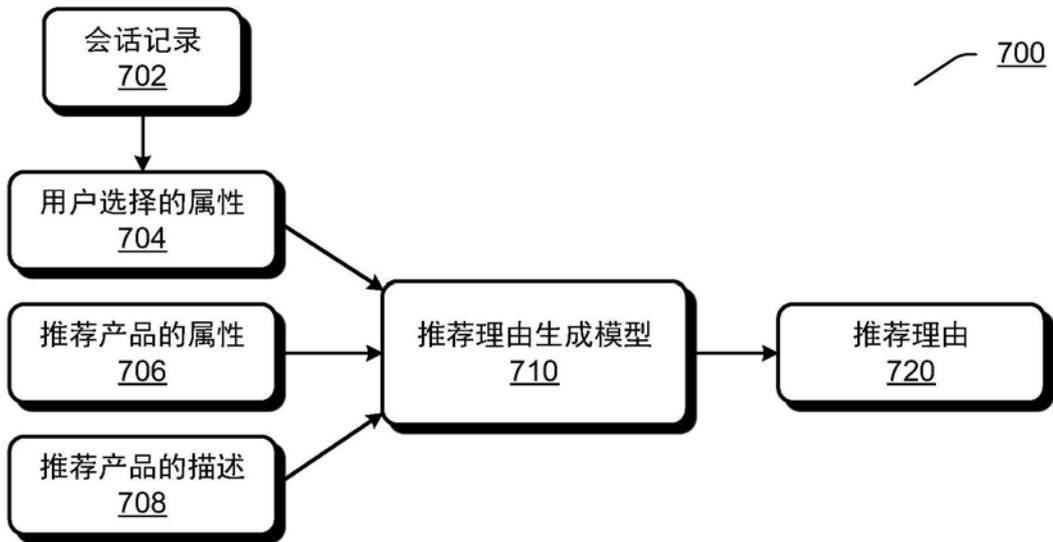


图7



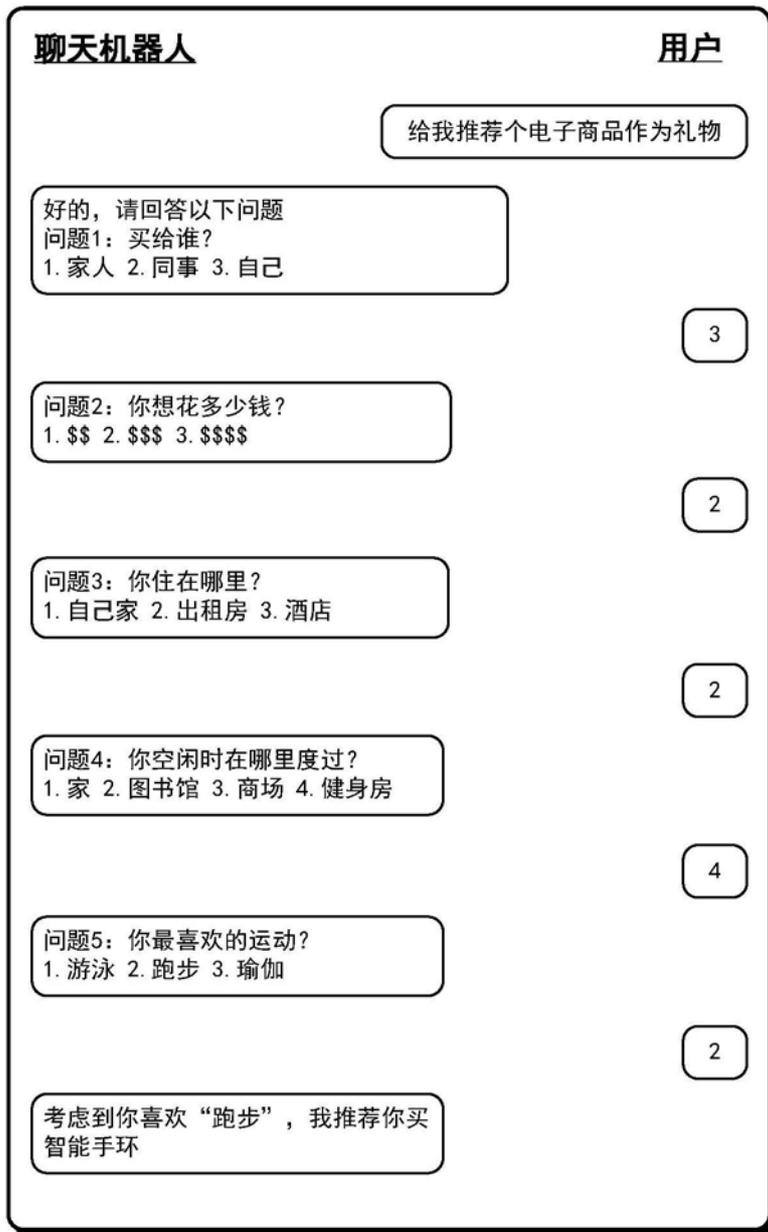


图9

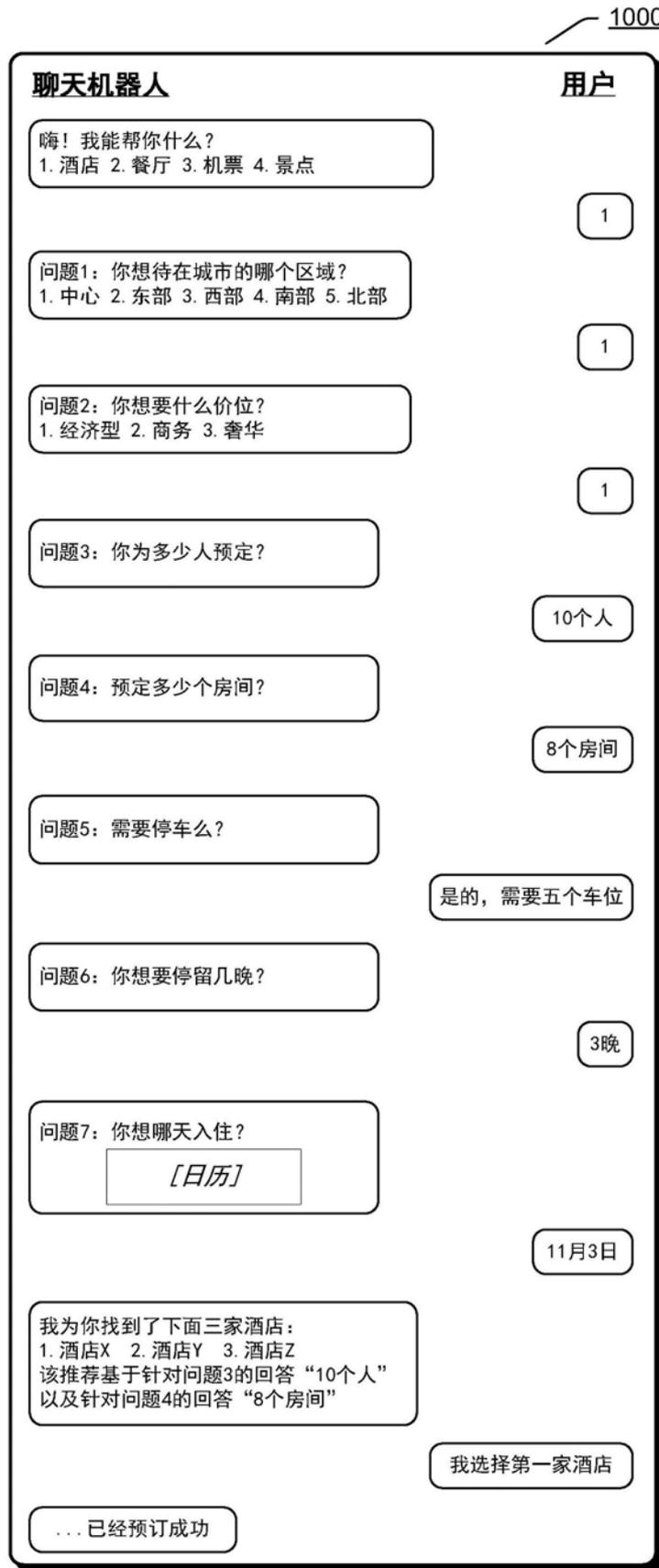


图10

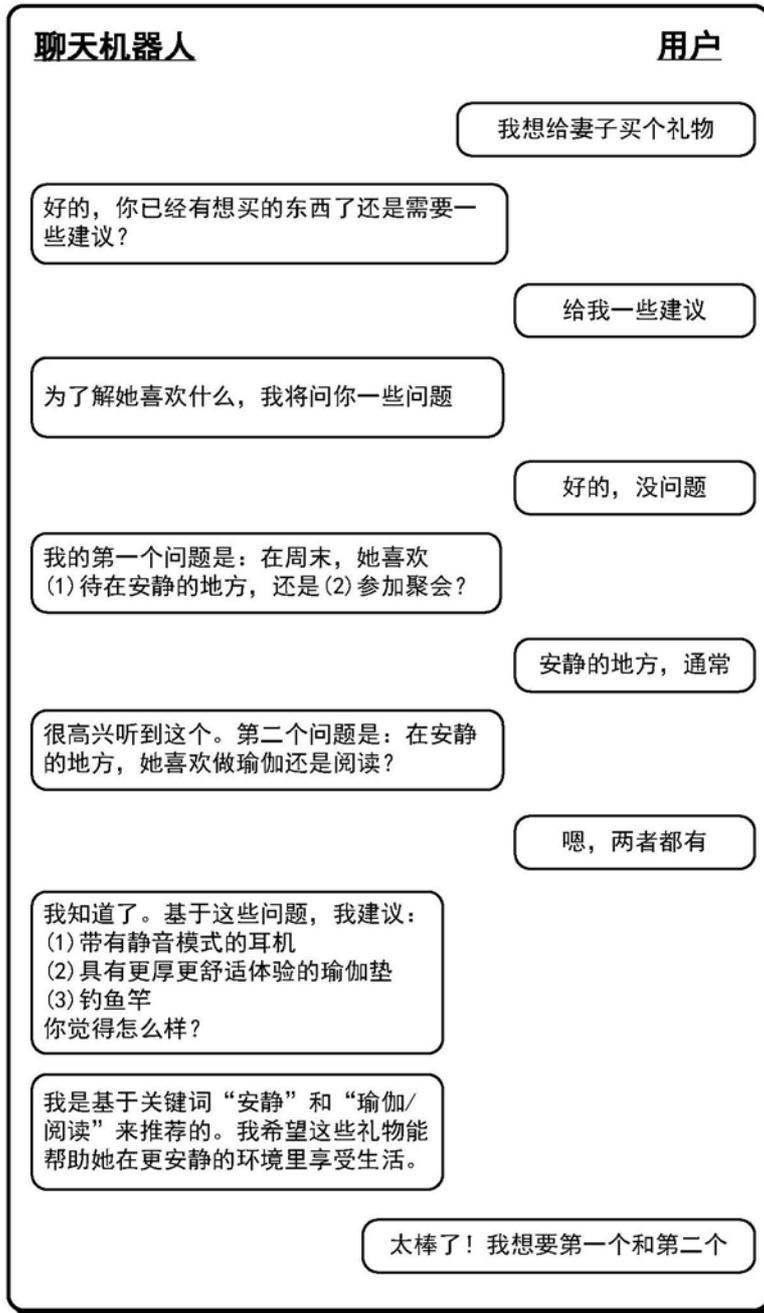


图11

1200

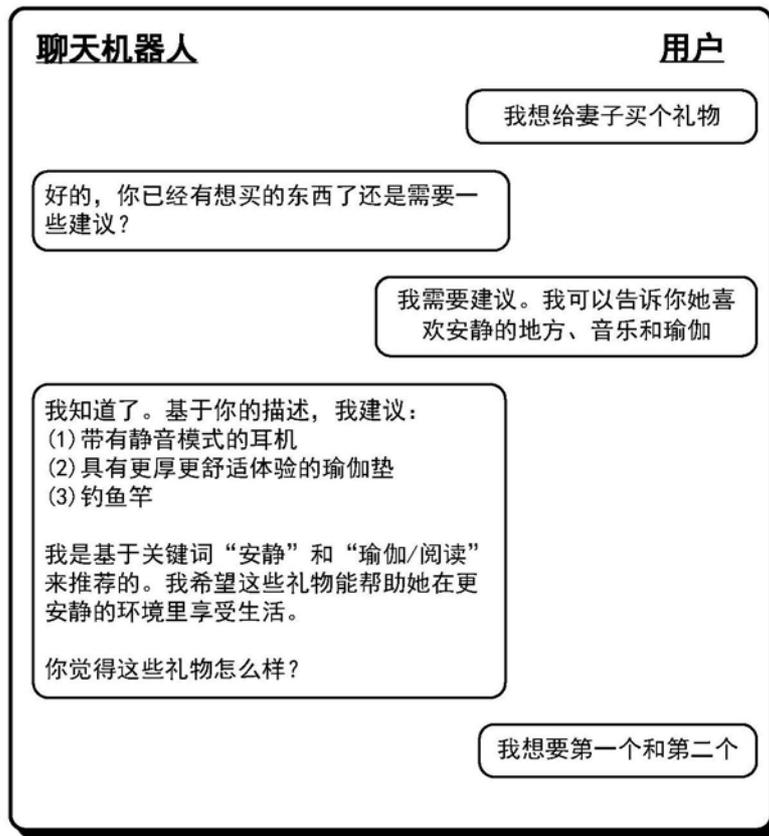


图12

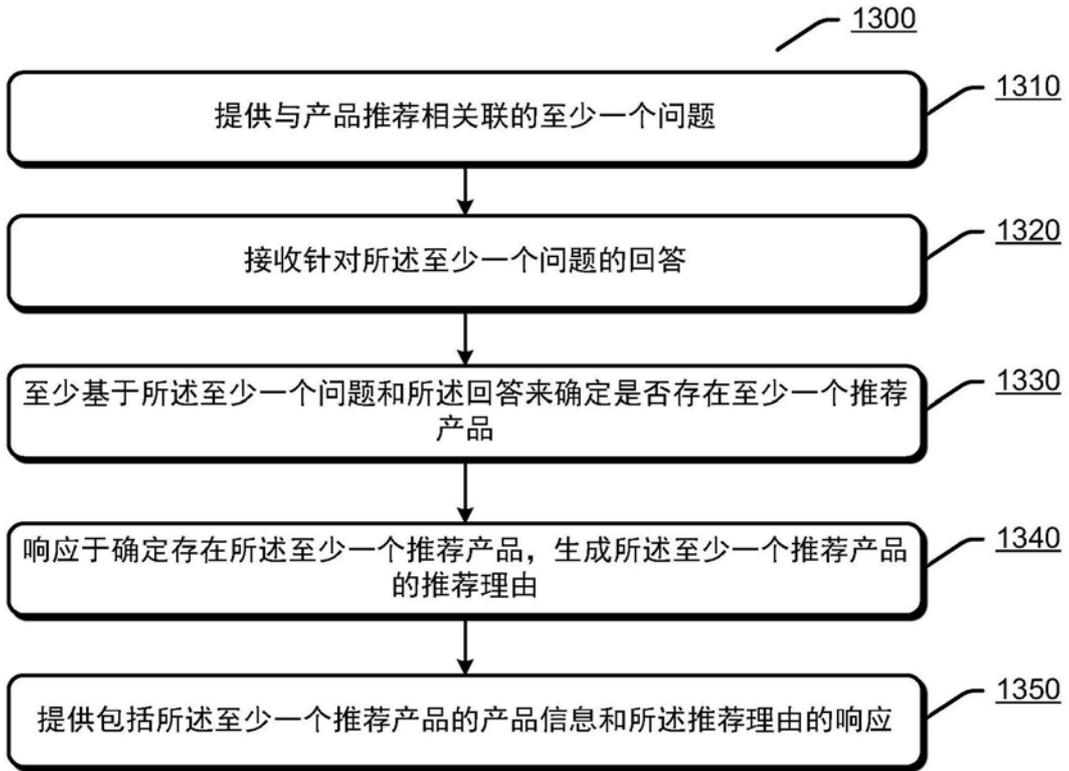


图13



图14



图15